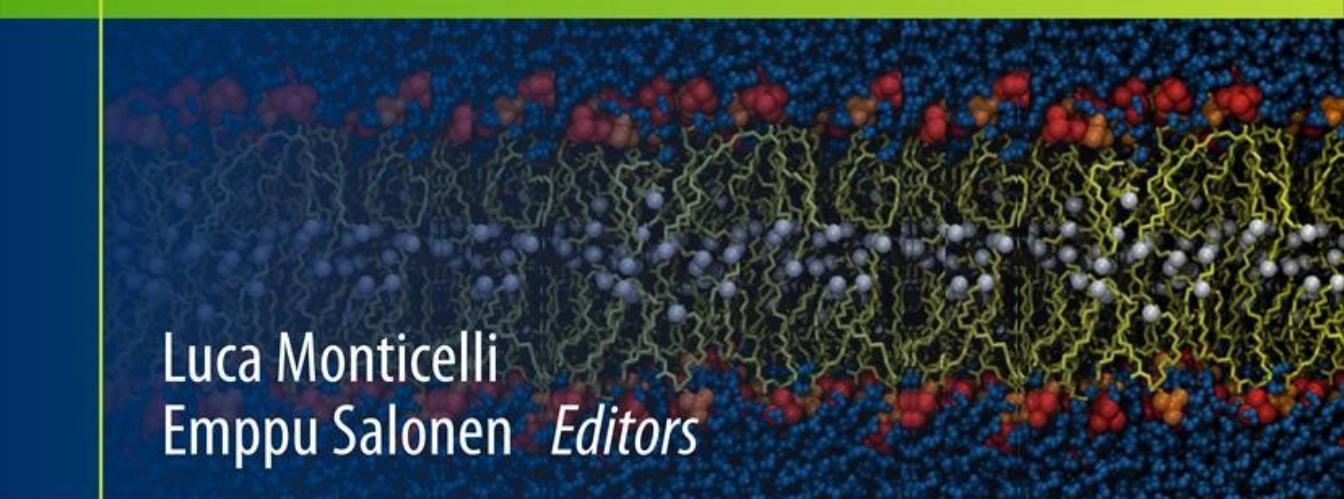


Methods in  
Molecular Biology 924

Springer Protocols



Luca Monticelli  
Emppu Salonen *Editors*

# Biomolecular Simulations

Methods and Protocols

 Humana Press

# METHODS IN MOLECULAR BIOLOGY™

*Series Editor*

John M. Walker

School of Life Sciences

University of Hertfordshire

Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:  
<http://www.springer.com/series/7651>



# **Biomolecular Simulations**

## **Methods and Protocols**

Edited by

**Luca Monticelli**

*INSERM, DSIMB, UMR-S 665, Paris, France*

**Emppu Salonen**

*Department of Applied Physics, School of Science,  
Aalto University, Espoo, Finland*

*Editors*

Luca Monticelli, Ph.D.  
INSERM, DSIMB, UMR-S 665  
Paris, France

Université Paris Diderot  
Sorbonne Paris Cité, UMR-S 665  
Paris, France

Emppu Salonen  
Department of Applied Physics  
School of Science  
Aalto University  
Espoo, Finland

ISSN 1064-3745                   ISSN 1940-6029 (electronic)  
ISBN 978-1-62703-016-8       ISBN 978-1-62703-017-5 (eBook)  
DOI 10.1007/978-1-62703-017-5  
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012946729

© Springer Science+Business Media New York 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

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

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Humana Press is a brand of Springer  
Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

---

## Preface

The field of molecular simulations has evolved, during the past 40 years, from picosecond studies of isolated macromolecules in vacuum to studies of complex, chemically heterogeneous biomolecular systems consisting of millions of atoms, with the simulation time scales spanning up to milliseconds. Much of the impressive progress in biomolecular simulations has been simply due to more powerful computers. However, tackling the sheer complexity of macromolecular systems of biological interest has never relied on Moore's law alone. Among the most important factors contributing to such enormous progress are the development of faster simulation algorithms and entirely novel approaches to molecular modelling, such as multiscale methods employing different levels of theory for different regions of the system studied. As the number of researchers in the field has increased, limitations inherent in each methodology have also been described, creating awareness of the pros and cons of different techniques, and contributing significantly to the progress in the field. Quantum mechanical (QM) calculations, while in principle providing a rigorous description of the molecules of interest, necessitate the use of some approximations in the underlying theory. Practical methodological solutions within these frameworks (for example, computationally efficient basis sets) have been crucial in QM calculations becoming a powerful tool in biomolecular modelling. Classical simulations employing empirical force fields have steadily become more versatile and accurate, thanks to more sophisticated potential energy functions even including explicit functions for hydrogen bonding and electronic polarization, and to more accurate parameterizations. The general level of accuracy of the classical simulations, combined with their speed, has made them perhaps the most widely used method in molecular modelling of biological systems. Coarse-graining the atomistic description of the molecules of interest has simplified the computations immensely while still retaining some of the key physicochemical properties of the system studied.

One major challenge in modelling biological systems is the very large span of length and time scales involved. Depending on the problem at hand, researchers look at motions occurring on time scales from femtoseconds ( $10^{-15}$  s) to hours, and covering distances from sub-atomic scale to cell size. Another challenge is the great chemical complexity and heterogeneity of biological systems. Even the smallest biological sample contains a wide variety of molecules. Computational studies – just as simple *in vitro* experiments – require much simplification. It is assumed that by understanding the properties and behaviour of simple model systems one can learn about the properties and behaviour of real, much more complex systems. When such assumption and other simplifications are justified, simulations can make a significant contribution to understanding biological systems' structure and functioning.

Excellent books are available to students and researchers who venture into the field of molecular modelling, covering both the basic foundations as well as more specialized aspects. With the present volume, we aim to present the foundations of well-established simulation techniques together with some of the recent developments in methods and practices. The latter rarely find ample coverage in traditional textbooks, but are being used more and more by researchers in the biological field. We also aim at giving

some practical examples on how to carry out simulations of some particular systems of great biological interest, and particularly systems including biological macromolecules.

The book consists of three sections, with the division based on the predominant classes of methods used in modelling at various length and time scales. The covered methodologies include electronic structure calculations, classical molecular dynamics simulations and coarse-grained techniques. Each section comprises a methodological and an application part. The former provides an introduction to the basic physics and chemistry underlying the computational models, and focuses particularly on recent developments. The application part illustrates examples on the four main classes of biological macromolecules, namely proteins, nucleic acids, lipids, and carbohydrates. This subdivision is in line with the traditions of the *Methods in Molecular Biology* series, with an introductory overview, theoretical foundations, and good practices of the methodologies used, followed by chapters illustrating their practical application in studies of biological macromolecules.

The target audience of the book includes both graduate students and researchers interested in computational modelling of biomolecular systems in physics, chemistry, and biology. The structuring of the different sections has been made so that after reading the first methodological chapters in the section in question, a non-expert reader can understand and appreciate the following application-oriented chapters. An expert on a given methodology can, in turn, jump directly to the chapters on state-of-the-art applications of the methodology covered. We hope that the readers will find this structure of the book useful and easy to approach.

The general structure of the book is as follows. The QM section contains reviews covering the most central contemporary methodologies of biomolecular modelling. The section begins with an overview of different electronic structure calculation methods with an emphasis on methodological issues related to the investigation of biological systems. The so-called ab initio molecular dynamics methodology for dynamic electronic structure calculations is then introduced. The remaining three chapters in the section address a more practical side of QM calculations of biological systems, that is, the hybrid QM and molecular mechanics methodology (QM/MM). These chapters feature a particular emphasis on studies of proteins, reflecting the most common domain of application of QM/MM methods.

In the section on atomistic simulations we cover the basic ideas and the most common techniques (molecular dynamics and classical force fields), and we also give space to a few recent developments that gained more and more importance in recent years: enhanced sampling algorithms, allowing for crossing energetic barriers and speed up the sampling; free energy calculations, that were rarely found in the literature only a decade ago due to the high computational cost, but are now accessible to a wide audience; polarizable force fields, which aim to increase the accuracy of classical empirical simulations with a tolerable computational overhead, and have become increasingly popular during recent years.

Coarse-graining has been perhaps the fastest evolving area in biomolecular modelling over the past few years, with new techniques and new force fields published monthly. Coarse-graining rarely finds any space in traditional textbooks, but it is becoming attractive for a very wide public. The diversity in the techniques developed is such that a thorough review would require an entire volume. The book covers some of the most exciting recent developments in the area, with applications to lipid membranes and membrane proteins in particular.

The increasing speed of molecular simulations hardware and software, and the development of force fields and methodologies make it possible to describe increasingly

complex biological systems and processes. Progress in several directions can be foreseen in the near future, with the development of faster and more accurate methodologies for electronic structure calculations, more refined classical force fields, and improved coarse-graining techniques. Multiscale modelling, that currently is one of the most challenging problems in molecular simulations, will require substantial theoretical and methodological development. Yet, the great interest in combining descriptions at different length scales bears expectations of a significant growth in this area.

Paris, France  
Espoo, Finland

Luca Monticelli  
Emppu Salonen



---

# Contents

Preface .....	v
Contributors .....	xi

## PART I QUANTUM MECHANICS CALCULATIONS

1 <i>Ab Initio</i> , Density Functional Theory, and Semi-Empirical Calculations .....	3
<i>Mikael P. Johansson, Ville R.I. Kaila, and Dage Sundholm</i>	
2 Ab Initio Molecular Dynamics .....	29
<i>Kari Laasonen</i>	
3 Introduction to QM/MM Simulations .....	43
<i>Gerrit Groenhof</i>	
4 Computational Enzymology .....	67
<i>Alessio Lodola and Adrian J. Mulholland</i>	
5 QM and QM/MM Simulations of Proteins .....	91
<i>Thomas Steinbrecher and Marcus Elstner</i>	

## PART II CLASSICAL MECHANICS: ATOMISTIC SIMULATIONS

6 Classical Molecular Dynamics in a Nutshell .....	127
<i>Susanna Hug</i>	
7 Enhanced Sampling Algorithms .....	153
<i>Ayori Mitsutake, Yoshiharu Mori, and Yuko Okamoto</i>	
8 Force Fields for Classical Molecular Dynamics .....	197
<i>Luca Monticelli and D. Peter Tieleman</i>	
9 Polarizable Force Fields .....	215
<i>Hanne S. Antila and Emppu Salonen</i>	
10 Electrostatics Interactions in Classical Simulations .....	243
<i>G. Andrés Cisneros, Volodymyr Babin, and Celeste Sagui</i>	
11 An Introduction to Best Practices in Free Energy Calculations .....	271
<i>Michael R. Shirts and David L. Mobley</i>	
12 Recipes for Free Energy Calculations in Biomolecular Systems .....	313
<i>Mahmoud Moradi, Volodymyr Babin, Celeste Sagui, and Christopher Roland</i>	
13 Molecular Docking Methodologies .....	339
<i>Andrea Bortolato, Marco Fanton, Jonathan S. Mason, and Stefano Moro</i>	
14 Simulation Studies of the Mechanism of Membrane Transporters .....	361
<i>Giray Enkavi, Jing Li, Paween Mahinthichaichan, Po-Chao Wen, Zhijian Huang, Saber A. Shaikh, and Emad Tajkhorshid</i>	
15 Molecular Dynamics Simulations of Lipid Bilayers: Simple Recipe of How to Do It .....	407
<i>Hector Martínez-Seara and Tomasz Róg</i>	

16	Simulations of Lipid Monolayers..... <i>Svetlana Baoukina and D. Peter Tieleman</i>	431
17	Simulating DNA by Molecular Dynamics: Aims, Methods, and Validation..... <i>Nicolas Foloppe, Marc Guérout, and Brigitte Hartmann</i>	445
18	Simulation of Carbohydrates, from Molecular Docking to Dynamics in Water ... <i>Nicolas Sapay, Alessandra Nurisso, and Anne Imbert</i>	469

### PART III MESOSCOPIC SIMULATIONS AND COARSE-GRAINED MODELS

19	Systematic Methods for Structurally Consistent Coarse-Grained Models .....	487 <i>W.G. Noid</i>
20	The Martini Coarse-Grained Force Field..... <i>Xavier Periole and Siewert-Jan Marrink</i>	533
21	Multiscale Molecular Modeling..... <i>Matej Praprotnik and Luigi Delle Site</i>	567
22	Coarse-Grained Models for Protein Folding and Aggregation .....	585 <i>Philippe Derreumaux</i>
23	Elastic Network Models: Theoretical and Empirical Foundations .....	601 <i>Yves-Henri Sanejouand</i>
24	An Introduction to Dissipative Particle Dynamics .....	617 <i>Zhong-Yuan Lu and Yong-Lei Wang</i>
25	Multiscale Molecular Dynamics Simulations of Membrane Proteins .....	635 <i>Syma Khalid and Peter J. Bond</i>
26	Vesicles and Vesicle Fusion: Coarse-Grained Simulations..... <i>Julian C. Shillcock</i>	659
	<i>Index</i> .....	699

---

## Contributors

- HANNE ANTILA • *Department of Chemistry, Aalto University, Espoo, Finland*
- VOLODYMYR BABIN • *Department of Physics, Center for High Performance Simulations (CHIPS), North Carolina State University, Raleigh, NC, USA*
- SVETLANA BAOUKINA • *Department of Biological Sciences, University of Calgary, Calgary, AB, Canada*
- PETER J. BOND • *Department of Chemistry, University of Cambridge, Cambridge, UK*
- ANDREA BORTOLATO • *Heptares Therapeutics Ltd, Hertfordshire, UK*
- G. ANDRÉS CISNEROS • *Department of Chemistry, Wayne State University, Detroit, MI, USA*
- PHILIPPE DERREUMAUX • *Laboratory of Theoretical Biochemistry, Paris 7 University, Paris, France*
- MARCUS ELSTNER • *Institute of Physical Chemistry, Karlsruhe Institute of Technology, Karlsruhe, Germany*
- GIRAY ENKAVI • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
- MARCO FANTON • *Dipartimento di Scienze Farmaceutiche, Università di Padova, Padova, Italy*
- NICOLAS FOLOPPE • *Vernalis (R&D) Ltd, Granta Park, Abingdon, Cambridge, UK*
- GERRIT GROENHOF • *Max Planck Institute for Biophysical Chemistry, Göttingen, Germany*
- MARC GUÉROULT • *INSERM, Paris, France Institut de Biologie Physico-Chimique, CNRS, Paris, France*
- BRIGITTE HARTMANN • *Laboratoire de biologie et pharmacologie appliquée (LBPA), CNRS, Paris, France*
- ZHIJIAN HUANG • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
- SUSANNA HUG • *Department of Applied Mathematics, University of Western Ontario, London, ON, Canada*
- ANNE IMBERTY • *Centre de Recherches sur les Macromolécules Végétales (CERMAV-CNRS), Grenoble cedex 9, France*
- MIKAEL P. JOHANSSO • *Department of Chemistry, University of Helsinki, Helsinki, Finland*
- VILLE R.I. KAILA • *Department of Chemistry, University of Helsinki, Helsinki, Finland*
- SYMA KHALID • *School of Chemistry, University of Southampton, Southampton, UK*
- KARI LAASONEN • *Department of Chemistry, Aalto University, Espoo, Finland*
- ALESSIO LODOLA • *Dipartimento Farmaceutico, Università degli Studi di Parma, Parma, Italy*
- JING LI • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
- ZHONG-YUAN LU • *State Key Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun, China*

- PAWEEN MAHINTHICHAICHAN • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
- SIEWERT-JAN MARRINK • *Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands*
- HECTOR MARTINEZ-SEARA • *Department of Physics, Tampere University of Technology, Tampere, Finland*
- JONATHAN S. MASON • *Heptares Therapeutics Ltd, Hertfordshire, UKLundbeck A/S, Copenhagen, Denmark*
- AYORI MITSUTAKE • *Department of Physics, Keio University, Yokohama, Kanagawa, Japan*
- DAVID L. MOBLEY • *Department of Pharmaceutical Sciences, University of California, Irvine, CA, USA*
- LUCA MONTICELLI • *INSERM, UMR-S665, Paris, France Univ Paris Diderot, Sorbonne Paris Cité, UMR-S 665, Paris, France Institut National de la Transfusion Sanguine (INTS), Paris, France*
- MAHMOUD MORADI • *Department of Physics, Center for High Performance Simulations (CHIPS), North Carolina State University, Raleigh, NC, USA*
- STEFANO MORO • *Dipartimento di Scienze Farmaceutiche, Università di Padova, Padova, Italy*
- YOSHIHARU MORI • *Department of Physics, Nagoya University, Nagoya, Aichi, Japan*
- ADRIAN J. MULHOLLAND • *Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, UK*
- WILLIAM G. NOID • *Department of Chemistry, The Pennsylvania State University, University Park, PA, USA*
- ALESSANDRA NURISSO • *Centre de Recherches sur les Macromolécules Végétales (CERMAV-CNRS), Grenoble cedex 9, France*
- YUKO OKAMOTO • *Department of Physics, Nagoya University, Nagoya, Aichi, Japan*
- XAVIER PERIOLE • *Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands*
- MATEJ PRAPROTKNIK • *National Institute of Chemistry, Ljubljana, Slovenia*
- TOMASZ RÓG • *Department of Physics, Tampere University of Technology, Tampere, Finland*
- CHRISTOPHER ROLAND • *Department of Physics, Center for High Performance Simulations (CHIPS), North Carolina State University, Raleigh, NC, USA*
- CELESTE SAGUI • *Department of Physics, Center for High Performance Simulations (CHIPS), North Carolina State University, Raleigh, NC, USA*
- EMPPU SALONEN • *Department of Applied Physics, Aalto University, Espoo, Finland*
- YVES-HENRI SANEJOUAND • *University of Nantes, Nantes, France*
- NICOLAS SAPAY • *Centre de Recherches sur les Macromolécules Végétales (CERMAV-CNRS), Grenoble cedex 9, France*
- SAHER A. SHAIKH • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
- JULIAN C. SHILLCOCK • *École polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland*

MICHAEL R. SHIRTS • *Department of Chemical Engineering, University of Virginia, Charlottesville, VA, USA*

LUIGI DELLE SITE • *Department of Mathematics and Computer Science, Freie Universität Berlin, Berlin, Germany*

THOMAS STEINBRECHER • *Institute of Physical Chemistry, Karlsruhe Institute of Technology, Karlsruhe, Germany*

DAGE SUNDHOLM • *Department of Chemistry, University of Helsinki, Helsinki, Finland*

EMAD TAJKHORSHID • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

D. PETER TIELEMAN • *Department of Biological Sciences, University of Calgary, Calgary, AB, Canada*

YONG-LEI WANG • *State Key Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun, China*

Po-CHAO WEN • *Department of Biochemistry, Center for Biophysics and Computational Biology, College of Medicine, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*



# **Part I**

## **Quantum Mechanics Calculations**

# Chapter 1

## ***Ab Initio, Density Functional Theory, and Semi-Empirical Calculations***

**Mikael P. Johansson, Ville R.I. Kaila, and Dage Sundholm**

### **Abstract**

This chapter introduces the theory and applications of commonly used methods of electronic structure calculation, with particular emphasis on methods applicable for modelling biomolecular systems. This chapter is sectioned as follows. We start by presenting *ab initio* methods, followed by a treatment of density functional theory (DFT) and some recent advances in semi-empirical methods. Treatment of excited states as well as basis sets are also presented.

**Key words:** Ab initio methods, Density functional theory, Wave-function methods, First-principles approach, Coupled-cluster methods, Second-order perturbation theory, Electronic excited states, Basis sets, Semi-empirical methods, Time-dependent density functional theory, Coupled-cluster linear response methods

---

### **1. Introduction**

Application of quantum chemical electronic structure calculations has become feasible in biomolecular modelling during recent decades due to development of new methods and significant increase in computing power. Development of density functionals since the 1980s has had a considerable impact on modelling biomolecular systems, and it is now possible to obtain results close to chemical accuracy for systems comprising several hundred atoms. Moreover, as we shall discuss here, application of *ab initio* methods on ever-larger systems has also become feasible during recent years, and can now in many cases be considered as an alternative to DFT modelling. The aim of this chapter is to briefly discuss quantum mechanical (QM) approaches that can be employed in computational studies of biomolecular systems and reactions. Due to the large variety of quantum chemical methods, we offer merely an overview, a glimpse, of available methods and protocols.

---

## 2. *Ab Initio* Methods

The starting point for most correlated *ab initio*, or wave-function (WF) calculations is the Hartree–Fock self-consistent field (HF-SCF) or mean-field approximation (1–3). The general assumption in the HF model is that the wave function can be expressed using one Slater determinant (SD) (4) or more generally, a linear sum of SDs with fixed coefficients called configuration state functions (CSFs). At the HF-SCF level, the one-particle functions (orbitals) are optimised by minimising the total HF electronic energy. Up front, we discourage large-scale applications at HF level, as the accuracy of the method is significantly lower compared to what can be obtained with density functional theory (DFT) which has the same, or even lower computational cost. The reason is of course that accounting for electron correlation is very important also when modelling biomolecules.

In WF theory, electron correlation is considered as an addition to the uncorrelated HF reference solution. The occupied and virtual HF orbitals ( $\phi_i$  and  $\phi_a$ ) and the corresponding orbital energies ( $\varepsilon_i$  and  $\varepsilon_a$ ) are used for estimating the correlation energy, defined as (5),

$$E_{\text{correlation}} = E_{\text{exact}} - E_{\text{HF}}. \quad (1)$$

$E_{\text{exact}}$  is the energy obtained by solving the Schrödinger equation in the employed basis set, and  $E_{\text{HF}}$  is the total energy of the Hartree–Fock calculation obtained with the same basis set.

In *ab initio* correlation calculations, the wave function is constructed as a linear combination of the SDs or CSFs that can be constructed by permuting the occupation numbers of the HF orbitals. The *ab initio* correlated wave functions can be written as,

$$\Psi = \sum_I C_I \Psi_I, \quad (2)$$

where  $C_I$  is the expansion coefficients of the SDs or CSFs ( $\Psi_I$ ). The difference between *ab initio* correlation approaches lies in the way the expansion coefficients are constructed and how large fraction of the complete set of SD's is considered in the calculation. The selected SD's to expand the wave function are obtained by replacing orbitals that are occupied in the HF wave function with unoccupied ones. In general, the replacement is done systematically by promoting one, two, three, etc., electrons from the occupied HF orbitals to the virtual space, corresponding to single (S), double (D), triple (T), etc., excitations. When all SD's are taken into account, the full configuration interaction (CI) model is obtained (6), the solutions of which correspond to the exact Schrödinger wave functions in the employed basis set.

At the CI level, the total electronic energy is minimised with respect to the state parameters  $C_I$ . In the related multi-configuration self-consistent-field (MCSCF) calculations, the  $C_I$  coefficients and the occupied orbitals are optimised until the energy minimum is reached. The non-integer orbital occupation numbers are calculated using the  $C_I$  coefficients (7). The MCSCF calculations are time-consuming, which restricts the size of the orbital space that can be afforded. The remaining correlation effects omitted in the MCSCF calculation can be estimated by employing second-order perturbation theory (PT2) or by using low-order CI corrections, called multi-reference CI (MRCI) (8). The most popular combination of MCSCF and perturbation theory is complete-active-space SCF calculations (CASSCF) (9) and second-order perturbation theory (CASPT2) originally developed by Roos et al. and implemented by others in a variety of versions (10). These state-of-the-art approaches are far from black-box methods, due to the difficulty in choosing the active space. This is currently limited to about 15 electrons in 15 orbitals due to the size of the CI calculations. As there are already ten d-electrons in transition metals, treating bioinorganic complexes with this method is very challenging. Despite this, the CASPT2 approach has been applied to treat excited states of many biomolecular systems (11–13). In addition, the electronic structure of porphyrin model systems (14, 15) and the oxoheme (16) has recently been treated with the CASPT2 methodology.

Before performing large-scale CASPT2 calculations, a thorough training in these advanced computational approaches is recommended. For this purpose, we highlight the European Summer School in Quantum Chemistry. The theory behind MCSCF perturbation theory calculations will not be further discussed here; comprehensive reviews have recently been published elsewhere (17, 18, 19).

At coupled-cluster (CC) levels of theory (20), an exponential ansatz is used for the wave function ( $\Psi$ ),

$$\Psi = e^{\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots} \Psi \quad (3)$$

where  $\Psi$  is the HF reference SD and  $\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots$  are the cluster operators containing S, D, T, etc. excitation operators and the corresponding cluster amplitudes, which are obtained by solving the CC equations (7). The cluster operators containing, for example, the single and double excitation amplitudes are given by,

$$\hat{T}_1 = \sum_{ia} t_i^a a_a^\dagger a_i \quad \hat{T}_2 = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_i a_j \quad (4)$$

where  $t_i^a$  and  $t_{ij}^{ab}$  are the single and double excitation amplitudes, respectively. The SDs considered in the CC calculations are selected by the creation and annihilation operators.  $a_a^\dagger$  creates an electron in the virtual orbital  $a$ , and  $a_i$  annihilates the electron in the  $i$ :th occupied orbital. The computational requirements of CI and CC

calculations are comparable. High-order CI and CC calculations, thus those considering SD's obtained by T, Q, etc. replacements from the HF reference are computationally very demanding. They are so far not applicable on actual biomolecular systems as they are limited to systems with <20 atoms (21). Truncated CI calculations involving S, D, T, etc. replacements are not size extensive (22, 23), a flaw which has drastically diminished their popularity.

The simplest *ab initio* approaches considering electron correlation effects are the second-order Møller–Plesset perturbation theory (MP2) (24) and the CC approximate singles and doubles (CC2) methods (25). The MP2 and CC2 expressions can be derived using several approaches, but it is probably most convenient to relate them as special cases to the coupled-cluster singles and doubles (CCSD) model. The MP2 method has fixed double excitation amplitudes and vanishing single excitation amplitudes. The MP2 double excitation amplitudes can be expressed using the orbital energies of the HF calculations and two-electron interaction energies as,

$$t_{ij}^{ab} = \frac{\langle ab||ij\rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \quad (5)$$

where  $\langle ab||ij\rangle$  denotes the antisymmetrised Coulomb integrals  $(ai|bj) - (aj|bi)$ , given in the Mulliken notation. The MP2 energy expression then becomes,

$$E_{MP2} = -\frac{1}{4} \sum_{ijab} t_{ij}^{ab} \langle ab||ij\rangle \quad (6)$$

The MP2 and CC2 calculations formally scale as  $N^5$ , where  $N$  is the size of the basis set. CC2 is an iterative method, where the single and double excitation amplitudes are obtained by solving the CC2 equations. The CC2 expression for the double excitation amplitudes is structurally identical to the one for the MP2 amplitudes. However, in the CC2 expression, the Coulomb integrals are transformed using the single excitation amplitudes. The CC2 double excitation amplitudes can be recalculated using given single excitation amplitudes, two-electron integrals, and orbital energies when needed. The CC2 double excitation amplitudes are therefore not stored rendering calculations on large molecules feasible (25, 26). The obtained CC2 energies are generally of the same quality as the MP2 ones, whereas the CC2 computational costs are larger. For molecules where correlation effects are small, the accuracy of ground-state structures and vibrational frequencies obtained in the MP2 and CC2 calculations are on average the same (27). However, for systems with large correlation effects the CC2 bond lengths are somewhat longer, the bond strengths weaker, and the vibrational frequencies lower than those obtained in MP2 calculations, and the MP2 model is to be preferred. The main use of the single excitation amplitudes in the CC2 model is to make it possible

to calculate properties of excited states and to introduce orbital polarisation effects. In general, it makes little sense to replace MP2 by CC2 in ordinary ground-state calculations. The CC2 method should mainly be used in calculations of excited-state properties.

The MP2 and CC2 models are in practice the only correlated *ab initio* computational methods that can be applied to large biomolecules today. This could well change in the near future as computers become faster and more efficient divide-and-conquer algorithms are being developed (28, 29, 30). Another reason for the success of the MP2 and CC2 methods is the efficient resolution-of-the-identity (RI) MP2 and CC2 algorithms that have been developed and implemented (26, 31, 32). The RI algorithms, also known as density-fitting methods, are regularly used in applications on biomolecular systems (33). A linear-scaling MP2 implementation already exists, so biomolecules consisting of more than a thousand atoms can indeed be investigated with present-day computer resources (34). Other *ab initio* correlation approaches that can be applied on large biomolecules are e.g., the molecular orbital fragmentation approaches (35–37).

## 2.1. Ground-State Calculations

The main advantage with *ab initio* correlation approaches is that the accuracy and reliability can be systematically improved by taking higher-order correlation effects into account. Thus, the reasons for deviations between calculated values and experimental results can be understood by employing methods that consider higher-order correlation effects. However, the huge computational costs of these methods makes biomolecular calculations unfeasible. The basis set size significantly affects the accuracy. The use of small basis sets in *ab initio* correlation calculations renders the obtained results unreliable despite a mathematically accurate treatment of electron correlation.

The simplest correlated WF method, MP2, *does* already account for one of the most important correlation effects for biomolecules, that is, dispersion interactions. Benchmark calculations on small molecules have shown that MP2 generally overestimates non-bonding interactions, however. This flaw can be corrected for by introducing semi-empirical scaling factors to the same-spin and opposite-spin contributions of the MP2 (38, 39) and CC2 correlation energies (40). In many applications, the spin-component scaled (SCS) and scaled opposite-spin (SOS) MP2 methods have been found to provide an accuracy comparable to the CCSD model perturbationally corrected for triple excitation, CCSD(T) (41). One should naturally be aware of the caveats of relying on an accuracy rooted in cancellation of errors.

The SOS-MP2 and SOS-CC2 can be made faster than the uncorrected versions as they can be formulated to formally scale as  $N^4$  (42). For example, the SOS-MP2 methodology has recently been applied to study peptide isomerisation for model systems comprising ~400 atoms (43).

## 2.2. Excited-State Calculations

Excited states can be studied at *ab initio* levels using a variety of related computational levels. The simplest and least accurate models are CI singles (CIS), CC singles (CCS), and the random-phase approximation (RPA), which in principle do not consider any electron correlation effects. The CIS and CCS calculations yield identical excitation energies, but the wave functions slightly differ due to the exponential ansatz of the CC model. RPA, also called HF linear response, forms a link between the linear response time-dependent DFT (TDDFT) models (44, 45) and the *ab initio*-correlated counterparts (46, 47). The excitation energies obtained at the CIS and CCS levels are generally inaccurate and therefore of limited value. Electron correlation effects have to be considered in calculations of excited-state properties. The most common *ab initio* correlation methods for studies of excited states of biomolecules are CC2 and the closely related CIS with perturbative corrections (CIS(D)) (48), its iterative variant (CIS( $D_\infty$ )) (49), and the algebraic diagrammatic construction through second order (ADC(2)) (50) methods. The excitation energies of these methods are obtained as poles of the response functions. The poles correspond to the eigenvalues of the Lagrangian stability matrix, which is the electronic Hessian for the variational ADC(2) method and the Jacobian for non-variational methods (27). The CASPT2, MRCI, and quantum Monte Carlo (QMC) (51, 52) methods can yield accurate excitation energies, but at high computational costs. The ADC(2) and CIS(D) methods have been shown to be approximations to the CC2 model (27, 53). For molecular gradients of excited states, the hermitian ADC(2) method is about a factor of three faster than the non-hermitian CC2.

The relative accuracy of the methods has not been very extensively assessed. However, available benchmark data show that the ADC(2) method gives singlet  $n \rightarrow \pi^*$  energies which are about 0.3 eV smaller than those obtained at the CC2 level. The singlet  $\pi \rightarrow \pi^*$  and singlet  $\pi \rightarrow \sigma^*$  energies calculated at the ADC(2) and CC2 levels agree well (40). The accuracy of the ADC(2), CC2, CIS ( $D_\infty$ ) excitation energies and the SCS and SOS counterparts is largely comparable (40, 53, 54).

The CC2 method has increasingly been applied on biomolecular model systems during recent years. For example, the isomerisation barriers for the ground and first excited state of retinal have been studied at the CC2-level (55) as well as spin and charge distribution in simple porphyrin models (56). A benchmark study of the excitation energies for biochromophores of rhodopsin, photoactive yellow protein, and the green fluorescent protein suggest, that the CC2 energies deviate by less than 0.15 eV from experimental energies, whereas for the commonly used TDDFT approach, the maximum deviation is 0.27 eV (57). Similar accuracies for CC2 have been reached in other benchmark sets (58). Furthermore, a recently introduced reduction of virtual space

(RVS) approximation within the CC2-framework, shows that virtual orbitals above a certain energy threshold can be omitted in the correlation calculation without significant loss in accuracy, whereas the computational cost is reduced by a factor of six. This renders CC2 treatments of large biomolecular systems possibly, such as a rhodopsin model comprising 165 atoms (59).

---

### 3. Density Functional Theory

While *ab initio* methods are appealing due to their rigorous and systematic treatment of electron correlation effects, they become prohibitively expensive with increasing system size. The vast majority of biologically relevant molecular systems are, for the moment, beyond a correlated *ab initio* wave-function treatment. The *de facto* work-horse for large-scale correlated quantum chemical studies has already for some time been DFT. Within DFT, instead of directly dealing with the  $N$ -dimensional wave function, one uses a short-cut *via* the electron density to obtain the quantum chemical properties of the system. While it in principle should be possible to reduce the dimensionality of the problem to only three spatial coordinates, as shown by Hohenberg and Kohn (60), this is not how the majority of DFT computations are carried out. Especially describing the kinetic energy of the electrons is greatly simplified by considering the system as a composition of one-electron orbitals within the Kohn–Sham framework (61). Thus, in practice, the dimensionality of DFT increases from three to  $N$  times three.

Accounting for electron correlation within wave-function (WF) theory is straightforward, with a clear, although increasingly cumbersome path towards a more complete and accurate description. Within DFT, this is not the case; developing functionals involves a fair degree of freedom, in contrast to the brute-force inclusion of higher-order terms in a WF treatment. While this freedom at times is seemingly random, a hierarchy of density functionals can still be discerned. Instead of including, say, more configurations or allowing for more complex excitations of the electrons, the different levels of DFT include an increasingly complex dependency on the information contained in the electron density.

#### 3.1. Functionals and Jacob's Ladder

The basic building block of most functionals is the local density approximation, LDA. The functional form only considers the value of the electron density at specific points. It is too simple and inaccurate for quantitative molecular studies, but as most present-day functionals are constructed as corrections to the “first level” approximation provided by LDA, it deserved some attention. We ignore here even simpler functional forms, like the Thomas–Fermi model (62, 63).

The LDA is exact for a uniform electron gas, and is applicable to slowly varying densities. The exchange/correlation (XC) energy functional thus has the following, seemingly simple form,

$$E_{\text{xc}}^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{\text{xc}}^{\text{unif}}(\rho) d\mathbf{r} \quad (7)$$

Above,  $\varepsilon_{\text{xc}}^{\text{unif}}$  gives the XC energy per electron in a uniform electron gas. An analytical expression for the exchange energy can be derived,

$$E_{\text{x}}^{\text{LDA}}[\rho] = -C_{\text{x}} \int \rho^{4/3}(\mathbf{r}) d\mathbf{r}, \quad C_{\text{x}} = \frac{3}{4} \left( \frac{3}{\pi} \right)^{1/3} \quad (8)$$

This is as far as the analytically derived train of functionals travels; already the correlation energy of the uniform electron gas is such a complex quantity that an explicit derivation of its density functional form has proven too difficult. Very good approximations that, using analytical functional forms, reproduce high-quality QMC results *do* exist (64). It is nevertheless sobering to reflect over the fact that already the LDA correlation functionals are simply *ad hoc* functionals with the sole anchor to physical reality being that the results they provide are in good agreement with what the “true” functional would provide. A few LDA correlation functionals are in use in modern computer codes, the most popular being two of the five functionals of Vosko, Wilk, and Nusair (65), VWN-3 and VWN-5, as well as a more recent alternative by Perdew and Wang, PW92 (66).

Although the values provided by the different functionals are very similar, they are not identical, which can lead to problems when attempts are made at reproducing results. As a further complication, this mild confusion propagates to the more complex functional forms, discussed below, as they in many cases are constructed as corrections to the LDA form. The choice of the LDA correlation functional is quite non-standard, with different program packages using different default settings. Unfortunately, it is not uncommon for a specific implementation of a higher-level functional to ignore the original authors’ specification of which LDA correlation functional to use; sometimes, this is not even clearly defined.

The electron density itself contains more information than its specific value at points in space. Considering increasing amounts of the information content of the density within the functional form has been described as climbing Jacob’s ladder of DFT, with each rung bringing the functional closer to perfection. Some of the most popular functionals for biological applications reside at the rung above LDA: the family of generalised gradient approximation (GGA) functionals include information about the gradient (of general order), that is, the change in electron density at specific points.

The added information is still local. The non-locality is derived from infinitesimally small surroundings of points in space. GGAs are nevertheless often referred to as semi-local approximation; in some literature, even non-local, which should be understood as merely being in contrast to the *local* density approximation.

Ascending the ladder, the functional form becomes more complex, which can be exploited in two ways: *(i)* a larger range of known physical features of the “true” functional can be fulfilled, and *(ii)* Empirical parameters can be introduced, selected so that the computed results reproduce a desired, usually experimental data set as faithfully as possible. Both paths can naturally be trod simultaneously. The non-empirical approach has the advantage of being more general, and from a puristic viewpoint, aesthetically more pleasing. The GGA form is, however, still too limited to provide quantitative predictions, especially for properties of biomolecules, without tuning the functional towards specific systems. This is done at the expense of knowingly breaking some constraints, with the motivation that in real-world applications, getting, say, the bonding energy correct is usually more relevant than describing the properties of the uniform electron gas faithfully. The pragmatic biomodeller, by necessity, includes semi-empirical functionals in the toolbox.

The most common non-empirical GGA is the Perdew–Burke–Ernzerhof functional, PBE (67). Its performance for biomolecules is unfortunately not very good. Of the huge family of semi-empirical functionals, we highlight the BP86 functional, composed of Becke’s exchange functional from 1988 (68) and the correlation functional by Perdew from 1986 (69). Its mature age, as far as GGAs go, could be seen as a disadvantage, and indeed, some of the inherent parameters could probably be fine-tuned to provide a better overall description for specific applications. It is, however, fairly well explored and used, and thus a large literature base documenting its performance is available. Of the more modern GGAs, we mention the Swart–Solà–Bickelhaupt functional, SSB-D (70), which includes an empirical dispersion correction, discussed in more detail in section 1.3.2. SSB-D has been specifically designed for several scenarios, which are highly relevant also in the modelling of biomolecules: spin-state energetics of transition metal compounds, reaction barriers, and hydrogen bonds are described well.

The introduction of GGAs presented a leap forward in accuracy over LDA, with especially the exchange energy being described better. For quantitative predictions, most GGAs in general are still not robust enough. Hartree–Fock theory can in principle provide the exact exchange energy when correlation effects are ignored. Thus, one could presume that an ideal exchange/correlation energy would be obtained by simply combining a density functional for correlation with the HF expression for exchange energy. This is

the basic idea of *hybrid functionals*, and was suggested already in connection with the LDA approximation (61). Completely discarding a density functional description of exchange in favour of the HF description is, however, not desirable. Instead, modern hybrid functionals describe exchange only partly with the HF expression, with the major constituent still being a DFT exchange functional.

For chemical applications, the most popular hybrid functional, B3LYP is based on Becke's three-parameter functional, B3 (71), in combination with the LYP correlation functional,

$$\begin{aligned} E_{xc}^{\text{B3LYP}} = & \alpha_0 E_x^{\text{HF}} + (1 - \alpha_0) E_x^{\text{LSDA}} + \alpha_x \Delta E_x^{\text{B88}} + (1 \alpha_c) E_c^{\text{VWN}} \\ & + \alpha_c E_c^{\text{LYP}} \end{aligned} \quad (9)$$

where  $\alpha_0$ ,  $\alpha_x$ , and  $\alpha_c$  are the three adjustable parameters of the original B3 formulation; in B3LYP, their values are set at  $\alpha_0 = 0.2$ ,  $\alpha_x = 0.72$ , and  $\alpha_c = 0.81$ . For the present purpose, the value of  $\alpha_0$  is most relevant, being the portion of exact exchange, here 20%. As noted in Eq. 9, the LDA part of correlation is specified as VWN. Indeed, there are at least three versions of B3LYP in circulation, using either the VWN-5, VWN-3, or VWN-1 form. As noted above, different programs can thus give slightly different B3LYP values, even if other settings like numerical integration grids were equal.

B3LYP has been hugely successful for chemical applications, and the accuracy it provides is, in most cases, far superior to standard non-hybrid GGAs. There are cases where B3LYP fails, and some other functional provides better values for a specific system, of course. Many of these cases are, however, well documented, due to the huge number of quantum chemical studies carried out with it each year. As a general purpose functional for studying biosystems, it is still a good choice. Being a hybrid functional, with the need of computing the HF-exchange part, it is significantly more costly to employ compared to non-hybrid GGAs. This does limit its applicability to a degree.

Continuing upwards the ladder of DFT, we arrive at the rung of meta-GGAs. In addition to the density and its gradient, meta-GGAs build upon the GGA description by further considering the kinetic energy density (most common) and/or the Laplacian of the density in their functional form. The general form of a meta-GGA is thus,

$$\varepsilon^{\text{mGGA}} = \varepsilon^{\text{mGGA}}(\rho_\alpha, \rho_\beta, \nabla \rho_\alpha, \nabla \rho_\beta, \nabla^2 \rho_\alpha, \nabla^2 \rho_\beta, \tau_\alpha, \tau_\beta) \quad (10)$$

$\tau$ , the kinetic energy density, depends on the Kohn–Sham orbitals,

$$\tau_\sigma(\mathbf{r}) = \frac{1}{2} \sum_i^{\text{occ}} |\nabla \Psi_{i\sigma}(\mathbf{r})|^2 \quad (11)$$

$\tau$  introduces some true non-locality into the functional form, due to the orbital dependency; meta-GGAs that only consider the

Laplacian  $\nabla^2\rho$  are still semi-local. With added flexibility in the functional form, meta-GGAs can again emulate the unknown true functional more completely. The partly non-local character of kinetic energy functionals also brings the possibility of describing new physics that is intractable for a purely local description. So far, meta-GGAs have not been extensively used for biomolecular simulations. This could change in the near future, especially as these functionals are computationally cheaper than hybrid functionals even with their more complex functional form. Many software packages already provide a choice of meta-GGAs, including hybrid meta-GGAs. An interesting newcomer, which could be promising for biochemical studies, is the non-hybrid revTPSS functional (72).

### **3.2. Non-Covalent Interactions Within DFT**

Non-covalent interactions are of special importance in the stability and formation of biological structures. Hydrogen bonds and weak dispersion forces accumulate in large structures and therefore should be accounted for. Hydrogen bonds are reasonably described within DFT. Beyond LDA level, most functionals perform quite well, the PBE and SSB-D functionals especially so. However, commonly used functionals fail to described long-range dispersion.

The dispersion interaction is the attraction between neutral fragments due to polarisation caused by quantum fluctuations of the electron density. It is also known as London forces, and especially within the biologically oriented literature as van der Waals forces. The lack of dispersion within standard functionals has long been acknowledged as one of the major shortcomings of contemporary DFT. The phenomenon is *non-local*, that is, no overlap of electron densities is needed. Due to this, standard semi-local DFT functionals cannot capture dispersion. Hybrid functionals incorporating non-local exchange bring no improvement either, as dispersion is a correlation effect. Also, the kinetic energy density incorporating meta-GGAs cannot properly describe an interaction between non-overlapping densities. Without added completeness in the functional form, there is no physically sound mechanism by which *any* (hybrid) GGA can describe dispersion; reparameterisation of existing functional forms, for example, cannot be generally successful. New physics needs to be introduced, for example, *via* the dipole moment of the exchange hole, as shown by Becke and Johnson (73, 74). In this XDM model, the authors argue that the instantaneous dipole moments that give rise to the dispersion interaction can be traced to the exchange-hole dipole moment, which can be computed from the occupied Kohn–Sham (or Hartree–Fock) orbitals; there is no need for time dependence or virtual orbitals. The model has been found to be promising for biological applications (75), but has not been extensively utilised in the literature.

Another approach incorporating new physics, involves climbing up the DFT ladder to the fifth rung, where also virtual,

unoccupied orbitals are taken into account. A typical functional where this is incorporated is the B2-PLYP functional by Grimme (76). The “pure” DFT part in B2-PLYP is based on the B88 exchange functional and the LYP correlation functional. In addition, HF exact exchange is added, and as a new ingredient, a second order perturbation term (PT2) is included. At WF theory level, MP2 is the computationally cheapest way of accounting for dispersion, and it can thus be considered a natural extension of the hybrid scheme, initiated by including HF exchange; the functional is thus an example of a *double hybrid functional*.

B2-PLYP does have some draw-backs compared to conventional functionals. The computational cost is higher, as the MP2-like term needs to be evaluated. As such, this does not add significantly to the computational time compared to standard hybrid functionals, as very efficient schemes for computing MP2 on top of Hartree–Fock have been devised. Instead, the major cost increase is indirect, due to the significantly higher basis set demand compared to normal DFT, caused by the necessity to describe the virtual space properly. This will be discussed in more detail in the section on basis sets. For the discussion in this section, it is, however, noteworthy that B2-PLYP is still lacking in the description of long-range dispersion interaction. The PT2 part is relatively small compared to the poorly performing LYP correlation functional.

Another, highly popular approach of including dispersion into DFT represents an almost opposite approach to the double hybrid functionals. Instead of including more costly ingredients from wave-function theory, dispersion can be approximated at a computationally negligible cost by the way of molecular mechanics. Classical MM force fields perform much better for dispersion than semi-local DFT. The  $R^{-6}$  term is simply one of the force field parameters. As dispersion is long range, it usually has a very small effect on the total density. This motivates the general form of dispersion including DFT and DFT-D,

$$E_{\text{DFT-D}} = E_{\text{KS-DFT}} + E_{\text{disp}} \quad (12)$$

The dispersion correction is thus simply added on top of a normal DFT calculation, modifying the potential energy surface, but not the electronic structure. This ideally results in better geometries and binding energies. Also gradients are easily implemented at DFT-D level, so improved vibrational frequencies can also be obtained.

The presently most common form of  $E_{\text{disp}}$  is due to Grimme (77), and has a relatively simple form,

$$E_{\text{disp}} = -s_6 \sum_{i=1}^{N_{\text{at}}-1} \sum_{j=i+1}^{N_{\text{at}}} \frac{C_6^{ij}}{R_{ij}^6} f_{\text{dmp}}(R_{ij}) \quad (13)$$

$N_{\text{at}}$  is the number of atoms,  $C_6$  are atom-specific dispersion coefficients, and  $C_6^{ij} = \sqrt{C_6^i C_6^j}$ . The  $C_6$  terms are computed from atomic ionisation potentials and static dipole polarisabilities, as  $C_6^a = 0.05 N \times I^a \alpha^a$ , with  $N = 2, 10, 18, 36, 54$ , for rows 1–5, respectively. The dispersion terms are assumed to be constant for all molecules and geometries.

The damping function  $f_{\text{dmp}}$  is compulsory to avoid near-singularities and attraction approaching infinity for small  $R$ . It also aids in keeping the van der Waals dispersion correction non-zero only for relevant interatomic distances, those neglected by normal DFT, that is, long-range interaction where the electron density overlap is small. Also, at small distances, the  $R^{-6}$  behaviour is not valid anymore, either. The form of the damping function in Ref. (77) is

$$f_{\text{dmp}}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/R_r - 1)}} \quad (14)$$

$R_r$  is the sum of the atomic van der Waals radii, which need to be fitted or computed. They are also assumed to be constant for all molecules, but this assumption is less crude for the atomic radii than for the  $C_6$  parameters.  $d$  is a “sufficiently large” damping parameter of 20, which gradually switches off the correction at decreasing distances.

Even with the damping function, there is an issue of double-counting correlation. This is partly remedied by the global scaling parameter  $s_6$ , which differs from functional to functional. For both BP86 and B3LYP,  $s_6 = 1.05$ , while for example for B2-PLYP, it is only 0.55, reflecting the fact that dispersion is partly accounted for already *via* the PT2 term of the double hybrid. The possible double counting of correlation is, however, still present. There is no clear-cut way of defining exactly what is missing in DFT, and thus, adding “something” on top can (will) fail, occasionally.

Some of the main weaknesses of the DFT-D scheme were addressed recently by Grimme et al. in connection with the third iteration of the dispersion correction scheme, denoted DFT-D3 (78, 79). By computing the pairwise dispersion coefficients and cutoff radii from first principles, an even higher accuracy could be obtained, with less empiricism. Other improvements include geometry dependence and inclusion of a  $R^{-8}$  term for the dispersion interaction. Also, Steinmann and Corminboeuf (80) recently presented a system-dependent dispersion correction scheme, denoted dDXDM, which reduces the empiricism of the earlier versions of DFT-D. The dispersion coefficients are computed using the XDM formalism by Becke and Johnson (73, 74), and the correction includes both  $R^{-8}$  and  $R^{-10}$  terms. The computational costs of both the DFT-D3 and dDXDM corrections are still negligible compared to those for obtaining the electronic DFT energy.

All in all, DFT-D is a simple and efficient, although empirical way of “fixing” DFT for the lack of dispersion. With functional-specific parameters, its various iterations can be combined with already existing functionals, with well-known behaviour. The DFT-D scheme has already found its way into several quantum chemistry program packages, and also the latest versions are starting to appear in updated versions. For biomolecular simulations, it is a very cost-efficient method, and as recently shown by Siegbahn et al., it is highly useful for accounting for the significant van der Waals effects not only in “traditional” situations but also, for example, in transition metal complexes (81). For example, as recently shown by Kaila et al., inclusion of dispersive forces, significantly affect binding energetics of oxygenous species in the active binuclear metal site of cytochrome *c* oxidase (82).

### 3.3. Excited-State Calculations

DFT studies on electronically excited states of biomolecules are performed using the linear response TDDFT approach. It is an efficient method in the sense that it provides accurate excitation energies for large molecules at a reasonable cost. The TDDFT equations are structurally similar to the RPA or linear response HF equations. The main difference is that the HF exchange contribution of the RPA expression is replaced by the exchange-correlation term of the DFT functional. The TDDFT equation giving the excitation energies as eigenvalues is written as (44, 45, 83, 84),

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \quad (15)$$

where  $I$  is the unit matrix and the  $A$  and  $B$  matrices are obtained as

$$(A + B)_{ia\sigma, jb\sigma'} = (\varepsilon_{a\sigma} - \varepsilon_{i\sigma})\delta_{ij}\delta_{ab}\delta_{\sigma\sigma'} + 2(i a\sigma | j b\sigma') + 2f_{ia\sigma, jb\sigma'}^{\text{xc}} - c_x\delta_{\sigma\sigma'}[(j a\sigma | i b\sigma) + (a b\sigma | i j\sigma)] \quad (16)$$

$$(A - B)_{ia\sigma, jb\sigma'} = (\varepsilon_{a\sigma} - \varepsilon_{i\sigma})\delta_{ij}\delta_{ab}\delta_{\sigma\sigma'} + c_x\delta_{\sigma\sigma'}[(j a\sigma | i b\sigma) - (a b\sigma | i j\sigma)] \quad (17)$$

$\varepsilon_{i\sigma}$  and  $\varepsilon_{a\sigma}$  are the orbital energies of the occupied and virtual Kohn–Sham orbitals,  $(pq\sigma | rs\sigma')$  is a two-electron integral in Mulliken notation and  $f_{ia\sigma, jb\sigma'}^{\text{xc}}$  is a matrix element of the exchange-correlation kernel in the adiabatic approximation. The molecular gradients and first-order properties of excited states can also be calculated (84). The resolution of the identity (RI-J) approximation (85) can be used to speed up TDDFT calculation when non-hybrid functionals are used.

The TDDFT method suffers from some serious problems that one should be aware of. The DFT potential has an incorrect asymptotic behaviour at long distances implying that charge-transfer states become too low in energy. The charge-transfer problem is a severe

drawback in applications on biomolecules because transitions to charge-transfer states appear in the chemically relevant part of the electronic excitation spectra. The self-interaction error of today's functionals shifts the ionisation potential down in energy, implying that excited states with excitation energies larger than the orbital energy of the highest occupied molecular orbital (HOMO) might be unreliable, because those states actually lie in the DFT continuum of the system (86, 87). This might be a serious problem when investigating excited state properties of anions and Rydberg states. However, the ionisation threshold can be shifted upwards by stabilising the system with counterions. The shape of the long-ranged DFT potential can be improved by mixing DFT and HF-exchange contributions (88). The larger the share of HF exchange, the closer the potential is to the exact  $1/r$  form. A larger amount of HF exchange can be included in the functional when different functionals are used depending on whether the electrons are close or farther away from each other (89, 90). The range-separated functionals shift the HOMO energy and the ionisation potential upwards in energy. The excitation energies of the low-lying states are less affected by the range separation, whereas problems with the charge-transfer and Rydberg states are reduced because the states are below the DFT ionisation threshold.

Computational studies on excited states of polyenes such as carotenoids are challenging, because for long polyene chains, the lowest excited state is a one-photon dipole-forbidden state, whereas the bright dipole-allowed transition has slightly larger energy. The commonly used TDDFT functionals and low-order *ab initio* approaches have difficulties to accurately describe excited states with a significant double excitation character. The low-order *ab initio* methods fail because their starting point consists of states constructed as single replacements from the ground-state configuration. The main reason for the failure of TDDFT in these cases is the inability of today's functionals to consider states with a strong double replacement character. However, by combining DFT with MRCI, the obtained excitation spectra of carotenoids were in better agreement with experimental data than the ones calculated at traditional DFT levels (91, 92, 93). Retinals also consist of a long hydrocarbon chain with alternating single and double bonds. However, recent computational studies at DFT and correlated *ab initio* methods show that the protonated Schiff bases and polyenes have completely different character of the electronic structure (87, 94). The lowest excited states of protonated Schiff base retinals are well described at DFT and TDDFT levels (55, 87, 94). Performance of different functionals has been assessed in benchmark sets, suggesting that TDDFT errors can be comparable to those of the CC2 method for well-behaving systems (95). TDDFT calculations have also been widely employed to study photobiological processes, often in combination with a hybrid quantum mechanics/molecular

mechanics (QM/MM) treatment, for example, in the green fluorescent protein (96, 97), photoactive yellow protein (98, 99), and rhodopsin (100, 101). An extensive rhodopsin model consisting of the retinal chromophore and its nearest residues has recently been studied at DFT and TDDFT levels (102).

---

## 4. Semi-Empirical Methods

Semi-empirical methods are based on the same theoretical framework as *ab initio* MO theory, but they aim at reducing the computational cost by neglecting or approximating time-consuming two-electron integrals. These integrals can be fitted to experimental data or replaced by analytical but approximate expressions. Thus, in contrast to *ab initio* methods, semi-empirical calculations require parametrisation of all elements involved in the molecular system.

Semi-empirical methods were widely popular in early days of quantum chemical calculations (103), but have remained important for quantum chemical treatment of computationally demanding systems. In the field of biomolecular modelling, the complexity and system size are still today often computationally too demanding for full DFT treatments. Moreover, large configurational sampling is important in free-energy calculations, which might also be outside the scope DFT methods (104). Although widely and successfully applied in biomolecular modelling, the empirical valence bond (EVB) method (105) and its variant multi-state EVB (MS-EVB) (106), are considered as reactive classical force fields rather than semi-empirical methods (107), although a similar SCF approach is employed to solve the EVB wave functions and respective energies.

Many semi-empirical methods are based on the Neglect of Diatomic Differential Overlap (NDDO) approximation (108), in which an SCF calculation is performed on the valence electrons, described by minimal basis sets. These comprise methods like the Modified Neglect of Differential Overlap (MNDO) (109), the Austin Method 1 (AM1) (110) and the Parameterised Model 3 (PM3) (111), of which the most recent representative is the PM6 parametrisation from 2006 (112). PM6 is parameterised based on a large experimental data set of over 9000 compounds (112) and includes currently 83 elements, up to bismuth. The other levels of approximations are Intermediate Neglect of Differential Overlap (INDO) and Complete Neglect of Differential Overlap (CNDO) (108). Despite the mature age of these approximations, for example, Zerner's Intermediate Neglect of Differential Overlap, ZINDO (113) has recently found new applications in the study of electronic coupling terms in biological electron transfer processes

(114, 115) as well as energetics of spin state (116) and optical properties (117, 118).

The accuracy of semi-empirical methods has similarly to DFT been enhanced by addition of a dispersion terms, with promising accuracies comparable to those obtained by DFT-D for problems where dispersion effects are significant (119). Parameters focusing on hydrogen bonding have also been introduced. The dispersion and H bonding corrected PM6, PM6-DH2, has recently been used to study, for example, ligand binding to HIV-1 protease (120). The method may thus have important implications in rational drug design, where a large number of target molecules are scanned to search for suitable substrate analogues.

A semi-empirical ansatz can be introduced within the theoretical framework of DFT. The Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) theory is obtained by replacing the quantum mechanical electron density by atomic partial charges (121). These semi-empirical DFT methods are computationally as efficient as other semi-empirical models and have during recent years also been extended to study, for example, optical and excited-state properties (122). In contrast to other semi-empirical methods, DFTB is not based on fitting to experimental data but rather to DFT calculations. SCC-DFTB methods, often in combination with a hybrid MM treatment of protein surroundings, have been employed to describe proton transfer in carbonic anhydrase and alcohol dehydrogenase (123, 124). As for DFT and the other semi-empirical methods, SCC-DFTB has difficulties in describing dispersion forces, but inclusion of empirical dispersion corrections (see above) is found to improve agreement with experiments (125).

---

## 5. Basis Sets

Basis sets are used to describe the wave function and density distribution in a molecule. Within KS-DFT, as in Hartree–Fock, their main use is in describing the molecular orbitals. In molecular calculations, they are most often atom centered. Plane-wave basis sets and local numerical basis functions are also employed. Basis sets are composed of a set of functions, which during the calculation are populated with the criterion of minimising the total electronic energy. For a perfect description within a given model, an infinite amount of different functions would be needed. In general, the smaller the basis set, that is, the fewer component functions it includes, the poorer is its ability to describe the true density and wave function. Basis set development has the goal of finding a suitable set of functions that can give a description of the molecular orbitals as accurately as possible (according to some, usually

energetic, criterion), given a predefined set of functions to choose from. In principle, the individual functions could have any mathematical form, whereas in practice, they resemble atomic orbitals (AOs).

Two main schemes are in contemporary use for describing the form of the individual AOs: Slater-type orbitals (STOs) and Gaussian-type orbitals (GTOs). STOs have some desirable properties, like a correct cusp behaviour at the nucleus and a correct exponential decay away from the nucleus. While GTOs lack both of these properties, they allow for easier integral evaluations and are presently the most common type by a wide margin, except in the case of semi-empirical methods, where STOs are used. The general form of a Cartesian GTO function can be expressed as

$$\phi(x, y, z, \alpha, i, j, k) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} \left[ \frac{(8\alpha)^{i+j+k} i! j! k!}{(2i)! (2j)! (2k)!} \right]^{\frac{1}{2}} x^i y^j z^k e^{-\alpha(x^2+y^2+z^2)} \quad (18)$$

where  $\alpha$  is the exponent that controls the width, that is, the tightness or diffuseness of the orbital function.  $i, j$ , and  $k$  are positive integers that control the type of function:  $i + j + k = 0$ , defines an  $s$ -type function, with  $i + j + k = 1$ , we obtain three different  $p$ -type functions, whereas  $i + j + k = 2$ , six different  $d$ -type functions are obtained, etc. Several older basis sets are of the Cartesian GTO form of Eq. 18, while newer basis sets are usually of spherical GTO type. The main difference is that with spherical GTOs, the number of functions for a given angular momentum corresponds to the correct number of different atomic orbitals. Instead of the six different  $d$ -functions of a Cartesian set, only five  $d$ -functions are present. A spherical basis set is thus somewhat more physically motivated, and is also more robust against linear dependency problems. When using a specific basis set with a specific program, one should verify that the scheme used within the program matches the scheme that the basis set was developed for. Both Cartesian and spherical basis sets are almost exclusively defined only with the exponent  $\alpha$  and the type of AO the exponent refers to (the sum of  $i + j + k$  above); the expansion into a specific number of component AOs is done internally by the program.

Different hierarchies of basis sets can be defined. In a *minimal basis set*, there are only enough functions to describe the neutral atom. For H and He, there is thus only 1  $s$ -function, for the first row atoms, 2  $s$ -functions ( $1s, 2s$ ) and a set of  $p$ -functions ( $2p_x, 2p_y, 2p_z$ ). Such basis sets are not useful for molecules. In a *double-zeta* (DZ) basis set, the number of basis functions are doubled, compared to the minimal basis set. A DZ basis set is already significantly better suited for molecular calculations, as it is able to describe the fact that the electron distribution can be different in different directions. Commonly, DZ-type basis sets are only of double-zeta

quality for the valence electrons, as these are mostly affected by chemical bonding. The core orbitals, for example,  $1s$  in carbon, usually change very little from their free atom form. Thus, doubling the amount of core orbitals is seldom worth the accompanied increased computational cost. In *triple-zeta* (TZ), *quadruple-zeta* (QZ), etc., basis sets, the same scheme of tripling, quadrupling, etc., the amount of functions is applied, allowing for a more accurate description of the true wave function and/or density. Again, functions are usually only added in the valence region.

In unpolarised basis sets, as described above, only functions of the type present in the free atom are used,  $s$  and  $p$  for carbon, for example. Higher angular momentum functions, popularly called *polarisation functions*, are, however, almost always important. They allow for deformation of the electron density around atoms, which otherwise would be impossible to describe. For example, the density around a hydrogen atom in a molecule, described by an unpolarised basis set with only  $s$ -functions, would always stay almost spherical. The basis set nomenclature for denoting polarisation functions is varied. For the Pople family, a star is commonly used, as in 6-31G vs. 6-31G\*. The letter P is also used, as in DZ vs. DZP, or in the Dunning family as cc-VDZ vs. cc-pVDZ.

For anions, highly excited electronic states, and loosely bound complexes, the HOMOs are usually spatially much more diffuse than “normal” orbitals. For describing such systems, it is necessary to add, or augment, the basis set with *diffuse functions*, functions with a very small  $\alpha$  exponent. The nomenclature again varies, with diffuse functions being denoted with, for example, + signs (6-31+G (d)), or the acronym aug, for augmented basis sets (aug-cc-pVDZ).

Several different basis set “families” have been devised over the years. Here, we highlight four:

- The *Dunning correlation consistent* family (cc-pVDZ, cc-pVTZ, aug-cc-pVQZ, ...) ([126](#))
- The *Pople* family (6-31G, 6-311G\*\*, 6-311G++(2d,2p), ...) ([127](#))
- The *Karlsruhe* family (SV(P), def-TZVP, def2-QZVPP, ...) ([128](#), [129](#))
- The *Jensen polarisation consistent* family (pc-0, pc-2, aug-pc-3, ...) ([130](#))

The *Dunning correlation consistent* sets are popular, and their performance is thus well understood. They have been devised to systematically converge the energy of correlated wave-function methods with increasing basis set size. Basis sets for many elements, up to Rn and of large size have been developed, along with a consistent set of diffuse functions. They are, however, not optimal for DFT calculations.

The *Pople* family of basis sets are also very popular, with well-known performance. Consistent sets of diffuse functions have been developed. Of the periodic table, mostly light elements up to Zn are included. However, care must also be taken not to get an unbalanced description of the system if it becomes necessary to mix basis set families.

To avoid a mixture of basis set families, the *Karlsruhe* family offers a good alternative to the Pople series. Basis sets up to quadruple-zeta quality have been developed for all elements up to Rn. The Karlsruhe set is also often more efficient than the Pople series, that is, similar quality results can be obtained with a smaller size of the basis set. Recently, the basis sets have also been augmented with sets of diffuse functions (131). We recommend the usage of the latest versions of the Karlsruhe sets, identified by the prefix “def2”.

Jensen has in the recent years developed the *polarisation consistent* family (130). These have been optimised for DFT calculations, and show a systematic convergence towards the basis set limit, allowing for efficient extrapolation schemes. The family includes diffuse functions, but at present time, mostly the upper part of the periodic table is covered. Partly due to the general contraction scheme of the basis functions, the Jensen family is somewhat more expensive compared to the Karlsruhe family.

Several other basis set families naturally exist, and we also note that systematic approaches for reaching basis set completeness (132) and optimal contraction schemes (133) have been proposed.

Basis sets developed for ground-state studies are often well suited for excited-state calculations. The smallest recommendable basis sets are of split valence quality having polarisation functions, denoted SV(P) or 6-31G\*. The SV(P) basis sets can be employed in geometry optimisations as well as in preliminary studies of excited states. The excitation energies of biomolecules and organic compounds calculated using SVP quality basis sets may be overestimated by 0.2–0.5 eV with qualitatively correct oscillator strengths.

Triple-zeta basis sets augmented with two sets of polarisation functions are recommended when more accurate excited-state properties are needed. For biomolecules or molecules embedded in proteins or solvents, very diffuse basis functions are often superfluous because the surrounding molecules prevent expansions of the wave function. Within TDDFT, the use of a doubly polarised triple-zeta basis set generally leads to smaller basis set errors than those of the employed functional. In basis set studies aiming at basis set limit values even larger basis set including diffuse functions are necessary.

## References

1. Hartree DR (1928) The wave mechanics of an atom with a non-Coulomb central field. Part I. theory and methods. Proc Cambridge Phil Soc 24:89–110
2. Slater JC (1930) Note on Hartree's method. Phys Rev 35:210–211
3. Fock V (1930) Näherungsmethoden zur Lösung des quantenmechanischen mehrkörperproblems. Z Phys 61:126–148
4. Slater JC (1929) The theory of complex spectra. Phys Rev 34:1293–1322
5. Löwdin PO (1962) Exchange, correlation, and spin effects in molecular and solid-state theory. Rev Mod Phys 34:80–87
6. Roothaan CCJ (1951) New developments in molecular orbital theory. Rev Mod Phys 23:69–89
7. Helgaker T, Jørgersen P, Olsen J (2000) Molecular electronic-structure theory. Wiley, Chichester
8. Lischka H, Dallos M, Shepard R (2002) Analytic MRCI gradient for excited states: formalism and application to the  $n - \pi^*$  valence- and  $n - (3s, 3p)$  Rydberg states of formaldehyde. Mol Phys 100:1647–1658
9. Roos BO, Taylor PR, Siegbahn PEM (1980) A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. Chem Phys 48:157–173
10. Andersson K, Malmqvist PÅ, Roos BO (1992) Second-order perturbation theory with a complete active space self-consistent field reference function. J Chem Phys 96:1218–1226
11. Ferré N, Olivucci M (2003) Probing the rhodopsin cavity with reduced retinal models at the CASPT2//CASSCF/AMBER Level of theory. J Am Chem Soc 125:6868–6869
12. Merchán M, Serrano-Andres L (2003) Ultra-fast internal conversion of excited cytosine via the lowest  $\pi\pi^*$  electronic singlet state. J Am Chem Soc 125:8108–8109
13. Sinicropi A, Andruriow T, Ferré N, Basosi R, Olivucci M (2005) Properties of the emitting state of the green fluorescent protein resolved at the CASPT2//CASSCF/CHARMM level. J Am Chem Soc 127:11534–11535
14. Ghosh A, Tangen E, Ryeng H, Taylor PR (2004) Electronic structure of high-spin iron (IV) complexes. Eur J Inorg Chem 2004:4555–4560
15. Ghosh A, Taylor PR (2005) Iron(IV) porphyrin difluoride does not exist: implications for DFT calculations on heme protein reaction pathways. J Chem Theory Comput 1:597–600
16. Jensen KP, Roos BO, Ryde U (2005) O<sub>2</sub>-binding to heme: electronic structure and spectrum of oxyheme, studied by multiconfigurational methods. J Inorg Biochem 99:45–54
17. Angeli C, Pastore M, Cimiraglia R (2007) New perspectives in multireference perturbation theory: the  $n$ -electron valence state approach. Theor Chem Acc 117:743–754
18. Schapiro I, Ryazantsev MN, Ding WJ, Huntress MM, Melaccio F, Andruriow T, Olivucci M (2010) Computational photobiology and beyond. Aust J Chem 63:413–429
19. Aquilante F, De Vico L, Ferré N, Ghigo G, Malmqvist PA, Neogrády P, Pedersen TB, Pitonák M, Reiher M, Roos BO, Serrano-Andrés L, Urban M, Veryazov V, Lindh R (2010) Software news and update MOLCAS 7: the next generation. J Comput Chem 31:224–247
20. Čížek J (1966) On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. J Chem Phys 45:4256–4266
21. Kallay M, Gauss J (2004) Calculation of excited-state properties using general coupled-cluster and configuration-interaction models. J Chem Phys 121:9257–9269
22. Bartlett RJ, Purvis GD (1978) Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem. Int J Quantum Chem 14:561–581
23. Pople JA, Binkley JS, Seeger R (1976) Theoretical models incorporating electron correlation. Int J Quantum Chem 10 S:1–19
24. Møller C, Plesset MS (1934) Note on an approximation treatment for many-electron systems. Phys Rev 46:618–622
25. Christiansen O, Koch H, Jørgensen P (1995) The second-order approximate coupled cluster singles and doubles model CC2. Chem Phys Lett 243:409–418
26. Hättig C, Weigend F (2000) CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. J Chem Phys 113:5154–5161
27. Hättig C (2006) In: Grotendorst J, Blugel S, Marx D (eds) Computational nanoscience: do it yourself! NIC series, John von Neumann Institute for Computing, Jülich

28. Hetzer G, Pulay P, Werner HJ (1998) Multipole approximation of distant pair energies in local MP2 calculations. *Chem Phys Lett* 290:143–149
29. Schütz M (2002) A new, fast, semi-direct implementation of linear scaling local coupled cluster theory. *Phys Chem Chem Phys* 4:3941–3947
30. Ziolkowski M, Jansik B, Kjærgaard T, Jørgensen P (2010) Linear scaling coupled cluster method with correlation energy based error control. *J Chem Phys* 133:014107
31. Feyereisen M, Fitzgerald G, Komornicki A (1993) Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem Phys Lett* 208:359–363
32. Weigend F, Häser M (1997) RI-MP2: first derivatives and global consistency. *Theoret Chem Acc* 97:331–340
33. Neese F, Schwabe T, Kossmann S, Schirmer B, Grimme S (2009) Assessment of orbital-optimized, spin-component scaled second-order many-body perturbation theory for thermochemistry and kinetics. *J Chem Theory Comput* 5:3060–3073
34. Doser B, Zienau J, Clin L, Lambrecht DS, Ochsenfeld C (2010) A linear-scaling MP2 method for large molecules by rigorous integral-screening criteria. *Z Phys Chem* 224:397–412
35. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M (1999) Fragment molecular orbital method: an approximate computational method for large molecules. *Chem Phys Lett* 313:701–706
36. Gordon MS, Freitag MA, Bandyopadhyay P, Jensen JH, Kairys V, Stevens WJ (2001) The effective fragment potential method: a QM-based MM approach to modeling environmental effects in chemistry. *J Phys Chem A* 105:293–307
37. Steinmann C, Fedorov DG, Jensen JH (2010) Effective fragment molecular orbital method: a merger of the effective fragment potential and fragment molecular orbital methods. *J Phys Chem A* 114:8705–8712
38. Grimme S (2003) Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J Chem Phys* 118:9095–9102
39. Jung Y, Lochan RC, Dutoi AD, Head-Gordon M (2004) Scaled opposite-spin second order Møller-Plesset correlation energy: an economical electronic structure method. *J Chem Phys* 121:9793–9802
40. Hellweg A, Grün SA, Hättig C (2008) Benchmarking the performance of spin-component scaled CC2 in ground and electronically excited states. *Phys Chem Chem Phys* 10:4119–4127
41. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M (1989) A fifth-order perturbation comparison of electron correlation theories. *Chem Phys Lett* 157:479–483
42. Rhee YM, Head-Gordon M (2007) Scaled second-order perturbation corrections to configuration interaction singles: efficient and reliable excitation energy methods. *J Phys Chem A* 111:5314–5326
43. Sumowski CV, Ochsenfeld C (2009) A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems. *J Phys Chem A* 113:11734–11741
44. Casida ME (1995) In: Chong DP (ed) Recent advances in density functional methods, Part I. World Scientific, Singapore
45. Bauernschmitt R, Ahlrichs R (1996) Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem Phys Lett* 256:454–464
46. Olsen J, Jørgensen P (1985) Linear and non-linear response functions for an exact state and for an MCSCF state. *J Chem Phys* 82:3235–3264
47. Koch H, Jørgensen P (1990) Coupled cluster response functions. *J Chem Phys* 93:3333–3344
48. Head-Gordon M, Rico RJ, Oumi M, Lee TJ (1994) A doubles correction to electronic excited-states from configuration-interaction in the space of single substitutions. *Chem Phys Lett* 219:21–29
49. Head-Gordon M, Oumi M, Maurice D (1999) Quasidegenerate second-order perturbation corrections to single-excitation configuration interaction. *Mol Phys* 96:593–574
50. Schirmer J (1982) Beyond the random-phase approximation: a new approximation scheme for the polarization propagator. *Phys Rev A* 26:2395–2416
51. Lester Jr. WA, Mitas L, Hammond B (2009) Quantum Monte Carlo for atoms, molecules and solids. *Chem Phys Lett* 478:1–10
52. Valsson O, Filippi C (2010) Photoisomerization of model retinal chromophores: insight from quantum Monte Carlo and multiconfiguration perturbation theory. *J Chem Theory Comput* 6:1275–1292

53. Hättig C (2005) Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv Quantum Chem* 50:37–60
54. Aquino AJA, Nachtigallova D, Hobza P, Truhlar DG, Hättig C, Lischka H (2011) The charge-transfer states in a stacked nucleobase dimer complex: A benchmark study. *J Comput Chem* 32:1217–1227
55. Send R, Sundholm D (2007) Coupled-cluster studies of the lowest excited states of the 11-*cis*-retinal chromophore. *Phys Chem Chem Phys* 9:2862–2867
56. Johansson MP, Sundholm D (2004) Spin and charge distribution in iron porphyrin models: a coupled cluster and density-functional study. *J Chem Phys* 120:3229–3236
57. Send R, Kaila VRI, Sundholm D (2011) Benchmarking the coupled-cluster approximate singles and doubles method on biochromophores. *J Chem Theory Comput* 7:2473–2484
58. Silva-Junior MR, Schreiber M, Sauer SPA, Thiel W (2008) Benchmarks for electronically excited states: time-dependent density functional theory and density functional theory based multireference configuration interaction. *J Chem Phys* 129:104103
59. Send R, Kaila VRI, Sundholm D (2011) Reduction of the virtual space for coupled-cluster calculations on large molecules and embedded systems. *J Chem Phys* 134:214114
60. Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. *Phys Rev* 136: B864–B871
61. Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. *Phys Rev* 140:A1133–A1138
62. Thomas LH (1927) The calculation of atomic fields. *Proc Cambridge Phil Soc* 23:542–548
63. Fermi E (1927) Un metodo statistico per la determinazione di alcune proprietà dell'atomo. *Rend Accad Naz Lincei* 6:602–607
64. Ceperly DM, Alder BJ (1980) Ground state of the electron gas by a stochastic method. *Phys Rev Lett* 45:566–569
65. Vosko SJ, Wilk L, Nusair M (1980) Accurate spin-dependent electron liquid correlation energies for local spin-density calculations—a critical analysis. *Can J Phys* 58:1200–1211
66. Perdew JP, Wang Y (1992) Accurate and simple analytic representation of the electron-gas correlation energy. *Phys Rev B* 45:13244–13249
67. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. *Phys Rev Lett* 77:3865–3868
68. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A* 38:3098–3100
69. Perdew JP (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys Rev B* 33:8822–8824
70. Swart M, Solà M, Bickelhaupt FM (2009) A new all-round density functional based on spin states and  $S_N^2$  barriers. *J Chem Phys* 131:094103
71. Becke AD (1993) A new mixing of Hartree–Fock and local density-functional theories. *J Chem Phys* 98:1372–1377
72. Perdew JP, Ruzsinszky A, Csonka GI, Constantin LA, Sun J (2009) Workhorse semilocal density functional for condensed matter physics and quantum chemistry. *Phys Rev Lett* 103:026403
73. Becke AD, Johnson ER (2005) Exchange-hole dipole moment and the dispersion interaction. *J Chem Phys* 122:154104
74. Becke AD, Johnson ER (2005) A density-functional model of the dispersion interaction. *J Chem Phys* 123:154101
75. Johnson ER, Becke AD (2006) Van der Waals interactions from the exchange hole dipole moment: application to bio-organic benchmark systems. *Chem Phys Lett* 432:600–603
76. Grimme S (2006) Semiempirical hybrid density functional with perturbative second-order correlation. *J Chem Phys* 124:034108
77. Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem* 27:1787–1799
78. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J Chem Phys* 132:154104
79. Grimme S, Ehrlich S, Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. *J Comput Chem* 32:1456–1465
80. Steinmann SN, Corminboeuf C (2010) A system-dependent density-based dispersion correction. *J Chem Theory Comput* 6:1990–2001
81. Siegbahn PEM, Blomberg MRA, Chen SL (2010) Significant van der Waals effects in transition metal complexes. *J Chem Theory Comput* 6:2040–2044
82. Kaila VRI, Oksanen E, Goldman A, Verkhovsky MI, Sundholm D, Wikström M (2011) A combined quantum chemical and crystallographic study on the oxidized binuclear center of cytochrome c oxidase. *Biochim Biophys Acta* 1807:769–778

83. Rappoport D, Furche F (2006) In: Marques MAL, Ullrich CA, Nogueira F, Rubio A, Burke K, Gross EKU (eds) Time-dependent density functional theory, Springer, Berlin
84. Furche F, Ahlrichs R (2002) Adiabatic time-dependent density functional methods for excited state properties. *J Chem Phys* 117:7433–7447
85. Bauernschmitt R, Häser M, Treutler O, Ahlrichs R (1997) Calculation of excitation energies within time-dependent density functional theory using auxiliary basis set expansions. *Chem Phys Lett* 264:573–578
86. Casida ME, Jamorski C, Casida KC, Salahub DR (1998) Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: characterization and correction of the time-dependent local density approximation ionization threshold. *J Chem Phys* 108:4439–4449
87. Lehtonen O, Sundholm D, Send R, Johansson MP (2009) Coupled-cluster and density functional theory studies of the electronic excitation spectra of *trans*-1,3-butadiene and *trans*-2-propeniminium. *J Chem Phys* 131:024301
88. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 98:5648–5652
89. Yanai T, Tew DP, Handy NC (2004) A new hybrid exchange? correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem Phys Lett* 393:51–57
90. Heyd J, Peralta JE, Scuseria GE, Martin RL (2005) Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *J Chem Phys* 123:174101
91. Cerón-Carrasco JP, Requena A, Marian CM (2010) Theoretical study of the low-lying excited states of  $\beta$ -carotene isomers by a multireference configuration interaction method. *Chem Phys* 373:98–103
92. Marian CM, Gilka N (2008) Performance of the density functional theory/multireference configuration interaction method on electronic excitation of extended  $\pi$ -systems. *J Chem Theory Comput* 4:1501–1515
93. Grimme S, Waletzke M (1999) A combination of Kohn-Sham density functional theory and multi-reference configuration interaction methods. *J Chem Phys* 111:5645–5655
94. Send R, Sundholm D, Johansson MP, Pawłowski F (2009) Excited-state potential-energy surfaces of polyenes and protonated Schiff bases. *J Chem Theory Comput* 5:2401–2414
95. Jacquemin D, Wathelet V, Perpète EA, Adamo C (2009) Extensive TD-DFT benchmark: singlet-excited states of organic molecules. *J Chem Theory Comput* 5:2420–2435
96. Thompson MJ, Bashford D, Noddleman L, Getzoff ED (2003) Photoisomerization and Proton Transfer in Photoactive Yellow Protein. *J Am Chem Soc* 125:8186–8194
97. Chiba M, Fedorov DG, Kitaura K (2007) Time-dependent density functional theory based upon the fragment molecular orbital method. *J Chem Phys* 127:104108
98. Vendrell O, Gelabert R, Moreno M, Lluch JM (2004) Photoinduced proton transfer from the green fluorescent protein chromophore to a water molecule: analysis of the transfer coordinate. *Chem Phys Lett* 396:202–207
99. Nemukhin AV, Topol IA, Burt SK (2006) Electronic excitations of the chromophore from the fluorescent protein asFP595 in solutions. *J Chem Theory Comput* 2:292–299
100. Wanko M, Hoffmann M, Strodel P, Koslowski A, Thiel W, Neese F, Frauenheim T, Elstner M (2005) Calculating absorption shifts for retinal proteins: computational challenges. *J Phys Chem B* 109:3606–3615
101. Altun A, Yokoyama S, Morokuma K (2008) Mechanism of spectral tuning going from retinal in vacuo to bovine rhodopsin and its mutants: multireference ab initio quantum mechanics/molecular mechanics studies. *J Phys Chem B* 112:16883–16890
102. Kaila VRI, Send R, Sundholm D (2012) The effect of protein environment on photoexcitation properties of retinal. *J Phys Chem B* 116:2249–2258
103. Dewar MJS (1975) Quantum organic chemistry. *Science* 187:1037–1044
104. Acevedo O, Jorgensen WL (2010) Advances in quantum and molecular mechanical (QM/MM) simulations for organic and enzymatic reactions. *Acc Chem Res* 43:142–151
105. Warshel A, Weiss RM (1980) An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J Am Chem Soc* 102:6218–6226
106. Schmitt UW, Voth GA (1998) Multistate empirical valence bond model for proton transport in water. *J Phys Chem B* 102:5547–5551
107. Jensen F, Norrby PO (2003) Transition states from empirical force fields. *Theor Chem Acc* 109:1–7
108. Pople JA, Santry DP, Segal GA (1965) Approximate self-consistent molecular orbital theory. I. Invariant procedures. *J Chem Phys* 43:S129–S135

109. Dewar MJS, Thiel W (1977) Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J Am Chem Soc* 99:4899–4907
110. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J Am Chem Soc* 107:3902–3909
111. Stewart JJP (1989) Optimization of parameters for semiempirical methods I: method. *J Comput Chem* 10:209–220
112. Stewart JJP (2007) Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J Mol Model* 13:1173–1213
113. Ridley J, Zerner M (1973) Intermediate neglect of differential overlap technique for spectroscopy—pyrrole and azines. *Theor Chem Acc* 32:111–134
114. Zheng XH, Stuchebrukhov AA (2003) Electron tunneling in proteins: implementation of ZINDO model for tunneling currents calculations. *J Phys Chem B* 107:6621–6628
115. Hayashi T, Stuchebrukhov AA (2010) Electron tunneling in respiratory complex I. *Proc Natl Acad Sci USA* 107:19157–19162
116. Ball DM, Buda C, Gillespie AM, White DP, Cundari TR (2002) Can semiempirical quantum mechanics be used to predict the spin state of transition metal complexes? an application of De Novo prediction. *Inorg Chem* 41:152–156
117. Gorelsky SI, Lever ABP (2001) Electronic structure and spectra of ruthenium diimine complexes by density functional theory and INDO/S. Comparison of the two methods. *J Organomet Chem* 635:187–196
118. Lalia-Kantouri M, Papadopoulos CD, Quirós M, Hatzidimitriou AG (2007) Synthesis and characterization of new Co(III) mixed-ligand complexes, containing 2-hydroxy-aryloximes and  $\alpha$ -diimines. Crystal and molecular structure of  $[\text{Co}(\text{saox})(\text{bipy})_2]\text{Br}$ . *Polyhedron* 26:1292–1302
119. Korth M, Pitoňák M, Řezáč J, Hobza P (2010) A transferable H-bonding correction for semiempirical quantum-chemical methods. *J Chem Theory Comput* 6:344–352
120. Fanfrlík J, Bronowska AK, Řezáč J, Přenosil O, Konvalinka J, Hobza P (2010) A reliable docking/scoring scheme based on the semiempirical quantum mechanical PM6-DH2 method accurately covering dispersion and H-bonding: HIV-1 protease with 22 ligands. *J Phys Chem B* 114:12666–12678
121. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B* 58:7260–7268
122. Niehaus TA, Suhai S, Della Sala F, Lugli P, Elstner M, Seifert G, Frauenheim T (2001) Tight-binding approach to time-dependent density-functional response theory. *Phys Rev B* 63:085108
123. Elstner M, Cui Q, Munih P, Kaxiras E, Frauenheim T, Karplus M (2003) Modeling zinc in biomolecules with the self consistent charge-density functional tight binding (SCC-DFTB) method: applications to structural and energetic analysis. *J Comput Chem* 24:565–581
124. Cui Q, Elstner M, Karplus M (2002) A theoretical analysis of the proton and hydride transfer in liver alcohol dehydrogenase (LADH). *J Phys Chem B* 106:2721–2740
125. Liu HY, Elstner M, Kaxiras E, Frauenheim T, Hermans J, Yang W (2001) Quantum mechanics simulation of protein dynamics on long timescale. *Protein Struct Funct Gen* 44:484–489
126. Dunning Jr TH (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J Chem Phys* 90:1007–1023
127. Hehre WJ, Stewart RF, Pople JA (1969) Self-consistent molecular-orbital methods. I. Use of gaussian expansions of slater-type atomic orbitals. *J Chem Phys* 51:2657–2664
128. Schäfer A, Horn H, Ahlrichs R (1992) Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J Chem Phys* 97:2571–2577
129. Schäfer A, Huber C, Ahlrichs R (1994) Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J Chem Phys* 100:5829–5835
130. Jensen F (2001) Polarization consistent basis sets: principles. *J Chem Phys* 115:9113–9125
131. Rappoport D, Furche F (2010) Property-optimized Gaussian basis sets for molecular response calculations. *J Chem Phys* 133:134105
132. Manninen P, Vaara J (2006) Systematic Gaussian basis-set limit using completeness-optimized primitive sets. A case for magnetic properties. *J Comput Chem* 27:434–445
133. Jensen F (2010) The optimum contraction of basis sets for calculating spin-spin coupling constants. *Theor Chem Acc* 126:371–382

# Chapter 2

## Ab Initio Molecular Dynamics

Kari Laasonen

### Abstract

In this chapter, an introduction to ab initio molecular dynamics (AIMD) has been given. Many of the basic concepts, like the Hellman–Feynman forces, the difference between the Car–Parrinello molecular dynamics and AIMD, have been explained. Also a very versatile AIMD code, the CP2K, has been introduced. On the application, the emphasis was on the aqueous systems and chemical reactions. The biochemical applications have not been discussed in depth.

**Key word:** Ab initio molecular dynamics, Car–Parrinello method, Simulations of liquid water, Proton transfer

---

### 1. Background

The atoms and molecules are always in movement. Even at very low temperatures the atoms move due to the quantum motion, but by far the more important movement is the thermal motion. Because they move we should be able to model this movement. This can be done using empirical potentials and describing atoms as point particles (empirical molecular dynamics, EMD). The situation is more delicate when the system is described using quantum mechanics. In quantum mechanics both the electrons and nuclei are described using the Schrödinger equation:

$$i\hbar \frac{\partial \Psi(r_i, R_I, t)}{\partial t} = H\Psi(r_i, R_I, t), \quad (1)$$

$$H = -\sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + H_e, \quad (2)$$

$$H_e = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 + \sum_{i < j} \frac{e^2}{|r_i - r_j|} - \sum_{i,I} \frac{e^2 Z_I}{|r_i - R_I|} + \sum_{I < J} \frac{e^2 Z_I Z_J}{|R_I - R_J|},$$

where  $R_I$  is atomic position,  $Z_I$  is atomic charge,  $M_I$  is atomic mass,  $r_i$  is electronic position, and  $m_e$  is electronic mass.

It is rather delicate to extract the classical Newtonian dynamics from the Schrödinger equation. An interested reader can find the details from a review by Marx and Hutter (1). In the Ehrenfest molecular dynamics the electronic excitations can be taken into account in the mean-field approximation. Even the simplest version where only the ground state wave function is included, this approach involves solution of the time-dependent Schrödinger equation. This approach has not been used much, but it can be implemented rather efficiently (2):

$$M_I \frac{d^2 R_I(t)}{dt^2} = -\nabla_I \langle \Psi_0 | H_e | \Psi_0 \rangle, \quad (3)$$

$$i\hbar \frac{\partial |\Psi_0\rangle}{\partial t} = H_e |\Psi_0\rangle. \quad (4)$$

A simpler approach where only the ground state wave function is included and the wave function is assumed to be always at the minimum is called the Born–Oppenheimer molecular dynamics (BOMD). This is the most commonly used version of ab initio molecular dynamics (AIMD):

$$M_I \frac{d^2 R_I(t)}{dt^2} = -\nabla_I \min_{\Psi} \langle \Psi_0 | H_e | \Psi_0 \rangle, \quad (5)$$

$$H_e |\Psi_0\rangle = E_0 |\Psi_0\rangle. \quad (6)$$

Note that  $H_e$  depends on the atomic positions.

The next step is to solve the electronic Schrödinger equation. In the general case this is not possible, and some approximations are needed. These methods are discussed more in the chapter by Johansson, Kaila, and Sundholm. In AIMD we are interested in single-determinant methods like Hartree–Fock (HF) or density functional theory (DFT) where the many-electron wave function can be written as  $\Psi_0 = \det |\psi_1, \psi_2, \dots, \psi_N|$  with some orthonormal single-particle orbitals;  $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ . Even with approximate methods (DFT or HF), the computational cost of AIMD is very high. One needs to do a medium size quantum chemical calculation at every molecular dynamics time step. This limits significantly the size of the system that can be studied.

There are several publications of AIMD explaining the computational and theoretical details, but they cannot be reviewed here. I recommend the review (1) which gives a relatively complete view of the field until 2000 and a book by the same authors (3). Also the lectures in the CECAM tutorial of CP2K give up to date information on AIMD (4). Additional sources of information are the www pages of the major AIMD codes (5, 6). For example, the CPMD www page has a rather long list of publications in this field. In this chapter, I take a more practical view and explain some basics of the AIMD and motivate its usefulness.

The AIMD can be solved with two main methods, the Car–Parrinello method and using directly the Born–Oppenheimer equation (Eq. 5). The drawback of Eq. 5 is that the wave function has to be optimized at every time step.

### 1.1. Forces

The force acting on atoms can be computed from the term

$$F_I = -\nabla_I \min_{\Psi} \langle \Psi_0 | H_e | \Psi_0 \rangle. \quad (7)$$

If the wave function does not have explicit atomic position dependence the forces can be computed from a simpler equation proposed by Hellman and Feynman,

$$F_I^{HF} = -\min_{\Psi} \langle \Psi_0 | \nabla_I H_e | \Psi_0 \rangle. \quad (8)$$

In a more general case the derivatives of the wave function ( $\nabla_I |\Psi_0\rangle \neq 0$ ) have to be taken into account which complicates the calculations. These are called the Pulay corrections (or Pulay forces).

## 2. Car–Parrinello Method

The Car–Parrinello (CP) method (or CPMD) is based on an elegant idea of adiabaticity. The electrons are much lighter than the atoms, and thus they should follow (adiabatically) the atoms. This means that the explicit wave function optimization is not necessary. On the other hand, the orthonormality constraint of the single-particle orbitals has to be maintained. The orthonormality can be preserved using the method of Lagrange multipliers. Constraint terms ( $\mathcal{G}_{ij} = \langle \psi_i | \psi_j \rangle - \delta_{ij} = 0$ ) with Lagrange multipliers ( $\Lambda_{ij}$ ) can be added to the energy

$$\sum_{ij} \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}). \quad (9)$$

During the dynamics the constraints  $\mathcal{G}_{ij} = 0$  have to be kept, and this defines the Lagrange multipliers. Computationally this orthogonalization is one of the most expensive parts of the code. It scales as  $N_e^2 N_{\text{basis}}$ , where the  $N_e$  is the number of electrons and  $N_{\text{basis}}$  is the number of basis functions. The same Lagrange multiplier technique is used to keep the bond constraints in empirical MD.

The CPMD consists of two dynamical equations, one for the nuclei and another for the electrons:

$$M_I \frac{d^2 R_I(t)}{dt^2} = -\nabla_I [\langle \Psi_0 | H_e | \Psi_0 \rangle + \sum_{ij} \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij})], \quad (10)$$

$$\mu \frac{d^2 \psi_i(t)}{dt^2} = -\frac{\partial}{\partial \psi_i} [\langle \Psi_0 | H_e | \Psi_0 \rangle + \sum_{ij} \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij})]. \quad (11)$$

The time evolution of both  $R(t)$  and  $\psi(t)$  can be solved from these equations. Note that Eq. 11 does not correspond to the true dynamics of the wave function which is described by the time-dependent Schrödinger equation. This is a fictitious dynamics that allows the electrons to follow the nuclei.

There is an interesting consequence of the Pulay forces in the Car–Parrinello dynamics. If the basis set depends on the atomic positions, also the overlap matrix will depend on them and the constraint term will have position dependence (1):

$$-\nabla_I \langle \psi_i | \psi_j \rangle \neq 0. \quad (12)$$

This term is missing from the publications since CPMD was originally developed for using the plane wave basis which does not depend on the atomic positions. Inclusion of this term will complicate the CP calculations using localized basis functions significantly, and BOMD (see Subheading 3) is more suitable when, for example, Gaussians are used as basis.

The CP method contains only one free parameter, the electron mass parameter  $\mu$ . Due to the fictitious dynamics,  $\mu$  does not need to be the electron mass. For sensible results, the starting wave function has to be well converged and the mass parameter much lighter than the lightest atom of the system (usually hydrogen or deuterium). The CPMD works well if the system has a stable electronic structure. This in practice means a large HOMO–LUMO gap (or band gap), which is the case with most chemically stable molecules.

The strength of the CPMD approach is that the wave function does not need to be optimized during the dynamics. The main drawback is that the (fictitious) electron dynamics is much faster than the nuclear motion, and a smaller time step for integration of the CPMD equations is needed. This time step depends on the mass parameter  $\mu$ . For hydrogen-containing systems, reasonable values are  $\mu = 300m_e$  and  $\Delta t = 0.1$  fs. If one replaces the hydrogen with deuterium, the mass and time step can be increased slightly. Still, this time step is an order of magnitude smaller than in empirical MD. Thus, there is a prize to pay for neglecting the wave function optimization. Another minor drawback is that the movement of the reasonably massive electrons slows down the atomic motion. (A similar effect has been seen in the Ehrenfest dynamics implementation in (2).) This will, for example, cause a red shift of the hydrogen vibration frequencies.

On the other hand, CPMD will produce very stable trajectories, and often,  $10^6$  CPMD steps can be run with a very small drift in energy (1, 4).

---

### 3. Born–Oppenheimer Dynamics

AIMD can be done using Eq. 5 directly. The efficiency of this approach depends critically on how effectively the new wave function can be solved. In standard quantum chemical calculations the wave function solution takes tens of iterations, but the MD approach has one advantage. The wave functions from the few previous time steps can be utilized to make an excellent guess for the new wave function. The time step in BOMD is limited by the atomic dynamics and is similar to the time step in EMD, 1–2 fs. Thus, if the number of the wave function optimization steps can be pushed down to ca. 10, BOMD is computationally comparable to CPMD. With the latest algorithms, this is the case.

The BOMD trajectories are sensitive to the wave function convergence criterion, and good convergence is needed to achieve the same stability as in CPMD. Convergence (the largest element of the wave function gradient) of the order of  $10^{-6}$  is needed. A good presentation of the stability of BOMD can be found in Hutter’s lectures in Ref. (4).

---

### 4. Quantum Chemistry

Besides dynamics, it is essential to choose the quantum chemical method to treat the electrons. Many quantum chemical codes like Gaussian (7) or Turbomole have the possibility to do AIMD, but they are not much used (they are quite slow). Usually AIMD is done with programs that have been designed for it. Such codes are CPMD, CP2K, NWChem, and to some extent VASP and CASTEP. The special AIMD codes do not provide many quantum chemical methods. By far the most used method is DFT using the generalized gradient approximation (GGA) (1). Also hybrid functionals, like B3LYP, can be used (8, 9), but they are computationally much more expensive. Very recently there has been improvement of the efficiency of the hybrid methods (10). The weak van der Waals interactions are not included in the GGA or hybrid functionals, but they can be added empirically to the calculations (11–13).

---

### 5. The Power of AIMD

AIMD calculations, are very expensive but they have one great advantage over the empirical methods—the electrons are included and thus the chemical bonds can break and form. The “force field” is also fully polarizable. Simply the used quantum chemical method sets the limits

of accuracy of the AIMD method. The numerical approximations like the basis set and pseudopotentials can be made quite small. Also the modern AIMD packages include most of the tools of empirical MD. These include the temperature and pressure control, and free energy calculations with different type of constraints. The latest AIMD codes can do similar simulations as empirical codes, but the system sizes and time scales are much smaller. This is of course a serious limitation in biochemical applications, and AIMD has not been much used in biochemistry. This does not mean that AIMD and especially QM/MM AIMD cannot be applied to biochemical problems. Only the number of suitable problems is rather limited. Also the state-of-the-art AIMD calculations are not easy to perform.

---

## 6. Practical AIMD Calculations

There are several computer codes that are designed to do AIMD. These include CPMD (5), CP2K (6), NWChem (14), and many others. Many of these packages are free, but they are not easy to install or use. The author of this chapter has long experience of using the CPMD and CP2K codes, and some comments of the usage of these codes are given below. AIMD simulations are very CPU time-consuming and very efficient parallel computers are needed. Naturally some test calculations can be done using a single PC, but for anything more serious, a supercomputer is needed. It is also important that the computer has fast communication between the processors. In this chapter I cannot go to the details of the code installations. For example, the latest version of CP2K can be downloaded from the Web. In the package, there are instructions for several computer architectures. The presentation of Guidon in Ref. (4) contains a lot of advice for installing CP2K, but very likely you will need also the help of the computer administrator.

For practical calculations one needs to worry about several parameters. These include the size of the system, simulation time, temperature, time step, quantum chemical model, basis set, and pseudopotentials. If the CPMD method is used also the effective electron mass has to be set. Of these the system size has to be set by the user. If aqueous systems are studied, the typical system size is ca. 100 waters with CPMD and up to 250 waters with CP2K. The accessible time scale is up to a few hundred picoseconds. Very often the simulations done are much shorter than 100 ps, but in my experience, reasonable equilibration requires usually around 25 ps. Often one needs to push the AIMD simulations as long as possible. Other solvents than water are harder to study since they contain more electrons. Systems like ionic liquids have been studied, but the system sizes and simulation times are very limited (15). In biochemistry several processes are much slower than this.

In some cases, constrained molecular dynamics can be used to study slow processes, but clearly, many biochemical problems are not accessible with AIMD.

The simulation temperature is often 300 K, but the temperature should be set to a bit higher, around 350 K to compensate for slow diffusion of water. The time step of a hydrogen- or deuterium-containing system in CPMD is 0.1–0.15 fs and in BOMD ca. 1 fs. As said earlier, the electrons are usually treated with DFT-GGA and the most used models are PBE or BLYP. I would recommend using the empirical van der Waals corrections in all calculations.

The basis set and pseudopotential combination is somewhat tricky. Some of the AIMD codes (CPMD, VASP, NWChem) use plane waves as the basis set, and they are very sensitive to the pseudopotentials. To keep the number of basis functions reasonably small, smooth pseudopotentials have to be used. For a nonexpert, the available pseudopotentials should be used, but care of their smoothness should be taken. In the CPMD package, the Troullier–Martins type pseudopotentials can be used. In VASP the PAW pseudopotentials are smooth and mostly reliable. The ultrasoft pseudopotentials proposed by Vanderbilt produce very smooth orbitals, and thus a low cutoff can be used, but the libraries of these pseudopotentials are not very complete.

The CP2K code uses Gaussians as basis. In the case of a Gaussian basis, the smoothness of the pseudopotentials is not an essential requirement, and CP2K has a rather complete pseudopotential library by Goedecker, Teter, and Hutter (16). Also all-electron calculations can be made.

Most of the used AIMD codes have a rather tricky interface. One of the most powerful AIMD codes, the CP2K, has a particularly complex interface, and the CP2K package does not have input examples. The presentations in the CP2K tutorial (4) are of some help, but unfortunately, the actual input examples are not included in the tutorial. This is annoying for someone who would like to do simple CP2K calculations, but the reason for this is that the CP2K is a complex code and the user should understand the input file, not just copy it from somewhere. To lower the barrier to use CP2K, I have included a simple example of 32 water molecules at the end of this chapter.

### 6.1. Equilibration

Because the AIMD time scale is quite short, a lot of care has to be paid in the system equilibration. If similar empirical MD can be done, it would work as good equilibration. This is a bit dangerous since then the equilibration is biased by the EMD force field and in short simulations, AIMD and EMD can seemingly agree very well. If EMD is not available (or it is not wanted to be used), a rather careful local optimization is needed before AIMD simulations. In any case, one needs to remember that molecular equilibration is a rather slow process compared to the AIMD time scale.

## 7. AIMD of Water

Liquid water has been the standard test for AIMD (17, 18). Naturally if AIMD (or more precisely the used DFT–GGA approach) fails to produce water properties correctly, it is not very useful for biochemical applications. The PBE or BLYP description of water is not perfect. The O–O pair correlation function is a bit too structured, and the diffusion coefficient is too small by a factor of 0.5. These results are somewhat worrying, and there have been several attempts to correct the DFT–GGA behavior (8, 9, 13). On the other hand, AIMD is the most reliable description of water when chemical reactions or strong solvation effects are considered. AIMD should be used for difficult cases like proton transfer reactions or chemical reactions in water.

## 8. AIMD of Reactions

In chemistry the chemical reactions are often of central importance. AIMD offers an excellent tool to study them. Unfortunately very seldom, the reactions are fast enough to happen in the AIMD (or even empirical MD) time scale. And even when they are fast enough several, events are needed to have a statistically meaningful description of the reaction. One important exception is the proton diffusion, which happens on the picosecond time scale and is accessible with AIMD. The proton diffusion has been studied in water (19), in a simple ion channel (20), ammonia (21) and water-methanol mixture (22). Also several other reactions can be studied with direct AIMD (there are a few examples in Ref. (1)).

With most reactions some method is needed to force the reaction to happen. In traditional quantum chemistry various transition-state search algorithms are used, but such approach is not useful in dynamical systems where the system is not in a local minimum. We need some biased dynamics, and the simplest method is to use a constraint to force the reaction to happen. The free energy of the reaction can be calculated by integrating the constraint force,  $F_g$  ( $g$  denoting some value of the constraint),

$$\Delta A = \int_{g_0}^{g_1} \langle F_g \rangle_T dg , \quad (13)$$

where  $\langle F_g \rangle_T$  means the average at temperature  $T$ . Also normal thermodynamic integration can be done.

The simplest constraint is to fix the distance (or angle, or torsional angle) between two atoms, but many of the codes offer relatively complex choices of constraints. One can, for example, use

the coordination numbers as constraints. The calculation of the free energy differences requires quite a lot of computer time since to integrate the force several constraint values are needed, and due to the molecular collisions, the instantaneous constraint forces are very noisy and reasonable averaging has to be done at each constraint value. Simulations of the order of 10 ps are needed for converged results for each constraint value (23–25).

It is useful to check if similar barriers are achieved when the reaction is studied in both ways. If the reaction  $A + B \rightarrow C$  is followed to some point, for example, to the transition state  $C^\#$ , the calculations can be reversed, and when ended to the original state, the same energy barrier should be achieved. In practice there is always some hysteresis, and this path reverse test gives some idea of the size of the hysteresis.

More sophisticated biased simulations can be done using different algorithms including metadynamics. This is a very interesting field, and for more details, I refer to the CECAM CP2K Tutorial (4) and the presentations by Iannuzzi.

To summarize, the reaction calculations are probably the most important AIMD calculations. They are not as difficult as they look, but some toy projects should be done to learn how these calculations work. Research projects are usually so demanding that very little experimentation can be done. Whatever constraint or method one is using, the thermal collisions will cause large fluctuations to the constraint, and long averaging is needed. One should be prepared to do a total of 100–200 ps AIMD simulations. The only positive thing is that many of these calculations can be run simultaneously.

## 9. AIMD in Biochemistry

Biochemical applications of AIMD are not very common due to the slowness of the biochemical processes. Among the pioneers in this field are Ursula Röthlisberger (26, 27) and Paolo Carloni, but recently, of course many other groups have entered this field. In biochemical applications it is almost impossible to study the relevant system fully with *ab initio* methods, thus, the QM/MM modeling has become popular. There are a few chapters on QM/MM methods and applications in this book, so here I do not go to details of this type of modeling. It is still useful to mention that in most of the QM/MM calculations, the quantum atoms do not follow any dynamics. Thus, the developed AIMD-QM/MM methods are very attractive (some examples are given in the lecture by Kuo in Ref. (4)). In those methods the whole system, both the QM and MM parts, can move simultaneously. The QM part will determine the computational cost of the calculations, and its size should be small.

Unfortunately, the QM/MM calculations are rather complicated since one needs to understand/model the QM, MM parts, and the boundary between them. I do not have experience of AIMD-QM/MM calculations, so it is safer that I do not go into details. This field is in rapid development the algorithms become faster and the programs easier to use. Especially the latter is important for a non-expert user.

---

## 10. Exotic AIMD

I want to mention a couple of less common AIMD methods. The excited state dynamics have interesting applications in biochemistry, but the path integral method will probably have less biochemical applications.

### **10.1. Quantum Motion of Nuclei: Path Integral Ab Initio Molecular Dynamics**

In the beginning, the Born–Oppenheimer approximation was used and the atoms were assumed as point particles. This is not necessary also the nuclear quantum motion can be modeled using the path integral molecular dynamics (PIMD) technique. The relevant equations are rather complex, and the details can be found in Refs. (1, 28). The PIMD has been implemented in CPMD and CP2K. The PIMD treatment is very interesting in the case of frustrated hydrogens, like in  $\text{CH}_5^+$  or protons in water (29). Again the calculations are rather time-consuming, but new phenomena can be modelled.

### **10.2. Excited States**

The static calculations of excited states are now quite routine. Most of the quantum chemical codes have routines to estimate the excited states. The time-dependent DFT (TD-DFT) is a very popular method and the computations are quite easy, but the results are not always very reliable. On the other hand, the molecular dynamics of molecules at excited states is not very developed, but it is an interesting direction for modeling. In Ref. (1), some basics of the excited state dynamics are described. Röthlisberger’s group has been active in this field, for one of their recent articles see Ref. (30). The methods capable of modelling excited states open possibilities for studying the very important photochemical reactions in biochemistry.

---

## 11. AIMD Has To Be Used: In Some Cases

I hope I have not been too pessimistic of the possibilities of AIMD in biochemistry. In my opinion it is a very valuable method for studying difficult topics like metal centers in proteins (31), chemical reactions in active sites, proton transfer reactions, properties of

excited states, etc. In my view any molecular modeling of proteins or biochemical systems is difficult since the time scales of these systems are long. In every case the system one wants to model has to be prepared carefully to be able to get meaningful results. Compared to EMD, the size and time limitations of AIMD are greater, but AIMD can solve problems EMD cannot.

The question is not whether one likes or does not like AIMD (or more generally quantum chemical methods) but that in some cases the QM methods have to be used. The information of atomic positions (and velocities) is not always enough, one also needs to know how the electrons behave.

## 12. Appendix: An Example Input for CP2K

Here is a simple input example for 32 water molecules. The POTENTIAL and BASIS\_SET can be found from the CP2K package in the directory cp2k/test/QS. In this directory there are also a lot of test inputs that are useful when new types of runs are planned. A word of warning: They cannot be used as they are but often they complement the manual. A very similar input can be found in the file cp2k/tests/QS/benchmark/H2O-32.inp. (See the actual coordinates from this file).

```
&GLOBAL
PREFERRED_FFT_LIBRARY FFTSG
PROJECT w32-test
RUN_TYPE MD
PRINT_LEVEL LOW
&END GLOBAL

&MOTION
&MD
ENSEMBLE NVT
STEPS 10000
Timestep 1.0
&THERMOSTAT
TYPE NOSE
&NOSE
TIMECON 310.0
&END NOSE
&END THERMOSTAT
TEMPERATURE 348.15
&END MD
&END MOTION
```

```
&EXT_RESTART
# EXTERNAL_FILE w32-test-1.restart # exist only after the first run
&END

&FORCE_EVAL
  METHOD Quickstep
  &DFT
    BASIS_SET_FILE_NAME ../cp2k_lib/BASIS_SET
    POTENTIAL_FILE_NAME ../cp2k_lib/POTENTIAL
    &MGRID
      CUTOFF 280
    &END MGRID

  &QS
    EPS_DEFAULT 1.0E-12
    EXTRAPOLATION ASPC
    EXTRAPOLATION_ORDER 3
  &END QS
  &SCF
    SCF_GUESS ATOMIC # for first time, next run SCF_GUESS RESTART
    EPS_SCF 0.4E-6 # this is important max is 1E-6
    MAX_SCF 25
    &OUTER_SCF
      EPS_SCF 1.0E-6
      MAX_SCF 2
    &END
    &OT
      MINIMIZER DIIS
      PRECONDITIONER FULL_ALL
      ENERGY_GAP 0.05
    &END
  &END SCF
  &XC
    &XC_FUNCTIONAL BLYP
  &END XC_FUNCTIONAL
  &END XC
&END DFT

&SUBSYS
  &CELL
    ABC 9.865 9.865 9.865
  &END CELL
  &COORD
```

```

0      2.280398    9.146539    5.088696
0      1.251703    2.406261    7.769908
0      1.596302    6.920128    0.656695
0      2.957518    3.771868    1.877387
. . . . find coordinates from cp2k/tests/QS/benchmark/H2O-32.inp
H      1.762019    9.820429    5.528454
H      3.095987    9.107088    5.588186
H      0.554129    2.982634    8.082024
H      1.771257    2.954779    7.182181
H      2.112148    6.126321    0.798136
H      1.776389    7.463264    1.424030
. . . .
&END COORD

&KIND O
  BASIS_SET TZVP-GTH
  POTENTIAL GTH-BLYP-q6
&END KIND

&KIND H
  BASIS_SET TZVP-GTH
  POTENTIAL GTH-BLYP-q1
  MASS 2.0000
&END KIND

&END SUBSYS
&END FORCE_EVAL

```

## References

- Marx D, Hutter J (2000) Ab initio molecular dynamics: theory and implementation. In: Grotendorst J (ed) NIC series volume 1: modern methods and algorithms of quantum chemistry proceedings. PDF file: <http://www.fz-juelich.de/nic-series/Volume1/>
- Andrade X, Castro A, Zueco D, Alonso JL, Echenique P, Falceto F, Rubio A (2009) Modified Ehrenfest formalism for efficient large-scale ab initio molecular dynamics. *J Chem Theory Comput* 5:728–742
- Marx D, Hutter J (2009) Ab initio molecular dynamics: basic theory and advanced methods, Oxford University Press, Oxford.
- 1st CP2K Tutorial: Enabling the Power of Imagination in MD Simulations. February 9, 2009. Link to the PDF files: <http://www.cp2k.org/index.php/cp2k-events>
- [www.cpmd.org](http://www.cpmd.org)
- [www.cp2k.org](http://www.cp2k.org)
- [www.gaussian.com](http://www.gaussian.com). See also: Schlegel HB, Millam JM, Iyengar SS, Voth GA, Scuseria GE, Daniels AD, Frisch MJ (2001) Ab initio molecular dynamics: propagating the density matrix with Gaussian orbitals. *J Chem Phys* 114:9758–9763
- Todorova T, Seitsonen AP, Hutter J, Kuo I-FW, Mundy CJ (2006) Molecular dynamics simulation of liquid water: hybrid density functionals. *J Phys Chem B* 110:3685–3691
- Guidon M, Schiffmann F, Hutter J, VandeVondele J (2008) Ab initio molecular dynamics using hybrid density functional. *J Chem Phys* 128:214104

10. Guidon M, Hutter J, VandeVondele J (2010) Auxiliary density matrix methods for Hartree-Fock exchange calculations. *J Chem Theory Comput* 6:2348–2364
11. Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem* 27:1787–1799
12. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* 132:154104
13. Lin I-C, Seitsonen AP, Coutinho-Neto MD, Tavernelli I, Rothlisberger U (2009) Importance of van der Waals interactions in liquid water. *J Phys Chem B* 113:1127–1131
14. [www.nwchem-sw.org](http://www.nwchem-sw.org); Valiev M, Bylaska EJ, Govind N, Kowalski K, Straatsma TP, van Dam HJJ, Wang D, Nieplocha J, Apra E, Windus TL, de Jong WA (2010) NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput Phys Commun* 181:1477
15. Kirchner B, Seitsonen AP (2007) Ionic liquids from Car–Parrinello simulations. 2. Structural diffusion leading to large anions in chloroaluminate ionic liquids. *Inorg Chem* 46:2751–2754
16. Goedecker S, Teter M, Hutter J (1996) Separable dual-space Gaussian pseudopotentials. *Phys Rev B* 54:1703–1710
17. Laasonen K, Sprik M, Parrinello M, Car R (1993) Ab initio liquid water. *J Chem Phys* 99:9080
18. Kuo I-FW, Mundy CJ, McGrath MJ, Siepmann JI, VandeVondele J, Sprik M, Hutter J, Chen B, Klein ML, Mohamed F, Krack M, Parrinello M (2004) Liquid water from first principles: investigation of different sampling approaches. *J Phys Chem B* 108:12990–12998
19. Tuckerman M, Laasonen K, Sprik M, Parrinello M (1995) Ab initio molecular dynamics simulation of solvation and transport of  $\text{H}_3\text{O}^+$  and  $\text{OH}^-$  ions in water. *J Chem Phys* 103:150
20. Sagnella D, Laasonen K, Klein ML (1996) Ab initio molecular dynamics study of proton transfer in poly-Glycine analog of the ion channel gramicidin A. *Biophys J* 71:1172
21. Liu Y, Tuckerman ME (2001) Protonic defects in hydrogen bonded liquids: structure and dynamics in ammonia and comparison with water. *J Phys Chem B* 105:6598–6610
22. Morrone JA, Haslinger KE, Tuckerman ME (2006) Ab initio molecular dynamics simulation of the structure and proton transport dynamics of methanol-water solutions. *J Phys Chem B* 110:3712–3720
23. Saukkoriipi J, Laasonen K (2008) Density functional studies of the hydrolysis of aluminum (chloro)hydroxide in water with CPMD and COSMO. *J Phys Chem A* 112:10873–10880
24. Doltsinis NL, Sprik M (2003) Theoretical  $\text{pK}_a$  estimates for solvated  $\text{P}(\text{OH})_5$  from coordination constrained Car–Parrinello molecular dynamics. *Phys Chem Chem Phys* 5:2612–2618
25. Timko J, Bucher D, Kuyucak S (2010) Dissociation of NaCl in water from ab initio molecular dynamics simulations. *J Chem Phys* 132:114510
26. Garrec J, Cascella M, Rothlisberger U, Fleurat-Lessard P (2010) Low inhibiting power of N···CO based peptidomimetic compounds against HIV-1 protease: insights from a QM/MM study. *J Chem Theory Comput* 6:1369–1379
27. Bucher D, Guidoni L, Maurer P, Rothlisberger U (2009) Developing improved charge sets for the modeling of the KcsA  $\text{K}^+$  channel using QM/MM electrostatic potentials. *J Chem Theory Comput* 5:2173–2179
28. Marx D, Tuckerman ME, Martyna GJ (1999) Quantum dynamics via adiabatic ab initio centroid molecular dynamics. *Comput Phys Commun* 118:166–184
29. Marx D, Parrinello M (1999) Molecular spectroscopy— $\text{CH}_5^+$ : the Cheshire cat smiles. *Science* 284:59; Marx D, Tuckerman ME, Hutter J, Parrinello M (1999) The nature of the hydrated excess proton in water. *Nature* 397:601–604
30. Tapavicza E, Tavernelli I, Rothlisberger U (2009) Ab initio excited state properties and dynamics of a prototype  $\sigma$ -bridged-donor-acceptor molecule. *J Phys Chem A* 113:9595–9602
31. Rovira C, Schulze B, Eichinger M, Evanseck, JD, Parrinello M (2001) Influence of the heme pocket conformation on the structure and vibrations of the Fe–CO bond in myoglobin: a QM/MM density functional study. *Biophys J* 81:435–445

# Chapter 3

## Introduction to QM/MM Simulations

Gerrit Groenhof

### Abstract

Hybrid quantum mechanics/molecular mechanics (QM/MM) simulations have become a popular tool for investigating chemical reactions in condensed phases. In QM/MM methods, the region of the system in which the chemical process takes place is treated at an appropriate level of quantum chemistry theory, while the remainder is described by a molecular mechanics force field. Within this approach, chemical reactivity can be studied in large systems, such as enzymes. In the first part of this contribution, the basic methodology is briefly reviewed. The two most common approaches for partitioning the two subsystems are presented, followed by a discussion on the different ways of treating interactions between the subsystems. Special attention is given on how to deal with situations in which the boundary between the QM and MM subsystems runs through one or more chemical bonds. The second part of this contribution discusses what properties of larger system can be obtained within the QM/MM framework and how. Finally, as an example of a QM/MM application in practice, the third part presents an overview of recent QM/MM molecular dynamics simulations on photobiological systems. In addition to providing quantities that are experimentally accessible, such as structural intermediates, fluorescence lifetimes, quantum yields and spectra, the QM/MM simulations also provide information that is much more difficult to measure experimentally, such as reaction mechanisms and the influence of individual amino acid residues.

**Key word:** Quantum mechanics, Molecular mechanics, QM/MM, Molecular dynamics

---

### 1. Introduction

In this chapter we present a short introduction into the development and application of computational techniques for modelling chemical reactions in the condensed phase. We start by reviewing the basic concepts of these methods. We then discuss how these methods can be used in practical computations and what kind of information can be obtained. We conclude this chapter with a short review of an application on a photobiological system, for which the simulations not only revealed the detailed sequence of events that follow photon absorption but also demonstrate how the biological environment controls the photochemical reaction.

## 2. QM/MM: Theory and Implementation

The size and complexity of a typical biomolecular system, together with the timescales that must be reached, necessitate the use of classical molecular dynamics for the nuclear degrees of freedom. In molecular dynamics (MD) simulations, Newton's equations of motion are solved numerically to obtain a trajectory of the dynamics of a molecule over a period of time (1). To model electronic rearrangements during a chemical reaction, a quantum mechanical description (QM) is required for those parts of the system that are involved in the reaction. For the remainder, a simple molecular mechanics force field model suffices (MM). The interactions in the system are thus computed within a hybrid QM/MM framework.

### 2.1. Molecular Mechanics

Molecular dynamics simulations of biological systems have come of age (2). Since the first application of MD on a small protein in vacuum more than three decades ago (3), advances in computer power, algorithmic developments and improvements in the accuracy of the used interaction functions have established MD as an important and predictive technique to study dynamic processes at atomic resolution (4). In the interaction functions, the so-called molecular mechanics force field, simple chemical concepts are used to describe the potential energy of the system (1):

$$V_{\text{MM}} = \sum_i^{N_{\text{bonds}}} V_i^{\text{bond}} + \sum_j^{N_{\text{angles}}} V_j + \sum_l^{N_{\text{torsions}}} V_l^{\text{torsion}} + \sum_i^{N_{\text{MM}}} \sum_{j>i}^{N_{\text{MM}}} V_{ij}^{\text{Coul}} + \sum_i^{N_{\text{MM}}} \sum_{j>i}^{N_{\text{MM}}} V_{ij}^{\text{LJ}}, \quad (1)$$

where  $N_{\text{MM}}$  is the number of atoms in the system. Bonds and angles ( $V^{\text{bond}}$ ,  $V^{\text{angle}}$ ) are normally modelled by harmonic functions, and torsions by periodic functions ( $V^{\text{torsion}}$ ). The pairwise electrostatic interaction between atoms with a partial charge ( $Q_i$ ) is given by Coulomb's law:

$$V_{ij}^{\text{Coul}} = \frac{e^2 Q_i Q_j}{4\pi\epsilon_0 R_{ij}}, \quad (2)$$

in which  $R_{ij}$  denotes the interatomic distance,  $e$  the unit charge and  $\epsilon_0$  the dielectric constant. Van der Waals interactions, for example the combination of short-range Pauli repulsion and long-range dispersion attraction, are most often modelled by the Lennard-Jones potential:

$$V_{ij}^{\text{LJ}} = \left( \frac{C_{12}^{ij}}{R_{ij}} \right)^{12} - \left( \frac{C_6^{ij}}{R_{ij}} \right)^6, \quad (3)$$

with  $C_{12}^{ij}$  and  $C_6^{ij}$  a repulsion and attraction parameter, respectively, which depend on the atomtypes of the atoms  $i$  and  $j$ .

Electrons are thus ignored in molecular mechanics force fields. Their influence is expressed by empirical parameters that are valid for the ground state of a given covalent structure. Therefore, processes that involve electronic rearrangements, such as chemical reactions, cannot be described at the MM level. Instead, these processes require a quantum mechanics description of the electronic degrees of freedom. However, the computational demand for evaluating the electronic structure places severe constraints on the size of the system that can be studied.

## 2.2. Hybrid Quantum Mechanics/Molecular Mechanics Models

Most biochemical systems, such as enzymes, are too large to be described at any level of *ab initio* theory. At the same time, the available molecular mechanics force fields are not sufficiently flexible to model processes in which chemical bonds are broken or formed. To overcome the limitations of a full quantum mechanical description on the one hand, and a full molecular mechanics treatment on the other hand, methods have been developed that treat a small part of the system at the level of quantum chemistry (QM), while retaining the computationally cheaper force field (MM) for the larger part. This hybrid QM/MM strategy was originally introduced by Warshel and Levitt (5) and is illustrated in Fig. 1. The justification for dividing a system into regions that are described at different levels of theory is the local character of most chemical reactions in condensed phases. A distinction can usually be made between a “reaction centre” with atoms that are directly involved in the reaction and a “spectator” region, in which the atoms do not directly participate in the reaction. For example, most reactions in solution involve the reactants and the first few solvation shells. The bulk solvent is hardly affected by the reaction, but can influence the reaction via long-range interactions. The same is true for most enzymes, in which the catalytic process is restricted to an active site located somewhere inside the protein. The rest of the protein provides an electrostatic background that may or may not facilitate the reaction.

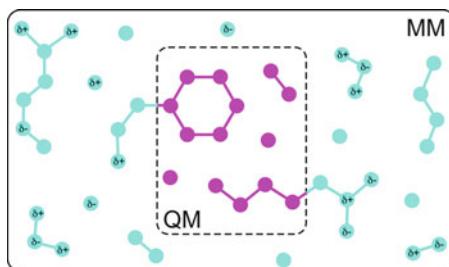


Fig. 1. Illustration of the QM/MM concept. A small region, in which a chemical reaction occurs and therefore cannot be described with a force field, is treated at a sufficiently high level of QM theory. The remainder of the system is modelled at the MM level.

The hybrid QM/MM potential energy contains three classes of interactions: interactions between atoms in the QM region, between atoms in the MM region and interactions between QM and MM atoms. The interactions within the QM and MM regions are relatively straightforward to describe, that is at the QM and MM level, respectively. The interactions between the two subsystems are more difficult to describe, and several approaches have been proposed. These approaches can be roughly divided into two categories: subtractive and additive coupling schemes.

### 2.3. Subtractive QM/MM Coupling

In the subtractive scheme, the QM/MM energy of the system is obtained in three steps. First, the energy of the total system, consisting of both QM and MM regions, is evaluated at the MM level. The QM energy of the isolated QM subsystem is added in the second step. Third, the MM energy of the QM subsystem is computed and subtracted. The last step corrects for including the interactions within the QM subsystem twice:

$$V_{\text{QM/MM}} = V_{\text{MM}}(\text{MM} + \text{QM}) + V_{\text{QM}}(\text{QM}) - V_{\text{MM}}(\text{QM}). \quad (4)$$

The terms QM and MM stand for the atoms in the QM and MM subsystems, respectively. The subscripts indicate the level of theory at which the potential energies ( $V$ ) are computed. The most widely used subtractive QM/MM scheme is the ONIOM method, developed by the Morokuma group (6, 7), and is illustrated in Fig. 2.

The main advantage of the subtractive QM/MM coupling scheme is that no communication is required between the quantum chemistry and molecular mechanics routines. This makes the implementation relatively straightforward. However, compared to the more advanced schemes that are discussed below, there are also disadvantages.

A major disadvantage is that a force field is required for the QM subsystem, which may not always be available. In addition, the force field needs to be sufficiently flexible to describe the effect of chemical changes when a reaction occurs.

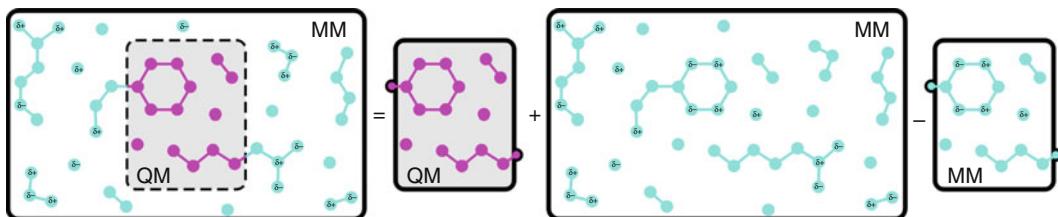


Fig. 2. Subtractive QM/MM coupling: The QM/MM energy of the total system (*left hand side of the equation*) is assumed to be equal to the energy of the isolated QM subsystem, evaluated at the QM level, plus the energy of the complete system evaluated at the MM level, minus the energy of the isolated QM subsystem, evaluated at the MM level. The last term is subtracted to correct for double counting of the contribution of the QM subsystem to the total energy. A prerequisite for the calculation is that a force field for the QM subsystem is available.

A further drawback of this method is absence of polarization of the QM electron density by the MM environment. This shortcoming can be particularly problematic for modelling biological charge transfer processes, since these are usually mediated by the protein environment. For a realistic description of such reactions a more consistent treatment of the interactions between the electrons and their surrounding environment is needed.

## **2.4. Additive QM/MM Coupling**

In additive schemes, the QM system is embedded within the larger MM system, and the potential energy for the whole system is a sum of MM energy terms, QM energy terms and QM/MM coupling terms:

$$V_{\text{QM/MM}} = V_{\text{QM}}(\text{QM}) + V_{\text{MM}}(\text{MM}) + V_{\text{QM-MM}}(\text{QM} + \text{MM}). \quad (5)$$

Here, only the interactions within the MM region are described at the force field level,  $V_{\text{MM}}(\text{MM})$ . In contrast to the subtractive scheme, the interactions between the two subsystems are treated explicitly:  $V_{\text{QM-MM}}(\text{QM} + \text{MM})$ . These interactions can be described at various degrees of sophistication.

### *2.4.1. Mechanical Embedding*

In the most basic approach, all interactions between the two subsystems are handled at the force field level. The QM subsystem is thus kept in place by MM interactions. This is illustrated in Fig. 3. Chemical bonds between QM and MM atoms are modelled by harmonic potentials ( $V^{\text{bond}}$ ), angles defined by one QM atom, and two MM atoms are described by the harmonic potential as well ( $V^{\text{angles}}$ ), while torsions involving at most two QM atoms are commonly modelled by a periodic potential function ( $V^{\text{torsion}}$ ). Non-bonded interactions, that is those between atoms separated by three or more atoms, are also modelled by force field terms: Van der Waals by the Lennard-Jones potential ( $V^{\text{LJ}}$ ) and electrostatics by the Coulomb potential ( $V^{\text{Coul}}$ ). In the most simple implementation of mechanical embedding, the electronic wave function is evaluated for an isolated QM subsystem. Therefore, the MM environment cannot induce polarization of the electron density in the QM region. For calculating the electrostatic interactions between the subsystems, one can either use a fixed set of charges for the QM region, for example, those given by the force field, or re-compute the partial charges on the QM atoms at every integration step of the simulation. In the second strategy, a least-squares fitting procedure is used to derive atomic charges that optimally reproduce the electrostatic potential at the surface of the QM subsystem (8, 9).

Lennard-Jones parameters are normally not updated. Therefore, problems may arise if during the simulation, changes occur in the chemical character of the atoms in the QM region, for example, in reactions that involve changes in the hybridization state of the

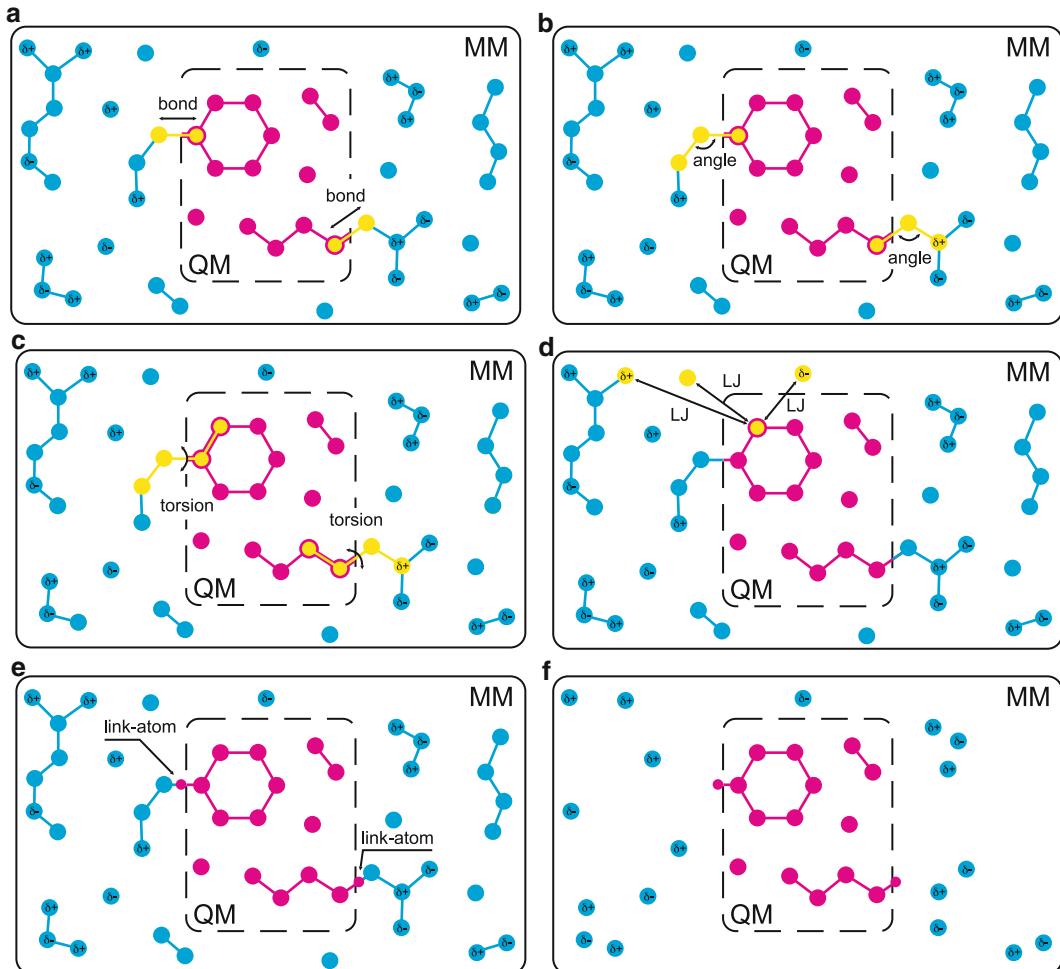


Fig. 3. Coupling between the QM and MM subsystems in the additive QM/MM schemes. The top panels (a)–(c) show bonded interactions between QM and MM atoms. These interactions are handled at the force field level (MM). Panel d shows the Van der Waals interactions between an atom in the QM region and three MM atoms. These interactions are modelled by the Lennard-Jones potential. Panel e illustrates the link atom concept. This atom caps the QM subsystem and is present only in the QM calculation. Panel f demonstrates how the electrostatic QM/MM interactions are handled. In the electrostatic embedding approach, the charged MM atoms enter the electronic Hamiltonian of the QM subsystem. In the mechanical embedding, partial MM charges are assigned to the QM atoms and the electrostatic interactions are computed by the pairwise Coulomb potential.

#### 2.4.2. Electrostatic Embedding

atoms. However, since the Lennard-Jones potential is a relatively short-ranged function, the error introduced by keeping the same parameters throughout the simulations is most likely not very large.

An improvement of mechanical embedding is to include polarization effects. In the electrostatic embedding scheme, the electrostatic interactions between the two subsystems are handled during the

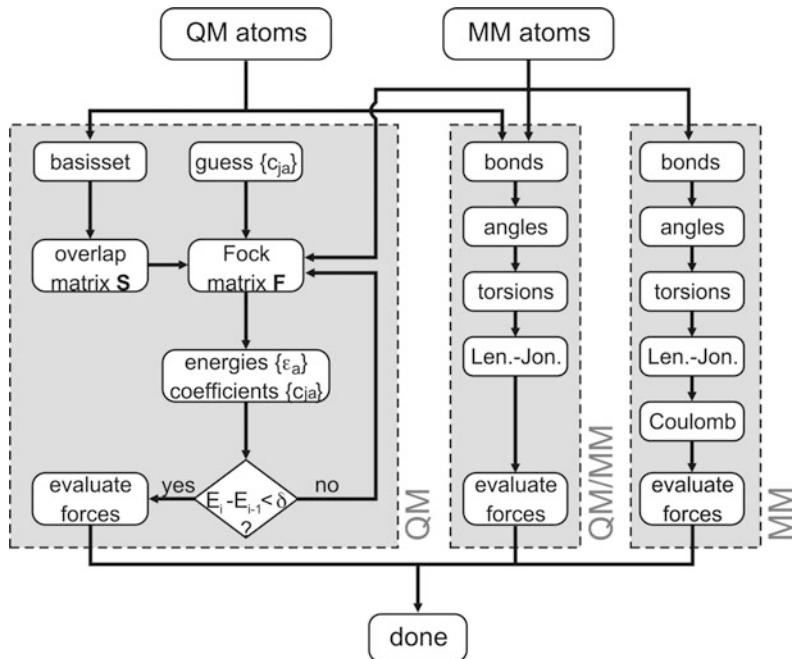


Fig. 4. Flow scheme of a QM/MM energy calculation within the electrostatic embedding scheme. Interactions between atoms in the MM subsystem are handled at the force field level (*third branch*). The QM atoms enter the self-consistent field routine, with the charged MM atoms included as point charges (*first branch*). Diagonalization of the augmented Fock matrix yields the energy of the QM atoms as well as the electrostatic interaction energy between the subsystems. All other interactions involving QM and MM atoms are described by the force field terms (*second branch*).

computation of the electronic wave function. The charged MM atoms enter the QM Hamiltonian as one-electron operators:

$$h_i^{\text{QM-MM}} = h_i^{\text{QM}} - \sum_J^M \frac{e^2 Q_J}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{R}_J|}, \quad (6)$$

where  $\mathbf{r}_i$  and  $\mathbf{R}_J$  are the positions of electron  $i$  and MM atom  $J$ ,  $h_i^{\text{QM}}$  is the original one-electron operator for the kinetic and nuclear attraction energy of electron  $i$  (10) and  $M$  is the number of MM atoms that have a partial charge  $Q_J$ . Thus, the electrons “see” these MM atoms as special nuclei with non-integer and possibly negative charges. Since the electronic Hamiltonian contains extra terms, the electrostatic embedding model requires modifications of the quantum chemistry software. Martin Field and co-workers were among the first researchers to implement this scheme (11) and developed an interface between the molecular mechanics program, Charmm (12) and the *semi-empirical* quantum chemistry package Mopac (13). Figure 4 shows a schematic overview of how the QM and MM routines are

interconnected in a practical implementation of electrostatic embedding. In the electrostatic coupling approach, the MM atoms can polarize the electrons in QM subsystem. However, the atomic charges of the MM atoms have been parametrized to provide a realistic description of an MM system, rather than of a physically correct charge distribution. Therefore, the question arises whether polarization induced by these MM charges is realistic or not. In principle, one would need to re-derive the charges for use in QM/MM frameworks. In reality, interactions between the systems are not only due to electrostatics between charged atoms, but also due to polarization, exchange, charge transfer, dispersion and Pauli repulsion. In force fields, only the *combination* of atomic charges and Lennard-Jones parameters provides a reasonable description of all these effects taken together, albeit in an implicit manner. Part of the interaction due to polarization of the QM region is thus already accounted for by the Lennard-Jones potential. Therefore, not only the MM charges, but also the Lennard-Jones parameters would need to be reparametrized for use in electrostatic embedding QM/MM simulations. However, in practice this is hardly done, and most workers use default force field parameters.

A further problem that may arise when using standard MM atomic charges to describe the charge distribution in the MM system, is the risk of over-polarization near the boundary. The point charges on the MM side of the interface may attract (or repel) the electrons too strongly, which could lead to electron density spilling out into the MM region. Such artefacts can become serious if large flexible basis set (e.g., with polarization and diffuse functions), or plane waves are used in the QM calculations. The electron spill out can be avoided by using smeared-out charges instead of the traditional point charges (14). A convenient way for smearing the charges is to use a Gaussian distribution centred at the MM atom:

$$\Omega_J^{\text{MM}}(\mathbf{r}) = \sqrt{\frac{Q_J^{\text{MM}}}{\pi\alpha^3}} \exp\left[-\frac{|(\mathbf{r} - \mathbf{R}_J)|^2}{2\alpha^2}\right], \quad (7)$$

where  $|\Omega_J^{\text{MM}}(\mathbf{r})|^2$  is the charge density at position  $\mathbf{r}$ , due to MM atom  $J$  at position  $\mathbf{R}_J$  and charge  $Q_J$ . The factor  $\alpha$  controls the width of the distribution and is a parameter that needs to be calibrated. In contrast to the point charge model, the Coulomb interaction between the QM electrons and the Gaussian charge distributions does not diverge if the electrons approach the MM atoms:

$$h_{ij}(\mathbf{r}_1) = Q_J \int \phi_i^*(\mathbf{r}_1) \frac{\text{erf}(|\mathbf{r}_1 - \mathbf{R}_J|/\alpha)}{|\mathbf{r}_1 - \mathbf{R}_J|} \phi_j(\mathbf{r}_1) d\mathbf{r}_1, \quad (8)$$

with  $\phi_i$  the molecular orbital and  $h_{ij}$  the one-electron integral describing the interaction of a single electron with MM atom  $J$ . Such renormalization of the coulomb interaction avoids the

unphysical attraction of the electrons to charged atoms at the boundary between the two subsystems.

#### *2.4.3. Polarization Embedding*

The next step in increasing the level of sophistication is to include the polarizability of the MM atoms. In the polarization embedding scheme both regions can mutually polarize each other. Thus, not only is the QM region polarized by the MM atoms, the QM region can also induce polarization in the MM system. Different approaches have been developed to model polarization of MM atoms. Among the most popular methods are the charge-on-a-spring model (15), the induced dipole model (16) and the fluctuating charge model (17).

To obtain the total QM/MM energy in the polarizable embedding approach, the MM polarizations need to be computed at every step of the self-consistent-field iteration of the QM wave function. Since the polarization is computed in a self-consistent manner as well, the QM/MM computation can become very cumbersome and demanding. As a compromise, Zhang and co-workers have suggested to include polarization only in a small shell of MM atoms around the QM region (18).

Although polarization embedding offers the most realistic coupling between the QM and MM regions, polarizable force field for biomolecular simulations are not yet available. Therefore, despite progress in the development of such force fields, QM/MM studies with polarizable MM regions have so far been largely restricted to non-biological systems (19).

#### **2.5. Capping Bonds at the QM/MM Boundary**

If the QM and MM subsystems are connected by chemical bonds, care has to be taken when evaluating the QM wave function. A straightforward cut through the QM/MM bond creates one or more unpaired electrons in the QM subsystem. In reality, these electrons are paired in bonding orbitals with electrons belonging to the atom on the MM side. A number of approaches to remedy the artefact of such open valences have been proposed.

##### *2.5.1. Link Atoms*

The most easy solution is to introduce a monovalent link atom at an appropriate position along the bond vector between the QM and MM atoms (Figs. 3e and 5). Hydrogen is most often used, but there is no restriction on the type of the link atom and even complete fragments, such as methyl groups, can be used to cap the QM subsystem. The link atoms are present only in the QM calculation, and are invisible for the MM atoms. In principle each link atom introduces three additional degrees of freedom to the system. However, in practice the link atom is placed at a fixed position along the bond in every step of the simulation, so that these additional degrees of freedom are removed again. At each step, the force acting on the link atoms are distributed over the QM and MM atoms of the bond according to the lever rule.

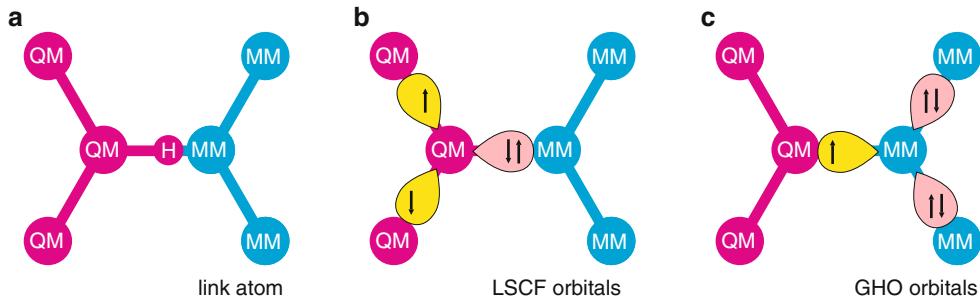


Fig. 5. Different approaches to cap the QM region: link atoms (a) and frozen orbitals (b,c). The hydrogen link atom (a) is placed at an appropriate distance along the QM/MM bond vector and is present only in the QM calculation. In the localized SCF method (b), a set of localized orbitals is placed on the QM atom. During the SCF iterations, the orbital pointing towards the MM atom is double-occupied and frozen, while the other orbitals are single-occupied and optimized. In the generalized hybrid orbital approach (c), a set of localized orbitals is placed on the MM atom. During the SCF interaction, the orbitals pointing towards the other MM atoms are double occupied and frozen, while the orbital pointing towards the QM atom is single-occupied and optimized.

### 2.5.2. Localized Orbitals

A popular alternative to the link atom scheme is to replace a chemical bond between the QM and MM subsystem by a double-occupied molecular orbital. This idea, which dates back to the pioneering work of Warshel and Levitt (5), assumes that the electronic structure of the bond is insensitive to changes in the QM region. The two most widely used approaches are the localized hybrid orbital method (20), which introduces orbitals at the QM atom (Fig. 5b), and the generalized hybrid orbital approach (21), which places additional orbitals on the MM atom (Fig. 5c).

In the localized self-consistent field (LSCF) method by Rivail and co-workers (20), the atomic orbitals on the QM atom of the broken bond are localized and hybridized. The hybrid orbital pointing towards the MM atom is occupied by two electrons. The other orbitals are each occupied by a single electron. During the SCF optimization of the QM wave function, the double-occupied orbital is kept frozen, while the other hybrid orbitals are optimized along with all orbitals in the QM region. The parameters in this method are the molecular orbital coefficients of the hybrid orbitals. In the original approach, these parameters are obtained by localizing orbitals in smaller model systems. This procedure thus assumes that the electronic structure of a chemical bond is transferable between different systems.

Alternatively, the coefficients of the frozen orbital can be obtained by performing a single point QM/MM calculation with a slightly enlarged QM subsystem. Any further broken bonds between the larger QM subsystem and the MM region are capped by link atoms in this calculation. The advantage of this so-called frozen orbital approach (22) is that no assumption is made on the electronic structure of the chemical bond. The disadvantage is that an electronic structure calculation has to be performed on a larger QM subsystem.

In the generalized hybrid orbital approach (GHO) of Gao and co-workers, hybrid orbitals are placed on the MM atom of the broken bond (21). In contrast to the LSCF scheme, the orbital pointing to the QM atom is optimized, while the others are kept double-occupied and frozen (Fig. 5).

In all localized orbital approaches, one or more parametrization steps are required. For this reason, the link atom is still the most widely adopted procedure for capping the QM region. Furthermore, studies that compared the accuracy of both methods showed that there is little advantage in using a localized orbitals rather than link atoms (23, 24).

In addition to capping the QM subsystem, one also needs to be careful if the MM atom at the other side of the chemical bond is charged. Since this atom is very near the QM subsystem, artefacts can easily arise due to the over-polarization effect, as discussed above. The easiest way to avoid this problem is to set the charges of that MM atom to zero. Alternatively, the charge can be shifted to MM atoms further away from the bond. The latter solution keeps the overall charge of the system constant.

### 3. QM/MM Applications

#### 3.1. Molecular Dynamics Simulations

The QM/MM method provides both potential energies and forces. With these forces, it is possible to perform a molecular dynamics simulation. However, because of the computational efforts required to perform *ab initio* calculations, the timescales that can be reached in QM/MM simulations is rather limited. At the *ab initio* or DFT level, the limit is in the order of few hundreds of picoseconds. With *semi-empirical* methods (e.g., AM1 (25), PM3 (26, 27), or DFTB (28)) for the QM calculation, the limit is roughly 100 times longer. Therefore, unless the chemical process under consideration is at least an order of magnitude faster than the timescale that can be reached, an unrestrained MD simulation is not the method of choice to investigate that process. Although the lack of sampling can be overcome by using enhanced sampling techniques, most researchers rely on energy minimization techniques to study chemical reactivity in condensed phases.

#### 3.2. Geometry Optimization

The traditional approach to study reactivity on a computer has been to characterize stationary points on the potential energy surface of the isolated system. The minima are identified as reactants and products, whereas the lowest energy saddle points that connect these minima are interpreted as the transition states. Extending this approach to QM/MM potential energy surfaces, however, is difficult, due to the much higher dimensionality of a typical QM/MM system. Since there are many more degrees of freedom that have to be optimized,

the geometry optimizer needs to be very efficient. Furthermore, the number of local minima in high dimensional systems is usually very large. At temperatures above zero, many of these minima are populated and there are also many paths connecting them. Therefore, even if the optimization can be carried out successfully, it may be difficult, if not impossible, to characterize all reaction pathways that are relevant for the process under study (29).

Despite these problems, optimizing the stationary points on the QM/MM potential energy surface is often the first step in exploring the reaction pathway. It usually gives important insights into the mechanism of the reaction, and the way by which it is controlled by the environment.

### 3.3. Free Energy Computation

To understand reactivity, one rather needs the *free energy* surface of the process. Computing free energies requires sampling of the underlying potential energy surface to generate ensembles. In equilibrium, the free energy difference  $\Delta G_{A \rightarrow B}$  between the reactant state ( $A$ ) and the product state ( $B$ ), both defined as areas on the free energy landscape, is determined by their populations  $p$ :

$$\Delta G_{A \rightarrow B} = -k_B T \ln \frac{p_B}{p_A}, \quad (9)$$

with  $k_B$  the Boltzmann constant, and  $T$  the temperature. However, for chemical reactions, the barriers separating the states  $A$  and  $B$  are high, and transitions are rare events. Therefore, it is not likely that both states are sampled sufficiently in a normal MD simulation to provide a reasonable estimate for  $\Delta G_{A \rightarrow B}$ .

#### 3.3.1. Umbrella Sampling

Equal sampling of  $A$  and  $B$  can be enforced by introducing a biasing potential that drives the system from state  $A$  into state  $B$ . After correcting for such biasing potential, the free energy can, in principle, be calculated with sufficient accuracy (30). A *single* simulation with a bias potential is not very efficient. Therefore, in practice, *several* independent simulations are carried out, each with a different biasing potential. These potentials are called umbrellas and are placed at different points along the reaction pathway. In each simulation, or window, the sampling is enhanced around the centre of the umbrella potential. Umbrella sampling yields a set of biased probability distributions. To generate the free energy profile for the entire pathway, the results of the various windows are combined and unbiased (31).

In QM/MM simulations, even the sampling of the windows can pose a problem due to the high demand on the computational resources for computing the wave function. As an approximation, the QM subsystem can be kept frozen in the windows. If also the charges on the QM atoms are kept fixed at each umbrella, no QM calculations are needed during the sampling of the remaining MM degrees of freedom. Thus, within this approach, the QM and

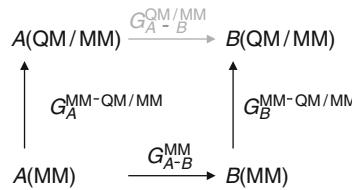


Fig. 6. Thermodynamic cycle for computing the free energy difference between states  $A$  and  $B$  at the QM/MM level ( $\Delta G_{A \rightarrow B}^{\text{QM/MM}}$ ). In the first step, the free energy difference between  $A$  and  $B$  is determined at the MM level ( $\Delta G_{A \rightarrow B}^{\text{MM}}$ ), either by thermodynamic integration or free energy perturbation. In the second step, the free energy required to transform the MM ensemble of  $A$  and  $B$  into the QM/MM ensemble ( $\Delta G_A^{\text{MM} \rightarrow \text{QM/MM}}$  and  $\Delta G_B^{\text{MM} \rightarrow \text{QM/MM}}$ ) are computed by free energy perturbation. The QM/MM free energy of converting  $A$  into  $B$  is calculated by adding up the free energy differences in going around the cycle from  $A(\text{QM/MM})$  to  $B(\text{QM/MM})$ . This procedure avoids computing the  $\Delta G_{A \rightarrow B}^{\text{QM/MM}}$  directly.

MM degrees of freedom are assumed to be uncoupled. Whether such assumption is valid, depends on the process at hand. Another issue concerns finding a suitable reaction path along which the umbrella sampling will be carried out.

### 3.3.2. Free Energy Perturbation

An alternative approach for extracting the free energy associated with the conversion between two states from QM/MM simulations is to use a combination of thermodynamic integration (32) and free energy perturbation (33). In thermodynamic integration (TI), the Hamiltonian is interpolated between the two states with a coupling parameter  $\lambda$ :

$$H(\mathbf{q}, \mathbf{p}, \lambda) = (1 - \lambda)H_A(\mathbf{q}, \mathbf{p}) + \lambda H_B(\mathbf{q}, \mathbf{p}), \quad (10)$$

where  $\mathbf{q}$  and  $\mathbf{p}$  are the positions and momenta of all atoms in the system. To obtain the free energy difference between state  $A$ , when  $\lambda = 0$ , and state  $B$ , when  $\lambda = 1$ , the system is sampled at fixed values of  $\lambda$  between 0 and 1, followed by integration over the ensemble averages of  $\langle \partial H / \partial \lambda \rangle_\lambda$  at these  $\lambda$  values with respect to  $\lambda$ :

$$\Delta G = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda. \quad (11)$$

An advantage of the TI approach is that the pathway connecting the two states does neither have to be physically meaningful nor possible. For example, the free energy cost of changing or even disappearing atoms, can be computed efficiently this way. Such non-physical transformations are usually only possible at the MM level. To get the free energy change at the QM/MM level, an additional step is required (34).

Because the free energy is a state function, its magnitude does not depend on the pathway taken. Therefore, one can always construct a so-called thermodynamic cycle, as shown in Fig. 6. For the free energy of a transformation at the QM/MM level, the quantity of interest is the

free energy associated with the top process ( $\Delta G_{A \rightarrow B}^{QM/MM}$ ). Since the cycle is closed (i.e. the  $\Delta G$ s add up to zero upon completing the cycle), this quantity can be computed as:

$$\Delta G_{A \rightarrow B}^{QM/MM} = \Delta G_{A \rightarrow B}^{MM} + \Delta G_B^{MM \rightarrow QM/MM} - \Delta G_A^{MM \rightarrow QM/MM}, \quad (12)$$

with the free energies defined in Fig. 6.

Thus, instead of calculating  $\Delta G_{A \rightarrow B}^{QM/MM}$  directly, which is often impossible, one can evaluate this free energy in three steps. First, the free energy of the process is calculated at the MM level, by means of thermodynamic integration (Eq. 11). In the second and third steps, the free energy associated with changing the potential energy landscape from MM to QM/MM is computed for the end states of the TI process ( $\Delta G^{MM \rightarrow QM/MM}$ ). One way of obtaining these two quantities is to make use of the free energy perturbation formalism that describes the free energy difference between two states as the overlap between the ensembles (33):

$$\begin{aligned} \Delta G^{MM \rightarrow QM/MM} &= G^{QM/MM} - G^{MM} \\ &= -k_B T \ln \left\langle \exp \left( -\frac{V^{QM/MM} - V^{MM}}{k_B T} \right) \right\rangle_{MM}, \end{aligned} \quad (13)$$

with  $k_B$  the Boltzmann constant,  $T$  the temperature,  $V^{MM}$  and  $V^{QM/MM}$  the potential energy at the QM/MM and MM levels, respectively. The Boltzmann factor is averaged over the ensemble generated at the MM level. Since many MM snapshots may be required to get a converged Boltzmann factor, sampling remains a critical issue.

### 3.4. Computational Spectroscopy

Spectroscopy in the visible and infrared spectral regions are among most important experimental techniques to probe the structure and dynamics of sub-picosecond photochemical processes. However, the interpretation of the spectra requires knowledge about the structure and dynamics of the system under study. Therefore, the full potential of this technique can only be realized when it is complemented by computational modelling. Many spectroscopic quantities can be computed accurately with quantum chemistry methods, but mostly for small model systems in isolation. Including the environment, as in QM/MM methods, therefore, may be required to obtain spectra that can be compared to experiment.

#### 3.4.1. UV/vis Absorption Spectra

This class of spectroscopic techniques probes the energy gaps between the different electronic states of the system. The absorption (or stimulated emission) spectra are sensitive to the structure, and structural changes can be traced in real time by time-resolved spectroscopic measurements (e.g. pump-probe). For small systems, the energy levels of the electronic states can be computed accurately

with high-end *ab initio* methods. Suitable methods are based on the complete active space self-consistent field method, such as CASSCF, RASSCF, and CASPT2 (35). However, these methods are too time and memory consuming for larger systems. Therefore, computing spectra of condensed phase systems requires a QM/MM approach. A realistic spectrum is obtained by evaluating the excitation energies in snapshots taken from classical MD trajectories. After averaging the excitation energies over the ensemble, the computed spectrum can be compared directly to the experimental spectrum (36).

#### 3.4.2. IR Absorption Spectra

Infrared spectroscopy probes transitions between vibrational states. The spectra provide a wealth of information about the structure of the system under study, but the assignment of the observed vibrational bands often requires modelling. A popular computational approach for computing vibrational spectra is the normal mode analysis (NMA). In this technique the matrix of second derivatives of the energy with respect to the nuclear displacements is evaluated and diagonalized. The resulting eigenvalues and eigenvectors are the intensities and vibrational modes of the system, respectively. Because this procedure is preceded by a rigorous energy minimization, the spectra are effectively calculated at zero Kelvin. Therefore, the width of the absorption bands, which reflects thermal averaging, are not accessible in the NMA approach.

An alternative approach to extract IR spectra from QM/MM simulations is to take the Fourier transform of the dipole-dipole autocorrelation function:

$$I(\omega) \propto \int_{-\infty}^{\infty} \langle \mu(t) \cdot \mu(0) \rangle \exp(-i\omega t) dt, \quad (14)$$

with  $I$  the intensity at the vibrational frequency  $\omega$ ,  $\mu(t)$  the system's dipole at time  $t$ . The major drawback of this method, however, is that the dipole moment needs to be sampled sufficiently. Therefore, this approach is most often used in conjunction with semi-empirical methods (37).

## 4. Case Study: QM/MM Simulations of a Photochemical Process

Photoactive yellow protein (PYP) is believed to be the primary photoreceptor for the photo-avoidance response of the salt-tolerant bacterium *Halorhodospira halophila* (38). PYP contains a deprotonated 4-hydroxycinnamic acid (or *p*-coumaric acid, PCA) chromophore linked covalently to the  $\gamma$ -sulphur of Cys69 via a thioester bond (Fig. 7). Upon absorbing a blue-light photon, PYP enters a fully reversible photocycle involving several intermediates on timescales ranging from a few hundred femtoseconds to seconds (38). Before the QM/MM work that was done to elucidate

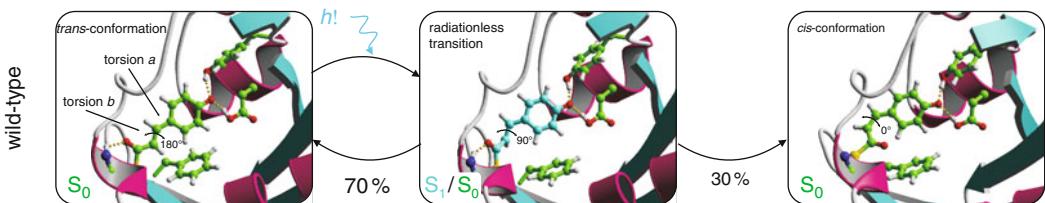


Fig. 7. Snapshots from excited-state trajectories of wild-type PYP, showing the chromophore (pca) in the active site pocket. The first snapshot is at the excitation. The second shows the configuration at the radiationless transition from  $S_1$  to  $S_0$ . The third snapshot shows the photoproduct, in which the carbonyl oxygen of the thioester linkage has flipped and is no longer hydrogen bonded to the backbone of Cys69.

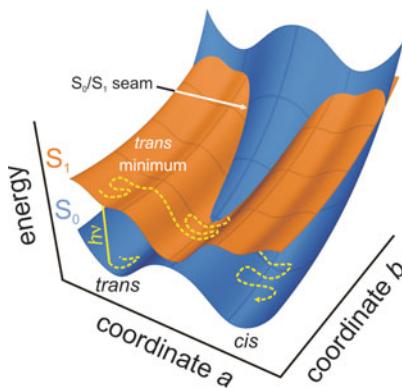


Fig. 8. Schematic overview of a photochemical reaction pathway (dashed line). After photon absorption, evolution takes place on the excited-state potential energy surface (red) until the system hits the  $S_1/S_0$  intersection seam. At the intersection, a radiationless transition to the ground state occurs (blue). After the decay, the system continues evolving in the ground state.

the mechanism by which photon absorption induces signalling, we briefly introduce the basic concepts of photochemistry.

#### 4.1. Photochemical Reactions

The central mechanistic feature in a photochemical reaction is the intersection seam between the potential energy surfaces of the excited ( $S_1$ ) and ground states ( $S_0$ , Fig. 8). Any point on the seam provides a funnel for efficient radiationless decay to the ground state. Just as a transition state separates reactants and products in ground-state chemistry, the seam separates the excited-state branch from the ground-state branch in a photochemical reaction. The crucial difference, however, is that while a transition state connects a reactant to a *single* product via a *single* reaction path, the seam connects the excited state and reacts to *several* products on the ground state via *several* paths. Just as ground-state reactivity is enhanced by a stabilization of the transition state, photoreactivity is also enhanced by stabilization of the seam.

##### 4.1.1. MD Simulations of Photochemical Processes

To model the dynamics of a photochemical reaction, the ground-state and excited-state potential energy surfaces must be described

accurately. After light absorption, the reaction starts in the excited state ( $S_1$ ), but ends in the ground state ( $S_0$ ). Therefore, it is essential to model the radiationless transition between the excited-and ground-state surfaces in a manner that is consistent with a quantum mechanical treatment of the complete system. Because we use Newton's equation of motion to compute molecular dynamics trajectories, the quantum mechanical character of the nuclei is ignored. As a consequence, population transfer from  $S_1$  to  $S_0$  cannot occur, and the classical trajectory is restricted to a single potential energy surface. Thus, in contrast to a full quantum mechanical approach, radiationless transitions do not take place spontaneously. Instead, a binary decision to jump to a different electronic surface must be made at every timestep in a single trajectory. The criterion for switching between electronic states must result in a distribution of state populations, whose average can be compared to observable quantities, such as quantum yield, lifetimes, etc.

In our simulations we allow hopping only at the intersection seam. In principle, this strict diabatic hopping criterion could lead to an underestimation of the population transfer probability. However, because of the high dimensionality of the seam, most trajectories can usually encounter such regions of high probability. The diabatic hopping model is clearly an approximation, but helps one to keep a proper physical insight, which is crucial in understanding complex systems.

#### **4.2. Chromophore in Vacuum**

To understand the intrinsic photochemical properties of the PYP chromophore, we have performed geometry optimizations of an isolated chromophore analogue at the CASSCF level of *ab initio* theory (39). In these optimizations, the complete  $\pi$  system of the chromophore was included in the active space, which thus consisted of 12 electrons in 11  $\pi$  orbitals. In addition to optimizing the local minima on the  $S_1$  potential energy surface and the barriers that separate them, we also located conical intersections in the vicinity of these minima. The optimizations revealed that there are two minima on  $S_1$ : the single-bond twisted minimum, in which the bond adjacent to the phenol ring is rotated by 90°, and the double-bond twisted minimum, in which the ethylenic bond is twisted at 90° (Fig. 9). In the isolated chromophore, there is almost no barrier for reaching the single-bond twisted  $S_1$  minimum from the Franck-Condon region, whereas there is a significant barrier to double-bond rotation. Thus, after photon absorption in vacuum, the main relaxation channel on  $S_1$  involves rotation of the single bond to 90°. We furthermore found that the  $S_1/S_0$  intersection seam lies very far away from this minimum. As a consequence, radiationless decay is not very efficient in vacuum. In subsequent QM/MM simulations, we have probed the effect of different environments on the photochemistry of the chromophore.

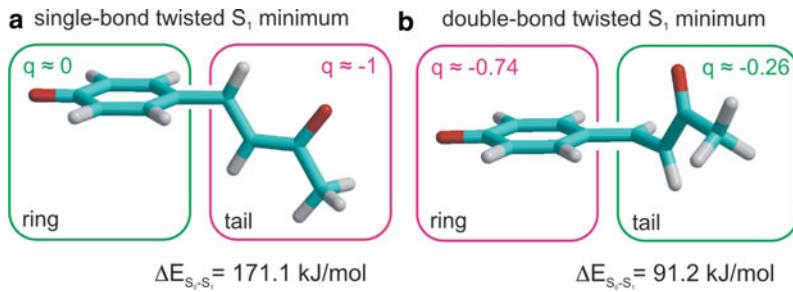


Fig. 9. Excited-state minimum energy configurations of a chromophore analogue. In both the single-bond twisted  $S_1$  minimum (**a**) and the double-bond twisted  $S_1$  minimum (**b**) there is a substantial energy gap between the ground and excited state. The distribution of the negative charge in these minima is opposite.

#### 4.3. Chromophore in Water

To examine the effect of an aqueous environment, we have performed 91 QM/MM excited-state dynamics simulations of a chromophore analogue in water (39). The chromophore was described at the CASSCF(6,6)/3-21G level of theory, while the water molecules were modelled by the SPC/E force field (40). The results of these simulations demonstrate that in water, radiationless decay is very efficient (39). The predominant excited-state decay channel involves twisting of the single bond (88%) rather than the double bond (12%). In contrast to vacuum, decay takes place very near these minima. Inspection of the trajectories revealed that decay is mediated by specific hydrogen bond interactions with water molecules. These hydrogen bonds are different for the single-and double-twisted  $S_1$  minima, which reflects the difference in charge distribution between these minima (41). In the single-bond twisted  $S_1$  minimum, the negative charge resides on the alkene moiety of the chromophore (Fig. 9). Three strong hydrogen bonds to the carbonyl oxygen stabilize this charge distribution to such an extent that the seam almost coincides with the single-bond twisted  $S_1$  minimum (Fig. 10). In the double-bond twisted  $S_1$  minimum, the negative charge is localized on the phenolate ring (Fig. 9). Transient stabilization of this charge distribution by two or more strong hydrogen bonds to the phenolate oxygen brings the seam very close to this  $S_1$  minimum (Fig. 10). Thus, in water, the ultra-fast excited-state decay is mediated by hydrogen bonds.

#### 4.4. Chromophore in the Protein

To find out how the protein mediates the photochemical process, we also carried out a series of QM/MM simulations of wild-type PYP (42). Fig. 7 shows the primary events after photoexcitation in the simulation. The chromophore rapidly decays to the ground state via a  $90^\circ$  rotation of the double bond (Fig. 7), rather than the single bond. During this photo-isomerization process, the hydrogen bonds between the chromophore's phenolate oxygen atom and the side chains of the highly conserved Tyr42 and

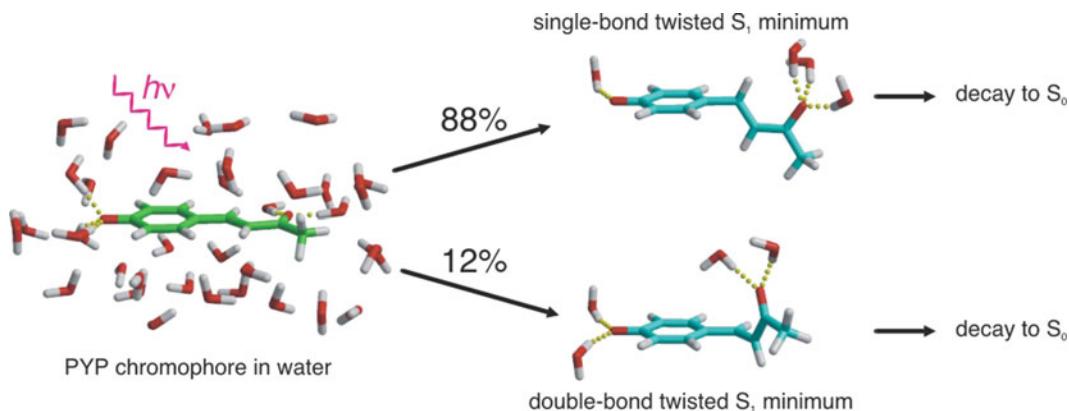


Fig. 10. In water the chromophore undergoes both single-and double-bond isomerization. Excited-state decay from these minima is very efficient due to stabilization of the chromophore's  $S_1$  charge distribution by specific hydrogen bond interactions.

Glu46 residues remain intact. Just as in water, these hydrogen bonds enhance excited-state decay from the double-bond twisted minimum.

Upon returning to the ground state, the chromophore either relaxes back to the original *trans* conformation ( $180^\circ$ ) or it continues isomerizing to a *cis* conformation ( $0^\circ$ ). In the latter case, the relaxation also involves a flip of the thioester linkage, which causes the carbonyl group to rotate  $180^\circ$ . During this rotation, the hydrogen bond between the carbonyl oxygen and the Cys69 backbone amino group is broken (Fig. 7). In total, 14 MD simulations were carried out, initiated from different snapshots from a classical ground-state trajectory. The fluorescence lifetime (200 fs) and isomerization quantum yield (30%) in the simulations agree well with experiments (400 fs (43) and 35% (38), respectively).

In the wild-type protein, no single-bond isomerization was observed. Thus, the protein not only provides the hydrogen bonds required for ultrafast decay but also controls which of the chromophore's bonds isomerizes upon photoexcitation. We identified the positive guanidinium moiety of Arg52 located just above the chromophore ring as the “catalytic” residue that enforces double-bond isomerization. The preferential electrostatic stabilization of the double-bond twisted  $S_1$  minimum by the positive Arg52 strongly favors double-bond isomerization over single-bond isomerization.

To elucidate the role of this arginine in the activation process in more detail, we performed excited-state dynamics simulations on the Arg52Gln mutant of PYP (44). This mutant can still enter the photocycle, albeit with a lower rate and quantum yield (45, 46). Without the positive Arg52, the predominant excited-state reaction in the mutant involves isomerization of a single bond in the

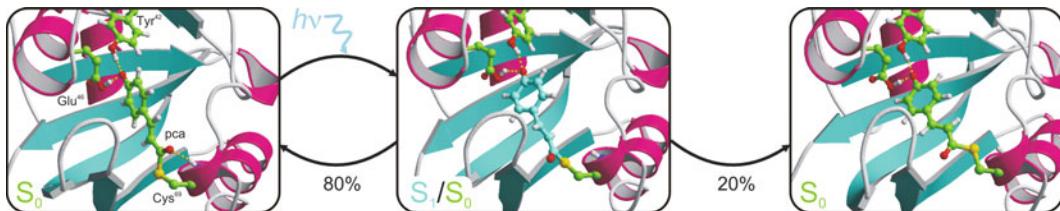


Fig. 11. Snapshots from an excited-state trajectory of the Arg52Gln mutant of PYP, showing the chromophore (pca) in the active site pocket. The first snapshot is at the excitation. The second snapshot shows the configuration at the radiationless transition from  $S_1$  to  $S_0$ . The third snapshot shows the photoproduct. In the mutant, isomerization takes place around the single bond. Like in the wild-type protein, the carbonyl oxygen of the thioester linkage flips, causing the break of the hydrogen bond to the backbone of Cys69.

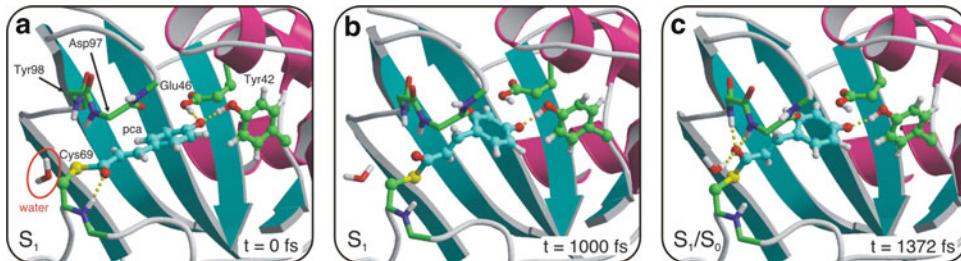


Fig. 12. Snapshots from an excited-state trajectory of the Arg52Gln mutant of PYP, demonstrating that three hydrogen bonds to the carbonyl moiety are essential for  $S_1$  decay at the single-bond twisted minimum. The first snapshot is at the excitation to  $S_1$ . The second snapshot shows the twisted configuration without hydrogen bonds to the carbonyl. The gap between  $S_1$  and  $S_0$  is far too high for decay at this configuration. However, the third snapshot shows two backbone amino groups and a bulk water that has moved into the chromophore pocket during the excited-state dynamics, donating the three hydrogen bonds that are required for efficient decay from the  $S_1$  minimum.

chromophore, rather than the double bond (Fig. 11) (47). This observation confirms that the role of Arg52 is to steer the initial events after photoabsorption to ensure rotation of the double rather than the single bond in the chromophore.

During the rotation of the single bond, the hydrogen bond between the carbonyl oxygen and Cys69 backbone amino group is broken. As shown in Fig. 12, new hydrogen bonds are rapidly formed between the carbonyl oxygen atom and the backbone amino groups of Tyr98 and Asp97. A water molecule from outside enters the chromophore pocket to donate a third hydrogen bond. With these three hydrogen bonds stabilizing the negative charge on the alkene moiety, the chromophore rapidly decays to  $S_0$ . Thus, the decay mechanism in the Arg52Gln mutant and in water are essentially the same.

Although single-bond isomerization does not result in the formation of the *cis* chromophore, a 180° flip of the thioester group and a rupture of the hydrogen bond to Cys69 was observed with a 20% quantum yield (Fig. 12). Together with the experimental

observation that the mutant has a photoactivation quantum yield of about 21% (46), this suggests that the key step to enter the photocycle is the oxygen flip rather than the double-bond isomerization.

To summarize, the simulations are consistent with experimental observations and have provided detailed structural and dynamic information at a resolution well beyond that achievable by other means. From the simulations, key amino acids have been identified and the mechanism by which they control the primary events in the photocycle of PYP. These are (i) double-bond photoisomerization, and (ii) the break of a hydrogen bond between the chromophore and the protein backbone. These events trigger a proton transfer from the protein to the chromophore, which ultimately leads to the signalling state of PYP (48).

---

## 5. Conclusion and Outlook

In this contribution we have reviewed the basic concepts of hybrid QM/MM simulation techniques. More elaborate discussions on the QM/MM method are available as review articles, see for instance references (49–53). In principle, QM/MM simulations can provide detailed structural information of chemical reactions in the condensed phase at an atomic resolution. In practice, the QM/MM methods still suffer from limitations in computational hardware, which restrict both system size and timescale of the processes that can be studied today. However, the expected increase in computer power, complemented by the development of more efficient electronic structure methods and new algorithms may soon enable the investigation of reactions in larger systems and at longer timescales. Therefore, QM/MM simulation has the potential to lead to a better understanding of chemical reactions, and the mechanisms by which in particular protein environments control these reactions. Ultimately, these simulations may not only enable accurate predictions of chemical properties but also become a standard tool for rational design of artificial molecular devices.

---

## Acknowledgements

The Volkswagenstiftung and Deutsche Forschungsgemeinschaft (SFB755) are acknowledged for their financial support. I am grateful to Dr. Mehdi Davari and Pedro Valiente for critically reading the manuscript.

## References

1. Jensen F (2001) Introduction to computational chemistry. Wiley, New York
2. Berendsen HJC (2001) Bio-molecular dynamics comes of age. *Science* 271:954–955
3. McCammon JA, Gelin BR, Karplus M, Wolynes PG (1976) Hinge-bending mode in lysozyme. *Nature* 262:325–326
4. Shaw D, Maragakis P, Lindorff-Larsen K, Piana S, Dror R, Eastwood M, Bank J, Jumper J, Salmon J, Shan Y, Wriggers W, (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346
5. Warshel A, Levitt M (1976) Theoretical studies of enzymatic reactions: dielectric, electrostatic and steric stabilization of carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227–249
6. Maseras F, Morokuma K (1995) IMOMM—a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition-states. *J Comput Chem* 16:1170–1179
7. Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) ONIOM: a multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels–Alder reactions and Pt(P(*t*-Bu)<sub>3</sub>)<sub>2</sub> + H<sub>2</sub> oxidative addition. *J Phys Chem* 100:19357–19363
8. Bayly C, Cieplak P, Cornell W, Kollman P (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges—the RESP model. *J Phys Chem* 97:10269–10280
9. Besler B, Merz K, Kollman P (1990) Atomic charges derived from semiempirical methods. *J Comput Chem* 11:431–439
10. Szabo A, Ostlund NS (1989) Modern quantum chemistry. Dover Publications, New York
11. Field MJ, Bash PA, Karplus M (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J Comp Chem* 11:700–733
12. Brooks B, Karplus M (1983) Harmonic dynamics of proteins—normal-modes and fluctuations in bovine pancreatic trypsin-inhibitor. *Proc Natl Acad Sci USA* 80:6571–6575
13. Dewar M (1983) Development and status of MNDO/3 and MNDO. *J Mol Struct* 100:41–50
14. Amara P, Field MJ (2003) Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor Chem Acc* 109:43–52
15. Lamoureux G, Roux B (2003) Modeling induced polarization with classical Drude oscillators: theory and molecular dynamics simulation algorithm. *J Chem Phys* 119:3025–3039
16. Warshel A, Sharma PK, Kato M, Xiang Y, Liu HB, Olsson MHM (2006) Electrostatic basis for enzyme catalysis. *Chem Rev* 106:3210–3235
17. Rappe AK, Goddard III WA (1991) Charge equilibration for molecular dynamics simulations. *J Phys Chem* 95:3358–3363
18. Zhang Y, Lin H, Truhlar D (2007) Self-consistent polarization of the boundary in the redistributed charge and dipole scheme for combined quantum-mechanical and molecular-mechanical calculations. *J Chem Theory Comput* 3:1378–1398
19. Hillier I (1999) Chemical reactivity studied by hybrid QM/MM methods. *J Mol Struct (Theochem)* 463:45–52
20. Assfeld X, Rivail J (1996) Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method. *Chem Phys Lett* 263:100–106
21. Gao J, Amara P, Alhambra C, Field M (1998) A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J Phys Chem A* 102:4714–4721
22. Philipp DM, Friesner RA (1999) Mixed ab initio QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide. *J Comput Chem* 20:1468–1494
23. Nicoll R, Hindle S, MacKenzie G, Hillier I, Burton N (2001) Quantum mechanical/molecular mechanical methods and the study of kinetic isotope effects: modelling the covalent junction region and application to the enzyme xylose isomerase. *Theor Chem Acc* 106:105–112, 10th International Congress of Quantum Chemistry, Nice, France, June 13–15, 2000
24. Rodriguez A, Oliva C, Gonzalez M, van der Kamp M, Mulholland A, (2007) Comparison of different quantum mechanical/molecular mechanics boundary treatments in the reaction of the hepatitis C virus NS3 protease with the NS5A/5B substrate. *J Phys Chem B* 111:12909–12915
25. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) The development and use of quantum-mechanical molecular models AM1—a new general-purpose quantum-mechanical molecular-model. *J Am Chem Soc* 107:3902–3909

26. Stewart JJP (1989) Optimization of parameters for semiempirical methods. I. Method. *J Comput Chem* 10:209–220
27. Stewart JJP (1989) Optimization of parameters for semiempirical methods. 2. Applications. *J Comput Chem* 10:221–264
28. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B* 58:7260–7268
29. Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. *J Phys Chem B* 109:15645–15650
30. Torrie GM, Valle JP (1977) Non-physical sampling distributions in Monte-Carlo free energy estimation—umbrella sampling. *J Comput Phys* 23:187–199
31. Roux B (1995) The calculation of the potential of mean force using computer-simulations. *Comp Phys Comm* 91:275–282
32. Kirkwood J (1935) Statistical Mechanics of Fluid Mixtures. *J Chem Phys* 3:300–313
33. Zwanzig R (1954) High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J Chem Phys* 22:1420–1426
34. Muller R, Warshel A (1995) Ab-initio calculations of free energy barriers for chemical-reactions in solution. *J Phys Chem* 99:17516–17524
35. Roos BO (1999) Theoretical studies of electronically excited states of molecular systems using multiconfigurational perturbation theory. *Acc Chem Res* 32:137–144
36. Schäfer LV, Groenhof G, Klingen AR, Ullmann GM, Boggio-Pasqua M, Robb MA, Grubmüller H (2007) Photoswitching of the fluorescent protein asFP595: mechanism proton pathways, and absorption spectra. *Angew Chemie Int Ed* 46:530–536
37. Kaminski S, Gaus M, Phatak P, von Stetten D, Elstner M, Mroginski M (2010) Vibrational Raman spectra from the self-consistent charge density functional tight binding method via classical time-correlation functions. *J Chem Theory Comput* 6:1240–1255
38. Hellingswerf KJ, Hendriks J, Gensch T (2003) Photoactive yellow protein, a new type of photoreceptor protein: will this “yellow lab” bring us where we want to go? *J Phys Chem A* 107:1082–1094
39. Boggio-Pasqua M, Robb M, Groenhof G (2009) Hydrogen bonding controls excited-state decay of the photoactive yellow protein chromophore. *J Am Chem Soc* 131:13580
40. Berendsen H, Grigera J, Straatsma T (1987) The missing term in effective pair potentials. *J Phys Chem* 91:6269–6271
41. Gromov EV, Burghardt I, Hynes JT, Köppel H, Cederbaum LS (2007) Electronic structure of the photoactive yellow protein chromophore: ab initio study of the low-lying excited singlet states. *J Photochem Photobiol A* 190:241–257
42. Groenhof G, Bouxin-Cademartory M, Hess B, De Visser S., Berendsen H, Olivucci M, Mark A, Robb M (2004) Photoactivation of the photoactive yellow protein: why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein. *J Am Chem Soc* 126:4228–4233
43. Mataga N, Chosrowjan H, Shibata Y, Imamoto Y, Tokunaga F (2000) Effects of modification of protein nanospace structure and change of temperature on the femtosecond to picosecond fluorescence dynamics of photoactive yellow protein. *J Phys Chem B* 104:5191–5199
44. Shimizu N, Kamikubo H, Yamazaki Y, Imamoto Y, Kataoka M (2006) The crystal structure of the R52Q mutant demonstrates a role for R52 in chromophore pK(a) regulation in photoactive yellow protein. *Biochemistry* 45:3542–3547
45. Changenet-Barret P, Plaza P, Martin MM, Chosrowjan H, Taniguchi S, Mataga N, Imamoto Y, Kataoka M (2007) Role of arginine 52 on the primary photoinduced events in the PYP photocycle. *Chem Phys Lett* 434:320–325
46. Takeshita K, Imamoto Y, Kataoka M, Mihara K, Tokunaga F, Terazima M (2002) Structural change of site-directed mutants of PYP: new dynamics during pR state. *Biophys J* 83:1567–1577
47. Groenhof G, Schäfer LV, Boggio-Pasqua M, Grubmüller H, Robb MA (2008) Arginine 52 controls photoisomerization in photoactive yellow protein. *J Am Chem Soc* in press JACS 130: 3250–3251
48. Groenhof G, Lensink MF, Berendsen HJC, Mark AE (2002) Signal transduction in the photoactive yellow protein. II. Proton transfer initiates conformational changes. *Proteins* 48:212–219
49. Gao J (1996) Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Acc Chem Res* 29:298–305
50. Monard G, Merz K (1999) Combined quantum mechanical/molecular mechanical

- methodologies applied to biomolecular systems. *Acc Chem Res* 32:904–911
51. Gao J, Truhlar D (2002) Quantum mechanical methods for enzyme kinetics. *Annu Rev Phys Chem* 53:467–505
52. Friesner R, Guallar V (2005) Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu Rev Phys Chem* 56:389–427
53. Senn H, Thiel W (2009) QM/MM methods for biomolecular systems. *Angew Chem Int Ed* 48:1198–1229

# Chapter 4

## Computational Enzymology

Alessio Lodola and Adrian J. Mulholland

### Abstract

Techniques for modelling enzyme-catalyzed reaction mechanisms are making increasingly important contributions to biochemistry. They can address fundamental questions in enzyme catalysis and have the potential to contribute to practical applications such as drug development.

**Key words:** QM/MM, Enzyme, Catalysis, Protein dynamics, Biomolecular simulation, Quantum mechanics/molecular mechanics

---

### 1. Introduction

Molecular modelling and simulations can explore mechanisms of biological catalysts (i.e., enzymes) at a level of detail that cannot be achieved experimentally (1–11). Modelling can unravel the mechanisms of enzyme-catalyzed reactions, identify the origins of catalysis, analyze effects of mutations and genetic variations, and help to develop structure–activity relationships (12–14). Since its origins (15, 16), computational enzymology has grown enormously, particularly in recent years. There has also been a significant improvement in the accuracy of computational methods. For example, it is now possible to achieve an unprecedented level of accuracy in calculations on enzyme-catalyzed reactions with combined quantum mechanics/molecular mechanics (QM/MM) methods (17). In the best cases, calculations can give activation energies that agree extremely well with experiment. High-level quantum chemical methods allow calculations of energy barriers, in the best cases, near “chemical accuracy” (1 kcal/mol) (18). Quantitative predictions at this level were only previously possible for very small molecules. Carefully parameterized empirical molecular simulation approaches also give excellent agreement with experiments for enzyme reactions (19).

Identifying the chemical mechanisms of enzymes solely from experiments is often difficult. Many mechanisms in the literature are probably wrong in important details, e.g., as more recent experiments and simulations have shown for hen egg-white lysozyme (20, 21). The physical origins of enzyme catalysis also continue to be hotly debated. Recent controversies have centered on “low-barrier” hydrogen bonds (22–25), so-called near-attack conformations (26, 27) enzyme dynamics (28, 29), quantum tunnelling (30–33), and entropic effects (34). The applicability of transition state theory to enzyme reactions has also been questioned (35). Molecular simulations are proving to be crucial in testing these proposals.

Transition states are central to understanding chemical reactivity and catalysis, but experiments cannot directly study them in enzymes because of their extremely short lifetimes, and because of the large size and complexity of enzymes. Molecular modelling can analyze transition states directly and identify interactions involved in catalysis (e.g., a conserved proline residue that specifically stabilizes the transition state for aromatic hydroxylation in the flavin dependent monooxygenases *para*-hydroxybenzoate hydroxylase and phenol hydroxylase (36, 37)). Such interactions may not be apparent (and may not exist) in available experimental structures. This type of knowledge can assist in ligand design, e.g., as a potential route to enhanced affinity. Also, in contrast to some experimental (e.g., structural) studies, which may require mutation of the enzyme or use of alternative (e.g., inefficient) substrates (e.g., to slow down the reaction, and prolong the lifetime of intermediates, to allow their spectroscopic or structural characterization), molecular modelling can study directly the “wild type” reaction (i.e., the reaction as it occurs in the naturally occurring enzyme). Computational enzymology interacts fruitfully with experiments, which can validate modelling approaches, which in turn can interpret experimental findings and suggest new experiments.

---

## 2. Materials and Methods

### 2.1. Choice of Enzyme Structure for Modelling

An enzyme structure from X-ray crystallography (ideally high resolution), is the usual starting point for modelling an enzyme-catalyzed reaction. A crystal structure of an enzyme alone, with no ligand bound at the active site, may not be useful, because it is difficult to predict binding modes and protein conformational changes associated with binding. NMR structure ensembles can give useful complementary information on dynamics and interactions, but generally do not define atomic positions precisely enough for mechanistic modelling. In some cases a homology model may be sufficiently reliable, but should be treated with great caution: the positions, relative orientations and packing of side chains may not

be modelled sufficiently accurately. An example of the successful use of homology modelling for mechanistic studies is an investigation of the substrate binding mode and reaction mechanism of a malarial protease with a novel active site, by automated docking, and molecular dynamics/reaction free energy profile simulations (38).

The structure used for modelling must accurately represent the reacting enzyme complex; a crystal structure of an enzyme-inhibitor complex is often a good choice. The inhibitor should resemble the substrate, product, transition state or an intermediate, in its bound conformation. One should remember that there can be local structural uncertainty due e.g., to protein dynamics, and conformational variability or disorder, even in high-resolution structures. It is usually not possible to determine crystal structures of active enzyme-substrate complexes, unless specialized conditions or variant substrates or enzymes are used. In some cases, structures of several complexes along the reaction pathway may be available (30). Calculations can then model the chemical and structural changes, unstable intermediates and transition state structures, providing a picture of the whole reaction in the enzyme. For many enzymes, the conformational changes that take place during the reaction are small, and so modelling based on a single structure of an enzyme complex can give a reliable picture of the reaction. It may, however, be necessary to consider the effects of conformational fluctuations in the protein (see below).

## 2.2. Effects of Protein Dynamics

Proteins have many conformational substates, and a single structure may not be truly representative for modelling a mechanism (39). Extensive conformational sampling may be needed to generate a representative ensemble of structures. To calculate free energy profiles, i.e., potentials of mean force (40), a simulation method must be capable of calculating trajectories of many picoseconds at least (or a similarly large number of configurations in a Monte Carlo simulation). Alternatively, molecular dynamics simulations can be used to generate multiple structural models for subsequent mechanistic calculations, to ensure wide sampling of possible enzyme configurations. If multiple different crystal structures of the same enzyme are available, these may be suitable as different starting models, and similarly help to examine the effects of structural variation on the reaction.

Protein dynamics are believed to be important to their biological functions in many cases (41). It is well known that many enzymes undergo large conformational changes during their reaction cycles (42). The possible relationship of dynamics to enzyme catalysis is more controversial. It has been proposed that protein dynamics contribute significantly to enhancing the rates of reaction in enzymes, but simulations indicate that the effect of protein dynamics in determining the rates of chemical reactions in enzymes is relatively small (43). Protein conformational changes

(e.g., involved with substrate binding or product release) can, however, be rate-limiting for the overall reaction in many enzymes (44), and in some cases are coupled to chemical changes (e.g., facilitating product release). Quantum effects such as nuclear tunnelling are important in reactions involving hydrogen transfer (30, 45, 46) and the effects of protein dynamics on reactions involving quantum tunnelling is an area of particularly active debate (5, 10, 28, 47).

### **2.3. Determining the Mechanism**

Determining the chemical mechanism is an essential first step in studying an enzyme-catalyzed reaction. This is not trivial: many “textbook” mechanisms are probably wrong. The first aim is to establish the identities and functions of catalytic residues; many mechanisms in the biochemical literature assign functions to residues which are probably incorrect. Next, any specific interactions that stabilize transition states or reaction intermediates should be identified and analyzed. A typical computational approach to modelling reactions is to optimize the structures of key species (such as transition state structures); preferably entire reaction pathways should be optimized or simulated.

### **2.4. Analyzing Catalysis**

To understand *catalysis*, i.e., to understand why a reaction in an enzyme reaction is faster than an uncatalyzed or nonenzymic reaction, the two reactions should be compared (although deciding on an appropriate “reference” reaction may be difficult (6)). Practical applications often have simpler aims such as predicting the effects of a mutation on activity or on the specificity of an enzyme for alternative substrates. Overall, understanding enzyme mechanism, specificity and catalysis presents a range of challenges, and different types of modelling or simulation methods are needed to investigate different types of question, as outlined below.

### **2.5. Methods for Modelling Enzymes and Enzyme-Catalyzed Reactions**

#### **2.5.1. Molecular Mechanics**

“Molecular mechanics” (MM) methods can model protein structure and dynamics well, but standard MM methods cannot be used directly to model chemical reactions, because of their simple functional forms (e.g., harmonic terms for bond stretching, and an inability to model changes in electronic distribution because of the invariant atomic point partial charge model). The simplicity of MM “force fields” (potential functions) allows long timescale (e.g., now up to millisecond) simulations of protein dynamics, and simulations of large proteins. Molecular dynamics simulations can study conformational changes (which are rate-limiting in many enzymes under typical conditions): e.g., simulations of the human scavenger decapping enzyme (DcpS) found a cooperative periodic opening and closing of the dimer, over tens of nanoseconds (48). Molecular dynamics simulations can also investigate substrate conformational behavior, which can help to develop mechanistic ideas,

but simulations of substrate complexes without consideration of the reaction can sometimes be misleading (49).

Computer programs for biomolecular dynamics simulations include AMBER (50), CHARMM (51), GROMOS (52), NAMD (53) and TINKER (54); these should not be confused with force fields, which may have the same or similar names. A force field consists of the energy function and the parameters. Current protein force fields use similar potential energy functions, in which bonds and valence angles are represented by harmonic terms, electrostatic interactions are represented by invariant point charges on atoms, a simple representation that cannot capture the full electrostatic properties of a molecule. Dispersion and exchange repulsion are included by a simple Lennard-Jones function (usually of the 12-6 variety). Current widely used all-atom MM force fields for proteins are OPLS/AA (55, 56), CHARMM22-27 (57), and AMBER (PARM99 (50, 58–60)). Force fields for other types of biological macromolecules (e.g., lipids, nucleic acids (61, 62)) consistent with these protein force fields have also been developed (e.g., the CHARMM27 (63, 64) and AMBER nucleic acid parameters (65) and CHARMM parameters for lipids (66)). Most biomolecular MM force fields have been developed to be compatible with simple point charge models of water, e.g., the TIP3P water model (67). Current standard biomolecular MM force fields only include electronic polarization in an average, invariant way. The next generation of protein MM force fields will probably include electronic polarization explicitly (68, 69).

Standard MM potential functions cannot be applied to model the breaking and making of bonds (and electronic reorganization) in a chemical reaction. Also, MM force field parameters are developed based on the properties of stable molecules, and so will usually not be applicable to transition states and intermediates. MM functions and parameters can be developed specifically for reactions (e.g., using different functional forms, such as Morse curves to allow bond breaking), which has been successful for organic reactions in solution (70). Such parameters are, however, applicable only to a particular reaction, or small class of reactions. Also, the form of the potential function imposes limitations, such as the neglect of electronic polarization.

### 2.5.2. Empirical Valence Bond Methods

In the empirical valence bond (EVB) method (6), a few resonance structures are chosen to represent the reaction. The energy of each resonance form is given by a simple empirical force field, with the potential energy given by solving the related secular equation. The EVB Hamiltonian is calibrated to reproduce experimental data for a reaction in solution, or ab initio results can be used (71). The surrounding protein and solvent are modelled by an empirical force field, with appropriate treatment of long-range electrostatics. The activation free energy of activation is calculated from free

energy perturbation simulations (72). The free energy surfaces can be calibrated by comparison with experimental data for reactions in solution. The EVB method allows the use of a non-geometrical reaction coordinate, which allows evaluation of nonequilibrium solvation effects (6). A mapping procedure gradually moves the system from the reactants to products. The simplicity of the EVB potential function allows extensive molecular dynamics simulations, giving good sampling (73). The EVB method is now widely used for studying reactions in condensed phases, particularly in enzymes (74–80).

### 2.5.3. Quantum Chemical Calculations on Small (Active Site) Models

In most enzymes, the chemical changes occurring in the reaction are confined to a relatively small region, the active site of the enzyme. One approach to the study of enzyme-catalyzed reactions is to study only the active site, using quantum chemical methods (this is sometimes called the “supermolecule” or cluster approach). Such models can represent important features of an enzyme reaction, and can identify likely mechanisms. The active site model should contain molecules representing the substrate(s) (and any cofactors) and enzyme residues involved in the chemical reaction or in binding substrate. Important functional groups (such as catalytic amino acid side chains) are represented by small molecules, e.g., acetate can represent an aspartate side chain, imidazole for histidine, etc.). The initial positions of these groups are usually coordinates taken from a crystal structure, or from a molecular dynamics simulation of an enzyme complex.

Quantum chemical calculations (i.e., methods that calculate molecular electronic structure using quantum mechanics, e.g., ab initio molecular orbital or density functional theory calculations) can give excellent results for reactions of small molecules. The best “ab initio” methods (such as CCSD(T)), which include correlation between electrons, can calculate rate constants for reactions involving very few atoms (in the gas phase) with small error bars, similar to experiments on these systems. Such calculations require very large computational resources, however, severely limiting the size of the system that can be treated. More approximate methods, (such as the semiempirical molecular orbital techniques AM1 and PM3), are computationally much cheaper, and can model larger systems (containing of the order of hundreds of atoms). Techniques (e.g., “linear-scaling” methods) have been developed that allow semiempirical electronic structure calculations on whole proteins (81–83). Semiempirical methods are, however, inaccurate for many applications (e.g., typical errors of over 10 kcal/mol for barriers and reaction energies, though specifically parameterized semiempirical methods can give improved accuracy for a particular reaction (40, 47)). Density functional theory (DFT) methods (e.g., applying the B3LYP functional) are considerably more accurate, while also allowing calculations on relatively large systems

(e.g., active site models of the order of 100 atoms), larger than is feasible with correlated ab initio calculations. Many DFT methods, however, lack important physical interactions, such as dispersion, which are important in the binding of ligands to proteins. Dispersion effects can also be important in the calculation of energy barriers (84). DFT often gives barrier heights that are too low by several kcal/mol, and it can be difficult to assess the accuracy of results, because DFT does not offer a route to their systematic improvement or testing.

Calculations on active site models can provide models of transition states and intermediates (see below). This has proved particularly useful for studying metalloenzymes, using DFT methods. In many metalloenzymes, all the important chemical steps take place at one metal center (or a small number of metal ions bound at one site). The metal also holds its ligands in place, limiting the requirement for restraints to maintain the correct active site structure. Calculations on small clusters can give useful mechanistic insight (85, 86): e.g., a mechanism can be ruled out if the calculated barriers for it are significantly higher than the experimentally derived activation energy, based on the accuracy of the computational method. The effects of the environment are usually either omitted, or included only in an approximate way (e.g., by continuum solvation methods, which cannot fully represent the heterogeneous electrostatic environment in an enzyme). It is useful to test the sensitivity of the results to the choice of, e.g., dielectric constant.

To calculate the energy barrier for a reaction in a cluster model, structures of the reactant, transition state, intermediates and products of the reaction should be optimized. Doing this while maintaining the correct orientations of the groups in the protein can be difficult. Small models may also lack some important functional groups. It is important to consider carefully which groups to include, striking a balance between computational feasibility and the desire for a larger, more extensive model. A larger model is not, however, always a better model: a larger model will be susceptible to greater conformational complexity: conformational changes, even outside the active site, may artificially affect relative energies along the reaction path). Also, charged groups can have unrealistically large effects on reaction energies. One should test the sensitivity of the results to the choice of model (and also to factors such as the choice of density functional).

#### *2.5.4. Combined Quantum Mechanics/Molecular Mechanics Methods for Modelling Enzyme Reactions*

“Hybrid” methods that combine quantum chemical methods with molecular mechanics allow more extensive calculations, on larger models of enzymes, than is possible with purely quantum chemical techniques. Such QM/MM methods are very important in computational enzymology. The QM/MM approach is simple: a small part of the system, at the active site, is treated quantum mechanically, i.e., by an electronic structure method of one of the types

discussed above, which allows modelling of the electronic rearrangements involved in the breaking and making of chemical bonds. The QM region contains the reacting groups of the enzyme, substrate and any cofactors. The rest of the system is treated by MM. QM/MM calculations can be carried out at ab initio (87) or semiempirical (88) molecular orbital, density functional (89) or approximate density functional (e.g., self-consistent charge density functional tight-binding, SCC-DFTB (90)) levels of QM electronic structure calculation. Different types of coupling between the QM and MM regions are possible (see below).

Many different QM/MM implementations are available, in several widely used programs. Reaction pathways and transition state structures can be optimized (91, 92). Molecular dynamics simulations are possible with cheaper QM/MM methods (such as semiempirical or SCC-DFTB level QM) (93). Free energy differences, such as activation free energies can be calculated, as can quantum effects such as tunnelling and zero-point corrections (5, 30, 40). High-level QM/MM calculations (e.g., ab initio or density functional level QM) are required for some systems and also have an important role in testing more approximate methods. The computational demands of high level (e.g., ab initio, (94)) QM/MM calculations (17) typically limit their application to “single point” energy calculations on structures optimized at lower levels (95). DFT/MM methods can be used for energy minimization/geometry optimization to generate reaction paths.

QM/MM methods can also be used in free energy perturbation simulations (96), e.g., to calculate relative binding affinities, and in molecular docking and scoring (97). QM/MM methods provide several advantages over MM methods in studies of ligands bound to proteins, including potentially a better physical description of a ligand (e.g., including electronic polarization), and avoiding the need for time-consuming MM parameterization for the ligand.

#### Interactions Between the QM and MM Regions

One of the main differences between various QM/MM methods is the type of QM/MM coupling employed i.e., in how the interactions (if any) between the QM and MM systems are treated (98). The simplest linking of QM and MM methods involves a straightforward “mechanical” embedding of the QM region in the MM environment, treating interactions between the QM and MM regions only by MM (i.e., the QM system is represented by (MM) point charges in its interaction with the MM environment). In calculations of this type, the QM/MM energy of the whole system,  $E_{\text{TOTAL}}^{\text{QM/MM}}$ , is calculated in a simple subtractive scheme. This simple subtractive approach can be applied to all combinations of theory levels (for example combining different levels of QM treatment (QM/QM) as opposed to QM with MM) and forms the basis for the (simplest form of the) multilayer

ONIOM (Our own N-layered Integrated molecular Orbital and molecular Mechanics) method (99). A QM/QM calculation involves a high and a low level of QM theory, with a small region treated by a high level and the *entire* model treated at the low level; polarization is included at the lower level of QM theory.

More intensive QM/MM calculations include polarization of the QM region by the MM environment. This is likely to be important for many enzymes, given their polar nature. QM/MM methods of this type include electrostatic interactions between the QM and MM regions in the QM calculation, thus modelling polarization of the QM system by the MM atoms, by directly including the MM atomic charges of the MM group in the QM calculation. The electronic structure calculation thus includes the effects of the MM atoms. A further level of complexity would involve polarization of the MM region also through the use of a polarizable MM force field, and potentially self-consistent polarization of the MM region through an iterative procedure. Models of this sort are vastly more computationally intensive and may not always yield better results (100). QM/MM methods that include polarization of the MM system have been developed for small molecular systems (101). Current standard MM force fields for biological macromolecules do not model changes in polarization, however.

In typical QM/MM calculations, the energy of the QM atoms,  $E_{QM}$ , is given by a molecular orbital or DFT method, and the energy of the atoms in the MM region,  $E_{MM}$ , is given by MM. A boundary term,  $E_{Boundary}$ , is usually necessary to account for the effects of the surroundings, e.g., to include the effects of parts of the protein that are not included in the simulation model. It may also be necessary to scale/reduce charges at the boundary of the simulation system: this represents the effects of dielectric screening in a crude sense, to avoid overestimating the effects of charged groups on the active site (102). The QM/MM interaction energy,  $E_{QM/MM}$  typically consists of terms due to electrostatic interactions and van der Waals interactions, and any bonded interaction terms. In many implementations, MM bonding terms (energies of bond stretching, angle bending, torsion angle rotation, etc.) are included for all QM/MM interactions which involve at least one MM atom (88). In an ab initio QM/MM calculation, the MM atomic charges are generally included directly through one-electron integrals. The treatment of QM/MM electrostatic interactions is a little less straightforward when semiempirical molecular orbital methods such as AM1 and PM3 are used, because they treat only valence electrons directly, including the core electrons together with the nucleus as an atomic “core.” In semiempirical QM/MM methods such as the AM1/CHARMM method of Field et al., the electrostatic interactions between QM and MM atoms are calculated by treating the MM atoms as if they were semiempirical atomic cores (88).

QM/MM van der Waals interactions (representing dispersion and exchange repulsion interactions between QM and MM atoms) are usually calculated by a molecular mechanics procedure (e.g., through Lennard-Jones terms), exactly as the corresponding interactions would be calculated between MM atoms not interacting through bonding terms. MM van der Waals parameters must therefore be chosen for each QM atom: these interactions are significant at short distances, and are important in determining QM/MM interaction energies and geometries. The van der Waals parameters are important in differentiating MM atom types in their interactions with the QM system, e.g., for MM atoms of the same charge, which would otherwise be indistinguishable to the QM system; van der Waals interactions are also important for interactions of the QM system with nearby MM atoms whose charges are close to zero. Often, standard MM van der Waals (Lennard-Jones) parameters optimized for similar MM groups are used for QM atoms in QM/MM calculations. This is convenient, but it is always important to consider whether the van der Waals parameters provide a reliable description of QM/MM interactions. Where necessary, the (MM) van der Waals parameters for the QM atoms can be optimized to reproduce experimental or high-level ab initio results (e.g., structures and interaction energies) for small molecular complexes (93). One limitation of current QM/MM approaches of this type is that the same van der Waals parameters are typically used for the QM atoms throughout a simulation: in modelling a chemical reaction, the chemical nature of the groups involved (treated by QM) may change, altering their interactions, and so the use of unchanging MM parameters may be inappropriate. Riccardi et al. have investigated the effects of van der Waals parameters in QM/MM (SCC-DFTB/CHARMM22) simulations (103). Different parameter sets gave very different results for gas-phase clusters and solvent structures around the solutes. However, condensed phase thermodynamic quantities (e.g., the calculated reduction potential and potential of mean force) were less sensitive to the van der Waals parameters. These authors concluded that work to improve the reliability of QM/MM methods for condensed phase energetic properties should focus on factors other than van der Waals interactions between QM and MM atoms, such as the treatment of long-range electrostatic interactions.

#### Treatment of Long-Range Electrostatic Interactions in QM/MM Simulations

To reduce computational requirements, the model may include only a part of the whole protein (for example, a rough sphere around the active site). In simulating a truncated protein system, it is necessary to include restraints or constraints in the boundary region to force the atoms belonging to it to remain close to their positions in the crystal structure. One common approach is the stochastic boundary molecular dynamics method (104, 105), in which the simulation system is divided into a reaction region,

a buffer region and a reservoir region. Typically, the whole simulation system may include all residues with an atom within a distance of e.g., 15–25 Å of an atom in the active site. The buffer region would contain atoms in the outer layer (e.g., outer 5 Å shell). Atoms in the reaction region are treated by standard Newtonian molecular dynamics, and are not subject to positional restraints. The protein heavy atoms in the buffer region are restrained to remain close to their (e.g., crystallographically determined) positions by harmonic forces, while a solvent deformable boundary potential prevents “evaporation” of water. Atoms in the buffer are subject to frictional and random forces (hence the term “stochastic”) to represent exchange of energy with the surroundings (reservoir region). Atoms in the reservoir region are usually not included because their presence (as fixed atoms) has been found to cause excessive rigidity of the protein.

Ideally, long-range electrostatic interactions should be included explicitly. Schemes for treatment of long-range electrostatic interactions in QM/MM simulations have been developed, to allow periodic boundary simulations (106). An alternative approach, for QM/MM calculations under spherical boundary conditions (107) is the generalized solvent boundary potential (GSBP) method (108). This retains the practical advantage of treating a truncated system, avoiding having to include the entire macromolecule in a periodic simulation. The effects of the bulk solvent and macromolecule atoms outside the simulation system are included at the Poisson-Boltzmann level. Simulations using the GSBP method were found to be more consistent with experimental data. Conventional stochastic boundary molecular dynamics simulations produced artifacts, depending on the treatment of electrostatic interactions. It was suggested that the commonly used interaction truncation schemes should not be applied if possible in QM/MM simulations, in particular for simulations that may involve extensive conformational sampling.

#### QM/MM Partitioning Methods and Schemes

Most QM/MM studies of enzymes require partitioning of covalently bonded molecules into QM and MM regions. Typically, some amino acid side chains participate in the reaction, and must therefore be included in the QM region. Other side chains will play binding roles, and a MM representation might be inadequate. Similarly, it may be more practical to treat only the reactive parts of large cofactors or substrates by quantum chemical methods. There are two general QM/MM partitioning techniques that can be employed: firstly special treatment of orbitals to satisfy the valence shell of the QM atom at the QM/MM junction, for example the local self-consistent field (LSCF) method (109, 110) or the generalized hybrid orbital (GHO) method (111). Alternatively a QM atom (or pseudoatom) can be added at the

QM/MM boundary, e.g., using a “link atom” or connection atom method.

The local self-consistent field (LSCF) method (112) uses a strictly localized bond orbital, also often described as a frozen orbital, for the QM atom at the frontier between QM and MM regions. The electron density of the orbital is calculated in advance, using small models, and does not change during the QM/MM calculation. The orbitals must be parameterized for each system and QM method. The LSCF method avoids the need for dummy atoms and provides a reasonable description of the chemical properties of the frontier bond. It has been applied at semiempirical (113) and ab initio (112) QM/MM levels.

The generalized hybrid orbital (GHO) method (114) uses hybrid orbitals as basis functions on the frontier atom of the MM fragment. It does not require extensive specific parameterization, unlike the LSCF method. It uses four hybrid orbitals for an  $sp^3$  carbon atom, one of which is included in the self-consistent field optimization of the QM region, while three are treated as auxiliary orbitals. The parameters for the frontier atom are optimized to reproduce properties of full QM systems. The localized orbitals can be transferred, without specific parameterization of the active orbital for each new system. A similar approach in DFT/MM calculations is to freeze the electron density at the QM/MM junction (115). The GHO method has been applied in QM/MM calculations at ab initio (116), SCC-DFTB (117) and density functional (118) QM levels.

The “dummy junction atom” or link atom method introduces so-called link atoms to satisfy the valence of the frontier atom in the QM system (119). The link atom is usually a hydrogen atom (88), but other atom types have been used, such as a halogen (120). The link atom approach has been criticized, e.g., because it introduces additional degrees of freedom associated with the link atom, and the fact that a C–H bond is clearly not exactly equivalent to a C–C bond. It is, however, simple and is widely used. The results can be sensitive to the positioning of the link atom, and also on exactly which MM atoms are excluded from the classical electrostatic field that interacts with the QM region. Comparison of the LSCF and link atom approaches for semiempirical QM/MM calculations, however, showed that the two methods gave similar results (121). It has been recommended that the link atom should interact with all MM atoms except for those closest to the QM atom to which the link atom is bonded. The link atom method can give good results, with a good choice of the boundary between QM and MM regions, e.g., across a carbon–carbon single bond, far from chemical changes, and also preferably not close to highly charged MM atoms.

Another method for dealing with the QM/MM boundary between covalently bonded atoms is the connection atom method

(110, 122), which uses a monovalent pseudoatom instead of a link atom. The parameters for the connection atom are optimized for the partitioned covalent bond. The connection atoms interact with the other QM atoms as a (specifically parameterized) QM atom, and with the other MM atoms as a standard carbon atom. This avoids the problem of a supplementary atom in the system, as the connection atom and the classical frontier atom are unified. However, the need to reparameterize for each type of covalent bond at a given level of quantum chemical theory is potentially laborious (123). The connection atom method has been developed for AM1 and PM3 (100), and DFT (122) QM/MM calculations. Tests indicated that it is more accurate than the standard link atom approach (100).

To overcome some of the problems that can arise with the single link atom method (e.g., an unphysical dipole), Brooks et al. have proposed a “double link atom” method (124). Also, their Gaussian delocalization method for MM atomic charges could simplify the calculation of energies and forces: e.g., even at short distances, the delocalized Gaussian MM method does not require the MM host atom charge to be excluded from the QM calculation, as would be necessary when treating it as a point charge. The delocalized Gaussian method can be combined with many QM/MM partitioning techniques, such as the link atom, frozen orbital, or pseudopotential methods. Tests of the delocalized Gaussian MM and double link atom methods on small model systems indicated that these methods gave better energetic properties than point atomic MM charge and single link atom methods.

Cui et al. have tested link atom QM/MM partitioning methods, for the SCC-DFTB QM method (125), including all the options available in the CHARMM program, which differ in their treatment of electrostatic interactions with the MM atoms close to the QM/MM frontier. They also proposed a divided frontier charge protocol, in which the partial charge associated with the MM atom bonded to the QM atom is distributed evenly on the other MM atoms in the same MM group. Tests of these various link atom schemes showed that QM/MM proton affinities and deprotonation energies are highly dependent on the particular link atom scheme employed: standard single link atom methods gave errors of up to 15–20 kcal/mol compared to pure QM calculations. Other schemes were found to give better results. Activation barriers and reaction energies were found, however, to be fairly insensitive to the choice of link atom scheme (e.g., within 2–4 kcal/mol) because of cancellation of errors. This is encouraging: the effect of using different link atom schemes in QM/MM simulations was found to be relatively small for chemical reactions in which the total charge does not change. Other technical details, such as the treatment of long-range electrostatics, are likely to play a more significant role in determining energetics generally, and should be treated carefully for reliable results.

## **2.6. Modelling Enzyme Reactions by Calculating Potential Energy Surfaces**

With QM or QM/MM methods, potential energy surfaces of enzyme reaction mechanisms can be explored at a level of accuracy that can enable discrimination between different mechanisms: e.g., if the barrier for a proposed mechanism is significantly larger than that derived from experiment (using transition state theory), within the limits of accuracy of the computational method and experimental error, then that mechanism can be considered to be unlikely. A mechanism with a calculated barrier comparable to the apparent experimental barrier (for that step, or failing that for the overall reaction) is more likely. However, to calculate rate constants also requires reliable methods to calculate enthalpies, energies, and free energies of reaction and activation, given the potential energy surface. Traditional approaches to modelling reactions (e.g., in the gas phase) rely on the identification of stationary points (reactants, products, intermediates, transition states) via geometry optimization, followed by computation of second derivatives to enable relatively simplistic evaluation of zero-point corrections, thermal and entropy terms. Algorithms developed for small molecules are often not suitable for large systems: e.g., direct calculation, storage and manipulation of Hessian matrices becomes extremely difficult. A basic means of modelling approximate reaction paths is the “adiabatic mapping” or “coordinate driving” approach. The energy of the system is calculated by minimizing the energy at a series of fixed (or restrained, e.g., by harmonic forces) values of a reaction coordinate, e.g., the distance between two atoms. This approach has been applied with success to many enzymes (1, 2), but it is only valid if one conformation of the protein can represent the state of the system at a particular value of the reaction coordinate. A single minimum energy structure of this conformation may adequately represent the several closely related structures making up the reacting conformational state. Minimizing the QM/MM potential energy of such a representative conformation along the reaction coordinate should give a reasonable approximation of the enthalpic component of the potential of mean force (the free energy profile) for the reaction. In contrast, simple calculations of potential energy surfaces are likely to be unsuccessful or misleading for enzyme reactions involving large movements of charge or large changes in solvation (e.g., particularly for solvent-exposed sites, where rearrangement of water molecules might involve an unrealistically large energy penalty where adiabatic mapping calculations are used, (126)).

Due to the complexity of protein internal motions, many conformational substates exist, and a single structure might not be truly representative. If this is the case, calculations including extensive sampling of the system to obtain configurationally averaged free-energy changes are needed, as opposed to energy minimizations, which do not include entropic effects and are sensitive to starting geometries (126). A more simple approach to investigating

conformational effects is to use molecular dynamics simulations, and/or to use multiple different crystal structures, to generate multiple models for mechanism (e.g., adiabatic mapping) calculations, to ensure wide sampling of possible enzyme configurations (127), with averaging, or Boltzmann-weighted averaging, of energy barriers.

Despite the limitations and drawbacks of the adiabatic mapping approach, it has been applied in many QM/MM applications. It has the advantage that it is simple to apply, and does not require intensive calculations, such as second derivative evaluations, or simultaneous treatment of several points on a pathway. It can be useful for initial scans of potential energy surfaces, and for generating approximate models of transition states and intermediates, in which some allowance is made for structural relaxation to chemical changes at the active site. It is suitable only for reactions involving small chemical and structural changes, involving a small number of groups. For some enzymes, this type of approach has been validated through a correlation of calculated QM/MM barriers with activation energies derived from experiment (37).

As mentioned before, approaches based purely on calculations of a potential energy surface may not account for significant conformational fluctuations of the protein. Conformational changes, even on a small scale, may be coupled to, or significantly affect, chemical changes. Fluctuations of the active site can greatly affect the energy barrier. In the case of fatty acid amide hydrolase, conformational fluctuations do not affect the general shape of the potential energy surfaces, but consistency between experimental and calculated barriers is observed only with a specific (and rarely occurring) arrangement of the enzyme-substrate complex (49). These findings indicate that investigation of different protein conformations is essential for a meaningful determination of the energetics of enzymic reactions for calculations of potential energy profiles or surfaces.

## **2.7. Calculating Free Energy Profiles for Enzyme-Catalyzed Reactions**

The rate constant of a reaction is actually related not to the potential energy barrier, but to the free energy barrier, according to transition state theory. The techniques above calculate potential energy barriers, for a particular conformation. Techniques that sample configurations along a reaction coordinate give a more sophisticated and extensive description, by taking account of multiple conformations and estimating entropic effects, and can be essential for modelling some types of enzyme reactions. Simulations of this type provide estimates of the free energy profile along a specific (reaction) coordinate, which is often referred to as the potential of mean force. Molecular dynamics and Monte Carlo methods in principle allow such sampling, but do not provide good sampling of high energy regions, such as in the vicinity of transition states. Conformational

sampling of processes of chemical change therefore requires specialized techniques, e.g., to bias the simulation to sample the transition state region. Umbrella sampling is such a method, which is widely used in molecular dynamics simulations e.g., with QM/MM techniques, to model enzymic reactions (27). In this technique, a biasing potential is applied to force the system to remain close to a specific value of a defined reaction coordinate. Often, an umbrella sampling simulation will begin with simulation of a transition state or reactant complex; an umbrella (e.g., harmonic) potential restrains the reaction coordinate to a value corresponding to e.g., the reactants. In other, subsequent simulations, the reference value of the restraint is changed by a small amount to sample other regions of the reaction coordinate. Often, the reaction coordinate is defined in terms of bond lengths, in which case a typical difference between the points would be 0.1–0.2 Å. The neighboring potentials should give overlapping distributions: this can be achieved by choosing an appropriate spacing of reaction coordinate values for different simulations, and an appropriate magnitude of the force constant of the restraint. The number of simulations is a balance between accuracy and efficiency. The reaction coordinate values during the (restrained) simulations are recorded. The effects of the restraining potentials are removed in the analysis and combined, typically by the weighted histogram analysis method (WHAM). This gives the unbiased potential of mean force along the reaction coordinate. It is important also to test for convergence with respect to length (and numbers) of simulations.

QM/MM umbrella sampling simulations are possible with low levels of QM theory, such as semiempirical molecular orbital methods (e.g., AM1 or PM3). Often, such methods are highly inaccurate for reaction barriers and energies. Their accuracy can be improved significantly by reparameterization for a specific reaction. For example, specifically parameterized semiempirical QM/MM methods have been used to investigate model reactions of glutathione-S-transferase (GST) enzymes. QM/MM umbrella sampling molecular dynamics simulations of the reaction of phenanthrene 9,10-oxide in a glutathione-S-transferase, identified a single amino acid as a likely determinant of stereospecificity in the epoxide ring opening (93). Similarly, specifically parameterized QM/MM methods have been applied to model the reaction between glutathione and 1-chloro-2,4-dinitrobenzene. The results of QM/MM umbrella sampling molecular dynamics simulations of this reaction in the M1-1 GST isoenzyme, in mutant enzymes, and in solution, agreed very well with experiment (128). QM/MM molecular dynamics simulations are much more computationally demanding than MM simulations, because of the computational expense of the evaluation of the QM forces. Typical QM/MM umbrella sampling applications have involved trajectories of picoseconds to nanoseconds (multiple simulations of 30 ps each, at each value of the

reaction coordinate, in the case of Bowman et al. (128)), with semiempirical QM methods. Approaches based on Monte Carlo simulations (129, 130) avoid the requirement for force calculations, and are also promising.

### 3. Notes

The choice of an appropriate method for the particular enzyme and questions of interest is vital. Careful testing and validation is important. Quantitative predictions of reaction rates or the effects of mutation remain very challenging, but for many enzymes, with appropriate methods, useful predictions can be made with some confidence. It is important to validate predictions from modelling by comparisons with experimental data. An example is comparison of calculated barriers for a series of alternative substrates with activation energies derived from experimental rates: demonstration of a correlation can validate mechanistic calculations (36, 131). Some enzymes have become important model systems in the development and testing of computational methods and protocols: these include chorismate mutase (1, 8, 10, 17, 27, 87), citrate synthase (3, 23, 132), P450<sub>cam</sub> (1, 2, 7, 11, 84), *para*-hydroxybenzoate hydroxylase (2, 17, 36), triosephosphate isomerase (5, 13, 102), fatty acid amide hydrolase (49, 95, 127, 129), and methylamine dehydrogenase (5, 7, 46, 47, 133).

### Acknowledgments

AJM is an EPSRC Leadership Fellow.

### References

1. Lonsdale R, Ranaghan KE, Mulholland AJ (2010) Computational enzymology. *Chem Commun* 46:2354–2372
2. Senn HM, Thiel W (2009) QM/MM methods for biomolecular systems. *Angew Chem Int Ed Engl* 48:1198–1229
3. van der Kamp MW, Mulholland AJ (2008) Computational enzymology: insight into biological catalysts from modelling. *Nat Prod Rep* 25:1001–1014
4. Mulholland AJ (2005) Modelling enzyme reaction mechanisms, specificity and catalysis. *Drug Discov Today* 10:1393–1402
5. Garcia-Viloca M, Gao J, Karplus M, Truhlar DG (2004) How enzymes work: analysis by modern rate theory and computer simulations. *Science* 303:186–195
6. Warshel A (2003) Computer simulations of enzyme catalysis: methods, progress, and insights. *Annu Rev Biophys Biomol Struct* 32:425–443
7. Friesner RA, Guallar V (2005) *Ab initio* quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu Rev Phys Chem* 56:389–427

8. Martí S, Roca M, Andrés J, Moliner V, Silla E, Tuñón I, Bertrán J (2004) Theoretical insights in enzyme catalysis. *Chem Soc Rev* 33:98–107
9. Himo F (2006) Quantum chemical modeling of enzyme active sites and reaction mechanisms. *Theor Chem Acc* 116:232–240
10. Warshel A, Sharma PK, Kato M, Xiang Y, Liu HB, Olsson MHM (2006) Electrostatic basis for enzyme catalysis. *Chem Rev* 106:3210–3235
11. Senn HM, Thiel W (2007) QM/MM studies of enzymes. *Curr Opin Chem Biol* 11:182–187
12. Mulholland AJ, Grant GH, Richards WG (1993) Computer modelling of enzyme catalysed reaction mechanisms. *Protein Eng* 6:133–147
13. Mulholland AJ, Karplus M (1996) Simulations of enzymic reactions. *Biochem Soc Trans* 24:247–254
14. Åqvist J, Warshel A (1993) Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem Rev* 93:2523–2544
15. Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227–249
16. Scheiner S, Lipscomb WN (1976) Catalytic mechanism of serine proteinases. *Proc Natl Acad Sci USA* 73:432–436
17. ClaeysSENS F, Harvey JN, Manby FR, Mata RA, Mulholland AJ, Ranaghan KE, Schultz M, Thiel S, Thiel W, Werner H-J (2006) High accuracy computation of reaction barriers in enzymes. *Angew Chem Int Ed* 45:6856–6859
18. Mulholland AJ (2007) Chemical accuracy in QM/MM calculations on enzyme-catalysed reactions. *Chem Cent J* 1:19
19. Braun-Sand S, Olsson MHM, Warshel A (2005) Computer modeling of enzyme catalysis and its relationship to concepts in physical organic chemistry. *Adv Phys Org Chem* 40:201–245
20. Vocadlo DJ, Davies GJ, Laine R, Withers SG (2001) Catalysis by hen egg white lysozyme proceeds via a covalent intermediate. *Nature* 412:3835–3838
21. Bowman AL, Grant IM, Mulholland AJ (2008) QM/MM simulations predict a covalent intermediate in the hen egg white lysozyme reaction with its natural substrate. *Chem Commun* 7:4425–4427
22. Cleland WW, Frey PA, Gerlt JA (1998) The low barrier hydrogen bond in enzymatic catalysis. *J Biol Chem* 273:25529–25532
23. Mulholland AJ, Lyne PD, Karplus M (2000) *Ab initio* QM/MM study of the citrate synthase mechanism: a low-barrier hydrogen bond is not involved. *J Am Chem Soc* 122:534–535
24. Schutz CN, Warshel A (2004) The low barrier hydrogen bond (LBHB) proposal revisited: the case of the Asp · His pair in serine proteases. *Proteins* 55:711–723
25. Molina PA, Jensen JH (2003) A predictive model of strong hydrogen bonding in proteins: the Nδ1–H–Oδ1 hydrogen bond in low-pH  $\alpha$ -chymotrypsin and  $\alpha$ -lytic protease. *J Phys Chem B* 107:6226–6233
26. Hur S, Bruice TC (2003) Just a near attack conformer for catalysis (chorismate to prephenate rearrangements in water, antibody, enzymes, and their mutants). *J Am Chem Soc* 125:10540–10542
27. Ranaghan KE, Mulholland AJ (2004) Conformational effects in enzyme catalysis: QM/MM free energy calculation of the ‘NAC’ contribution in chorismate mutase. *Chem Commun* 10:1238–1239
28. Olsson MH, Parson WW, Warshel A (2006) Dynamical contributions to enzyme catalysis: critical tests of a popular hypothesis. *Chem Rev* 106:1737–1756
29. Hammes-Schiffer S, Watney JB (2006) Hydride transfer catalyzed by *Escherichia coli* and *Bacillus subtilis* dihydrofolate reductase: Coupled motions and distal mutations. *Phil Trans Roy Soc B* 361:1365–1373
30. Masgrau L, Roujeinikova A, Johannissen LO, Horth P, Basran J, Ranaghan KE, Mulholland AJ, Sutcliffe MJ, Scrutton NS, Leys D (2006) Atomic description of an enzyme reaction dominated by proton tunneling. *Science* 312:237–241
31. Hatcher E, Soudackov AV, Hammes-Schiffer S (2007) Proton-coupled electron transfer in soybean lipoxygenase: dynamical behavior and temperature dependence of kinetic isotope effects. *J Am Chem Soc* 129:187–196
32. Limbach HH, Lopez JM, Kohen A (2006) Arrhenius curves of hydrogen transfers: tunnel effects, isotope effects and effects of pre-equilibria. *Phil Trans Roy Soc B* 361:1399–1415
33. Nagel ZD, Klinman JP (2006) Tunneling and dynamics in enzymatic hydride transfer. *Chem Rev* 106:3095–3118

34. Villa J, Strajbl M, Glennon TM, Sham YY, Chu ZT, Warshel A (2000) How important are entropic contributions to enzyme catalysis? *Proc Natl Acad Sci USA* 97:11899–11904
35. Lonsdale R, Harvey JN, Manby FR, Mulholland AJ (2011) Comment on “A stationary-wave model of enzyme catalysis” by Carlo Canepa. *J Comput Chem* 32:368–369
36. Ridder L, Harvey JN, Rietjens IMCM, Vervoort J, Mulholland AJ (2003) Ab initio QM/MM modeling of the hydroxylation step in *p*-hydroxybenzoate hydroxylase. *J Phys Chem B* 107:2118–2126
37. Ridder L, Mulholland AJ, Rietjens IMCM, Vervoort J (2000) A quantum mechanical/molecular mechanical study of the hydroxylation of phenol and halogenated derivatives by phenol hydroxylase. *J Am Chem Soc* 122:8728–8738
38. Bjelic S, Åqvist J (2004) Prediction of structure, substrate binding mode, mechanism, and rate for a malaria protease with a novel type of active site. *Biochemistry* 43:14521–14528
39. Zhang YK, Kua J, McCammon JA (2003) Influence of structural fluctuation on enzyme reaction energy barriers in combined quantum mechanical/molecular mechanical studies. *J Phys Chem B* 107:4459–4463
40. Gao JL, Truhlar DG (2002) Quantum mechanical methods for enzyme kinetics. *Annu Rev Phys Chem* 53:467–505
41. Karplus M, Gao YQ, Ma JP, van der Vaart A, Yang W (2005) Protein structural transitions and their functional role. *Phil Trans R Soc A* 363:331–355
42. Fersht A (1999) Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding. Freeman, New York
43. Olsson MHM, Warshel A (2004) Solute solvent dynamics and energetics in enzyme catalysis: the S(N)2 reaction of dehalogenase as a general benchmark. *J Am Chem Soc* 126:15167–15179
44. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipayiou G, Eisenmesser EZ, Kern D (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol* 11:945–949
45. Kohen A, Cannio R, Bartolucci S, Klinman JP (1999) Enzyme dynamics and hydrogen tunneling in a thermophilic alcohol dehydrogenase. *Nature* 399:496–499
46. Sutcliffe MJ, Masgrau L, Roujeinikova A, Johannissen LO, Hothi P, Basran J, Ranaghan KE, Mulholland AJ, Leys D, Scrutton NS (2006) Hydrogen tunnelling in enzyme-catalysed H-transfer reactions: flavoprotein and quinoprotein systems. *Phil Trans Roy Soc B* 361:1375–1386
47. Ranaghan KE, Mulholland AJ (2010) Computer simulations of quantum tunnelling in enzyme-catalysed hydrogen transfer reactions. *Interdiscip Sci* 2:78–97
48. Pentikäinen U, Pentikäinen OT, Mulholland AJ (2008) Cooperative symmetric to asymmetric conformational transition of the apo-form of scavenger decapping enzyme revealed by simulations. *Proteins* 70:498–508
49. Lodola A, Mor M, Zurek J, Tarzia G, Piomelli D, Harvey JN, Mulholland AJ (2007) Conformational effects in enzyme catalysis: reaction via a high energy conformation in fatty acid amide hydrolase. *Biophys J* 92: L20–L22
50. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods R (2005) The AMBER biomolecular simulation programs. *J Comput Chem* 26:1668–1688 (see <http://amber.scripps.edu/>)
51. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217, (See also <http://www.charmm.org>)
52. Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF (1999) The GROMOS biomolecular simulation program package. *J Phys Chem A* 103:3596–3607, (see <http://www.igc.ethz.ch/gromos/gromos.html>)
53. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802, (see <http://www.ks.uiuc.edu/Research/namd/>)
54. Ponder JW, Richards FM (1987) An efficient newton-like method for molecular mechanics energy minimization of large molecules. *J Comput Chem* 8:1016–1024, (see <http://dasher.wustl.edu/tinker/>)
55. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
56. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLSAA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 105:6474–6487

57. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
58. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic molecules. *J Am Chem Soc* 117:5179–5197
59. MacKerell AD (2005) Empirical force fields for proteins: current status and future directions. *Ann Rep Comp Chem* 1:91–111
60. Ponder JW, Case DA (2003) Force fields for protein simulations. *Adv Protein Chem* 66:27–75
61. Cheatham TE III (2005) Molecular modeling and atomistic simulation of nucleic acids. *Ann Rep Comp Chem* 1:75–90
62. Cheatham TE III (2004) Simulation and modeling of nucleic acid structure, dynamics and interactions. *Curr Opin Struct Biol* 14:360–367
63. Foloppe N, MacKerell AD (2000) All-atom empirical force field for nucleic acids: I parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* 21:86–104
64. MacKerell AD, Banavali NK (2000) All-atom empirical force field for nucleic acids: II application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* 21:105–120
65. Cheatham TE, Cieplak P, Kollman PA (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn* 16:845–862
66. Feller SE, Yin DX, Pastor RW, MacKerell AD (1997) Molecular dynamics simulation of unsaturated lipid bilayers at low hydration: parameterization and comparison with diffraction studies. *Biophys J* 73:2269–2279
67. Jorgensen WL, Tirado-Rives J (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc Natl Acad Sci USA* 102:6665–6670
68. Kim BC, Young T, Harder E, Friesner RA, Berne BJ (2005) Structure and dynamics of the solvation of bovine pancreatic trypsin inhibitor in explicit water: a comparative study of the effects of solvent and protein polarizability. *J Phys Chem B* 109:16529–16538
69. Shaw KE, Woods CJ, Mulholland AJ (2010) Compatibility of quantum chemical methods and empirical (MM) water models in quantum mechanics/molecular mechanics liquid water simulations. *J Phys Chem Lett* 1:219–223
70. Lim D, Jenson J, Repasky MP, Jorgensen WL (1999) Solvent as catalyst: computational studies of organic reactions in solution. In: Truhlar DG, Morokuma K (eds) *Transition state modeling for catalysis*. American Chemical Society, Washington DC
71. Bentzien J, Muller RP, Florian J, Warshel A (1998) Hybrid *ab initio* quantum mechanics molecular mechanics calculations of free energy surfaces for enzymatic reactions: the nucleophilic attack in subtilisin. *J Phys Chem B* 102:2293–2301
72. Warshel A (1997) Computer modeling of chemical reactions in enzymes and solutions. John Wiley & Sons, New York
73. Villa J, Warshel A (2001) Energetics and dynamics of enzymatic reactions. *J Phys Chem B* 105:7887–7907
74. Warshel A, Sharma PK, Chu ZT, Åqvist J (2007) Electrostatic contributions to binding of transition state analogues can be different from the corresponding contributions to catalysis: phenolates binding to the oxyanion hole of ketosteroid isomerase. *Biochemistry* 46:1466–1476
75. Bjelic S, Åqvist J (2006) Catalysis and linear free energy relationships in aspartic proteases. *Biochemistry* 45:7709–7723
76. Trobro S, Åqvist J (2005) Mechanism of peptide bond synthesis on the ribosome. *Proc Natl Acad Sci USA* 102:12395–12400
77. Sharma PK, Xiang Y, Katom M, Warshel A (2005) What are the roles of substrate-assisted catalysis and proximity effects in peptide bond formation by the ribosome? *Biochemistry* 44:11307–11314
78. Hammes-Schiffer S (2004) Quantum-classical simulation methods for hydrogen transfer in enzymes: a case study of dihydrofolate reductase. *Curr Opin Struct Biol* 14:192–201
79. Liu HB, Warshel A (2007) Origin of the temperature dependence of isotope effects in enzymatic reactions: the case of dihydrofolate reductase. *J Phys Chem B* 111:7852–7861

80. Sharma PK, Chu ZT, Olsson MHM, Warshel A (2007) A new paradigm for electrostatic catalysis of radical reactions in vitamin B12 enzymes. *Proc Natl Acad Sci USA* 2007 (104):9661–9666
81. Van derVaart A, Gogonea V, Dixon SL, Merz KM (2000) Linear scaling molecular orbital calculations of biological systems using the semiempirical divide and conquer method. *J Comput Chem* 21:1494–1504
82. Khandogin J, York DM (2004) Quantum descriptors for biological macromolecules from linear-scaling electronic structure methods. *Proteins* 56:724–737
83. Khandogin J, Musier-Forsyth K, York DM (2003) Insights into the regioselectivity and RNA-binding affinity of HIV-1 nucleocapsid protein from linear-scaling quantum methods. *J Mol Biol* 330:993–1004
84. Lonsdale R, Harvey JN, Mulholland AJ (2010) Inclusion of dispersion effects significantly improves accuracy of calculated reaction barriers for cytochrome p450 catalyzed reactions. *J Phys Chem Lett* 1:3232–3237
85. Himo F, Siegbahn PE (2003) Quantum chemical studies of radical-containing enzymes. *Chem Rev* 103:2421–2456
86. Siegbahn PE, Himo F (2009) Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *J Biol Inorg Chem* 14:643–651
87. Woodcock HL, Hodoscek M, Sherwood P, Lee YS, Schaefer HF, Brooks BR (2003) Exploring the quantum mechanical/molecular mechanical replica path method: a pathway optimization of the chorismate to prephenate Claisen rearrangement catalyzed by chorismate mutase. *Theor Chem Acc* 109:140–148
88. Field MJ, Bash PA, Karplus M (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J Comput Chem* 11:700–733
89. Lyne PD, Hodoscek M, Karplus M (1999) A hybrid QM-MM potential employing Hartree-Fock or density functional methods in the quantum region. *J Phys Chem A* 103:3462–3471
90. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J Phys Chem B* 105:569–585
91. Marti S, Moliner V (2005) Improving the QM/MM description of chemical processes: a dual level strategy to explore the potential energy surface in very large systems. *J Chem Theory Comput* 1:1008–1016
92. Prat-Resina X, Bofill JM, Gonzalez-Lafont A, Lluch JM (2004) Geometry optimization and transition state search in enzymes: different options in the microiterative method. *Int J Quantum Chem* 98:367–377
93. Ridder L, Rietjens IMCM, Vervoort J, Mulholland AJ (2002) Quantum mechanical/molecular mechanical free energy simulations of the glutathione S-transferase (M1-1) reaction with phenanthrene 9,10-oxide. *J Am Chem Soc* 124:9926–9936
94. Ranaghan KE, Ridder L, Szefczyk B, Sokalski WA, Hermann JC, Mulholland AJ (2004) Transition state stabilization and substrate strain in enzyme catalysis: ab initio QM/MM modelling of the chorismate mutase reaction. *Org Biomol Chem* 2:968–980
95. Lodola A, Mor M, Hermann JC, Tarzia G, Piomelli D, Mulholland AJ (2005) QM/MM modelling of oleamide hydrolysis in fatty acid amide hydrolase (FAAH) reveals a new mechanism of nucleophile activation. *Chem Commun* 35:4399–4401
96. Riccardi D, Schaefer P, Cui Q (2005)  $pK_a$  calculations in solution and proteins with QM/MM free energy perturbation simulations: a quantitative test of QM/MM protocols. *J Phys Chem B* 109:17715–17733
97. Raha K, Merz KM (2004) A quantum mechanics-based scoring function: study of zinc ion-mediated ligand binding. *J Am Chem Soc* 126:1020–1021
98. Bakowies D, Thiel W (1996) Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J Phys Chem* 100:10580–10594
99. Svensson M, Humberg S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) ONIOM: A multilayered Integrated MO + MM Method for geometry optimizations and single point energy predictions. A test for Diels – Alder reactions and  $\text{Pt}(\text{P}(\text{t}-\text{Bu})_3)_2 + \text{H}_2$  oxidative addition. *J Phys Chem* 100:19357–19363
100. Antes I, Thiel W (1999) Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. *J Phys Chem A* 103:9290–9295
101. Jensen L, van Duijnen PT (2005) The first hyperpolarizability of p-nitroaniline in 1,4-dioxane: a quantum mechanical/molecular mechanics study. *J Chem Phys* 123:Art. No. 074307
102. Cui Q, Karplus M (2002) Quantum mechanics/molecular mechanics studies of triosephosphate isomerase-catalyzed reactions: effect of geometry and tunneling on proton-transfer

- rate constants. *J Am Chem Soc* 124:3093–3124
103. Riccardi D, Li GH, Cui Q (2004) Importance of van der Waals interactions in QM/MM simulations. *J Phys Chem B* 108:6467–6478
  104. Poulsen TD, Garcia-Viloca M, Gao JL, Truhlar DG (2003) Free energy surface, reaction paths, and kinetic isotope effect of short-chain acyl-coa dehydrogenase. *J Phys Chem B* 107:9567–9578
  105. Brooks CL III, Karplus M, Pettitt BM (1988) Proteins, a theoretical perspective of dynamics, structure and thermodynamics. Wiley, New York
  106. Nam K, Gao JL, York DM (2005) An efficient linear-scaling ewald method for long-range electrostatic interactions in combined QM/MM calculations. *J Chem Theory Comput* 1:2–13
  107. Schaefer P, Riccardi D, Cui Q (2005) Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules. *J Chem Phys* 123:Art. No. 014905
  108. Im W, Berneche S, Roux B (2001) Generalized solvent boundary potentials for computer simulations. *J Chem Phys* 114:2924–2937
  109. Monard G, Loos M, Théry V, Baka K, Rivail J-L (1996) Hybrid classical quantum force field for modelling very large molecules. *Int J Quantum Chem* 58:153–159
  110. Assfeld X, Rivail J-L (1996) Quantum chemical computations on parts of large molecules: The *ab initio* local self consistent field method. *Chem Phys Lett* 263:100–106
  111. Gao J, Amara P, Alhambra C, Field MJ (1998) Method for the treatment of boundary atoms in combined QM/MM calculations. *J Phys Chem A* 102:4714–4721
  112. Ferre N, Assfeld X, Rivail J-L (2002) Specific force field parameters determination for the hybrid *ab initio* QM/MM LSCF method. *J Comput Chem* 23:610–624
  113. Antonczak S, Monard G, Ruiz-Lopez MF, Rivail J-L (1998) Modeling of Peptide Hydrolysis by Thermolysin. A Semiempirical and QM/MM Study. *J Am Chem Soc* 120:8825–8833
  114. Garcia-Viloca M, Gao JL (2004) Generalized hybrid orbital for the treatment of boundary atoms in combined quantum mechanical and molecular mechanical calculations using the semiempirical parameterized model 3 method. *Theor Chem Acc* 111:280–286
  115. Wesolowski TA, Warshel A (1993) Frozen density functional approach for *ab-initio* calculations of solvated molecules. *J Phys Chem* 97:8050–8053
  116. Pu JZ, Gao JL, Truhlar DG (2004) Generalized hybrid orbital (GHO) method for combining *ab initio* hartree–fock wave functions with molecular mechanics. *J Phys Chem A* 108:632–650
  117. Pu JZ, Gao JL, Truhlar DG (2004) Combining self-consistent-charge density-functional tight-binding (SCC-DFTB) with molecular mechanics by the generalized hybrid orbital (GHO) Method. *J Phys Chem A* 108:5454–5463
  118. Pu JZ, Gao JL, Truhlar DG (2005) Generalized hybrid-orbital method for combining density functional theory with molecular mechanicals. *Chemphyschem* 6:1853–1865
  119. Amara P, Field MJ (2003) Evaluation of an *ab-initio* quantum mechanical molecular mechanical hybrid-potential link-atom method. *Theor Chem Acc* 109:43–52
  120. HyperChem Users Manual (2002) HyperCube, Inc: Waterloo, Ontario, Canada
  121. Reuter N, Dejaegere A, Maigret B, Karplus M (2000) Frontier bonds in QM/MM methods: a comparison of different approaches. *J Phys Chem A* 104:1720–1735
  122. Zhang Y, Lee T-S, Yang W (1999) A pseudo-bond approach to combining quantum mechanical and molecular mechanical methods. *J Chem Phys* 110:46–54
  123. Monard G, Prat-Resina X, Gonzalez-Lafont A, Lluch JM (2003) Determination of enzymatic reaction pathways using QM/MM methods. *Int J Quant Chem* 93:229–244
  124. Das D, Eurenus KP, Billings EM, Sherwood P, Chatfield DC, Hodosek M, Brooks BR (2002) Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. *J Chem Phys* 117:10534–10547
  125. Konig PH, Hoffmann M, Frauenheim T, Cui Q (2005) A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method. *J Phys Chem B* 109:9082–9095
  126. Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On possible pitfalls in *ab initio* quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. *J Phys Chem B* 109:15645–15650
  127. Lodola A, Sirirak J, Fey N, Rivara S, Mor M, Mulholland AJ (2010) Structural fluctuations in enzyme-catalyzed reactions: determinants

- of reactivity in fatty acid amide hydrolase from multivariate statistical analysis of quantum mechanics/molecular mechanics paths. *J Chem Theor Comput* 6:2948–2960
128. Bowman AL, Ridder L, Rietjens IMCM, Vervoort J, Mulholland AJ (2007) Molecular determinants of xenobiotic metabolism: QM/MM simulation of the conversion of 1-chloro-2,4-dinitrobenzene catalyzed by M1-1 glutathione S-transferase. *Biochemistry* 46:6353–6363
129. Acevedo O, Jorgensen WL (2010) Advances in quantum and molecular mechanical (QM/MM) simulations for organic and enzymatic reactions. *Acc Chem Res* 43:142–151
130. Woods CJ, Manby FR, Mulholland AJ (2008) An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *J Chem Phys* 128:Art. No. 014109
131. Ridder L, Mulholland AJ (2003) Modeling biotransformation reactions by combined quantum mechanical/molecular mechanical approaches: from structure to activity. *Curr Top Med Chem* 3:1241–1256
132. van der Kamp MW, Zurek J, Manby FR, Harvey JN, Mulholland AJ (2010) Testing high-level QM/MM methods for modeling enzyme reactions: acetyl-CoA deprotonation in citrate synthase. *J Phys Chem B* 114:11303–11314
133. Ranaghan KE, Masgrau L, Scrutton NS, Sutcliffe MJ, Mulholland AJ (2007) Analysis of classical and quantum paths for deprotonation of methylamine by methylamine dehydrogenase. *Chemphyschem* 8:1816–1835

# Chapter 5

## QM and QM/MM Simulations of Proteins

Thomas Steinbrecher and Marcus Elstner

### Abstract

Molecular dynamics simulations of biomolecules have matured into powerful tools of structural biology. In addition to the commonly used empirical force field potentials, quantum mechanical descriptions are gaining popularity for structure optimization and dynamic simulations of peptides and proteins. In this chapter, we introduce methodological developments such as the QM/MM framework and linear-scaling QM that make efficient calculations on large biomolecules possible. We identify the most common scenarios in which quantum descriptions of peptides and proteins are employed, such as structural refinement, force field development, treatment of unusual residues, and predicting spectroscopic and excited state properties. The benefits and shortcomings of QM potentials, in comparison to classical force fields, are discussed, with special emphasis on the sampling problems of protein conformational space. Finally, recent examples of QM/MM calculations in light-sensitive membrane proteins illustrate typical applications of the reviewed methods.

**Key words:** Molecular dynamics, QM/MM, Protein conformations, Theoretical spectroscopy, Structure refinement

---

### 1. Introduction

The simulation of structure and dynamics of biological systems via molecular mechanics has matured into routine application over the last decades with empirical force fields becoming robust and reliable tools for the description of molecular structure and energies (1, 2). These molecular mechanics (MM) force fields (3, 4) model covalent chemistry by sets of harmonic or trigonometric potentials for the bond length, bond angle, and dihedral angle. In addition to these bonded terms, the force fields contain nonbonded contributions, typically modeled by the interaction of fixed atomic partial charges and van der Waals interactions, described by the 12-6 Lennard-Jones potential.

Several limitations of MM force fields are immediately apparent, for example, the use of springs in order to describe bonds or, more

general, the complete abandonment of a molecular description based on electronic density or wave function, instead of relying on empirically parametrized, arbitrary potentials. This leads to the high computational efficiency of these methods but, on the other hand, renders them incapable of treating properties like bond-breaking in chemical reactions, excited states dynamics, or the calculation of spectroscopic properties, without resorting to combined QM/MM models. The contribution of G. Groenhof in this volume introduces the QM/MM framework and discusses its use for excited states dynamics, and the contribution of A. Mulholland discusses the use of QM/MM methods for catalytic applications. In this contribution, we will elucidate some aspects of the computation of spectroscopic properties.

Additionally, we will focus on a second area of applications, where the MM limitations are less obvious and which is right at the heart of the application regime they have been designed for. It is the computation of energies, structures, and dynamics for complex biological molecules. In quantum chemistry there is the widespread, opinion that force fields represent a deficient way of describing molecules and that with increasing computer power, they should be gradually replaced by first approximate and then more advanced quantum methods like density functional theory (DFT). This view, however, overlooks two quite important issues: The first one is the importance of proper sampling, that is, the necessity to perform extended MD simulations in order to explore sufficiently large parts of the relevant phase space. While single-point energy calculations using QM methods for large molecules are possible today, true exhaustive conformational sampling is not, which makes their application questionable. The second issue is concerned with calculation accuracy. There is the common belief that QM (e.g., DFT) methods are in general superior to MM methods when it comes to computing molecular energies. This however is not necessarily true in many cases. First of all, HF and DFT methods lack a description of dispersion interactions (the “attractive” contribution of the vdW forces), which are of utmost importance in macromolecules. This alone makes these methods incapable of recognizing stable complex structures at all. Second, even with empirical vdW corrections applied, the accuracy is sensitively dependent on the chosen simulation parameters, for example, the type of DFT functional used and the basis set quality. To give just one example, for small, and medium-sized basis sets, HF and DFT show errors of 4–20 kJ/mol per hydrogen bond. Having in mind the marginal stability of proteins, that is, that the folded state is stabilized with respect to the unfolded one by maybe no more than a few kJ/mole, it is completely unclear how methods with such errors should be able to predict the thermodynamic properties rightly, assuming that these errors may add up for all the hydrogen bonds involved in the

stabilization of proteins. Therefore, the relative accuracy and applicability of MM vs. QM approaches should be much more carefully discussed in the case of biomolecules.

Nevertheless, QM methods offer one great advantage, which is their greater transferability compared to MM models. Since force fields are parametrized to reproduce certain condensed phase properties, they are expected to work very well for a wide range of problems, as far as the systems studied are reasonably close to those used in parametrization. The meaning of “close” is somewhat ambiguous here and involves not only chemical nature but also environment; for example, it is not clear that force fields can be applied to peptides and DNA fragments in gas phase or to protein folding since nonequilibrium configurations not sufficiently considered in the parametrization could turn out to be important in these cases. At least to some degree, the ability of force fields to fold proteins has been demonstrated impressively in a recent publication (5). Central to the transferability problem of MM parameters is the lack of polarization in standard MM force fields. This affects different properties to a different degree, and we will discuss this issue in more detail in the context of QM/MM applications below.

In a first part of this work, we will focus on applications of QM methods to compute total energies, energy differences, and structures of peptides and proteins. This has become quite a wide field in the last years, and we are not able to cover all important applications but rather discuss a selection of studies which we believe to be exemplary in order to understand the limits and merits of force field vs. QM descriptions. The second part of this chapter deals with another field of application, where QM instead of MM treatments are a necessity since they involve more than the calculation of energies and forces for electronic ground states. Some examples of this application regime are covered elsewhere in this volume, like excited states dynamics (see the contribution of G. Groenhof) or (bio-) catalytic questions (see the article of A. Mulholland). Here, we will discuss the computation of spectroscopic properties of biomolecules in a little more detail.

Therefore, after introducing QM methods, we will focus on energies and structures of peptides and proteins computed with QM methods in the following part, while we turn to the application of QM/MM approaches to compute spectroscopic properties in the second part of this chapter.

## **1.1. QM Methods and Their Combinations in Multi-Scale Frameworks**

### **1.1.1. QM Approaches**

Quantum chemistry methods can be roughly partitioned into three classes, into (i) the so-called post-HF methods, which include correlation effects on top of a HF calculation, (ii) standard SCF methods like HF and DFT, and (iii) semi-empirical (SE) methods, which can be derived from either HF or DFT by approximating and neglecting many of the costly integrals and which usually apply a

minimal basis set. Based on HF, a variety of very successful methods have been developed, starting several decades ago from CNDO, MNDO, AM1, and PM3 to the more refined methods like OMx, PM6, or PDDG-PM3, which perform surprisingly well for many properties like geometries and heats of formations of large sets of molecules. Alternatively, semiempirical methods can be derived from DFT, as for example, the self-consistent-charge density-functional tight-binding method (SCC-DFTB) (6, 7). Concerning the accessible system sizes, the three methods are separated from each other by approximately an order of magnitude. To give a rough estimate, post-HF methods can treat systems with up to 50–100 atoms, HF and DFT systems with up to 500–1,000 atoms, and SE methods work on systems with up to 20,000 atoms. Of course, these are numbers for up-to-date desktop computers; using high-performance computing, somewhat larger systems can be treated. DFT and SE methods usually exhibit a  $O(N^3)$  scaling; that means that the required CPU time increases with the third power of the number of atoms  $N$  in the system. HF shows an  $O(N^4)$  scaling, while post HF methods like MP2 or CCSD show a scaling behavior with the fifth power of  $N$  or even higher. Due to this scaling, for example, SE methods can treat only about 10 times as many atoms as DFT, even though they are about 1,000 times faster. Performing MD simulations for a certain number of atoms, however, the SE methods can take full advantage of their better computational efficiency, and MD simulations which are three orders of magnitude longer than those of DFT are possible. Therefore, SE methods remain an important workhorse in the toolbox of quantum chemistry when large system sizes and, in particular, long MD sampling times are needed, as discussed in more detail below.

To treat system sizes which are not accessible by any of the above quantum methods, two different approaches have been developed in the last decades: on the one hand, the combination of QM methods with computationally cheaper methods in so-called multiscale approaches and, on the other hand, algorithms which overcome the unfavorable  $O(N^x)$  scaling properties, the so-called linear scaling or  $O(N)$  methods.

### 1.1.2. Multiscale Methods for Biological Applications

The combination of QM with molecular mechanics (MM) methods in so-called QM/MM approaches is discussed in detail in the contribution of G. Groenhof to this volume. Building on these QM/MM approaches, it is possible to combine several different simulation techniques in so-called multiscale approaches; for a recent review, see (8, 9). There is a variety of possibilities to combine different levels of theory into a multi-method framework, which allows to tackle problems occurring in very large systems and on a multitude of timescales. One can, for example, combine two different QM methods into a QM/QM'/MM method or can combine QM/MM methods with continuum methods. In the last

years, we have combined SCC-DFTB with MM and continuum electrostatic methods; for recent reviews, see (10, 11). Some of the applications using such methods are discussed below.

### 1.1.3. Linear-Scaling Schemes for Large QM Systems

Even with the seemingly never-ending increase in computational resources and algorithmic efficiency, there exists strong demand for faster QM models to enable longer dynamics simulations or quicker single-point results. As mentioned above, QM calculations have, apart from being demanding tasks per se, quite an unfavorable scaling behavior with respect to the number of atoms (more precisely: basis functions) treated, leading to additional problems in the treatment of large systems. The two main bottlenecks are the assembly of the Fock matrix which involves calculating the two electron integrals and therefore scales as  $O(N^4)$  in HF and  $O(N^3)$  in DFT-GGA and the matrix diagonalization step which scales as  $O(N^3)$ . Semiempirical parametrized models greatly speed up the integral evaluation and therefore scale as  $O(N^3)$  overall. Even this is prohibitive when biomolecules, which are 10–1,000 times larger than typical small molecule systems, are considered. For this reason, a variety of linear-scaling algorithms, the so-called  $O(N)$ -methods, have been developed and validated over the penultimate decade which continue to play an important role in QM simulations of biomolecules (12, 13).

A variety of different approaches exist, from divide and conquer (14–16), the fragment orbital approach (17–19), molecular fractionation with conjugated caps (20–22), and the multilevel fragment-based approach (23) to Stewart’s localized molecular orbital methods MOZYME,(24) to name only some of them which have been applied to biomolecules. A more detailed review is found in Ref. (13, 25). All these have in common that they avoid diagonalizing the matrices for the full system by breaking it down into multiple parts (which can range from complete molecules down to individual bonds), solving the QM problem of the part only and constructing a description of the full system from the individual solutions. For example, the divide-and-conquer approach applied to proteins may select each amino acid residue as a fragment. The Roothaan–Hall equation is solved for each of the fragments, including a buffer region of additional atoms within a given distance, and the complete Fock matrix of the system is assembled from the fragment results according to special combination rules (26). Since for large systems all fragments are of comparable size and their number increases linearly with system size, this results in linear scaling of computation time. Additionally, as all fragment computations are independent of each other, this approach very easily lends itself to parallelization.

Without efficient  $O(N)$ -methods, many of the simulations described below would not have been feasible. They allow treatment of large systems at an accuracy very close to that of a

hypothetical full-system QM calculation.  $O(N)$ -methods are better suited for applications including semiempirical Hamiltonians since, as presented, they only provide efficient scaling for the matrix diagonalization step, not the second potential bottleneck of a QM calculation, integral evaluation. They are, however, rivaled in practical applications by combined QM/MM-calculations, which also allow simulation of large systems, often at lower total computational effort than linear-scaling QM (27, 28).

#### 1.1.4. The Sampling Problem

Tackling large systems is only one part of the problem, the other is the question of how to deal with the usually huge phase space which is sampled by large biological structures. In many cases, it is not at all sufficient to determine energies of one structure only for a biomolecule, neither may a single starting conformation be sufficient to compute reaction energies or minimum energy pathways.

The issue of phase-space sampling or equilibration of the conformational ensemble is one of the most central problems in all molecular simulations. The goal of a molecular dynamics simulation is usually to produce a set of structural snapshots over a given time from which average system properties can be determined according to the ergodicity theorem. If the simulation is *converged*, that is, the conformational ensemble generated reasonably similar to that of a hypothetical simulation of infinite length, then the time-averaged system properties for one simulated particle should be equal to the thermodynamic averages over large numbers of particles. This direct connection between microscopic simulation data and macroscopic experimental results makes molecular simulation an enormously valuable tool of the theoretical chemist.

However, in practice, two main obstacles have to be taken into account: First, every usable potential energy function for molecules must make severe approximations, and second, finite computational resources limit the duration of simulations. Therefore, the simulated potential energy landscape explored by the trajectory is not the true physical one, and the system is limited to a region of phase space close to its starting structure. Unfortunately, since more accurate theoretical descriptions of molecules tend to be computationally more expensive, trying to overcome one of these obstacles just serves to make the other one more challenging. Quantum mechanical dynamics simulations lie squarely on one side of this dilemma, using elaborate, and presumably fairly exact, calculations at the expense of speed, while more empirical methods like molecular-mechanics force fields sacrifice some accuracy to facilitate longer simulation times.

The sampling problem is especially significant for biomolecules like proteins, which are large systems with exceptionally intricate potential energy landscapes, where large scale, slow conformational

changes not only are common but also tend to be of functional significance. The experience with molecular dynamics simulations over the last decades has shown that sub-nanosecond length simulations of proteins can have significant sampling problems and that current methods are still orders of magnitude away from convergence for many important types of problems. Therefore, selecting and preequilibrating a good starting structure, utilizing enhanced sampling techniques, and critically questioning the convergence of results are crucially important for QM and QM/MM simulations of proteins.

Biological reactions occur on a variety of timescales, ranging from the ultrafast photochemical reactions occurring within several 100 fs up to the millisecond or even second timescale reactions found in catalysis or protein folding. Direct molecular dynamics simulations with DFT methods can be performed up to several tens of picoseconds, and while SE methods can readily approach the nanosecond regime, that still covers only a fraction of the relevant range. The SE methods are sufficient to describe reactions with effective free energy barriers of a few kcal/mole, such as the transitions between different conformations of the alanine dipeptide (see discussions below). The barriers for most chemical reactions, however, are much higher and cannot be overcome during a brute-force, unbiased MD simulation. One way of dealing with this problem is to employ reaction path techniques, that is, to calculate the path of minimal energy between a given reactant and product. This approach reduces the number of solvable problems to those where initial and final states are known and clearly defined. Furthermore, minimum energy path methods neglect the effect of thermal fluctuations, which has been shown to be a dramatic simplification (29, 30). In general, free energy techniques like FEP or umbrella sampling have to be applied (31), which even further increase the demand for extensive MD simulations, far beyond what is feasible using, for example, DFT level of theory. This is the reason that SE calculations enjoyed increasing popularity in recent years, although they show an overall lower accuracy compared to DFT, and thus, their application to new systems generally requires careful testing.

Taking all these into account, often neither linear scaling nor QM/MM approaches are the optimal solution to a given problem, simply because the necessary sampling times cannot be achieved by either. In these cases, using a “better” potential energy surface may be much less relevant than accounting for entropic contributions properly. These difficulties illustrate a fundamental fact of computational modeling, namely, that no single method (QM or classical) can be the ultimate solution for all problems, which most of the time require an appropriate combination of available methods.

## 2. QM and QM/MM Potentials to Improve Structure, Dynamics, and Property Predictions

### 2.1. Peptide Simulations

#### 2.1.1. Benchmark Calculations

In the quest to apply quantum mechanical calculations to the study of whole proteins, QM simulations of peptides are a natural first milestone (32). Peptides, short linear polymers built from several amino acid building blocks, are small compared to biomolecules but larger than typical organic compounds. They already contain most of the chemical peculiarities that define proteins, such as a limited elementary composition, strong intramolecular interactions, variable and flexible side chains, and the peptidic bond, as their defining structural feature. It is therefore common to see quantum simulation approaches intended for the study of proteins first applied to model peptides, where a limited system size allows for thorough tests. Much attention has been put to the question of how well the peptide-backbone potential energy surface is described in various levels of quantum calculations (33), especially in comparison to classical, molecular-mechanics force fields. Additionally, predicting spectroscopic properties of peptides from simulation has become an important application.

Typical structural studies aim at calculating the stability, relative energies, and barriers between various peptide conformations (34). As an example, an early study using a fully quantum mechanical molecular dynamics (QM MD) simulation describes the conformational dynamics of an alanine dipeptide analogue on the picosecond timescale using Born-Oppenheimer dynamics (35). A main focus was on the interconversion of the two minimum energy conformers  $C7_{eq}$  and  $C5$ , which contain an internal hydrogen bond in a seven- and five-membered ring, respectively and fall broadly into the  $\beta$ -sheet region of the protein Ramachandran plot. A density functional description gave the relative energies and transition barrier height for the two conformers in good agreement to high-level post-HF calculation (36, 37).

Multiple interconversions between the two peptide conformers occurred within less than 100 ps on a much faster timescale than classical force field molecular dynamics simulations had predicted (35). This observation already indicates that QM MD simulations, with their explicit treatment of polarization and correlation effects, are able to predict properties of peptides that empirical potential functions may misrepresent.

#### 2.1.2. Peptide Spectroscopy and Secondary Structure Formation

The ability of QM methods to compute molecular properties based solely on structural models makes their application in predicting and analyzing spectroscopic data highly appealing. Information from a wide field of experimental methods, ranging from vibrational (IR), VCD, Raman, and UV/VIS to NMR and EPR spectroscopy, can be combined with quantum mechanical calculations. In the following,

we will focus on vibrational spectra, a field that serves as a prime example application of QM methods to peptides, with recent advances into treating larger systems (38) and physiologically relevant compounds (39).

Accurately calculating the vibrational spectra of a chemical compound requires fairly elaborate computations that require a well-optimized molecular geometry, determination of the Hessian matrix eigenvalues and the atomic polar tensors, and a variety of other properties (40). It is therefore common to restrict QM calculations to small models and transfer the spectral parameters from these to larger molecules (41). Still, computed vibrational frequencies, even from high-level QM calculations, show systematic errors which can be remedied by applying empirical scaling factors (42), at a cost to transferability. In general, hybrid DFT methods like B3LYP have been less dependent on scaling factors than HF-based methods (43).

Extensive work has been done on using theoretical analysis of infrared and vibrational circular dichroism spectra of peptides with the aim of secondary structure determination. The Keiderling group has been especially prolific here, studying the effects of distortion on  $\beta$ -sheets, the spectral characteristics useful to distinguish between  $\alpha$ - and  $\beta_{10}$ -helices, and the effect of solvent interactions on the amide I band in hairpins (44–51). In these studies, theoretical models establish the crucial link between observable spectra of macroscopic systems and the atomistic details on how different conformational states of molecules contribute to them.

Calculations in vacuum can be sufficient for some of these studies if qualitative information on spectra interpretation or energetic parameters is all that is needed. For example, the Scheffler group used periodic DFT calculations of polypeptide chains in gas phase to investigate cooperativity effects and structure formation in some detail (52–54). However, early studies recognized the importance of including a description of environmental effects, especially solvation. Since analytical second derivatives for implicit solvent models are often not available in custom computational tools, the addition of several explicit key water molecules is a common approximation (43). Alternatively, an implicit solvent model may be used for geometry optimization only, to generate starting structures for later vacuum single-point calculations (48). The Dannenberg group has investigated various aspects of the formation of peptide secondary structure formation using DFT and hybrid methods. In particular, the effect of incremental binding enthalpies in gas phase and solution has been computed in detail (55–58).

### 2.1.3. Condensed Phase Simulations and Solvent Effects

An impediment to conducting QM studies of peptides is the choice of an appropriate treatment for the solvent environment peptides are normally found in. Approaches range from complete neglect of solvent, that is, vacuum simulations (59–61), simple continuum

models with maybe a few selected explicit water molecules (62–65), to full explicit solvation, either with a QM/MM multiscale model (66–69) or full quantum treatment of the solvated system (70, 71). An accurate representation of solvent is especially important when zwitterionic peptides are simulated, compared to capped neutral peptide analogues (72, 73).

While today's computational resources are sufficient for high-level QM MD simulations of peptides (74), the main focus of QM peptide simulations has been on the performance of semiempirical methods, especially in view of future applications to proteins, with QM MD simulations serving as validation benchmarks (34, 43, 75–78). A comparison of the semiempirical SCC-DFTB, AM1, and PM3 methods with DFT, HF, and MP2 calculations on glycine and alanine peptide analogues with 2–9 peptide bonds shows that discrepancies in the predicted structures and relative energies occur even among the high-level QM calculations, complicating their use as references. The AM1 and PM3 methods, which are parametrized for small molecules, had difficulties in predicting secondary structures and their stabilities (79), but SCC-DFTB performed acceptably (80). When the effects of explicit solvation were included in dynamics simulations of small dipeptides, SCC-DFTB-based QM/MM simulations outperformed various contemporary molecular-mechanics force fields in mappings of the peptidic  $\phi/\psi$ -space (81). The success of semiempirical methods is fortunate since the amount of sampling necessary in these studies, on the order of several nanoseconds, precludes the use of higher-level QM approaches. Several further studies suggest a benefit in treating peptides quantum mechanically in dynamics simulations (82–86).

However, the central “dogma” of computational chemistry, that a more expensive method should yield a more accurate result, does not necessarily hold for peptide simulations using QM methods and newer classical force fields (87, 88). The  $\phi/\psi$ -phase space maps, NMR coupling constants and solvent radial distribution functions computed with a variety of semiempirical methods were found to perform comparable to those of the newest *Amber* force field (89) in large-scale REMD simulations. Interestingly, both QM and MM results were in only moderate agreement to experiments, indicating that there is room for improvement in both theoretical approaches. However, since the aim here is the accurate prediction of relatively small differences in conformational stability, errors in the range of fractions of a kcal/mol can already significantly affect results. With this consideration, both classical as well as quantum mechanical models can be considered quite successful in the description of peptides.

### 2.1.4. Beyond Simple Model Structures: Side Chain Effects and Nonnatural Peptides

Up to now, we have only looked at small model peptides like glycine and alanine and focused on the peptide backbone structure. Including larger flexible side chains complicates the picture significantly (90–92), especially if aromatic residues are introduced, since dispersion interactions between side chain and backbone become important (93). Consequently, in a comparison of various ab initio methods, only those functionals like M06-2X, which are known to describe the mid-range correlation effects properly, or functionals like TPSS-D, which contain an empirical dispersion correction, could successfully describe structures and energetics for small peptides containing an aromatic side chain (94). Standard DFT-GGA functionals like PBE or BLYP, and even hybrid functionals like B3LYP, are known to fail for the description of systems which are stabilized by the attractive part of the VdW interactions, the so-called dispersion forces. In the last years, many DFT-GGA functionals have been augmented by an additional force field term (basically carrying a  $-1/r^6$  term), which improves the description of VdW systems significantly. These functionals are denoted with an additional “–D.”

For a Gly-Phe-Ala tripeptide, SCC-DFTB-D dynamics yielded a gas-phase potential energy surface corroborating experiments (95), and similar results were obtained for Phe-Gly-Phe, which includes an intramolecular  $\pi - \pi$ -stacking interaction (96).

So far, only examples of the 21 genetically coded  $\alpha$ -L-amino acids have been considered. However, one significant additional benefit of ab initio simulations is that novel compounds can be treated without the need to validate additional parameters (97), as would be the case in an MM treatment. For example, SCC-DFTB QM and QM/MM simulations readily gave minimum energy structures and relative conformational energies for  $\beta$ - and  $\alpha/\beta$ -mixed peptides in good agreement to ab initio calculations (98).

## 2.2. Protein Structure and Dynamics

When discussing quantum mechanical treatments of proteins, two cases must be distinguished, namely, QM/MM models in which only a (usually small) part of the biomolecule form the QM system, embedded in an MM environment, and full QM models describing the whole macromolecule. This following section will deal mainly with the latter case, while the chapters of A. Mulholland and G. Groenhof as well as a later section of this chapter deal with explicit QM/MM simulations. For selected applications and reviews, see Refs. (10, 99–104).

An attempt to study entire proteins using a quantum mechanical approach faces the same obstacles as the peptide simulations above, only to an even larger extent: Classical empirical force field, which can be finely tuned to describe protein dynamics, offers strong competition; important types of interactions, like dispersion forces, are not always described well by QM potentials, and exhaustive sampling of phase space is computationally more expensive and

requires orders of magnitude longer timescales than for small molecules (105). The combination of these factors makes it unsurprising that full QM molecular dynamics of proteins have not been widely adopted, outside of specialized applications for example, in X-ray structure refinement and protein structure prediction. We will consider two examples of such applications below, after a look at why pure QM dynamics simulations of proteins are rare.

### 2.2.1. QM Dynamics of Whole Proteins

While ab initio level geometry optimizations of macromolecules have been possible for years (106–113), we are not aware of any high-level QM MD calculations on proteins. Instead, semiempirical methods in combination with efficient scaling algorithms have to be used (16, 86, 114–121). Liu et al. have reported a 350 ps-length dynamics simulation of the small protein *crambin* using the approximate density functional method SCC-DFTB, with an additional empirical correction for long-range van der Waals forces (122). The protein was embedded in explicit solvent (which was described by the classical TIP3P model). The QM MD yielded stable simulations, with an average all-atom rmsd of 0.86 Å, comparable to that of classical MD simulations with three different force fields, with comparable protein flexibility in all trajectories. The inclusion of long-range dispersion interactions proved to be essential to maintain a stable protein structure; a test without the vdW correction resulted in large structural distortion after less than 100 ps simulation time. An analysis of structural parameters comparing QM MD, three classical MD simulations, and a high-resolution X-ray structure showed the QM results agreeing better to experiment than the classical ones; especially the backbone peptide unit planarity was significantly better reproduced. Additionally, the QM simulation found atomic partial charges of backbone atoms fluctuating by about 10% and some charge transfer between the protein termini, two effects that classical force fields neglect.

Available computational resources have increased substantially in the decade since this study and comparable simulations of small proteins on the multi-nanosecond timescale would be feasible today. Nevertheless, few such attempts have been made. The reason for this may be that semiempirical models like DFTB are not necessarily more accurate in describing proteins than molecular-mechanics force field while at the same time being computationally much more demanding. For example, the crucially important dispersion interaction requires elaborate computations on levels better than MP2 perturbation theory and is, as discussed above, only described using an empirical Lennard-Jones term in semiempirical models. Additionally, an important source of uncertainty in protein simulations comes from insufficient sampling of conformational space which can at best be reduced by generating longer simulation trajectories. So, while the increased transferability and explicit treatment of charge transfer and polarization in QM models of proteins

are clearly desirable, at present, instead of attempting QM dynamics simulations of proteins, a given amount of computer time is probably better spent in producing a classical MD simulation with an order of magnitude more sampling. Nevertheless, QM treatments of whole proteins have found important niches in some applications of protein structure studies.

### *2.2.2. Structure Optimization and QM Decoy Scoring*

X-ray crystallography, a crucial technique that has supplied modern biochemistry with the bulk of structural data on biomolecules, uses a refinement process to build a plausible molecular structure fitting for the measured electron densities. Since electron densities can be ambiguous with respect to the position of hydrogen atoms and the elemental nature of atoms, simplified molecular mechanics models are normally used to provide geometric restraints and amino acid residue topologies (123). The idea of instead using quantum mechanical potentials in this fitting process, either directly or as an additional step of refinement, has been advanced (124, 125), because a QM potential can readily treat unusual residues like cofactors, includes correct treatment of electrostatics, and allows residues to deviate from predefined template structures depending on their chemical environment. The approach has been shown to permit structural optimization of the very large (150,000 atoms) photosystem-I complex at high-level DFT calculations (126), but applications for molecules containing a few thousand atoms treated at a semiempirical level of theory are more typical.

In a large assessment of the PM6 method as a tool for protein modelling, several proteins, ranging in size from 250 to 15,000 atoms, were subjected to unconstrained structure optimization (127). Utilizing a localized molecular orbital technique, a full macromolecule could be energy minimized in about 1 CPU day. Special care had to be taken to use a minimizer suitable for systems with thousands of degrees of freedom and Cartesian coordinates proved to be better suited to describe the complex topologies of proteins than internal ones. Initially, the PM6 method was shown to reproduce single amino acid residue geometries and secondary structure elements faithfully, with slight geometric distortions in the case of  $\beta$ -turns. For proteins taken from the pdb database, the main findings were that PM6 resulted in stable structures close to their starting geometries and is suitable to model the surroundings of protein-bound metal ions. It was possible to correct several small errors in the X-ray data and compute heats of formation,  $pK_a$  values, and other physical properties.

So semiempirical methods are clearly suitable for protein structure optimization tasks. However, some problems might hinder their adoptions as routine modeling tools. First, even high-resolution X-ray structures have coordinate uncertainties in the range of 0.5 Å, and it is hard to validate if a given QM-optimized structure is truly closer to the “correct” geometry than the original crystal structure.

Furthermore, solvation effects are important in proteins, and side chain protonation states can vary; two effects QM (as well as other) modeling tools have difficulties to address. Finally, restricting the refinement process to geometry optimizations, due to computational costs, allows only small deviations from the starting geometry. This limits the optimization to within the original potential energy basin in phase space and makes larger conformational changes, which would be necessary to correct significant mistakes in X-ray structures, inaccessible.

Another intriguing application of quantum mechanics to polypeptides was advanced recently (128, 129), in which semiempirical energy calculations were used as a scoring function for protein structure prediction. The study used the PM3 Hamiltonian, augmented by an empirical dispersion correction and continuum solvent model, with a divide-and-conquer approach to rank several decoy sets against the corresponding native structures. Decoy sets are ensembles of slightly misfolded protein structures, and consistently ranking native structures better than decoys is an important characteristic of a good scoring function. If structures are geometry optimized before calculating the energy score, the QM scoring function could slightly outperform the classical force field in ranking the sets. The advantage of the QM approach became more pronounced when the different contributions to the total energy (solvation, dispersion, and PM3 energy term) were modified by optimized weighting factors. Scoring native X-ray structures was more reliable than NMR-derived ones. The good performance of a semiempirical Hamiltonian not even optimized for peptides with a relatively simple protocol suggests significant potential of this approach for future applications.

### *2.2.3. Improving Classical Force Fields*

It has been mentioned several times so far that classical empirical molecular mechanics force fields can produce results of comparable accuracy to quantum mechanical theory. The often good agreement of QM and MM models is unsurprising since one of the main fundaments of molecular mechanics is deriving parameters by fit to QM test calculations (89, 130–132). A promising attempt to devise better theoretical models is the incorporation of additional quantum mechanical information into the framework of classical molecular mechanics calculations.

One obvious starting point for such a scheme is the derivation of atomic partial charges for molecular dynamics (133–136). Since partial charges are not observable quantities, QM calculations are usually needed to obtain them, but in the process of assigning fixed charges to atoms, changes in the environment are neglected or averaged over, effectively prohibiting atomic polarization, a major flaw in traditional force fields. An alternative, besides the use of polarizable force fields, is to use a QM model on parts of the system where polarization plays a major role. One application of such a

QM/MM approach, D-RESP, involves refitting the atomic partial charges computed for the QM part of the system “on-the-fly” according to the RESP fit procedure (137) and was shown to be computationally efficient and well suited to accurately reproduce the electrostatic field around solvated molecules (138).

Another scheme to include QM data directly into the improvement of empirical potentials is called force matching, where force field parameters are optimized by matching them to forces computed from QM or QM/MM simulations (139–142). In this way, polarization, temperature, and pressure effects are explicitly influencing the parameters, which should capture the properties of condensed phase, solvated molecules. Maurer et al. suggest a two-step fitting process, in which first a QM/MM dynamics simulation is run for several picoseconds, with the compound to parametrize set as the QM region of the system (143). Then a set of atomic partial charges is derived as a best fit to the electrostatic field the QM atoms generate. A least-squares optimization is conducted to fit the bonded terms of the force field to the difference of nonbonded (Lennard-Jones and electrostatics) and true QM forces. This approach determines all the parameters required by a common force field potential except van der Waals parameters. Simulations of small test compounds showed that force-matching optimized parameter sets gave excellent results in describing their structure, dipole moments, and solvent distribution functions. However, one potential problem of the force matching scheme is that due to the necessarily short QM/MM trajectories, the conformational dynamics of flexible compounds are not fully explored. This leads to problems in determining parameters influenced by slowly sampled degrees of freedom, for example, dihedral angles. Nevertheless, the appeal of a method capable of generating accurate force field parameters for arbitrary compounds in a straightforward way is highly appealing.

### **2.3. Biomolecular Complexes**

Since ab initio QM theory, in contrast to empirical force field models, allows for the treatment of molecules containing any element in arbitrary chemical configuration, it has high predictive power for new and unusual compounds. This desirable property is of somewhat limited use in the field of biomolecular modeling since the majority of systems studied are built only from the 20 standard amino acids, five nucleotides, and a small variety of common cofactors and inorganic ions. These building blocks combine in a very well-behaved manner, forming stable covalent or hydrogen bonds and van der Waals complexes, with little of the subtleties of chemical bonding that, for example, transition metal complexes or solid state crystals exhibit. (Note that this characterization obviously does not hold for the wide field of bioinorganic chemistry, which will not be treated here.) Nevertheless, in the cases where unusual components or complexes do occur, QM models are much better suited to describe them than parametrized ones. Two cases will be

distinguished here: first, naturally occurring cofactors and binding situations that benefit from a quantum chemist's perspective and second, describing small drug-like molecules binding to protein target receptors.

### 2.3.1. Cofactors and Unusual Ligands

Simulations of biomolecular complexes composed of nonstandard building blocks form a very diverse area of study and QM simulations in this field range from electrostatic effects in ion channels (144, 145) to binding of unusual ligands (146–149). We will discuss work on two central model systems of biophysics only, namely, the bacterial *photosynthetic reaction center* (PRC) and *bacteriorhodopsin*. This necessarily incomplete and subjective selection was made since both systems involve functional light-activated excited states and have been extensively studied spectroscopically, two topics that will be revisited in Sect. 5.3 below.

*Bacteriorhodopsin* (BR) is a membrane-bound proton pump with a retinal cofactor bound as a Schiff base to a lysine side chain. It serves as a key subject of interest for studies of membrane protein structure and photochemistry in biomolecules and bioenergetics. BR, its homologues, and the structurally related eucaryotic *rhodopsin* exhibit a light-driven conformational change, and the corresponding photocycle has been studied extensively by means of spectroscopy. The combination of photoexcitation, large-scale conformational change, proton transfer events, and complex protein environment makes the detailed function of the cofactor an extraordinarily challenging topic to model.

Babitzki et al. have recently provided a detailed examination of the suitability of different levels of theory to describe the retinal binding pocket in BR (150). The difficulty here is to maintain a delicate network of hydrogen bonds connecting several charged residues, structural water molecules, and the cofactor which stabilizes the binding site and influences the chromophore spectrum. A purely classical MD treatment of the system resulted in dislocation of important water molecules and formation of a salt bridge artifact; treating retinal alone quantum mechanically in a QM/MM approach did not rectify this problem. To give a measure of the subtle effects influencing such structural models, it was shown that the distorted structure from plain MD simulations lay only a few kcal/mol below the correct one from X-ray crystallography. The authors concluded that the lack of electronic polarization in the MM part of the binding site was mainly responsible and suggested the use of QM-derived polarization adapted partial charges as a remedy since a full QM treatment of the binding site was not feasible due to its size. Their improved model, which used quantum mechanical information but no actual QM MD calculations, was successfully able to yield a conformational ensemble of BR structures in good agreement with experimental data that could be used for subsequent predictions of vibrational spectra.

As a second example, we will consider QM/MM studies of the bacterial *PRC*. This protein complex, involved in photosynthesis, contains a large amount of cofactors, several chlorophylls, carotenoids, quinones, and a non-heme iron ion, which form its light-driven long-range electron transfer chain. As one of the first membrane proteins for which high quality X-ray structural data was available, the PRC has attracted immense interest of theoreticians for a long time. Reimers and Hush have conducted an extensive series of theoretical and modelling studies of the PRC, touching topics as diverse as structural optimization, mutation effects, and excitation spectra (151–157). In more recent times, DFT studies on spin density distributions in the two chlorophyll molecules, forming the so-called *special pair* where initial charge separation occurs, could shed light on the origin of an asymmetry in the electron transfer pathway (158). Likewise, at the end point of the charge transfer cascade, QM/MM simulations could elucidate the influence of the environment on the bound quinone IR and EPR spectra (159, 160).

The remarkable variety among the research goals and methods in the studies cited here should serve as an example as to how different the potential subjects of quantum mechanical calculations on proteins can be.

### 2.3.2. Quantum Mechanical Drug Design

The computational drug design problem breaks down into two basic tasks (161, 162): the *docking* and the *scoring* problems. The first of these concerns itself with predicting possible binding poses of small molecules in binding sites, and it can be solved by using a variety of available fast docking algorithms (163–170). The second problem is tasked with computing the free enthalpy of binding for such poses, and it is here that QM models may offer advantages over the typically used simple, empirical scoring functions. Together with the usual drawbacks of high computational cost and limited potential for conformational sampling, a QM scoring function promises a more exact treatment of internal energies and electrostatic interactions, describing phenomena like polarization and protein-ligand charge transfer effects which are neglected in cruder potentials. An additional bonus is that QM models do not require system-specific parameterization and atom-typing.

Several studies have tested the suitability of semiempirical Hamiltonians for the prediction of binding strength (171–176), while higher-level ab initio treatments are rarer (177–179). In general, both the AM1 and PM3 parameterizations appear capable of predictive binding free energy computations and distinguishing between true and decoy binding poses. However, the performance of QM scoring rarely exceeds that of commonly used docking tools or MM-based free energy methods. On average, calculated binding free enthalpies show mean errors of ca. 2 kcal/mol, which would be classified as good but not overwhelming accuracy in the drug design field. Additionally, QM scoring methods appear to be

more dependent on good, high-resolution starting structures and correct consideration of moderating water molecules. A promising future direction is to use weighting approaches that contain QM scoring data as one ingredient in fitted potential functions (173). Overall, the widespread adoption of quantum mechanical scoring tools in computational drug design appears to depend on further improvements of the proposed protocol.

An alternative approach has been advanced recently, in which the expensive QM calculations are restricted only to the preparation of the receptor binding site, and the actual placement and scoring of compounds is performed using classical docking tools. For example, in a study of ligands binding to human nuclear receptors, Fischer et al. used a linear-scaling QM calculation to obtain accurate electrostatic potentials of the receptor site, which was used as one component in a classical scoring function. The QM scoring function was shown to yield binding free energies in better correlation to experimental data when compared to a classical force field-based scoring function (180).

---

### 3. Protein Properties: Spectroscopy and Excited States

The discussion so far concentrated on the calculation of energetic and structural properties of biomolecules while now we turn to the field of theoretical spectroscopy. On the experimental side, spectroscopic techniques address a variety of problems. As mentioned above, IR and NMR can elucidate the structure and dynamics of polypeptides and proteins, and IR spectroscopy can determine the protonation states of titrable amino acid side chains, to give just a few examples. Spectroscopic techniques are in particular helpful for the investigation of ligands and protein cofactors, which we will discuss in a little more detail below. In the following, we want to focus on the calculation of spectroscopic properties of protein active sites, in particular of protein cofactors. Since this is quite a wide field, we will limit the discussion to proteins studied by our group in the last years and related work.

More general views concerning excited states dynamics, calculation of excitation energies, and exciton spectra can be found in recent reviews (181–183). All well-known quantum chemistry packages (among them GAMESS, NWChem, Jaguar, ORCA, Turbomole, Gaussian, etc.) support not only the calculation of energies and forces to perform geometry optimizations but also the computation of a variety of spectroscopical properties like IR, Raman, NMR, ESR, and UV/vis spectra. Additionally, when combined into QM/MM packages, these techniques are readily available also for more complex systems. In principle, only a QM/MM geometry optimization is required in order to compute the desired

spectroscopic data for the system of interest. However, as many studies in the recent years have shown, several critical issues should be kept in mind.

### 3.1. Issues for QM/MM

#### *Modeling of Spectroscopic Properties*

##### *3.1.1. Determining the Ground State:*

*One Global Minimum vs. Conformational Sampling*

Most theoretical approaches start from the crystal structure and perform a (QM/MM) geometry optimization which leads to the nearest local minimum, usually assumed to either correspond to the global minimum or be at least close to it. It is then further assumed that any experimental spectroscopic data can be interpreted on the basis of properties computed using this model structure, for example, in the case of optical spectroscopy, it is assumed that the experimental absorption maximum can be estimated from the vertical excitation energy for this particular geometry. This may accidentally indeed be the case but in general may sensitively depend on the details of the molecular structure. These assumptions seem to be appropriate, for example, for the bacteriorhodopsin (bR) ground state, where the retinal chromophore is tightly bound by hydrogen bonding to a rigid H-bonded network consisting of retinal, Asp85/Asp212, three water molecules, and Arg82; see Fig. 1.

Performing QM/MM MD simulations (in the order of 1 ns) in order to sample the excitation spectrum reveals a Gaussian distribution of excitation energies, with the maximum coinciding with the vertical excitation energy for a QM/MM-minimized structure (184). Therefore, for this system, it seems at first sight that the complex fluctuates around a geometry which is nicely approximated by the optimized geometry. However, this simple picture may not be

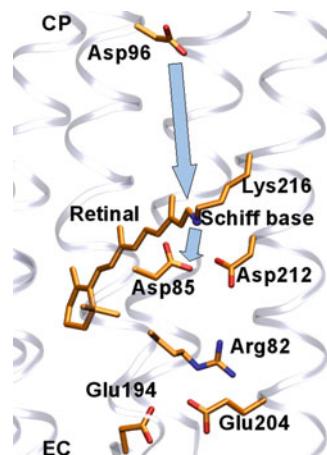


Fig. 1. The binding pocket of bR showing the Schiff base retinal and important side chains along the proton transfer pathway. After illumination, the Schiff base proton is transferred to Asp85, in a second step the proton between Glu194 and Glu204 is released to the bulk. In later steps of the photocycle, the Schiff base is reprotonated from Asp96 and the proton from Asp85 is transferred to Glu194/204.

completely true. More extended MD simulations using a force field showed that three local minima of the retinal chromophore are sampled during dynamics. Interestingly, the three structures seem to lead to indistinguishable spectroscopic properties (150, 185). This example shows that one has to be quite careful since the optimized local minimum starting from the crystal structure may not be a good approximation in general. If the structure of the chromophore and the binding pocket allow for more conformational flexibility, a single evaluation of the spectra for optimized geometries is no longer sufficient. In this case, dynamical effects can be incorporated on different levels. The most simple approach is to compute the spectroscopic data for various conformations along the MD trajectory, thereby including the effects of fluctuations of chromophore and protein. However, several issues occur here: First of all, the detailed molecular structure of the chromophore may be of importance, that is, it may not be an accurate approach to apply the force field for the active site, in particular for excitation energies and NMR, as described in more detail below. In case of optical properties, the spectra can be sampled along QM/MM trajectories, that is, vertical excitation energies are computed for selected snapshots (184). For vibrational spectra, the so-called instantaneous normal mode analysis (INMA) can be applied, where a normal mode analysis is applied for the selected snapshots to compute for example, IR (160) or Raman (186–188) intensities. Note, that in this case, imaginary eigenvalues of the Hessian occur since it is not computed for an optimized structure. Instead of using INMA, the time correlation functions of dipole moments and polarizabilities can be used instead to compute IR and Raman spectra. A detailed analysis of these methods has been put forward by Tavan and coworkers (160, 189, 190), but for sufficient sampling, full DFT methods may still be too slow and well-calibrated semiempirical methods may be useful (191–193).

The general problem of sufficient sampling (see also Sect. 1.1.4) poses severe limitations to practical computations. DFT/MM definitely cannot sample timescales above a few picoseconds. Semiempirical methods like AM1, PM3, or SCC-DFTB can reach the nanosecond regime, which is sufficient for some applications. However, if local minima are separated by barriers of just a few kcal/mole, longer timescales have to be accessed using MM in combination with efficient sampling techniques (150). Unfortunately, in the case of bR, it has been shown that standard force field parameters are not able to reproduce the details of the hydrogen-bonded structure around the chromophore, which is essential to reproduce the spectroscopic properties. Refining the atomic charges in order to incorporate polarization effects, however, could solve this problem (150), as discussed above. This shows again how important it can be to include protein polarizability to some degree. Similar problems have been shown for other proteins

as well, where the standard fixed charge model seems to have problems with the proper description of intra-protein hydrogen bonds (194, 195).

Another important issue in using proper starting conformations is the determination of rotamer states, which may not be unambiguously resolved by crystal structure analysis alone. Since side chain rotamers can be separated by large energy barriers and have a huge impact on structure and spectroscopic properties, they should be investigated quite carefully, as we have shown recently for the orientation of Ser198 in rhodopsin (196).

### 3.1.2. QM Accuracy

Of central importance is the accuracy of the QM method itself. For excited states, this concerns first the description of ground state geometry and second the calculation of accurate excitation energies. For conjugated molecules (as biological chromophores), the excitation energy depends sensitively on the bond alternation (difference between the bond length of successive single and double bonds). Unfortunately, different computational models can differ substantially in this value, leading to large variations in the computed excitation energies, as discussed, for example, in detail for retinal in bR (197). Usually, hybrid DFT methods like B3LYP or post-HF methods like MP2 compute this value reliably, while, for example, DFT-GGA or plain HF under- or overestimates the bond alternation, leading to quite large red or blue shifts in the computed excitation energies.

For the calculation of vertical excitation energies of chromophores containing several tens of atoms, in principle, a multitude of QM approaches can be applied like HF/CIS, TD-DFT, MRCI, and CASSCF/CASPT2 approaches (198, 199). TD-DFT unfortunately fails for biological chromophores. On the one hand, the general accuracy of TD-DFT is in the order of 0.5 eV (198), but for certain low-lying electronic states of  $\pi$  character, the accuracy can be much better which, at first sight, looks quite promising. On the other hand, TD-DFT is completely unsuitable for charge transfer excitations, which are quite characteristic for biologically relevant chromophores (199). In many cases, the use of hybrid functionals can ameliorate this problem, but unfortunately, this still does not turn TD-DFT into a useful method for biochromophores. It still leads to a wrong description of excited states potential energy surfaces (200), but even more importantly, the response to external electrostatic fields is not properly treated by TD-DFT. Since the protein environment determines the vertical absorption energy via electrostatic interactions, this property is crucial for an accurate determination of excitation energies for different chromophores in their natural surrounding. This has been investigated in detail in Ref. (197). Therefore, TD-DFT cannot be considered to be predictive for excited states calculations in proteins, and agreement with experiment is only found accidentally. Since a similar problem is also

apparent for CIS-approaches, only correlated multireference methods like CASPT2 or MRCI can be regarded sufficiently reliable to compute vertical excitation energies.

The situation is different for vibrational spectroscopy, where DFT methods can be applied due to their reasonable accuracy for vibrational properties. However, deviations in the order of 10–100 cm<sup>-1</sup> can occur, which makes a direct comparison to experiment often difficult. In the NMA and INMA approach, the accuracy can be improved significantly by applying frequency scaling techniques (186, 187). For the dynamical approaches, however, where correlation functions are evaluated along MD trajectories, scaling is not possible. In this case, special functionals have to be selected which have been shown to lead to reliable results for special purposes: For example, Tavan and coworkers identified the BP86 functional to be sufficiently accurate to simulate the spectra of biological chromophores in the mid-IR regime (150, 160).

### *3.1.3. Limitation of Standard QM/MM Approaches: Charge Transfer, Polarization, Dispersion*

In the last years, it has become apparent that the standard MM framework is not sufficient to model the environment of the active site when accurate spectroscopic data is required. A main limitation is the static atomic charge model, which represents the electrostatic environment of the active site. Computing atomic charges using QM models for the protein environment, Tavan and coworkers have shown that the calculated IR spectra can be significantly improved. For vertical excitation energies, the results can vary in the order of 0.05–0.1 eV, depending on the charge model used (e.g., charges from CHARMM or AMBER force fields). Buss and coworkers therefore used QM-derived charges for the binding pocket of the retinal chromophore in rhodopsin while neglecting the remainder of the protein. It has been shown in detail that ground-state polarized charges have a sizable influence of excitation energies (201–203). However, not only ground state polarization matters; it is also the response of the environment to the dipole moment of retinal in excited states, which leads to an additional red shift. In total, polarization in bR leads to a red shift of about 0.14 eV (201, 202), which is quite significant and questions the use of standard QM/MM approaches for this purpose. Similar results have been reported also by other studies (203).

While polarization is one factor missing in standard QM/MM approaches, the change of the VdW interaction due to excitation is another one. Since the excited state wave function is usually more diffuse than the ground state, it is clear that VdW interactions should change accordingly. These effects cannot be treated in a simple fashion; they require the incorporation of the relevant (aromatic) residues in the QM treatment of the excited states, which becomes costly at the CASPT2 or MRCI level of theory. In bR, there are three aromatic residues close to the chromophore, which contribute to an estimated additional dispersion red shift of about 0.07 eV (204). Therefore, in

bR polarization and dispersion effects lead to an excitation red shift of about 0.2 eV altogether, which is completely neglected in standard QM/MM approaches. This should be kept in mind when estimating the accuracy of excitation energies using standard QM/MM methods.

### **3.2. Specific Applications**

#### *3.2.1. Determining the Factors Responsible for Color Tuning in bR vs. SRII*

Bacteriorhodopsin (see discussion above) and sensory rhodopsin II are both trans-membrane proteins, which contain the same retinal cofactor and have a very similar 3-D structure. However, they differ in the maximum absorption wavelength by about 70 nm (0.3 eV). In the immediate vicinity of the chromophore, they differ by ten amino acids, which have been examined with respect to their impact on excitation energies by using mutation studies. Additionally, there is a slight difference in the hydrogen-bonded network in which the retinal chromophore is embedded. Experimental studies have focused on these two factors in order to explain the color tuning between the two proteins. In our calculations (184), we used a standard QM/MM approach (no polarization, no dispersion) since we were only interested in the relative shift and both binding pockets differ only in polar amino acids; that is, polarization and dispersion effects were expected to be small with respect to the shift. Nevertheless, to obtain an accurate estimate, a high-level MRCI method had to be applied for the QM region. With these calculations, we could investigate the effect of mutations in the binding pocket and hydrogen-bonding effects individually. Single amino acid mutations contribute typically less than 0.1 eV to the shift; in total, the difference due to the binding pocket was estimated to be in the order of 40 %. Another 40 % results from the different hydrogen-bonding interactions, and the remaining contribution is due to the remainder of the protein. The effect of all other amino acids has been estimated by using a perturbation analysis, that is, “switching off” each amino acid one by one and computing the effect on the excitation energy.

#### *3.2.2. Glu181 Protonation State in Rhodopsin*

The protonation state of the glutamate 181 in rhodopsin was debated over decades, since different experimental techniques as well as different calculations led to contradictory results. The protonation state of this amino acid has an important role for the function of the protein. UV/Vis and resonance Raman experiments concluded that Glu181 is protonated since its substitution by a charge neutral amino acid did not lead to significant spectral shifts, in contrary to the expectation. FTIR experiments, on the other hand, did not find any change of protonation state of this residue for later intermediates; for which, however, Glu181 is known to be unprotonated. Therefore, these experiments lead to the conclusion that Glu181 is unprotonated. Several computational studies supported one or the other conclusion. The problem is that any

computational model has an error of about 0.1–0.5 eV (or even more), depending on the particular choice of QM and MM methods. Therefore, in some calculations, the protonated, in other the unprotonated Glu181, fits better to the experimental data. In order to draw any conclusion, one has to carefully calibrate the methodology, that is, compute excitation energies for a variety of rhodopsins in order to get a good estimate of the error. It further turned out that the inclusion of MM polarization is crucial since different protein environments exhibit different polarizabilities due to the different aromatic residues present. The calculations showed that the difference in the vertical excitation energy of Rh wild type with unprotonated Glu181 to the neutral Glu181Q mutant is very small, that is, this mutation does not have a significant effect on the absorption and resonance Raman spectra (196). Therefore, these experiments cannot be used to determine the protonation state, neither is this possible by the calculations. Therefore, the only decisive experiments are the FTIR ones, which concluded that Glu181 is unprotonated (196).

### 3.2.3. The bR O State

For the O state of the bR photocycle, there is no wild-type crystal structure available at this time. Therefore, we tried to model the O state structure using templates from other intermediates and mutant structures (205). It was again interesting to see that MM polarization played a crucial role when comparing vertical excitation energies from different bR intermediates. These structures differ by the protonation state of an aspartate close to the retinal chromophore, which seems to induce a significant amount of polarization in the binding pocket. A second prerequisite was the placement of water molecules in the binding pocket, which seems to have an important role in stabilizing it. This again shows how much the modelling details matter for the simulation of spectroscopic data, that is, the computed spectra are highly sensitive to fine details of the structural models and the representation of the charges in the binding pocket.

## 4. Conclusions

In this contribution, we have discussed the use of QM methods for the simulation of biological structures. Although QM is usually thought to provide a higher-level description of biological systems compared to MM force fields, recent simulations have shown that this is not always the case.

VdW interactions and hydrogen bonds are central to the stability of biological structures. Small basis set calculations lead to BSSE errors, which may be huge compared to the intrinsic accuracy of force field methods. Further, vdw forces pose severe difficulties to

QM methods and are only accurately covered at that high QM levels, which are clearly not applicable to more than 100 atoms at best. Of course, there are practical work-arounds like the addition of empirical force field terms to HF or DFT, which may represent a practical solution for many applications. Nevertheless, the sampling problem, which is already quite substantial using MM methods, becomes prohibitive for QM methods in many cases. Therefore, there are clear regimes of application for QM methods. These are mostly in a QM/MM framework, plus a few cases where a full QM description of larger structures may lead to meaningful results, but a routine application of QM to the simulation of large-scale structures and dynamics is not in sight in the next years.

MM methods, on the other hand, provide a quite reliable prediction of structures and dynamics of large biological systems. However, they are problematic in many cases, often due to the use of a fixed charge model, thereby neglecting polarization. This may affect their use in QM/MM simulations, where MM polarizability can be important, but also in pure MM MD simulations, where internal hydrogen bonds may not always be represented at the desired accuracy. MM parameters are constantly being refined, but one may argue that a substantial improvement will require a modification of the basic force field form. More accurate force fields will probably use polarizable models, either introducing atomic polarizabilities, a Drude oscillator model, or a charge equilibration scheme. But it may be that the functional forms will have to go even further and make use of more sophisticated quantum mechanical information, as used, for example, in the Xpol model or in the SIBFA force field (206, 207).

These more involved models may lead to a drastically improved description of biological structure, however, at cost of significantly higher computational costs. On the other hand, linear-scaling methods in combination with SE methods will make a pure QM description cheaper. Therefore, in the near future, there will be a variety of methods in the regime between standard MM and semi-empirical QM, which can fill the gap of three orders of magnitude in computing costs and give an improved description of large and complex systems.

## References

- Karplus M, McCammon J (2002) Molecular dynamics simulations of biomolecules. *Nat Struct Biol* 9:646–652
- van Gunsteren WF et al. (2006) Biomolecular modeling: goals, problems, perspectives. *Angew Chemie Int Ed* 45:4064–4092
- MacKerell A et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
- van Gunsteren WF, Billeter SR, Eising AA, Hunenberger PH, Kruger P, Mark AE, Scott WRP, Tironi IG (1996) Biomolecular simulation: the GROMOS manual and user guide. vdf Hochschulverlag, ETH Zurich
- Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W (2010) Atomic-level characterization of the

- structural dynamics of proteins. *Science* 330:341–346
6. Porezag D, Frauenheim T, Kohler T, Seifert G, Kaschner R (1995) Construction of tight-binding-like potentials on the basis of density-functional theory—application to carbon. *Phys Rev B* 51:12947–12957
  7. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B* 58:7260–7268
  8. Grotendorst J, Attig N, Bluegel S, Marx D (eds) (2009) Multiscale simulation methods in molecular sciences. Publication series of the John von Neumann Institute for computing
  9. Lee T, York D (eds) (2008) Models for biocatalysis: modern techniques and applications. Springer, Berlin
  10. Riccardi D, Schaefer P, Yang Y, Yu H, Ghosh N, Prat-Resina X, Konig P, Li G, Xu D, Guo H, Elstner M, Cui Q (2006) Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological processes. *J Phys Chem B* 110:6458–6469
  11. Elstner M, Cui Q (2009) Multi-scale methods for the description of chemical events in biological systems. Publication Series of the John von Neumann Institute for Computing
  12. Gogonea V, Suarez D, van der Vaart A, Merz K (2001) New developments in applying quantum mechanics to proteins. *Curr Op Struct Biol* 11:217–223
  13. Goedecker S (1999) Linear scaling electronic structure methods. *Rev Mod Phys* 71:1085–1123
  14. Yang W (1991) Direct calculation of electron-density in density-functional theory. *Phys Rev Lett* 66:1438–1441
  15. Yang W, Lee T (1995) A density-matrix divide-and-conquer approach for electronic-structure calculations of large molecules. *J Chem Phys* 103:5674–5678
  16. Dixon S, Merz K (1996) Semiempirical molecular orbital calculations with linear system size scaling. *J Chem Phys* 104:6643–6649
  17. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M (1999) Fragment molecular orbital method: an approximate computational method for large molecules. *Chem Phys Lett* 313:701–706
  18. Nakano T, Kaminuma T, Sato T, Fukuzawa K, Akiyama Y, Uebayasi M, Kitaura K (2002) Fragment molecular orbital method: use of approximate electrostatic potential. *Chem Phys Lett* 351:475–480
  19. Fedorov D, Kitaura K (2006) The three-body fragment molecular orbital method for accurate calculations of large systems. *Chem Phys Lett* 433:182–187
  20. Gao A, Zhang D, Zhang J, Zhang Y (2004) An efficient linear scaling method for ab initio calculation of electron density of proteins. *Chem Phys Lett* 394:293–297
  21. He X, Zhang J (2005) A new method for direct calculation of total energy of protein. *J Chem Phys* 122:031103
  22. Zhang D, Zhang J (2003) Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *J Chem Phys* 119:3599–3605
  23. Rezac J, Salahub D (2010) Multilevel fragment-based approach (MFBA): a novel hybrid computational method for the study of large molecules. *J Comp Theor Comput* 6:91–99
  24. Stewart J (1996) Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations. *Int J Quant Chem* 58:133–146
  25. Ochsenfeld C, Kussmann J, Lambrech D (2007) Linear-scaling methods in quantum chemistry. Wiley-VCH, New York
  26. He X, Merz K (2010) Divide and conquer Hartree–Fock calculations on proteins. *J Comp Theor Comput* 6:405–411
  27. Titmuss S, Cummins P, Bliznyuk A, Rendell A, Gready J (2000) Comparison of linear-scaling semiempirical methods and combined quantum mechanical/molecular mechanical methods applied to enzyme reactions. *Chem Phys Lett* 320:169–176
  28. Zhang R, Lev B, Cuervo J, Noskov S, Salahub D (2010) A guide to QM/MM methodology and applications. *Adv Quant Chem* 59:353–400
  29. Zhang Y, Kua J, McCammon J (2003) Influence of structural fluctuation on enzyme reaction energy barriers in combined quantum mechanical/molecular mechanical studies. *J Phys Chem B* 107:4459–4463
  30. Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. *J Phys Chem B* 109:15645–15650
  31. Kollman P (1993) Free-energy calculations—applications to chemical and biochemical phenomena. *Chem Rev* 93:2395–2417

32. March N, Matthai C (2010) The application of quantum chemistry and condensed matter theory in studying amino-acids, protein folding and anticancer drug technology. *Theo Chem Acc* 125:193–201
33. van Mourik T (2004) First-principles quantum chemistry in the life sciences. *Phil Trans Royal Soc A* 362:2653–2670
34. Beachy M, Chasman D, Murphy R, Halgren T, Friesner R (1997) Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. *J Am Chem Soc* 119:5908–5920
35. Wei D, Guo H, Salahub D (2001) Conformational dynamics of an alanine dipeptide analog: an ab initio molecular dynamics study. *Phys Rev E* 64:011907
36. Frey R, Coffin J, Newton S, Ramek M, Cheng V, Momany F, Schafer L (1992) Importance of correlation-gradient geometry optimization for molecular conformational analyses. *J Am Chem Soc* 114:5369–5377
37. Ramek M, Cheng V, Frey R, Newton S, Schaffer L (1991) The case of glycine continued—some contradictory SCF results. *J Mol Struct Theochem* 81:1–10
38. Nagai Y, Nakanishi T, Okamoto H, Takeda K, Furukawa Y, Usui K, Mihara H (2005) IR study on stacking manner of peptide nanorings in peptide nanotubes. *Jpn J Appl Phys* 44:7654–7661
39. Abdali S, Niehaus T, Jalkanen K, Cao X, Nafie L, Frauenheim T, Suhai S, Bohr H (2003) Vibrational absorption spectra, DFT and SCC-DFTB conformational study and analysis of [Leu]enkephalin. *Phys Chem Chem Phys* 5:1295–1300
40. Jalkanen K, Jurgensen V, Claussen A, Rahim A, Jensen G, Wade R, Nardi F, Jung C, Degtyarenko I, Nieminen R, Herrmann F, Knapp-Mohammady M, Niehaus T, Frimand K, Suhai S (2006) Use of vibrational spectroscopy to study protein and DNA structure, hydration, and binding of biomolecules: a combined theoretical and experimental approach. *Int J Quant Chem* 106:1160–1198
41. Bour P, Sopkova J, Bednarova L, Malon P, Keiderling T (1997) Transfer of molecular property tensors in Cartesian coordinates: a new algorithm for simulation of vibrational spectra. *J Comp Chem* 18:646–659
42. Scott A, Radom L (1996) Harmonic vibrational frequencies: an evaluation of Hartree–Fock, Moller–Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J Phys Chem* 100:16502–16513
43. Jalkanen K, Suhai S (1996) N-acetyl-L-alanine N'-methylamide: a density functional analysis of the vibrational absorption and vibrational circular dichroism spectra. *Chem Phys* 208:81–116
44. Kubelka J, Keiderling T (2001) Differentiation of beta-sheet-forming structures: ab initio-based simulations of IR absorption and vibrational CD for model peptide and protein beta-sheets. *J Am Chem Soc* 123:12048–12058
45. Kubelka J, Silva R, Keiderling T (2002) Discrimination between peptide 3(10)- and alpha-helices. Theoretical analysis of the impact of alpha-methyl substitution on experimental spectra. *J Am Chem Soc* 124:5325–5332
46. Kubelka J, Huang R, Keiderling T (2005) Solvent effects on IR and VCD spectra of helical peptides: DFT-based static spectral simulations with explicit water. *J Phys Chem B* 109:8231–8243
47. Silva R, Yasui S, Kubelka J, Formaggio F, Crisma M, Toniolo C, Keiderling T (2002) Discriminating 3(10)- from alpha helices: vibrational and electronic CD and IR absorption study of related Aib-containing oligopeptides. *Biopolymers* 65:229–243
48. Bour P, Kubelka J, Keiderling T (2002) Ab initio quantum mechanical models of peptide helices and their vibrational spectra. *Biopolymers* 65:45–59
49. Bour P, Keiderling T (2003) Empirical modeling of the peptide amide I band IR intensity in water solution. *J Chem Phys* 119:11253–11262
50. Bour P, Keiderling T (2004) Structure, spectra and the effects of twisting of beta-sheet peptides. A density functional theory study. *J Mol Struct Theochem* 675:95–105
51. Bour P, Keiderling T (2005) Vibrational spectral simulation for peptides of mixed secondary structure: method comparisons with the TrpZip model hairpin. *J Phys Chem B* 109:23687–23697
52. Ireta J, Neugebauer J, Scheffler M, Rojo A, Galvan M (2003) Density functional theory study of the cooperativity of hydrogen bonds in finite and infinite alpha-helices. *J Phys Chem B* 107:1432–1437
53. Ireta J, Neugebauer J, Scheffler M, Rojo A, Galvan M (2005) Structural transitions in the polyalanine alpha-helix under uniaxial strain. *J Am Chem Soc* 127:17241–17244

54. Ireta J, Scheffler M (2009) Density functional theory study of the conformational space of an infinitely long polypeptide chain. *J Chem Phys* 131:085104
55. Wiczorek R, Dannenberg J (2005) Enthalpies of hydrogen-bonds in alpha-helical peptides. An ONIOM DFT/AM1 study. *J Am Chem Soc* 127:14534–14535
56. Salvador P, Asensio A, Dannenberg JJ (2007) The effect of aqueous solvation upon alpha-helix formation for polyalanines. *J Phys Chem B* 111:7462–7466
57. Tsai MIH, Xu Y, Dannenberg JJ (2009) Ramachandran revisited. DFT energy surfaces of diastereomeric trialanine peptides in the gas phase and aqueous solution. *J Phys Chem B* 113:309–318
58. Plumley JA, Dannenberg JJ (2010) The importance of hydrogen bonding between the glutamine side chains to the formation of amyloid VQIVYK Parallel beta-sheets: an ONIOM DFT/AM1 Study. *J Am Chem Soc* 132:1758–1759
59. Kaminsky J, Jensen F (2007) Force field modeling of amino acid conformational energies. *J Comp Theor Comput* 3:1774–1788
60. Csontos J, Palermo N, Murphy R, Lovas S (2008) Calculation of weakly polar interaction energies in polypeptides using density functional and local Moller-Plesset perturbation theory. *J Comp Chem* 29:1344–1352
61. Wang Z Wu C Lei H Duan Y (2007) Accurate ab initio study on the hydrogen-bond pairs in protein secondary structures. *J Comp Theor Comput* 3:1527–1537
62. Han W, Jalkanen K, Elstner M, Suhai S (1998) Theoretical study of aqueous N-acetyl-L-alanine N'-methylamide: structures and Raman, VCD, and ROA spectra. *J Phys Chem B* 102:2587–2602
63. Thar J, Zahn S, Kirchner B (2008) When is a molecule properly solvated by a continuum model or in a cluster ansatz? a first-principles simulation of alanine hydration. *J Phys Chem B* 112:1456–1464
64. Hudaky I, Hudaky P, Perczel A (2004) Solvation model induced structural changes in peptides. A quantum chemical study on Ramachandran surfaces and conformers of alanine diamide using the polarizable continuum model. *J Comp Chem* 25:1522–1531
65. Wang Z, Duan Y (2004) Solvation effects on alanine dipeptide: a MP2/cc-pVTZ//MP2/6-31G\*\* study of (Phi, Psi) energy maps and conformers in the gas phase, ether, and water. *J Comp Chem* 25:1699–1716
66. Jono R, Watanabe Y, Shimizu K, Terada T (2010) Multicanonical ab Initio QM/MM molecular dynamics simulation of a peptide in an aqueous environment. *J Comp Chem* 31:1168–1175
67. Kwac K, Lee K, Han J, Oh K, Cho M (2008) Classical and quantum mechanical/molecular mechanical molecular dynamics simulations of alanine dipeptide in water: comparisons with IR and vibrational circular dichroism spectra. *J Chem Phys* 128:105106
68. Nemukhin A, Grigorenko B, Bochenkova A, Kovba V, Epifanovsky E (2004) Structures of the peptide-water complexes studied by the hybrid quantum mechanical-molecular mechanical (QM/MM) technique. *Struct Chem* 15:3–9
69. Han W, Elstner M, Jalkanen K, Frauenheim T, Suhai S (2000) Hybrid SCC-DFTB/molecular mechanical studies of H-bonded systems and of N-acetyl-(L-Ala)(n) N'-methylamide helices in water solution. *Int J Quant Chem* 78:459–479
70. Hugosson H, Laio A, Maurer P, Rothlisberger U (2006) A comparative theoretical study of dipeptide solvation in water. *J Comp Chem* 27:672–684
71. Degtyarenko I, Jalkanen K, Gurtovenko A, Nieminen R (2007) L-alanine in a droplet of water: a density-functional molecular dynamics study. *J Phys Chem B* 111:4227–4234
72. Jalkanen K, Elstner M, Suhai S (2004) Amino acids and small peptides as building blocks for proteins: comparative theoretical and spectroscopic studies. *J Mol Struct Theochem* 675:61–77
73. Jalkanen KJ, Degtyarenko IM, Nieminen RM, Cao X, Nafie LA, Zhu F, Barron LD (2008) Role of hydration in determining the structure and vibrational spectra of L-alanine and N-acetyl L-alanine N'-methylamide in aqueous solution: a combined theoretical and experimental approach. *Theor Chem Acc* 119:191–210
74. Gaigeot M (2009) Unravelling the conformational dynamics of the aqueous alanine dipeptide with first-principle molecular dynamics. *J Phys Chem B* 113:10059–10062
75. Headgordon T, Headgordon M, Frisch M, Brooks C, Pople J (1991) Theoretical study of blocked glycine and alanine peptide analogs. *J Am Chem Soc* 113:5989–5997
76. Deng Z, Polavarapu P, Ford S, Hecht L, Barron L, Ewig C, Jalkanen K (1996) Solution-phase conformations of N-acetyl-N'-methyl-L-alaninamide from vibrational Raman optical activity. *J Phys Chem* 100:2025–2034

77. Rommelmohle K, Hofmann H (1993) Conformation dynamics in peptides—quantum-chemical calculations and molecular-dynamics simulations on N-acetylalanyl-N'-methylamide. *J Mol Struct Theochem* 104:211–219
78. Elstner M, Jalkanen K, Knapp-Mohammady M, Frauenheim T, Suhai S (2000) DFT studies on helix formation in N-acetyl-(L-alanyl) (n)-N'-methylamide for n=1–20. *Chem Phys* 256:15–27
79. Mohle K, Hofmann H, Thiel W (2001) Description of peptide and protein secondary structures employing semiempirical methods. *J Comp Chem* 22:509–520
80. Elstner M, Jalkanen K, Knapp-Mohammady M, Frauenheim T, Suhai S (2001) Energetics and structure of glycine and alanine based model peptides: approximate SCC-DFTB, AM1 and PM3 methods in comparison with DFT, HF and MP2 calculations. *Chem Phys* 263:203–219
81. Hu H, Elstner M, Hermans J (2003) Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine “dipeptides” (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. *Protein Struct Funct Genet* 50:451–463
82. Seabra G, Walker R, Elstner M, Case D, Roitberg A (2007) Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the amber molecular dynamics package. *J Phys Chem A* 111:5655–5664
83. Torras J, Seabra G, Deumens E, Trickey S, Roitberg A (2008) A versatile AMBER-Gaussian QM/MM interface through PUPIL. *J Comp Chem* 29:1564–1573
84. Elstner M, Frauenheim T, Suhai S (2003) An approximate DFT method for QM/MM simulations of biological structures and processes. *J Mol Struct Theochem* 632:29–41
85. Nemukhin A, Grigorenko B, Bochenkova A, Topol I, Burt S (2002) A QM/MM approach with effective fragment potentials applied to the dipeptide-water structures. *J Mol Struct Theochem* 581:167–175
86. Elstner M, Frauenheim T, Kaxiras E, Seifert G, Suhai S (2000) A self-consistent charge density-functional based tight-binding scheme for large biomolecules. *Phys Stat Sol B* 217:357–376
87. Seabra G, Walker R, Roitberg A (2009) Are current semiempirical methods better than force fields? a study from the thermodynamics perspective. *J Phys Chem A* 113:11938–11948
88. Echenique P, Alonso J (2008) Efficient model chemistries for peptides. I. General framework and a study of the heterolevel approximation in RHF and MP2 with Pople split-valence basis sets. *J Comp Chem* 29:1408–1422
89. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Protein Struct Funct Genet* 65:712–725
90. Kang Y (2006) Conformational preferences of non-prolyl and prolyl residues. *J Phys Chem B* 110:21338–21348
91. Huang Z, Yu W, Lin Z (2006) First-principle studies of gaseous aromatic amino acid histidine. *J Mol Struct Theochem* 801:7–20
92. Makshakova O, Ermakova E (2010) Computational study of hydrogen-bonding complex formation of helical polypeptides with water molecule. *J Mol Struct Theochem* 942:7–14
93. Riccardi D, Li G, Cui Q (2004) Importance of van der Waals interactions in QM/MM Simulations. *J Phys Chem B* 108:6467–6478
94. Valdes H, Pluhackova K, Pitonak M, Rezac J, Hobza P (2008) Benchmark database on isolated small peptides containing an aromatic side chain: comparison between wave function and density functional theory methods and empirical force field. *Phys Chem Chem Phys* 10:2747–2757
95. Valdes H, Spiwok V, Rezac J, Reha D, Abo-Riziq A, de Vries M, Hobza P (2008) Potential-energy and free-energy surfaces of glycyl-phenylalanyl-alanine (GFA) tripeptide: experiment and theory. *Chemistry Eur J* 14:4886–4898
96. Valdes H, Pluhackova K, Hobza P (2009) Phenylalanyl-glycyl-phenylalanine tripeptide: a model system for aromatic-aromatic side chain interactions in proteins. *J Comp Theor Comput* 5:2248–2256
97. McKinney B, Urban J (2010) Fluoroolefins as peptidic mimetics. 2. A computational study of the conformational ramifications of peptide bond replacement. *J Phys Chem A* 114:1123–1133
98. Zhu X, Yethiraj A, Cui Q (2007) Establishing effective simulation protocols for beta- and alpha/beta-mixed peptides. I. QM and QM/MM models. *J Comp Theor Comput* 3:1538–1549
99. Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions—dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme. *J Mol Biol* 103:227–249

100. Field M, Bash P, Karplus M (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J Comp Chem* 11:700–733
101. Gao J, Xia X (1992) A priori evaluation of aqueous polarization effects through monte-carlo QM-MM simulations. *Science* 258:631–635
102. Shurki A, Warshel A (2003) Structure/function correlations of proteins using MM, QM/MM, and related approaches: methods, concepts, pitfalls, and current progress. *Protein Sim* 66:249–313
103. Oldfield E (2005) Quantum chemical studies of protein structure. *Phil Trans Royal Soc B* 360:1347–1361
104. Senn H, Thiel W (2007) QM/MM methods for biological systems. *Top Curr Chem* 268:173–290
105. Carloni P, Rothlisberger U, Parrinello M (2002) The role and perspective of a initio molecular dynamics in the study of biological systems. *Acc Chem Res* 35:455–464
106. Challacombe M, Schwegler E (1997) Linear scaling computation of the Fock matrix. *J Chem Phys* 106:5526–5536
107. Van Alsenoy C, Yu C, Peeters A, Martin J, Schafer L (1998) Ab initio geometry determinations of proteins. 1. Crambin. *J Phys Chem A* 102:2246–2251
108. Scuseria G (1999) Linear scaling density functional calculations with Gaussian orbitals. *J Phys Chem A* 103:4782–4790
109. Tsuda K, Kaneko H, Shimada J, Takada T (2001) Ab initio MO studies of interaction mechanisms of Protein Kinase C with cell membranes. *Comp Phys Comm* 142:140–143
110. Sato F, Yoshihiro T, Era M, Kashiwagi H (2001) Calculation of all-electron wavefunction of hemoprotein cytochrome c by density functional theory. *Chem Phys Lett* 341:645–651
111. Inaba T, Tahara S, Nisikawa N, Kashiwagi H, Sato F (2005) All-electron density functional calculation on insulin with quasi-canonical localized orbitals. *J Comp Chem* 26:987–993
112. Nakanishi I, Fedorov D, Kitaura K (2007) Molecular recognition mechanism of FK506 binding protein: an all-electron fragment molecular orbital study. *Protein Struct Funct Genet* 68:145–158
113. Inaba T, Tsunekawa N, Hirano T, Yoshihiro T, Kashiwagi H, Sato F (2007) Density functional calculation of the electronic structure on insulin hexamer. *Chem Phys Lett* 434:331–335
114. Pichierri F (2005) Insights into the interplay between electronic structure and protein dynamics: the case of ubiquitin. *Chem Phys Lett* 410:462–466
115. Exner T, Mezey P (2003) Ab initio quality properties for macromolecules using the ADMA approach. *J Comp Chem* 24:1980–1986
116. Gogonea V, Merz K (1999) Fully quantum mechanical description of proteins in solution. Combining linear scaling quantum mechanical methodologies with the Poisson-Boltzmann equation. *J Phys Chem A* 103:5171–5188
117. van der Vaart A, Suarez D, Merz K (2000) Critical assessment of the performance of the semiempirical divide and conquer method for single point calculations and geometry optimizations of large chemical systems. *J Chem Phys* 113:10512–10523
118. Van der Vaart A, Gogonea V, Dixon S, Merz K (2000) Linear scaling molecular orbital calculations of biological systems using the semi-empirical divide and conquer method. *J Comp Chem* 21:1494–1504
119. Zhang D, Zhang J (2004) Full ab initio computation of protein-water interaction energies. *J Theor Comp Chem* 3:43–49
120. Dixon S, Merz K (1997) Fast, accurate semi-empirical molecular orbital calculations for macromolecules. *J Chem Phys* 107:879–893
121. Elstner M (2006) The SCC-DFTB method and its application to biological systems. *Theo Chem Acc* 116:316–325
122. Liu H, Elstner M, Kaxiras E, Frauenheim T, Hermans J, Yang W (2001) Quantum mechanics simulation of protein dynamics on long timescale. *Protein Struct Funct Genet* 44:484–489
123. Engh R, Huber R (1991) Accurate bond and angle parameters for X-ray protein-structure refinement. *Acta Cryst A* 47:392–400
124. Yu N, Yennawar H, Merz K (2005) Refinement of protein crystal structures using energy restraints derived from linear-scaling quantum mechanics. *Acta Cryst D* 61:322–332
125. Yu N, Li X, Cui G, Hayik S, Merz K (2006) Critical assessment of quantum mechanics based energy restraints in protein crystal structure refinement. *Prot Sc* 15:2773–2784
126. Canfield P, Dahlbom M, Hush N, Reimers J (2006) Density-functional geometry optimization of the 150 000-atom photosystem-I trimer. *J Chem Phys* 124:024301
127. Stewart J (2009) Application of the PM6 method to modeling proteins. *J Mol Model* 15:765–805

128. Wollacott A, Merz K (2007) Assessment of semiempirical quantum mechanical methods for the evaluation of protein structures. *J Comp Theor Comput* 3:1609–1619
129. He X, Fusti-Molnar L, Cui G, Merz K (2009) Importance of Dispersion and Electron Correlation in ab Initio Protein Folding. *J Phys Chem B* 113:5290–5300
130. Morozov A, Misura K, Tsemekhman K, Baker D (2004) Comparison of quantum mechanics and molecular mechanics dimerization energy landscapes for pairs of ring-containing amino acids in proteins. *J Phys Chem B* 108:8489–8496
131. Masia M (2008) Ab initio based polarizable force field parametrization. *J Chem Phys* 128:184107
132. Mackerell A, Feig M, Brooks C (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comp Chem* 25:1400–1415
133. Okiyama Y, Watanabe H, Fukuzawa K, Nakano T, Mochizuki Y, Ishikawa T, Ebina K, Tanaka S (2009) Application of the fragment molecular orbital method for determination of atomic charges on polypeptides. II. Towards an improvement of force fields used for classical molecular dynamics simulations. *Chem Phys Lett* 467:417–423
134. Thomas A, Milon A, Brasseur R (2004) Partial atomic charges of amino acids in proteins. *Protein Struct Funct Genet* 56:102–109
135. Dupradeau F, Pigache A, Zaffran T, Savineau C, Lelong R, Grivel N, Lelong D, Rosanski W, Cieplak P (2010) The R.ED. tools: advances in RESP and ESP charge derivation and force field library building. *Phys Chem Chem Phys* 12:7821–7839
136. Duan Y, Wu C, Chowdhury S, Lee M, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comp Chem* 24:1999–2012
137. Bayly C, Cieplak P, Cornell W, Kollman P (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges—the RESP model. *J Phys Chem* 97:10269–10280
138. Laio A, VandeVondele J, Rothlisberger U (2002) D-RESP: dynamically generated electrostatic potential derived charges from quantum mechanics/molecular mechanics simulations. *J Phys Chem B* 106:7300–7307
139. Ercolessi F, Adams J (1994) Interatomic potentials from 1st-principles calculations—the force-matching method. *Europhys Lett* 26:583–588
140. Laio A, Bernard S, Chiarotti G, Scandolo S, Tosatti E (2000) Physics of iron at Earth's core conditions. *Science* 287:1027–1030
141. Izvekov S, Parrinello M, Burnham C, Voth G (2004) Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: a new method for force-matching. *J Chem Phys* 120:10896–10913
142. Akin-Ojo O, Song Y, Wang F (2008) Developing ab initio quality force fields from condensed phase quantum-mechanics/molecular-mechanics calculations through the adaptive force matching method. *J Chem Phys* 129:064108
143. Maurer P, Laio A, Hugosson H, Colombo M, Rothlisberger U (2007) Automated parametrization of biomolecular force fields from quantum mechanics/molecular mechanics (QM/MM) simulations through force matching. *J Comp Theor Comput* 3:628–639
144. Bucher D, Guidoni L, Maurer P, Rothlisberger U (2009) Developing improved charge sets for the modeling of the KcsA K<sup>+</sup> channel using QM/MM electrostatic potentials. *J Comp Theor Comput* 5:2173–2179
145. Bucher D, Rothlisberger U (2010) Molecular simulations of ion channels: a quantum chemist's perspective. *J Gener Phys* 135:549–554
146. McMahon B, Stojkovic B, Hay P, Martin R, Garcia A (2000) Microscopic model of carbon monoxide binding to myoglobin. *J Chem Phys* 113:6831–6850
147. Marti M, Crespo A, Capece L, Boechi L, Bikiel D, Scherlis D, Estrin D (2006) Dioxygen affinity in heme proteins investigated by computer simulation. *J Inorg Biochem* 100:761–770
148. Friesner R, Baik M, Gherman B, Guallar V, Wirstam M, Murphy R, Lippard S (2003) How iron-containing proteins control dioxygen chemistry: a detailed atomic level description via accurate quantum chemical and mixed quantum mechanics/molecular mechanics calculations. *Coord Chem Rev* 238:267–290
149. Xie H, Wu R, Zhou Z, Cao Z (2008) Exploring the interstitial atom in the FeMo cofactor of nitrogenase: insights from QM and QM/MM calculations. *J Phys Chem B* 112:11435–11439
150. Babitzki G, Denschlag R, Tavan P (2009) Polarization effects stabilize Bacteriorhodopsin's

- chromophore binding pocket: a molecular dynamics study. *J Phys Chem B* 113:10483–10495
151. Hutter M, Reimers J, Hush N (1998) Modeling the bacterial photosynthetic reaction center. 1. Magnesium parameters for the semi-empirical AM1 method developed using a genetic algorithm. *J Phys Chem B* 102:8080–8090
  152. Hutter M, Hughes J, Reimers J, Hush N (1999) Modeling the bacterial photosynthetic reaction center. 2. A combined quantum mechanical molecular mechanical study of the structure of the cofactors in the reaction centers of purple bacteria. *J Phys Chem B* 103:4906–4915
  153. Reimers J, Hughes J, Hush N (2000) Modeling the bacterial photosynthetic reaction center 3: interpretation of effects of site-directed mutagenesis on the special-pair midpoint potential. *Biochemistry* 39:16185–16189
  154. Hughes J, Hutter M, Reimers J, Hush N (2001) Modeling the bacterial photosynthetic reaction center. 4. The structural, electrochemical, and hydrogen-bonding properties of 22 mutants of *Rhodobacter sphaeroides*. *J Am Chem Soc* 123:8550–8563
  155. Reimers J, Shapley W, Hush N (2003) Modelling the bacterial photosynthetic reaction center. V. Assignment of the electronic transition observed at 2200 cm<sup>-1</sup> in the special-pair radical-cation as a second-highest occupied molecular orbital to highest occupied molecular orbital transition. *J Chem Phys* 119:3240–3248
  156. Reimers J, Shapley W, Rendell A, Hush N (2003) Modelling the bacterial photosynthetic reaction center. VI. Use of density-functional theory to determine the nature of the vibronic coupling between the four lowest-energy electronic states of the special-pair radical cation. *J Chem Phys* 119:3249–3261
  157. Reimers J, Hush N (2003) Modeling the bacterial photosynthetic reaction center. VII. Full simulation of the intervalence hole-transfer absorption spectrum of the special-pair radical cation. *J Chem Phys* 119:3262–3277
  158. Yamasaki H, Takano Y, Nakamura H (2008) Theoretical investigation of the electronic asymmetry of the special pair cation radical in the photosynthetic type-II reaction center. *J Phys Chem B* 112:13923–13933
  159. Sinnecker S, Flores M, Lubitz W (2006) Protein-cofactor interactions in bacterial reaction centers from *Rhodobacter sphaeroides* R-26: effect of hydrogen bonding on the electronic and geometric structure of the primary quinone. A density functional theory study. *Phys Chem Chem Phys* 8:5659–5670
  160. Nonella M, Mathias G, Eichinger M, Tavan P (2003) Structures and vibrational frequencies of the quinones in *Rb. sphaeroides* derived by a combined density functional/molecular mechanics approach. *J Phys Chem B* 107:316–322
  161. Drews J (2000) Drug discovery: a historical perspective. *Science* 287:1960–1964
  162. Steinbrecher T, Labahn A (2010) Towards accurate free energy calculations in ligand protein-binding studies. *Curr Med Chem* 17:767–785
  163. Moustakas D, Lang P, Pegg S, Pettersen E, Kuntz I, Brooijmans N, Rizzo R (2006) Development and validation of a modular, extensible docking program: DOCK 5. *J Comp Aided Mol Design* 20:601–619
  164. T.J.A. Ewing IK (1997) Critical evaluation of search algorithms for automated molecular docking and database screening. *J Comp Chem* 18:1175–1189
  165. Morris G, Goodsell D, Halliday R, Huey R, Hart W, Belew R, Olson A (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comp Chem* 19:1639–1662
  166. Rarey M, Kramer B, Lengauer T (1996) A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 261:470–489
  167. Kramer B, Rarey M, Lengauer T (1999) Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Protein Struct Funct Genet* 37:228–241
  168. Jones G, Willett P, Glen R, Leach A, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 267:727–748
  169. Friesner R, Banks J, Murphy R, Halgren T, Klicic J, Mainz D, Repasky M, Knoll E, Shelley M, Perry J, Shaw D, Francis P, Shenkin P (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 47:1739–1749
  170. Halgren T, Murphy R, Friesner R, Beard H, Frye L, Pollard W, Banks J (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47:1750–1759
  171. Villar R, Gil M, Garcia J, Martinez-Merino V (2005) Are AM1 ligand-protein binding enthalpies good enough for use in the rational design of new drugs? *J Comp Chem* 26:1347–1358

172. Nikitina E, Sulimov V, Zayets V, Zaitseva N (2004) Semiempirical calculations of binding enthalpy for protein-ligand complexes. *Int J Quant Chem* 97:747–763
173. Raha K, Merz K (2005) Large-scale validation of a quantum mechanics based scoring function: predicting the binding affinity and the binding mode of a diverse set of protein-ligand complexes. *J Med Chem* 48:4558–4575
174. Vasilyev V, Bliznyuk A (2004) Application of semiempirical quantum chemical methods as a scoring function in docking. *Theo Chem Acc* 112:313–317
175. Li J, Reynolds C (2009) A quantum mechanical approach to ligand binding - Calculation of ligand-protein binding affinities for stromelysin-1 (MMP-3) inhibitors. *Can J Chem* 87:1480–1484
176. Peters M, Raha K, Merz K (2006) Quantum mechanics in structure-based drug design. *Curr Op Drug Disc* 9:370–379
177. Wang M, Wong C (2007) Rank-ordering protein-ligand binding affinity by a quantum mechanics/molecular mechanics/Poisson-Boltzmann-surface area model. *J Chem Phys* 126:026101
178. Heady L, Fernandez-Serra M, Mancera R, Joyce S, Venkitaraman A, Artacho E, Skylaris C, Ciacchi L, Payne M (2006) Novel structural features of CDK inhibition revealed by an ab initio computational method combined with dynamic simulations. *J Med Chem* 49:5141–5153
179. Fukuzawa K, Kitaura K, Uebayasi M, Nakata K, Kamimura T, Nakano T (2005) Ab initio quantum mechanical study of the binding energies of human estrogen receptor alpha with its ligands: an application of fragment molecular orbital method. *J Comp Chem* 26:1–10
180. Fischer B, Fukuzawa K, Wenzel W (2008) Receptor-specific scoring functions derived from quantum chemical models improve affinity estimates for in-silico drug discovery. *Protein Struct Funct Genet* 70:1264–1273
181. Neugebauer J (2009) Subsystem-based theoretical spectroscopy of biomolecules and biomolecular assemblies. *ChemPhysChem* 10:3148–3173
182. Schapiro I, Ryazantsev MN, Ding WJ, Huntress MM, Melaccio F, Andruniow T, Olivucci M (2010) Computational photobiology and beyond. *Austr J Chem* 63:413–429
183. Mata RA (2010) Application of high level wavefunction methods in quantum mechanics/molecular mechanics hybrid schemes. *Phys Chem Chem Phys* 12:5041–5052
184. Hoffmann M, Wanko M, Strodel P, Koenig PH, Frauenheim T, Schulten K, Thiel W, Taj-khorshid E, Elstner M (2006) Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II. *J Am Chem Soc* 128:10808–10818
185. Babitzki G, Mathias G, Tavan P (2009) The infrared spectra of the retinal chromophore in bacteriorhodopsin calculated by a DFT/MM approach. *J Phys Chem B* 113:10496–10508
186. Mroginski MA, Mark F, Thiel W, Hildebrandt P (2007) Quantum mechanics/molecular mechanics calculation of the Raman spectra of the phycocyanobilin chromophore in alpha-c-phycocyanin. *Biophys J* 93:1885–1894
187. Mroginski MA, von Stetten D, Escobar FV, Strauss HM, Kaminski S, Scheerer P, Guenther M, Murgida DH, Schmieder P, Bongards C, Gaertner W, Mailliet J, Hughes J, Essen LO, Hildebrandt P (2009) Chromophore structure of cyanobacterial phytochrome Cph1 in the Pr state: reconciling structural and spectroscopic data by QM/MM calculations. *Biophys J* 96:4153–4163
188. Mroginski MA, Kaminski S, Hildebrandt P (2010) Raman spectra of the phycobilin cofactor in phycoerythrocyanin calculated by QM/MM methods. *ChemPhysChem* 11:1265–1274
189. Schmitz M, Tavan P (2004) Vibrational spectra from atomic fluctuations in dynamics simulations. I. Theory, limitations, and a sample application. *J Chem Phys* 121:12233–12246
190. Schmitz M, Tavan P (2004) Vibrational spectra from atomic fluctuations in dynamics simulations. II. Solvent-induced frequency fluctuations at femtosecond time resolution. *J Chem Phys* 121:12247–12258
191. Kaminski S, Gaus M, Phatak P, von Stetten D, Elstner M, Mroginski MA (2010) Vibrational Raman spectra from the self-consistent charge density functional tight binding method via classical time-correlation functions. *J Chem Theor Comput* 6:1240–1255
192. Yu H, Cui Q (2007) The vibrational spectra of protonated water clusters: a benchmark for self-consistent-charge density-functional tight binding. *J Chem Phys* 127:234504
193. Phatak P, Ghosh N, Yu H, Cui Q, Elstner M (2008) Amino acids with an intermolecular proton bond as proton storage site in bacteriorhodopsin. *Proc Natl Acad Sci USA* 105:19672–19677
194. Xu Z, Mei Y, Duan L, Zhang D (2010) Hydrogen bonds rebuilt by polarized protein-specific charges. *Chem Phys Lett* 495:151–154

195. Duan LL, Mei Y, Zhang QG, Zhang JZH (2009) Intra-protein hydrogen bonding is dynamically stabilized by electronic polarization. *J Chem Phys* 130:115102
196. Fraehmcke JS, Wanko M, Phatak P, Mroginski MA, Elstner M (2010) The protonation state of glu181 in rhodopsin revisited: interpretation of experimental data on the basis of QM/MM calculations. *J Phys Chem B* 114:11338–11352
197. Wanko M, Hoffmann M, Strodel P, Koslowski A, Thiel W, Neese F, Frauenheim T, Elstner M (2005) Calculating absorption shifts for retinal proteins: Computational challenges. *J Phys Chem B* 109:3606–3615
198. Dreuw A, Head-Gordon M (2005) Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem Rev* 105:4009–4037
199. Dreuw A (2006) Quantum chemical methods for the investigation of photoinitiated processes in biological systems: theory and applications. *ChemPhysChem* 7:2259–2274
200. Wanko M, Garavelli M, Bernardi F, Niehaus T, Frauenheim T, Elstner M (2004) A global investigation of excited state surfaces within time-dependent density-functional response theory. *J Chem Phys* 120:1674–1692
201. Wanko M, Hoffmann M, Frauenheim T, Elstner M (2008) Effect of polarization on the opsin shift in rhodopsins. I. A combined QM/QM/MM model for bacteriorhodopsin and pharaonis sensory rhodopsin II. *J Phys Chem B* 112:11462–11467
202. Wanko M, Hoffmann M, Fraehmcke J, Frauenheim T, Elstner M (2008) Effect of polarization on the opsin shift in rhodopsins. II. Empirical polarization models for proteins. *J Phys Chem B* 112:11468–11478
203. Soederhjelm P, Husberg C, Strambi A, Olivucci M, Ryde U (2009) Protein influence on electronic spectra modeled by multipoles and polarizabilities. *J Chem Theor Comput* 5:649–658
204. Wanko M (2008) Ph.D. thesis, Technical University of Braunschweig, Germany
205. Phatak P, Fraehmcke JS, Wanko M, Hoffmann M, Strodel P, Smith JC, Suhai S, Bondar AN, Elstner M (2009) Long-distance proton transfer with a break in the bacteriorhodopsin active site. *J Am Chem Soc* 131:7064–7078
206. Xie W, Gao J (2007) Design of a next generation force field: the X-POL potential. *J Chem Theor Comput* 3:1890–1900
207. Gresh N, Cisneros GA, Darden TA, Piquemal JP (2007) Anisotropic, polarizable molecular mechanics studies of inter- and intramolecular interactions and ligand-macromolecule complexes. A bottom-up strategy. *J Chem Theor Comput* 3:1960–1986

## **Part II**

### **Classical Mechanics: Atomistic Simulations**

# Chapter 6

## Classical Molecular Dynamics in a Nutshell

Susanna Hug

### Abstract

This chapter provides an overview of the various techniques that are commonly used in classical molecular dynamics simulations. It describes suitable algorithms for the integration of Newton's equation of motion over many time steps for systems containing a large number of particles, different choices of boundary conditions as well as available force fields for biological systems, that is, the mathematical description of the interactions of atoms and molecules with each other. It also illustrates algorithms used to simulate systems at constant temperature and/or pressure and discusses their advantages and disadvantages. It presents a few methods to save CPU time and a summary of popular software for biomolecular molecular dynamics simulations.

**Key words:** Molecular dynamics simulation, Numerical integration, Boundary conditions, Force field, Thermostat, Barostat, Computer simulation

---

### 1. Introduction

Biological phenomena typically involve many different time and length scales. For example, in the process of protein folding, time scales range from tens of femtoseconds (bond vibrations (1)) to seconds and more (folding of an entire protein complex (2)). The length scale ranges from a few Ångströms on the atomic level to the size of a folded protein (1–100 nm). Different methods are normally employed to describe the dynamics of biological macromolecules depending on the length and time scale of the process of interest. Methods based on quantum mechanics provide the most accurate results; however, they are computationally expensive, and thus they are limited to short time scales (a few tens of picoseconds) and small length scales (nanometers), involving relatively few atoms. Theory and applications of these techniques are introduced in the first section of the present book. Methods based on classical mechanics (usually referred to as molecular mechanics) allow the description of larger systems (up to millions of atoms or more) on longer time

scales (typically up to microseconds). These are the subject of the second section of the book. To access longer time and length scales, models based on classical mechanics need to be simplified. Several ways to simplify the description of molecular systems have been devised, and we will refer to them as mesoscopic or coarse-grained methods; these are the subject of the book's third section.

Molecular mechanics includes a wide range of simulation techniques. Methods like energy minimization, normal mode analysis, and Monte Carlo simulations do not rely on equations of motion, and therefore do not provide direct information on the dynamics of the system. In methods such as Brownian dynamics (BD), dissipative particle dynamics (DPD), and molecular dynamics (MD), the positions and momenta of all particles in a system are calculated as a function of time by integrating an equation of motion in discrete time steps. Among the different molecular mechanics techniques, molecular dynamics is probably the most commonly used in the field of biological macromolecules. In molecular dynamics (MD), atomic motion is simulated by solving Newton's equation of motion simultaneously for all atoms in the system. MD simulations can be used to obtain both equilibrium and transport properties of a system. Examples of systems that can be simulated by classical MD simulations are ligand binding to enzymes, the self-assembly of lipids in relatively small aggregates, small molecules adsorbing to a surface, and the influence of molecules on the properties of a lipid bilayer.

In classical MD simulations, electrons are not treated explicitly. Forces between atoms are calculated using an empirically derived force field, that is, a collection of mathematical functions and parameters that describe the interaction between different types of atoms in different chemical environments (see the chapters by Tieleman and by Salonen). The separation between electronic and nuclear motions (known as Born–Oppenheimer approximation) is justified by the very different time scale of nuclear and electronic motions (due to the high ratio between the masses of nuclei and electrons). In classical force fields, electronic interactions are taken into account implicitly in an average, effective way. By treating the atoms in a classical way, computation time is saved and larger systems can be simulated for a longer time span. Therefore, in classical MD, a coarse-grained model of an atom is used, where the nucleus and the electrons are combined in one particle and atoms are modeled as inert spheres evolving according to Newton's equation of motion.

The aim of this chapter is to give the reader an overview of the various techniques that are commonly used in MD simulations. We will provide some details, but we encourage the reader to take a look at the many excellent books such as Allen and Tildesley (3), Frenkel and Smit (4), and Rapaport (5) to mention a few. They also provide source codes and pseudocodes in different programming languages.

---

## 2. Integration

Newton's equation of motion for a system consisting of  $N$  particles interacting via a potential  $U(\vec{r}_i)$ ,  $i = 1, \dots, N$  can be formulated as the following second-order differential equation:

$$\ddot{\vec{F}}_i = m_i \ddot{\vec{r}}_i \quad (1)$$

or equivalently, as two first-order differential equations:

$$\begin{cases} \dot{\vec{r}}_i = \frac{\vec{p}_i}{m_i}, \\ \dot{\vec{p}}_i = \vec{F}_i \end{cases} \quad (2)$$

where the forces are given as

$$\vec{F}_i = -\frac{\partial U(\vec{r}_i)}{\partial \vec{r}_i} \quad (3)$$

plus the initial conditions. Here,  $m_i$  stands for the mass of the particle and  $\dot{\vec{r}}_i$  denotes the derivative of  $\vec{r}_i$  with respect to time  $t$ . Therefore, in three dimensions for each time step, either a set of  $3N$  second-order differential equations or an equivalent set of  $6N$  first-order differential equations needs to be solved.

This set of ordinary differential equations is transformed into difference equations. An integrator is defined by the chosen form and solution method of the difference equations. There are several requirements for such an integrator. First, the algorithm should be consistent, that is, in the limit of the step size  $h \rightarrow 0$ , it should reproduce the original differential equations. It should also be accurate, meaning that the numerical solution should be close to the actual solution, which for longer time intervals also requires it to be stable. Another important criterion is efficiency; the calculation of the forces is very computationally expensive, and therefore, being able to use a large time step saves computation time. Because Newton's equations of motion are time-reversible, the algorithm should be time-reversible as well. The algorithm should also be symplectic (4). This means that in the transformation from the previous time step to the next time step, the volume in phase space is preserved and the total energy of the system is conserved. In Hamiltonian systems, symplecticness is expressed by Liouville's theorem. In an ensemble containing different members with a phase space vector  $\vec{x}(\{\vec{p}_i\}, \{\vec{r}_i\})$ ,  $i = 1, \dots, N$  containing all the positions and momenta corresponding to a particular microstate, the probability to find a particular member in a small volume  $d\vec{x}$  around a point  $\vec{x}$  in the phase space at time  $t$  is denoted by  $f(\vec{x}, t)d\vec{x}$ , where  $f(\vec{x}, t)$  is the phase space probability density or phase space distribution function. The total number of members in an ensemble is  $\int d\vec{x}f(\vec{x}, t)$ . Liouville's theorem states that the total number of

systems in an ensemble is constant. From this theorem, the Liouville equation

$$\frac{df}{dt} = \frac{\partial}{\partial t} f(\vec{x}, t) + \dot{\vec{x}} \cdot \nabla_{\vec{x}} f(\vec{x}, t) = 0 \quad (4)$$

is found, which implies that the phase space distribution function is constant along any trajectory in phase space. A symplectic algorithm is accurate enough (6), but an accurate algorithm is not necessarily symplectic (7). Accurate algorithms have a good short-term energy conservation, but they may have a long-term energy drift. A symplectic algorithm does not necessarily have a good short-term energy conservation (4) but will have little long-term energy drift. Although symplecticness is a criterion that an integrator for MD simulations should satisfy, it does not guarantee a good algorithm. Finally, the algorithm should also be easy to implement. Many methods usually used to solve ordinary differential equations, such as Runge–Kutta methods, are not suitable for MD simulations, because they are not symplectic and therefore have a long-term drift in the total energy (8, 9). Second-order algorithms such as the Verlet algorithm (10) and algorithms derived from it, such as the velocity Verlet algorithm (11) and the leapfrog algorithm (12), perform well for complex systems (described below). They are reasonably accurate, satisfy the other criterions mentioned above, and can also easily be modified for different equations of motion. There exist higher-order algorithms with higher accuracy and therefore better short-term energy conservation, of which, the most popular class in MD simulations are predictor–corrector algorithms (4, 13). These algorithms, however, are not time-reversible nor are they symplectic, and therefore, they have a long-term energy drift (4).

## **2.1. Derivation of Algorithms from the Liouville Formulation using the Trotter Expansion**

Trotter expansion (14, 15) is a tool used to develop algorithms from equations of motion. Especially for complex systems, it is hard to find a scheme which is symplectic, that is, the volume in phase space is preserved and the growth in the error of energy conservation is bounded. Examples of algorithms found by this method are the reversible reference system propagator (r-RESPA) technique (16, 17), a multiple time-step (MTS) integration method, or a symplectic algorithm to implement the Nosé–Hoover Chain equations (see below) to simulate the NVT ensemble (15). As an illustrative example, the method for a Hamiltonian system is described below.

By introducing the phase space vector  $\vec{x}(\{\vec{p}_i\}, \{\vec{r}_i\})$ ,  $i = 1, \dots, N$  for a system consisting of  $N$  particles, where all the momenta  $\vec{p} \equiv \vec{p}_1, \dots, \vec{p}_N$  and coordinates  $\vec{r} \equiv \vec{r}_1, \dots, \vec{r}_N$  are collected, Newton's equation of motion in the formulation of first-order

differential equations (Eq. 2) can be written in the form of an operator equation

$$\dot{\vec{x}} = iL\vec{x}, \quad (5)$$

where the  $iL$  is a generalization of the Liouville operator, which has the formal solution

$$\vec{x}(t) = \exp(iLt)\vec{x}(0). \quad (6)$$

In a system with Hamiltonian given by

$$\mathcal{H}\left(\{\vec{p}_i\}, \{\vec{r}_i\}\right) = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m_i} + U\left(\{\vec{r}_i\}\right), \quad (7)$$

the Liouville operator becomes

$$iL = \sum_{i=1}^N \left[ \frac{\partial \mathcal{H}}{\partial \vec{p}_i} \cdot \frac{\partial}{\partial \vec{r}_i} - \frac{\partial \mathcal{H}}{\partial \vec{r}_i} \cdot \frac{\partial}{\partial \vec{p}_i} \right] = \sum_{i=1}^N \frac{\vec{p}_i}{m_i} \cdot \frac{\partial}{\partial \vec{r}_i} + \sum_{i=1}^N \vec{F}_i \cdot \frac{\partial}{\partial \vec{p}_i} \quad (8)$$

$$= \sum_{i=1}^N \dot{\vec{r}}_i \cdot \frac{\partial}{\partial \vec{r}_i} + \sum_{i=1}^N \dot{\vec{p}}_i \cdot \frac{\partial}{\partial \vec{p}_i} \equiv iL_1 + iL_2. \quad (9)$$

The formal solution can be determined only for a few simple cases explicitly, for example, for the assumption  $iL = iL_1$  or  $iL = iL_2$ . If

$$iL = iL_1 = \sum_{i=1}^N \dot{\vec{r}}_i(0) \cdot \frac{\partial}{\partial \vec{r}_i} \quad (10)$$

is inserted into the formal solution and a Taylor expansion is used, the following result is found:

$$\vec{x}(t) = \exp(iL_1 t)\vec{x}(0) = \exp\left(\sum_{i=1}^N \dot{\vec{r}}_i t \frac{\partial}{\partial \vec{r}_i}\right) \vec{x}(0) \quad (11)$$

$$= \vec{x}(0) + iL_1 t \vec{x}(0) + \frac{(iL_1)^2}{2!} \vec{x}(0) + \dots \quad (12)$$

$$= \sum_{n=0}^{\infty} \left( \sum_{i=1}^N \left( \frac{(\dot{\vec{r}}_i(0)t)^n}{n!} \frac{\partial^n}{\partial \vec{r}_i^n} \right) \vec{x}(0) \right) \quad (13)$$

$$= x\left[\{\dot{\vec{p}}_i(0)\}, \left\{\left(\vec{r}_i + \dot{\vec{r}}_i(0)t\right)\right\}\right], \quad (14)$$

where the momenta are at  $t = 0$  and to the position the velocity multiplied by time  $t$  is added. Hence, the effect of the application of  $\exp(iL_1 t)$  is a simple shift of the coordinates. It can be shown in a similar manner that the application of  $\exp(iL_2)$ , where

$$iL_2 = \sum_{i=1}^N \dot{\vec{p}}(0) \cdot \frac{\partial}{\partial \vec{p}_i} \quad (15)$$

leads to a simple shift in the momenta. Now to find the effect of the total Liouville operator, the Trotter theorem (18, 19) can be used, which states

$$\exp(iLt) = \exp[(iL_1 + iL_2)t] \quad (16)$$

$$= \lim_{P \rightarrow \infty} \left[ \exp\left(\frac{iL_2 t}{2P}\right) \exp\left(\frac{iL_1 t}{P}\right) \exp\left(\frac{iL_2 t}{2P}\right) \right]^P. \quad (17)$$

Defining  $\frac{t}{P} = \Delta t$  for finite and large  $P$  leads to the following approximation (20):

$$\exp(iL\Delta t) \approx \left[ \exp\left(iL_2 \frac{\Delta t}{2}\right) \exp(iL_1 \Delta t) \exp\left(iL_2 \frac{\Delta t}{2}\right) \right] + \mathcal{O}(\Delta t^3) \quad (18)$$

$$\exp(iLP\Delta t) \approx \prod_{k=1}^P \exp\left(iL_2 \frac{\Delta t}{2}\right) \exp(iL_1 \Delta t) \exp\left(iL_2 \frac{\Delta t}{2}\right) + \mathcal{O}(t\Delta t^2), \quad (19)$$

and the formal solution of the Liouville equation can be replaced by a discretized scheme, by using the above approximation. Hence, by applying the operator  $\exp(iL_2 \frac{\Delta t}{2}) \exp(iL_1 \Delta t) \exp(iL_2 \frac{\Delta t}{2})$ , the system advances one time step.

The total effect of applying this operator to the coordinates and the momenta is found to be

$$\vec{p}_i(0) \rightarrow \vec{p}_i(0) + \frac{\Delta t}{2} \left( \vec{F}_i(0) + \vec{F}_i(\Delta t) \right) \quad (20)$$

$$\vec{r}_i(0) \rightarrow \vec{r}_i(0) + \Delta t \dot{\vec{r}}_i \left( \frac{\Delta t}{2} \right) \quad (21)$$

$$= \vec{r}_i(0) + \Delta t \dot{\vec{r}}_i(0) + \frac{\Delta^2 t}{2m_i} \vec{F}_i(0), \quad (22)$$

which is equivalent to the *velocity Verlet algorithm* (11) defined as

$$\vec{v}_i \left( t + \frac{\Delta t}{2} \right) = \vec{v}_i(t) + \frac{1}{2m_i} \vec{F}_i(t) \Delta t \quad (23)$$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i \left( t + \frac{\Delta t}{2} \right) \Delta t \quad (24)$$

$$\vec{v}_i(t + \Delta t) = \vec{v}_i \left( t + \frac{\Delta t}{2} \right) + \frac{1}{2m_i} \vec{F}_i(t + \Delta t) \Delta t. \quad (25)$$

Switching the terms in the sum of the Liouville operator and using the Trotter theorem (18, 19) again lead to

$$\exp(iLt) = \exp[(iL_2 + iL_1)t] \quad (26)$$

$$= \lim_{P \rightarrow \infty} \left[ \exp\left(\frac{iL_1 t}{2P}\right) \exp\left(\frac{iL_2 t}{P}\right) \exp\left(\frac{iL_1 t}{2P}\right) \right]^P. \quad (27)$$

The total effect of applying this operator is

$$\vec{p}_i(0) \rightarrow \vec{p}_i(0) + \Delta t \vec{F}_i \left( \frac{\Delta t}{2} \right) \quad (28)$$

$$\vec{r}_i(0) \rightarrow \vec{r}_i(0) + \frac{\Delta t}{2} \left[ \dot{\vec{r}}_i(0) + \dot{\vec{r}}_i(\Delta t) \right] \quad (29)$$

$$= \vec{r}_i(0) + \Delta t \vec{v}_i(0) + \frac{\Delta^2 t}{2m_i} \vec{F}_i \left( \frac{\Delta t}{2} \right). \quad (30)$$

These equations are equal to the *position Verlet algorithm* (16, 21, 22) defined as

$$\vec{r}_i \left( t + \frac{\Delta t}{2} \right) = \vec{r}_i(t) + \frac{\Delta t}{2} \vec{v}_i(t) \quad (31)$$

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{1}{m} \vec{F}_i \left( t + \frac{\Delta t}{2} \right) \Delta t \quad (32)$$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i \left( t + \frac{\Delta t}{2} \right) + \frac{\Delta t}{2} \vec{v}_i(t + \Delta t). \quad (33)$$

The application of each operator as one step can easily be turned into instructions in a computer code. This direct translation technique is very useful in more complex systems, for example, for non-Hamiltonian or quantum systems.

A transformation from the original phase point  $\vec{x}_0$  to the phase point a time step later  $\vec{x}_{\Delta t}$  is a symplectic transformation if and only if it satisfies

$$J^T T J = T, \quad (34)$$

where  $J$  is the Jacobian

$$J = \begin{vmatrix} \frac{\partial(r_{\Delta t}^1, \dots, r_{\Delta t}^n)}{\partial(r_0^1, \dots, r_0^n)} & \frac{\partial(r_{\Delta t}^1, \dots, r_{\Delta t}^n)}{\partial(p_0^1, \dots, p_0^n)} \\ \frac{\partial(p_{\Delta t}^1, \dots, p_{\Delta t}^n)}{\partial(r_0^1, \dots, r_0^n)} & \frac{\partial(p_{\Delta t}^1, \dots, p_{\Delta t}^n)}{\partial(p_0^1, \dots, p_0^n)} \end{vmatrix}, \quad (35)$$

and  $T$  is the  $6N$ -dimensional matrix

$$T = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \quad (36)$$

where  $I$  is the  $3N \times 3N$  identity matrix. For the velocity and position Verlet integrator, the Jacobian of the transformation for a time step is the  $6N \times 6N$  identity matrix, which satisfies Eq. 34. Therefore, these integrators are symplectic.

## 2.2. The Verlet Algorithm

The Verlet algorithm (10) is obtained by adding together the Taylor expansions of  $\vec{r}(t + \Delta t)$  and  $\vec{r}(t - \Delta t)$  about time  $t$ . It reads as follows:

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \frac{1}{m} \vec{F}_i(t)(\Delta t)^2. \quad (37)$$

The Verlet algorithm uses the position and force at time  $t$  and the position at time  $(t - \Delta t)$  to calculate the new position at time  $(t + \Delta t)$ . Therefore, at  $t = 0$ , the position at time  $(-\Delta t)$  is needed. This problem can be solved by using a Taylor expansion about  $\vec{r}(t)$  or a different algorithm for the first time step. The velocity, which does not appear explicitly in the Verlet algorithm, can be obtained from the finite difference formula:

$$\vec{v}_i(t) = \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i(t - \Delta t)}{2\Delta t}. \quad (38)$$

In this algorithm, the velocity term is always a step behind the position term.

## 2.3. The Leapfrog Algorithm

The leapfrog algorithm (12) is a modification of the Verlet algorithm, given by

$$\vec{v}_i\left(t + \frac{\Delta t}{2}\right) = \vec{v}_i\left(t - \frac{\Delta t}{2}\right) + \frac{1}{m_i} \vec{F}_i(t)\Delta t \quad (39)$$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i\left(t + \frac{\Delta t}{2}\right)\Delta t. \quad (40)$$

Eliminating the velocity from this algorithm shows that it is algebraically equivalent to the Verlet algorithm. The leapfrog algorithm uses the position and the force at time  $t$  and the velocity at half a time step ( $t - \frac{\Delta t}{2}$ ) to update the positions and velocities. The name of this algorithm refers to the fact that the velocity leaps over the coordinate to give the next half-step value of the velocity, which is then used to calculate the new positions. The velocity at a full time step is obtained as follows:

$$\vec{v}_i = \frac{\vec{v}_i\left(t + \frac{\Delta t}{2}\right) + \vec{v}_i\left(t - \frac{\Delta t}{2}\right)}{2}. \quad (41)$$

The advantage of programming the equations of motion in the form of the leapfrog algorithm is that the velocity appears explicitly. However, it might be desirable to obtain the velocity at the same time step as the position directly from the algorithm without having to perform further calculations, then the velocity or position Verlet algorithm (described above) might be the integrator of choice.

---

### 3. Interatomic Interactions

Interactions can be divided in bonded interactions (bond stretching, bond rotation, angle bending) and nonbonded interactions (electrostatic and van der Waals interactions). These interactions are modeled by a force field (an appropriate mathematical function and the associated parameters). These parameters are determined as much as possible from experimental data and supplemented with data obtained from *ab initio* calculations (23). The potential of a certain atom depends on its binding partners and its chemical environment (e.g., aqueous or organic solvent) therefore, the force fields usually have several atom types describing different situations (23). The force fields for biomolecules are either all atom, that is, all atoms are explicitly represented, or united atom, that is, only heavy (non-hydrogen) atoms and polar hydrogen atoms are included explicitly and the nonpolar hydrogen atoms are included in the carbon atom to which they are bound by an increased van der Waals radius. Some commonly used force fields for biological systems among others are CHARMM (24, 25), GROMOS (26, 27), AMBER (28, 29), or OPLS (30, 31). CHARMM and OPLS are all-atom force fields, making it easier to treat hydrogen bonds. GROMOS is a united-atom force field, which saves computation time by not including all hydrogen atoms explicitly. There are limitations of classical force fields, for example, by assigning just one point charge to each atom, multiple moments of an atom cannot be represented. Therefore, force fields for biomolecules using more computationally expensive polarizable models instead of the fixed-point charge models are being developed. An example is the AMOEBA force field developed by Ponder et al. (32, 33). To model bonded interactions, two-body potentials are used to describe the bond stretching, for example, a harmonic potential, the Morse potential (34), or the FENE potential (35). To model angle bending and bond rotation, three-body and four-body potentials are needed, respectively. To model the bonds in semiconductors, three-body potentials (e.g., Stillinger–Weber potential (36)) or other many-body potentials (e.g., Tersoff potential (37)) are used. To model nonbonded interactions in liquids, two-body potentials are usually sufficient. A more detailed discussion of the treatment of interactions in classical MD simulations can be found in the chapters covering force fields by Luca Monticelli and Peter Tielemans and electrostatics by Celeste Sagui et al.

---

## 4. Simulations at Constant Temperature and Pressure: Thermostats and Barostats

Direct integration of Newton's equations of motion produces the microcanonical ensemble, that is, the number of particles  $N$ , the volume  $V$ , and the total energy  $E$  are kept constant (NVE ensemble). Experiments on the other hand are often conducted at constant temperature and/or constant pressure. It is therefore desirable to perform MD simulations in ensembles other than the microcanonical, such as the canonical ensemble (NVT ensemble), where the total energy is allowed to change in order to keep the temperature constant, or the isothermal-isobaric ensemble (NPT ensemble), where in addition to the temperature, the pressure is kept constant by allowing the volume to change.

Thermostats can be broadly divided into global and local (38). The global ones act instantaneously with the same strength on all particles. Examples of such thermostats are the Nosé–Hoover (39, 40) and the Berendsen weak-coupling method (41). Local thermostats include the Langevin (42), Andersen (43), and the DPD methods (44).

### 4.1. The Nosé–Hoover Thermostat

Nosé (39) proposed an extended system approach introducing an additional coordinate  $s$  representing a thermal reservoir. Associated with this coordinate  $s$  are the potential energy term  $U_s = gk_B T \ln s$  and the kinetic energy term  $K_s = Qs^2/2$ , where  $Q$  represents an effective “mass.” It can be shown that in order to obtain the system average of a canonical ensemble, the parameter  $g$  has to be chosen equal to the number of the degrees of freedom plus 1 (45). The Lagrangian of this extended system is

$$\mathcal{L}_s = K + K_s - U(\{\vec{r}_i\}) - U_s, \quad (42)$$

where  $K = 1/2 \sum_i m_i \dot{\vec{r}}_i^2$ . The equations of motion can be obtained using the Euler–Lagrange equation. Hoover developed a different set of equation based on Nosé’s idea, but independent of time scaling (40).

Hoover (40) and later Nosé (46) demonstrated that their equations of motion would sample the phase space of a canonical ensemble. However, Ciccotti and Kalibaeva (47) have introduced a method to find the correct statistical ensemble of a system with extended variable dynamics by using the approach of Tuckerman et al. (48, 49) to formulate the statistical mechanics of non-Hamiltonian dynamical system, which yields a partition and distribution function for the Nosé–Hoover dynamics completely different from those functions for a canonical ensemble.

This problem is solved in the Nosé–Hoover chain algorithm (50, 51), which samples the correct canonical ensemble. In this approach,

several heat baths are coupled to each other, each represented by an additional coordinate  $s_k$ . Through this succession of thermostats coupled to each other, the kinetic energy fluctuation of one thermostat is controlled through the coupling to the next thermostat.

#### 4.2. The Berendsen Weak-Coupling Algorithm

In the method proposed by Berendsen et al. (41), the system is weakly coupled to a heat bath with the desired temperature  $T_d$ . When the current kinetic temperature  $T$  of the system deviates from the desired temperature, it is corrected by

$$\frac{dT}{dt} = \frac{T_d - T}{t_T}, \quad (43)$$

where the time constant  $t_T$  determines the rate of this correction.

So the velocities are rescaled at each time step by the scaling factor

$$\chi = \left[ 1 + \frac{\Delta t}{t_T} \left( \frac{T_d}{T} - 1 \right) \right]^{1/2}, \quad (44)$$

where  $\Delta t$  is the simulation time step. By rescaling the velocities, the fluctuations of kinetic energy are suppressed.

It has been shown (52) that this thermostat violates the equipartition theorem and hence does not generate a canonical ensemble. Over the simulation time, there is a gradual loss of the vibrational, internal kinetic energy and an increase in translational, external kinetic energy. In general, the kinetic energy of zero and lower-frequency motions (e.g., global translation and rotation) is increasing, while the kinetic energy associated with high-frequency motions (i.e., internal degrees of freedom) is decreasing. This leads to what is referred to as a “flying ice cube” (52). This problem is fixed in the velocity rescale thermostat described below.

An example where the Berendsen thermostat leads to wrong results has been pointed out by Leyssale and Vignoles (53). They simulated small nanodiamond clusters and compared simulations thermostated with the Berendsen method with simulations using the Gaussian isokinetic thermostat of Evans et al. (54) and the stochastic thermostat of Andersen described below.

#### 4.3. The Parrinello–Bussi Thermostat

The Parrinello–Bussi thermostat (55) is an improved version of the Berendsen weak-coupling thermostat, and it samples the correct canonical ensemble. In this thermostat, the system is first evolved one time step as it would without any thermostat. Next, the kinetic energy is evolved one time step as well by using stochastic dynamics. Finally, the velocities are rescaled according to the new kinetic energy. The only condition on the stochastic dynamics is that it does not alter the canonical distribution  $\bar{P}(K_d) dK_d \propto K_d^{N_f/2-1} \exp(-\beta K_d) dK_d$ , where  $K_d$  is the target value for the kinetic energy,  $N_f$  is the number of degrees of freedom, and  $\beta$  is the inverse of the temperature.

If a first-order differential equation in  $K$  is chosen, the most general form under the appropriate conditions is found to be

$$dK = \left[ D(k) \frac{\partial \log(\bar{P})}{\partial K} + \frac{\partial D(K)}{\partial K} \right] dt + \sqrt{2D(K)} dW, \quad (45)$$

where  $D(K)$  is an arbitrary positive definite function of  $K$  and  $dW$  is a Wiener process. If the canonical distribution is used with the average kinetic energy  $\bar{K} = N_f/(2\beta)$ , we obtain

$$dK = \left[ \frac{N_f D(K)}{2K_d K} (K_d - K) - \frac{D(k)}{K} + \frac{\partial D(K)}{\partial K} \right] dt + \sqrt{2D(K)} dW. \quad (46)$$

The choice of  $D(K)$  can affect the performance of the thermostat. Bussi et al. proposed

$$D(K) = \frac{2KK_d}{N_f t_T}, \quad (47)$$

which then leads to the following expression for the stochastic dynamics of the kinetic energy  $K$ :

$$dK = (K_d - K) \frac{dt}{t_T} + 2 \sqrt{\frac{KK_d}{N_f}} \frac{dW}{\sqrt{t_T}}. \quad (48)$$

#### 4.4. The Andersen Thermostat

To keep the temperature constant in real systems, it needs be coupled to a heat bath at the desired temperature. In this method proposed by Andersen (43), stochastic collisions of randomly selected particles with an imaginary heat bath take place at certain time intervals, that is, a new velocity from the Maxwell–Boltzmann distribution corresponding to the desired temperature is assigned to the chosen particle. Between these collisions, the system behaves according to Newtonian dynamics at a constant energy. The collisions take the system from one constant-energy surface to another, ensuring that all the important regions of the phase space are visited. This combination leads to an irreducible Markov chain (56). It can be shown that this thermostat produces a canonical ensemble (43). Therefore, the Andersen thermostat is a good choice for studying static quantities, since it is quite simple. However, since random velocities are used, the dynamics that is generated is unphysical. Therefore, the Andersen thermostat should not be used to determine dynamical quantities, for example, diffusion coefficients.

#### 4.5. The Langevin Thermostat

Langevin dynamics (42) describes the movements of particles in a viscous medium by adding a friction and a noise term to Newton's equation of motion:

$$m_i \ddot{\vec{r}}_i = \vec{F}_i - m_i \Gamma_i \dot{\vec{r}}_i + \vec{\xi}_i(t), \quad (49)$$

where  $\Gamma_i$  is the friction constant and  $\vec{\xi}_i$  describes Gaussian noise with its first and second moments given as  $\langle \vec{\xi}_i(t) \rangle = 0$  and  $\langle \vec{\xi}_i(t) \vec{\xi}_j(t') \rangle = 6k_B T \Gamma \delta_{ij} \delta(t - t')$ .

The noise term models Brownian motion when the particle randomly collides with solvent molecules of temperature  $T$ . The magnitude of the friction constant determines how fast the temperature is relaxed. For large inverse friction constants compared to the time scales in the simulated system, Langevin dynamics resembles a MD simulation with a stochastic heat bath. The Langevin thermostat does not conserve momentum. In real systems, however, momentum cannot vanish, but is transported away. This transport of momentum is responsible for long-range hydrodynamic interactions. Since momentum is not conserved, the hydrodynamic interactions are screened and not described correctly. This is not always a problem, and its importance should be considered on a case-by-case basis. The Langevin thermostat has been very successfully and extensively used in particular in polymer simulations. The Langevin thermostat is generally a good choice in systems where hydrodynamic is not important.

#### **4.6. Dissipative Particle Dynamics (DPD Thermostat)**

In DPD (44, 57, 58), a local friction and a noise term are added to Newton's equations of motion as in Langevin dynamics, but the friction force dampens the velocity differences of neighboring particles, instead of the velocities of the particles directly as is done in the Langevin thermostat and the random force acts on pairs of neighboring particles to fulfill Newton's third law. In that way, locality and momentum conservation is given, a requirement to model hydrodynamics at large length and time scales. The friction or dissipative force and the random force are split up in particle-pair forces. An introduction to DPD is presented in the book's section on mesoscopic techniques.

#### **4.7. The Andersen Barostat**

The barostat proposed by Andersen (43) introduces an external variable  $V$ , the volume of the simulations box, to which the system is coupled. This creates the same effect as a piston has in a real system. This new variable  $V$  has the kinetic energy

$$K_V = \frac{1}{2} Q \dot{V}^2, \quad (50)$$

where  $Q$  stands for the “mass” of the piston, and the potential energy  $U_V = P_d V$ , where  $P_d$  is the desired pressure. By introducing the scaled molecular position  $\vec{s} = \vec{r}/V^{1/3}$  and velocity  $\dot{\vec{s}} = \vec{v}/V^{1/3}$ , the Lagrangian of the extended system can be expressed as

$$\begin{aligned} \mathcal{L}_V(\vec{s}, \dot{\vec{s}}, V, \dot{V}) &= K + K_V - U - U_V \\ &= \frac{1}{2} V^{2/3} \sum_{i=1}^N m_i \dot{\vec{s}}_i^2 + \frac{1}{2} Q \dot{V}^2 - U(V^{1/3} \vec{s}) - P_d V. \end{aligned} \quad (51)$$

The equations of motion are obtained from this Lagrangian for the extended system by using the Euler–Lagrange equation. They generate an isobaric–isoenthalpic ensemble.

Another barostat is the *Nosé–Hoover barostat*, an extension of Andersen’s approach for molecular systems (59), where the equations of motion are independent of time scaling (40).

#### 4.8. The Parrinello–Rahman Barostat

In the Parrinello–Rahman (60, 61) barostat, the simulation box cannot just change its size, but also its shape, by using scaled coordinates  $\vec{r}_i = H\vec{s}_i$ , where  $H = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$  is the transformation matrix and  $\vec{b}_i$  are so-called coulomb vectors representing the sides of the box. The volume of the box is  $V = |H| = \vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)$ . The squared distance between two particles  $i$  and  $j$  is given by  $r_{ij}^2 = (\vec{s}_i - \vec{s}_j)^T G(\vec{s}_i - \vec{s}_j)$  with the metric tensor  $G = H^T H$ . The “potential energy” associated to the volume of the box is  $U_V = P_d V$ , and the extra “kinetic energy” term is  $K_V = \frac{1}{2} QTr(\dot{H}^T \dot{H})$  with  $P_d$  being the desired pressure and  $Q$  the “mass” of the box. The equations of motion can again be obtained from the Lagrangian of the extended system

$$\begin{aligned} \mathcal{L}_V = K + K_V - U - U_V &= \frac{1}{2} \sum_{i=1}^N m_i \dot{\vec{s}}_i^T G \dot{\vec{s}}_i + \frac{1}{2} QTr(\dot{H}^T \dot{H}) \\ &\quad - \sum_{i=1}^N \sum_{j>i}^N u(r_{ij}) - P_d V \end{aligned} \quad (52)$$

by using the Euler–Lagrange equation. This barostat is useful for simulating solids; since the simulation box can change its shape, phase changes resulting in systems with a different unit cell are possible. When simulating liquids, problems might arise due to the changes in the box shape (13).

#### 4.9. The Berendsen Algorithm

In this approach proposed by Berendsen (41), pressure is controlled by adding an extra term to the equation of motion

$$\frac{dP}{dt} = \frac{P_d - P}{t_p}, \quad (53)$$

where  $P_d$  is desired pressure and  $t_p$  is a time constant. At each time step, the volume of the box is scaled by the factor

$$\chi = 1 - \beta_T \frac{\Delta t}{t_p} (P_d - P), \quad (54)$$

where  $\beta_T$  is the isothermal compressibility and  $\Delta t$  is the simulation time step and the positions are scaled by a factor  $\chi^{1/3}$ . The scaling at each time step can lead to oscillations of the pressure in more ordered systems. The Berendsen thermostat can be useful for pre-equilibrating a system; however, it is not recommended in production MD runs. It is easy to implement together with the Berendsen thermostat.

## 5. Boundary Conditions

To set up a simulation, one must choose the appropriate boundary conditions. In an *open system*, particles are able to move without limits. Often, this is not a good choice as typically the particle density is required to be constant. The other extreme is *closed boundaries*. In the case of closed boundaries, a large portion of particles will reside close to the surfaces, and that may lead to a situation in which surface effects dominate physical behavior. There are, however, cases in which closed boundaries are the most natural choice such as simulations of self-gravitating systems (62).

The most common choice is *periodic boundary conditions* (PBC). When PBC are applied, the *image of the simulation box* is repeated infinitely, see Fig. 1. The shaded box in Fig. 1 is called the *central box*. This is the actual simulation box that is being replicated to form an infinite lattice filling the whole space; Fig. 1 shows this in two dimensions. In each of the replicated boxes, the periodic images of the molecules will move exactly the same way as the original ones in the central box. When a molecule leaves the central box, one of its images will enter through the opposite face. Thus, it is enough to keep track of only the molecules in the central box. The nearest

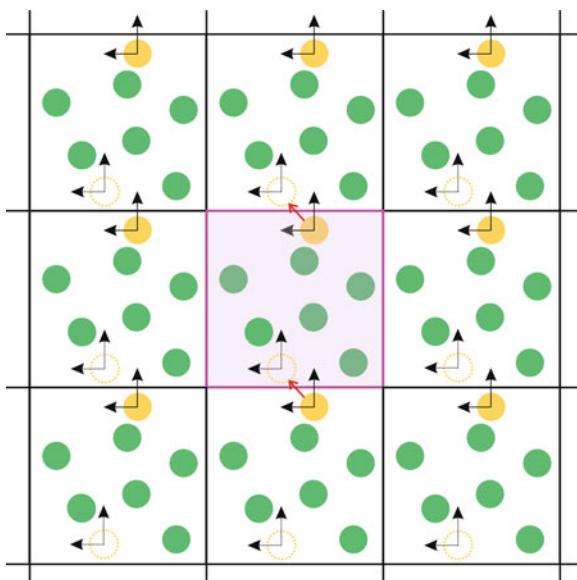


Fig. 1. Periodic boundary conditions: Schematic illustration of (*cubic*) periodic boundary conditions in a two dimensional system. A snapshot of the whole system and the movement of the *light grey particle* during the next time step are shown. Velocity components are represented by *black arrows*, the central box is shaded *grey*. Reprinted with permission (63).

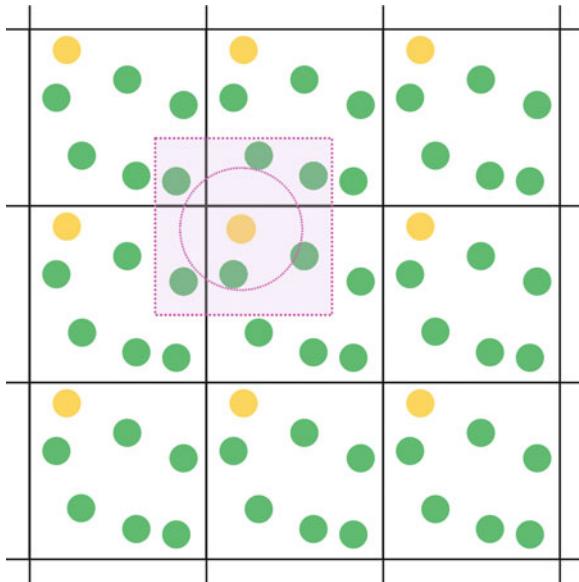


Fig. 2. Nearest image convention. Schematic illustration of the nearest image convention in a two dimensional system. Force acting on the *light grey particle* is calculated. Therefore only those images of the *grey particles* that lay inside the *shaded box* are considered. Spherical cutoff is drawn by a *dashed line*. Reprinted with permission (63).

image convention, see Fig. 2, is a technique to approximate the short-ranged forces in a system with PBC. To calculate the forces acting on a particular particle, only the particles inside the original simulation box are considered. PBC aim to make the space isotropic: PBC model bulk systems and no molecules lay on the surface since there are no surfaces present. Because of their simple geometries, cubic and rectangular boxes are the most common choices. Other space-filling geometries that can reduce system sizes, such as the rhombic dodecahedron (64) and the truncated octahedron (65, 66), are also possible.

PBC will not, however, make the system genuinely infinite. Long-ranged potentials, such as those caused by electric monopoles, may cause interactions between a particle and its own periodic images. This may lead to artificial periodicity (67, 68). Another artifact is the absence of long-wavelength fluctuations (4); they cannot be seen in these systems if the wavelength exceeds the linear length of the central box. For these reasons, the vicinity of continuous phase transitions, where macroscopic fluctuations over large distances occur (69, 70), is practically inaccessible to simulation. Additionally, angular momentum is not preserved by PBC (71). Despite these problems, it is generally thought, and in most cases verified by practice, that PBC have little or no effect on the equilibrium properties and structure of fluids if interactions are short-ranged, and we are far enough from phase transitions.

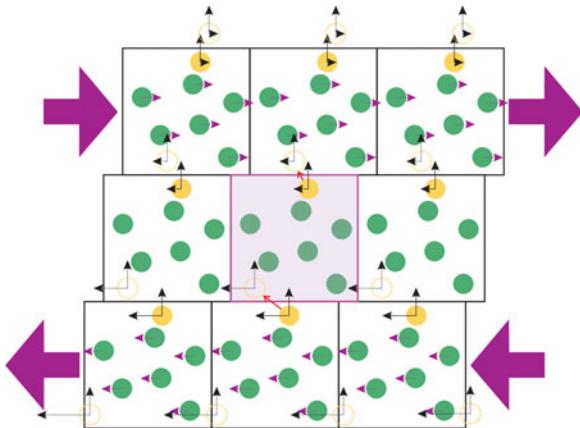


Fig. 3. Lees–Edwards boundary conditions. Schematic illustration of the Lees–Edwards boundary conditions. A snapshot of the whole system and the movement of the *light grey particle* during the next time step are shown. Components of velocity (due to shear) are depicted with (grey) arrows. The central box is shaded grey. Reprinted with permission (73).

Simulating shear conditions is an important application of MD. When a shear is applied, we must modify the boundary conditions, and the most common choice are the so-called *Lees–Edwards boundaries* (72). The Lees–Edwards boundaries are periodic, but the image cells in the plane above and below the original simulation box move in opposite directions, see Fig. 3. A particle that leaves the box is replaced on the other side with a shift, and its velocity is changed according to the shear rate. When we compare this to the illustration of the PBC in Fig. 1, we notice that the central boxes (shaded again grey) are identical. The difference comes from the periodic images which are considered to move relative to the central box according to the applied shear. All the periodic images positioned in the first row of boxes above the central box are given an extra velocity  $\gamma L$  (depicted with the grey arrows) in the shear direction. Here,  $\gamma$  is the shear strength and  $L$  the box length. Similarly, images in the box below acquire negative extra velocity  $-\gamma L$ . Images in the second row get  $+2\gamma L$  and  $-2\gamma L$ , and correspondingly, the more remote layers of boxes move proportionally faster to the central one. Hence, particles in each of the periodic image boxes will move exactly as those in the central box, when compared to their own surroundings; however, when compared to boxes in the rows below and above them, there will be a velocity difference of  $\pm\gamma L$  in the shear direction between all the particles. The directions perpendicular to shear are treated as with regular PBCs.

As a result of applying the Lees–Edwards boundary conditions with a constant shear strength  $\gamma$ , a steady linear velocity profile, such as in the solution of the Navier–Stokes equation for a bulk liquid under shear, will gradually develop into the system.

---

## 6. Saving CPU Time with Neighbor Lists

The most time-consuming part in an MD simulation is the calculation of the forces acting between particles. In a system containing  $N$  particles, the computation of interactions consists of  $N(N - 1)/2$  terms. Since interactions are very small beyond a certain distance, cut-off methods are often used for Lennard-Jones and electrostatics potentials; this means that forces are calculated only for atoms within a certain threshold distance. To reduce the computational cost of non-bonded interactions, neighbor lists or cell lists are often used. As for speeding up electrostatic interactions, we refer the reader to the chapter by Sagui.

### 6.1. Verlet Neighbor List

The so-called Verlet lists (10, 13) provide a method to reduce the time needed for force calculation. The principle is simple: Let us assume that  $r_{cut}$  is the usual short-range cutoff, then the particles for which  $r > r_{cut}$  do not contribute to the force acting on the particle under consideration. If we now introduce a second cutoff, the so-called Verlet skin  $r_v$  such that  $r_v > r_{cut}$ , we can reduce the force calculation to involve only particles for which  $r < r_v$ . For each of the particles, we construct a list of neighbors that are inside  $r_v$ . This is the aforementioned Verlet list. We update the Verlet neighbor lists only when the maximum displacement of a particle is larger than  $|r_v - r_{cut}|$ . Other criteria can also be used.  $r_v$  is a free parameter, and the optimal value for a particular system depends for example on the density. see Fig. 4 for an illustration. Details and implementation are provided in Refs. (4, 13).

### 6.2. Cell List or Cell-Linked Lists

In this method the simulation box is divided in equilateral cells of linear size larger than  $r_{cut}$ . Then, for the calculation of the interactions of each particle with the others, only the cells containing the particle and the neighboring cells need to be considered, see Fig. 5. Expensive distance calculations can be avoided, and time is saved. The cell list scales as  $\mathcal{O}(N)$  thus providing a significant speed-up. A number of variations and improvements have been introduced by various authors (74–76).

---

## 7. Constraints

Another very common way to speed up all-atom MD simulations is to treat molecules as rigid or semirigid units, with fixed bond lengths (and sometimes also fixed bond angles). The reason for this is that bond vibrations have a very high frequency which would require the simulation time step to be very short. Those bond vibrations are typically not of interest; hence, constraint algorithms are applied to eliminate those fast degrees of freedom.

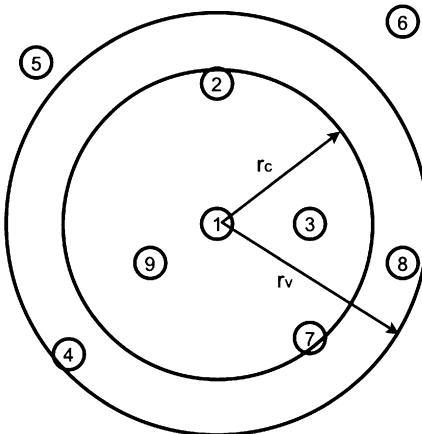


Fig. 4. Neighbor list: At this time step particle 1 interacts with particles 2,3,7 and 9, but particle 4 and 8 are also on the neighbor list of particle 1, although they do not interact with particle 1.

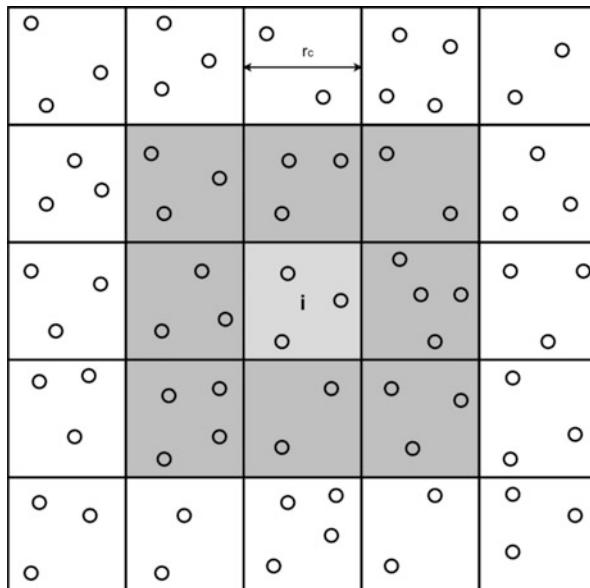


Fig. 5. Cell list: for a particle in cell  $i$  only the particles in the neighboring cells (grey) need to be considered for the force calculation.

Constraint techniques can be classified into three categories: (1) those based on Lagrange multipliers, (2) on the application of separate internal coordinate systems, and (3) on explicit constraint forces. In the following, we discuss the most common ones used in biomolecular simulations.

### 7.1. Iterative Constraint Algorithms: SHAKE and RATTLE

These two methods are based on the use of Lagrange multipliers. Constraints can be defined in the form of a set of algebraic equations, such as

$$\sigma_\alpha(\vec{r}_i, \vec{r}_j) = r_{ij}^2 - d_{ij}^2 = 0, \quad (55)$$

where  $\alpha = 1, \dots, K$  for  $K$  constraints, and  $r_i$  is the position of atom  $i$ . A new Lagrangian can be defined by adding the constraints, which should be equal to zero:

$$\mathcal{L}' = \mathcal{L} - \sum_\alpha \lambda_\alpha \sigma_\alpha(\{\vec{r}_i\}), \quad (56)$$

where  $\lambda_\alpha$  stands for the Lagrange multipliers, which need to be determined. Using the Euler–Lagrange equations

$$\frac{\partial}{\partial t} \frac{\partial \mathcal{L}'}{\partial \dot{\vec{r}}_i} = \frac{\partial \mathcal{L}'}{\partial \vec{r}_i}, \quad (57)$$

we can derive the new equations of motion for the constrained system

$$m_i \ddot{\vec{r}}_i = -\frac{\partial U}{\partial \vec{r}_i} - \sum_\alpha \lambda_\alpha \frac{\partial \sigma_\alpha}{\partial \vec{r}_i} \equiv \vec{F}_i + \sum_\alpha \vec{G}_i(\alpha). \quad (58)$$

The last term is a sum over the constraint forces.

Since  $\sigma_\alpha = 0$ , it follows that  $\partial_t \sigma_\alpha = 0$  and  $\partial_{tt} \sigma_\alpha = 0$ . The latter condition is needed to determine the Lagrange multipliers. The set of positions and momenta at a certain time step later determined by an integration algorithm is a function of the undetermined Lagrange multipliers  $\lambda_\alpha$ . Hence, the constraint algorithm needs to find the values for  $\lambda_\alpha$  which satisfy all the constraints simultaneously. SHAKE (77) is an iterative algorithm which cyclically adjusts the coordinates to satisfy each constraint one by one until a specified tolerance is reached. RATTLE (78) is a modification of the SHAKE algorithm making it suitable for the popular velocity Verlet integrator. SETTLE (79) solves the constraints for rigid water molecules analytically and is very commonly used in biomolecular simulations. It has been demonstrated (79) that for systems with few constraints, that is, water, SETTLE is of higher accuracy and three to nine times faster than RATTLE.

### 7.2. Linear Constraint Solver (LINCS)

LINCS (80) is another Lagrange multiplier-based method. It is a LINCS and is based on earlier work of Edberg, Evans, Morriss, and Baranyai (81, 82). LINCS resets the constraints themselves as opposed to the derivatives of the constraints with the aim of providing a more accurate and faster method.

The matrix formulation of Newton's equation of motion for a system with  $N$  particles reads

$$\frac{d^2\vec{r}}{dt^2} = M^{-1}\vec{F}, \quad (59)$$

where  $\vec{r}(t)$  is a  $3N$  vector containing the positions,  $\vec{F}$  is a force vector of length  $3N$ , and  $M$  is a  $3N \times 3N$  diagonal matrix containing the masses of the particles. Constraints of the form  $\sigma_\alpha(\vec{r}) = 0$ ,  $\alpha = 1, \dots, K$  are applied to the system. It has been shown (22) that the constrained system can be described by the following  $3N$  second-order differential equations,

$$\frac{d^2\vec{r}}{dt^2} = (I - TB)M^{-1}\vec{F} - T \frac{dB}{dt} \frac{d\vec{r}}{dt} \text{ with } T = M^{-1}B^T(BM^{-1}B^T)^{-1}. \quad (60)$$

$T$  is the transformation matrix of dimension  $3N \times K$  to transform from constrained to Cartesian coordinates.  $B$  is a  $K \times 3N$  matrix containing the directions of the constraints with  $B_{\alpha i} = \frac{\partial \sigma_\alpha}{\partial \vec{r}_i}$ .  $I - TB$  is the projection matrix, that sets the constraints equal to zero.  $BM^{-1}\vec{F}$  is a vector of dimension  $K$  containing the second derivatives of the bond lengths in direction of the bonds. Since the projection of the new bond onto the old directions of the bond is set to the desired length, the resulting bond is too long. Therefore, the positions need to be corrected. The part that uses most of the CPU time is the inversion of the constraint coupling matrix  $BM^{-1}B^T$ . This has to be done every time step, since  $B$  depends on the positions. To simplify this inversion, a  $K \times K$ , matrix  $S$  is introduced,

$$S = \text{Diag}\left(\sqrt{\frac{1}{m_{l_1}} + \frac{1}{m_{l_2}}, \dots, \frac{1}{m_{K_1}} + \frac{1}{m_{K_2}}}\right), \quad (61)$$

that is,  $S$  is the inverse square root of the diagonal of  $BM^{-1}B^T$  and  $(BM^{-1}B^T)^{-1} = SS^{-1}(BM^{-1}B^T)^{-1}S^{-1}S = S(I - A)^{-1}S$ .  $A$  is a symmetric sparse matrix with zeros on the diagonal. Therefore, to find the inverse, we can use  $(I - A)^{-1} = I + A + A^2 + \dots$  if all eigenvalues of  $A$  are smaller than one. This is true for molecules with only bond constraints. However, if the angles are constrained as well, there might be eigenvalues larger than one. In this case, a different method to invert the constraint coupling matrix has to be used. The question arises of where to truncate the series expansion. For molecules with bond angles near  $90^\circ$ , such as proteins, the error is negligible if the terms up to the 4th order are included. The fact that the inverse of the constraint coupling matrix is approximated by series expansion makes this algorithm easy to parallelize, and hence, it has become very popular. It has been shown that LINCS performs three to four times faster than

SHAKE (80). LINCS can only be used for bond constraints and isolated angle constraints (e.g., proton angle in OH); for other angle constraints, different constraint methods, for example, SHAKE or RATTLE, need to be used.

---

## 8. Software for Biomolecular Simulations

In the following, we list some of the popular software packages for classical MD:

- **LAMMPS** (83) (Large-scale Atomic/Molecular Massively Parallel Simulator) is an open source code that can be used to model soft materials (e.g., biomolecules or polymers), solid-state systems (e.g., metals or semiconductors) as well as coarse-grained systems. The code, written in C++, is easy to modify and extend, and it runs efficiently on multiple processors. Web: <http://lammps.sandia.gov>.
- **NAMD** (84) (NAnoscale Molecular Dynamics), another very popular open source code. It is designed for biomolecular simulations, and it performs particularly well in large-scale parallel simulations. Web: [www.ks.uiuc.edu/Research/namd](http://www.ks.uiuc.edu/Research/namd).
- **GROMACS** (Groningen Machine for Chemical Simulations) (85–88) is a very popular open source MD code for biomolecular systems. Web: [www.gromacs.org](http://www.gromacs.org).
- **Amber** (89, 90) and **CHARMM** (Chemistry at HARvard Macromolecular Mechanics) (24, 91) are licensed packages for MD simulations of biomolecular simulations. Web: <http://ambermd.org> and [www.charmm.org](http://www.charmm.org).

On a single processor, GROMACS performs two to three times faster than Amber and CHARMM (86). Amber and NAMD perform comparably on up to 64 processors. If the number of processors is increased, NAMD performs better than Amber. GROMACS seems to reach its maximum performance with 64 processors. LAMMPS scales well up to 512 processors (92).

- **ESPResSo** (Extensible Simulation Package for Research on Soft matter) (93) is an open source, parallelized code for MD simulations of coarse-grained atomistic or bead-spring models. It can also perform lattice-Boltzmann simulations. Its performance is not quite as high as the above-mentioned codes; however, it has several unique features. Web: [http://espressomd.org/wiki/Main\\_Page](http://espressomd.org/wiki/Main_Page).

- **GPU computing.** Over the past couple of years, graphics processing units (GPUs) have been employed for MD simulations (94–97). GPUs were originally developed to manipulate computer graphics, for example, in computer games; however, GPU hardware and software have become also accessible for general-purpose computing. GPUs provide an inexpensive and a very efficient alternative. Since MD simulations are computationally expensive, there have been efforts to use GPUs for these calculations. Although there are not many fully developed packages at this time, Amber, NAMD, LAMMPS, and Gromacs have at least beta versions available. Another notable MD package is **HOOMD-blue**. Web: <http://codeblue.umich.edu/hoomd-blue/>

## Acknowledgments

I would like to thank Markus Miettinen for providing some of the figures and Mikko Karttunen for reading the manuscript.

## References

1. Cramer CJ (2004) Essentials of computational chemistry: theories and models. Wiley, Chichester, England
2. Jackson SE (1998) How do small single-domain proteins fold? *Fold Des* 3:R81–R91
3. Allen MP, Tildesley DJ (2007) Computer simulation of liquids. Oxford University Press, Oxford
4. Frenkel D, Smit B (2002) Understanding molecular simulation: from algorithms to applications. Academic Press, San Diego, CA
5. Rapaport DC (2004) The art of molecular dynamics simulation. Cambridge University Press, Cambridge.
6. Zhong G, Marsden JE (1988) Lie–Poisson Hamilton–Jacobi theory and Lie–Poisson integrators. *Phys Lett A* 133:134–139
7. Miller RH (1991) A horror story about integration methods. *J Comput Phys* 93:469–476
8. Lasagni FM (1988) Canonical Runge–Kutta methods. *Z Angew Math Phys* 39:952–953
9. Candy J, Rozmus W (1991) A symplectic integration algorithm for separable Hamiltonian functions. *J Comput Phys* 92:230–256
10. Verlet L (1967) Computer experiments on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules. *Phys Rev* 159:98–103
11. Swope WC, Andersen HC, Berens PH, Wilson KR (1982) A computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules—application to small water clusters. *J Chem Phys* 76:637–649
12. Hockney RW, Goel SP, Eastwood JW (1974) Quiet high-resolution computer models of a plasma. *J Comput Phys* 14:148–158
13. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Oxford University Press, Oxford
14. Tuckerman ME, Martyna GJ (2000) Understanding modern molecular dynamics: techniques and applications. *J Phys Chem B* 104:159–178
15. Tuckerman ME, Alejandre J, Lopez-Rendon R, Jochim AL, Martyna GJ (2006) A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal–isobaric ensemble. *J Phys A: Math Gen* 39:5629–5651
16. Tuckerman M, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular-dynamics. *J Chem Phys* 97:1990–2001
17. Han G, Deng Y, Glimm J, Martyna G (2007) Error and timing analysis of multiple time-step integration methods for molecular dynamics. *Comput Phys Commun* 176:271–291

18. Trotter HF (1959) On the product of semi-groups of operators. *Proc Amer Math Soc* 10:545–551
19. Creutz M, Gocksch A (1989) Higher-order hybrid Monte-Carlo algorithms. *Phys Rev Lett* 63:9–12
20. Strang G (1968) On construction and comparison of difference schemes. *SIAM J Numer Anal* 5:506–517
21. Barth E, Schlick T (1998) Overcoming stability limitations in biomolecular dynamics. I. Combining force splitting via extrapolation with Langevin dynamics in LN. *J Chem Phys* 109:1617–1632
22. Batcho P, Schlick T (2001) Special stability advantages of position-Verlet over velocity-Verlet in multiple-time step integration. *J Chem Phys* 115:4019–4029
23. Leach AR (2001) Molecular modelling: principles and applications. Pearson Education Limited, Essex, England
24. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217
25. Nilsson L, Karplus M (1986) Empirical energy functions for energy minimization and dynamics of nucleic-acids. *J Comput Chem* 7:591–616
26. Daura X, Mark A, van Gunsteren W (1998) Parametrization of aliphatic CHn united atoms of GROMOS96 force field. *J Comput Chem* 19:535–547
27. Schuler LD, Daura X, van Gunsteren WF (2001) An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase *J Comput Chem* 22:1205–1218
28. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P (1984) A new force-field for molecular mechanical simulation of nucleic-acids and proteins. *J Am Chem Soc* 106:765–784
29. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J Am Chem Soc* 117:5179–5197
30. Jorgensen WL, Tirado-Rives J (1988) The OPLS potential functions for proteins—energy minimizations for crystals of cyclic-peptides and crambin. *J Am Chem Soc* 110:1657–1666
31. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 105:6474–6487, Symposium on molecular dynamics—the next millennium, New York, June 02–03, 2000
32. Ren PY, Ponder JW (2002) Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J Comput Chem* 23:1497–1506
33. Ren PY, Ponder JW (2003) Polarizable atomic multipole water model for molecular mechanics simulation. *J Phys Chem B* 107:5933–5947
34. Morse PM (1929) Diatomic molecules according to the wave mechanics. ii. vibrational levels. *Phys Rev* 34:57–64
35. Kremer K, Grest G (1990) Dynamics of entangled linear polymer melts—a molecular-dynamics simulation. *J Chem Phys* 92:5057–5086
36. Stillinger FH, Weber TA (1985) Computer-simulation of local order in condensed phases of silicon. *Phys Rev B* 31:5262–5271
37. Tersoff J (1988) New empirical-approach for the structure and energy of covalent systems. *Phys Rev B* 37:6991–7000
38. Lowe C (1999) An alternative approach to dissipative particle dynamics. *Europhys Lett* 47:145–151
39. Nosé S (1984) A molecular-dynamics method for simulations in the canonical ensemble. *Mol Phys* 52:255–268
40. Hoover WG (1985) Canonical dynamics—equilibrium phase-space distributions. *Phys Rev A* 31:1695–1697
41. Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR (1984) Molecular-dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
42. Schneider T, Stoll E (1978) Molecular-Dynamics study of a 3-dimensional one-component model for distortive phase-transitions. *Phys Rev B* 17:1302–1322
43. Andersen HC (1980) Molecular-dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72:2384–2393
44. Soddeemann T, Dünweg B, Kremer K (2003) Dissipative particle dynamics: a useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys Rev E* 68:046702
45. Frenkel D, Smit B (1996) Understanding molecular simulations: from algorithms to applications. Academic Press, London, UK
46. Nosé S (1986) An extension of the canonical ensemble molecular-dynamics method. *Mol Phys* 57:187–191
47. Ciccotti G, Kalibaeva G (2004) Molecular dynamics of complex systems: non-Hamiltonian,

- constrained, quantum-classical. In: Karttunen M, Vattulainen I, Lukkarinen A (ed) Novel methods in soft matter simulations. Lecture notes in physics, vol 640, pp. 150–189. International summer school on novel methods in soft matter simulations (SOFTSIMU 20020), Helsinki, Finland, May 31–June 06, 2002
48. Tuckerman ME, Mundy CJ, Martyna GJ (1999) On the classical statistical mechanics of non-Hamiltonian systems. *Europhys Lett* 45:149–155
  49. Tuckerman ME, Liu Y, Ciccotti G, Martyna GJ (2001) Non-Hamiltonian molecular dynamics: generalizing Hamiltonian phase space principles to non-Hamiltonian systems. *J Chem Phys* 115:1678–1702
  50. Martyna GJ, Klein ML, Tuckerman M (1992) Nosé–Hoover chains—the canonical ensemble via continuous dynamics. *J Chem Phys* 97:2635–2643
  51. Tuckerman ME, Berne BJ, Martyna GJ, Klein ML (1993) Efficient molecular-dynamics and hybrid Monte-Carlo algorithms for path-integrals. *J Chem Phys* 99:2796–2808
  52. Harvey SC, Tan RKZ, Cheatham T (1998) The flying ice cube: velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J Comput Chem* 19:726–740
  53. Leyssale JM, Vignoles GL (2008) Molecular dynamics evidences of the full graphitization of a nanodiamond annealed at 1500 K. *Chem Phys Lett* 454:299–304
  54. Evans DJ, Morriss GP (1990) Statistical mechanics of nonequilibrium liquids. Academic Press, London, UK
  55. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126:014101
  56. Kampen NGV (1981) Stochastic processes in physics and chemistry. North-Holland, Amsterdam
  57. Nikunen P, Karttunen M, Vattulainen I (2003) How would you integrate the equations of motion in dissipative particle dynamics simulations? *Comput Phys Commun* 153:407–423
  58. Müller M, Katsov K, Schick M (2006) Biological and synthetic membranes: What can be learned from a coarse-grained description? *Phys Rep* 434:113–176
  59. Nosé S, Klein ML (1983) Constant pressure molecular-dynamics for molecular-systems. *Mol Phys* 50:1055–1076
  60. Parrinello M, Rahman A (1981) Polymorphic transitions in single-crystals—a new molecular-dynamics method. *J Appl Phys* 52:7182–7190
  61. Parrinello M, Rahman A, Vashishta P (1983) Structural transitions in superionic conductors. *Phys Rev Lett* 50:1073–1076
  62. Ispolatov I, Karttunen M (2003) Collapses and explosions in self-gravitating systems. *Phys Rev E* 68:036117
  63. Miettinen MS (2010) Computational modeling of cationic lipid bilayers in saline solutions. Ph.D. thesis, Aalto University School of Science and Technology, Finland
  64. Wang SS, Krumhans JA (1972) Superposition assumption. II. High-density fluid argon. *J Chem Phys* 56:4287–4290
  65. Adams DJ (1979) Computer-simulation of ionic systems—distorting effects of the boundary-conditions. *Chem Phys Lett* 62:329–332
  66. Adams, D. J. (1980) The problem of the long-range forces in the computer simulation of condensed media. In: Ceperley, D. M. (ed) NRCC Workshop Proc., Berkeley 9, p. 13
  67. Mandell MJ (1976) Properties of a periodic fluid. *J Stat Phys* 15: 299–305
  68. Impey RW, Madden PA, Tildesley DJ (1981) On the calculation of the orientational correlation parameter G2. *Mol Phys* 44:1319–1334
  69. Luckhurst GR, Simpson P (1982) Computer-simulation studies of anisotropic systems. VIII. The Lebwohl–Lasher model of nematogens revisited. *Mol Phys* 47:251–265
  70. Mouritsen OG, Berlinsky AJ (1982) Fluctuation-induced 1st-order phase-transition in an anisotropic planar model of N2 on graphite. *Phys Rev Lett* 48:181–184
  71. Lenstra D, Mandel L (1982) Angular-momentum of the quantized electromagnetic-field with periodic boundary-conditions. *Phys Rev A* 26:3428–3437
  72. Lees AW, Edwards SF (1972) Computer study of transport processes under extreme conditions. *J Phys Part C Solids* 5:1921–1928
  73. Miettinen MS (2004) From molten globule to swollen coil: simulations of a lone polymer chain in an explicit solvent. Master's thesis, Helsinki University of Technology, Finland
  74. Mattson W, Rice BM (1999) Near-neighbor calculations using a modified cell-linked list method. *Comput Phys Commun* 119:135–148
  75. Heinz TN, Hünenberger PH (2004) A fast pairlist-construction algorithm for molecular simulations under periodic boundary conditions *J Comput Chem* 25:1474–1486
  76. Gonnet P (2007) A simple algorithm to accelerate the computation of non-bonded interactions in cell-based molecular dynamics simulations. *J Comput Chem* 28:570–573
  77. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of cartesian equations of motion of a system with

- constraints—molecular-dynamics of N-alkanes. *J Comput Phys* 23:327–341
78. Andersen HC (1983) RATTLE—a velocity version of the SHAKE algorithm for molecular-dynamics calculations. *J Comput Phys* 52:24–34
79. Miyamoto S, Kollman PA (1992) SETTLE—an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* 13: 952–962
80. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 18:1463–1472
81. Edberg R, Evans DJ, Morrise GP (1986) Constrained molecular-dynamics-simulations of liquid alkanes with a new algorithm. *J Chem Phys* 84:6933–6939
82. Baranyai A, Evans DJ (1990) New algorithm for constrained molecular-dynamics simulation of liquid benzene and naphthalene. *Mol Phys* 70:53–63
83. Plimpton S (1995) Fast parallel algorithms for short-range molecular-dynamics. *J Comput Phys* 117:1–19
84. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
85. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS—a message-passing parallel molecular-dynamics implementation. *Comput Phys Commun* 91:43–56
86. Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model* 7:306–317
87. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26:1701–1718
88. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447
89. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. *Comput Phys Commun* 91:1–41
90. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
91. Brooks BR, Brooks III CL, Mackerell Jr. AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30:1545–1614
92. Hein JI, Reid F, Smith L, Bush I, Guest M, Sherwood P (2005) On the performance of molecular dynamics applications on current high-end systems. *Philos Trans Roy Soc A* 363:1987–1998
93. Limbach HJ, Arnold A, Mann BA, Holm C (2006) ESPResSo—an extensible simulation package for research on soft matter systems. *Comput Phys Comp* 174:704–727
94. Stone JE, Phillips JC, Freddolino PL, Hardy DJ, Trabuco LG, Schulten K (2007) Accelerating molecular modeling applications with graphics processors. *J Comput Chem* 28:2618–2640
95. Van Meel JA, Arnold A, Frenkel D, Zwart SFP, Bellemann RG (2008) Harvesting graphics power for MD simulations. *Mol Simulat* 34:259–266
96. Hardy DJ, Stone JE, Schulten K (2009) Multi-level summation of electrostatic potentials using graphics processing units. *Parallel Comput* 35:164–177
97. Anderson JA, Lorenz CD, Travesset A (2008) General purpose molecular dynamics simulations fully implemented on graphics processing units. *J Comput Phys* 227:5342–5359

# Chapter 7

## Enhanced Sampling Algorithms

Ayori Mitsutake, Yoshiharu Mori, and Yuko Okamoto

### Abstract

In biomolecular systems (especially all-atom models) with many degrees of freedom such as proteins and nucleic acids, there exist an astronomically large number of local-minimum-energy states. Conventional simulations in the canonical ensemble are of little use, because they tend to get trapped in states of these energy local minima. Enhanced conformational sampling techniques are thus in great demand. A simulation in generalized ensemble performs a random walk in potential energy space and can overcome this difficulty. From only one simulation run, one can obtain canonical-ensemble averages of physical quantities as functions of temperature by the single-histogram and/or multiple-histogram reweighting techniques. In this article we review uses of the generalized-ensemble algorithms in biomolecular systems. Three well-known methods, namely, multicanonical algorithm, simulated tempering, and replica-exchange method, are described first. Both Monte Carlo and molecular dynamics versions of the algorithms are given. We then present various extensions of these three generalized-ensemble algorithms. The effectiveness of the methods is tested with short peptide and protein systems.

**Key words:** Monte Carlo, Molecular dynamics, Generalized-ensemble algorithm, Replica-exchange method, Simulated tempering, Multicanonical algorithm

---

### 1. Introduction

Conventional Monte Carlo (MC) and molecular dynamics (MD) simulations of biomolecules are greatly hampered by the multiple-minima problem. The canonical fixed-temperature simulations at low temperatures tend to get trapped in a few of a huge number of local-minimum-energy states which are separated by high energy barriers. One way to overcome this multiple-minima problem is to perform a simulated annealing (SA) simulation (1), and it has been widely used in biomolecular systems (see, e.g., Refs. (2–8) for earlier applications). The SA simulation mimics the crystal-making process, and temperature is lowered very slowly from a sufficiently high temperature to a low one during the SA simulation. The Boltzmann weight factor is dynamically changed, and so the

thermal equilibrium is continuously broken. Hence, although the global-minimum potential energy or the value close to it may be found, accurate thermodynamic averages for fixed temperatures cannot be obtained.

A class of simulation methods, which are referred to as the *generalized-ensemble algorithms*, overcome both above difficulties, namely, the multiple-minima problem and inaccurate thermodynamic averages (for reviews see, e.g., Refs. (9–16)). In the generalized-ensemble algorithm, each state is weighted by an artificial, non-Boltzmann probability weight factor so that a random walk in potential energy space may be realized. The random walk allows the simulation to escape from any energy barrier and to sample much wider conformational space than by conventional methods. Unlike SA simulations, the weight factors are fixed during the simulations so that the eventual reach to the thermal equilibrium is guaranteed. From a single simulation run, one can obtain accurate ensemble averages as functions of temperature by the single-histogram (17, 18) and/or multiple-histogram (19, 20) reweighting techniques (an extension of the multiple-histogram method is also referred to as the *weighted histogram analysis method* (WHAM) (20)).

One of the most well-known generalized-ensemble algorithms is perhaps the *multicanonical algorithm* (MUCA) (21, 22) (for reviews see, e.g., Refs. (23, 24)). The method is also referred to as *entropic sampling* (25–27) and *adaptive umbrella sampling* (28) of the potential energy (29). MUCA can also be considered as a sophisticated, ideal realization of a class of algorithms called *umbrella sampling* (30). Also closely related methods are *Wang-Landau method* (31, 32), which is also referred to as *density of states Monte Carlo* (33), and *meta dynamics* (34) (see also Ref. (35)). MUCA and its generalizations have been applied to spin systems (see, e.g., Refs. (36–42)). MUCA was also introduced to the molecular simulation field (43). Since then, MUCA and its generalizations have been extensively used in many applications in protein and other biomolecular systems (44–78). Molecular dynamics version of MUCA has also been developed (29, 50, 54) (see also Refs. (50, 79) for the Langevin dynamics version). MUCA has been extended so that flat distributions in other variables instead of potential energy may be obtained (see, e.g., Refs. (37, 38, 49, 55, 57, 70, 76)). This can be considered as a special case of the multidimensional (or, multivariable) extensions of MUCA, where a multidimensional random walk in potential energy space and in other variable space is realized (see, e.g., Refs. (49, 55, 56, 72, 78)). In this article, we just present one of such methods, namely, the *multibaric-multithermal algorithm* (MUBATH) where a two-dimensional random walk in both potential energy space and volume space is realized (72–75).

While a simulation in multicanonical ensemble performs a free 1D random walk in potential energy space, that in *simulated*

*tempering* (ST) (80, 81) (the method is also referred to as the *method of expanded ensemble* (80)) performs a free random walk in temperature space (for a review, see, e.g., Ref. (82)). This random walk, in turn, induces a random walk in potential energy space and allows the simulation to escape from states of energy local minima. ST and its generalizations have also been applied to chemical physics field and biomolecular systems (51, 52, 83–91).

MUCA and ST are powerful, but the probability weight factors are not a priori known and have to be determined by iterations of short trial simulations. This process can be nontrivial and very tedious for complex systems with many degrees of freedom.

In the *replica-exchange method* (REM) (92–94), the difficulty of weight factor determination is greatly alleviated. (A closely related method was independently developed in Ref. (95). Similar methods in which the same equations are used but emphasis is laid on optimizations have been developed (96, 97). REM is also referred to as *multiple Markov chain method* (98) and *parallel tempering* (82). Details of literature about REM and related algorithms can be found in recent reviews (10, 15, 99).) In this method, a number of noninteracting copies (or, replicas) of the original system at different temperatures are simulated independently and simultaneously by the conventional MC or MD method. Every few steps, pairs of replicas are exchanged with a specified transition probability. The weight factor is just the product of Boltzmann factors, and so it is essentially known.

REM has already been used in many applications in protein systems (100–115). Other molecular simulation fields have also been studied by this method in various ensembles (116–120). Moreover, REM was introduced to the quantum chemistry field (121). The details of molecular dynamics algorithm for REM, which is referred to as the *replica-exchange molecular dynamics* (REMD) have been worked out in Ref. (101), and this led to a wide application of REM in the protein folding and related problems (see, e.g., Refs. (122–144)).

However, REM also has a computational difficulty: As the number of degrees of freedom of the system increases, the required number of replicas also greatly increases, whereas only a single replica is simulated in MUCA and ST. This demands a lot of computer power for complex systems. Our solution to this problem is: Use REM for the weight factor determinations of MUCA, which is much simpler than previous iterative methods of weight determinations, and then perform a long MUCA production run. The method is referred to as the *replica-exchange multicanonical algorithm* (REMUCA) (105, 109, 110). In REMUCA, a short replica-exchange simulation is performed, and the multicanonical weight factor is determined by the multiple-histogram reweighting techniques (19, 20). Another example of a combination of REM and ST is the *replica-exchange simulated tempering* (REST) (86). In REST,

a short replica-exchange simulation is performed, and the simulated tempering weight factor is determined by the multiple-histogram reweighting techniques (19, 20).

We have introduced two further extensions of REM, which we refer to as *multicanonical replica-exchange method* (MUCAREM) (105, 109, 110) (see also Refs. (133, 134)) and *simulated tempering replica-exchange method* (STREM) (87) (see also Ref. (135) for a similar idea). In MUCAREM, a replica-exchange simulation is performed with a small number of replicas each in multicanonical ensemble of different energy ranges. In STREM, on the other hand, a replica-exchange simulation is performed with a small number of replicas in “simulated tempering” ensemble of different temperature ranges.

Finally, one is naturally led to a multidimensional (or, multivariable) extension of REM, which we refer to as the *multidimensional replica-exchange method* (MREM) (103) (see also Refs. (117, 145)). (The method is also referred to as *generalized parallel sampling* (146), *Hamiltonian replica-exchange method* (108), and *Model Hop ping* (147).) Some other special cases of MREM can be found in, e.g., Refs. (132, 148–153). Another special realization of MREM is *replica-exchange umbrella sampling* (REUS) (103), and it is particularly useful in free energy calculations (see also Ref. (104) for a similar idea). In this article, we just present one of such methods, namely, the REM in the isobaric-isothermal ensemble, where not only temperature values but also pressure values are exchanged in the replica-exchange processes (11, 118, 120, 128, 129). (The results of the first such application of the two-dimensional replica-exchange simulations in the isobaric-isothermal ensemble were presented in Ref. (11).)

In this article, we describe the generalized-ensemble algorithms mentioned above. Namely, we first review the three familiar methods: REM, ST, and MUCA. We then describe various extensions of these methods (103, 154–157). Examples of the results by some of these algorithms are then presented.

## 2. Generalized-Ensemble Algorithms

### 2.1. Replica-Exchange Method

Let us consider a system of  $N$  atoms of mass  $m_k$  ( $k = 1, \dots, N$ ) with their coordinate vectors and momentum vectors denoted by  $q \equiv \{q_1, \dots, q_N\}$  and  $p \equiv \{p_1, \dots, p_N\}$ , respectively. The Hamiltonian  $H(q, p)$  of the system is the sum of the kinetic energy  $K(p)$  and the potential energy  $E(q)$ :

$$H(q, p) = K(p) + E(q), \quad (1)$$

where

$$K(p) = \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m_k}. \quad (2)$$

In the canonical ensemble at temperature  $T$ , each state  $x = (q, p)$  with the Hamiltonian  $H(q, p)$  is weighted by the Boltzmann factor:

$$W_B(x; T) = \exp(-\beta H(q, p)), \quad (3)$$

where the inverse temperature  $\beta$  is defined by  $\beta = 1/k_B T$  ( $k_B$  is the Boltzmann constant). The average kinetic energy at temperature  $T$  is then given by

$$\langle K(p) \rangle_T = \left\langle \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m_k} \right\rangle_T = \frac{3}{2} N k_B T. \quad (4)$$

Because the coordinates  $q$  and momenta  $p$  are decoupled in Eq. 1, we can suppress the kinetic energy part and can write the Boltzmann factor as

$$W_B(x; T) = W_B(E; T) = \exp(-\beta E). \quad (5)$$

The canonical probability distribution of potential energy  $P_{\text{NVT}}(E; T)$  is then given by the product of the density of states  $n(E)$  and the Boltzmann weight factor  $W_B(E; T)$ :

$$P_{\text{NVT}}(E; T) \propto n(E) W_B(E; T). \quad (6)$$

Because  $n(E)$  is a rapidly increasing function and the Boltzmann factor decreases exponentially, the canonical ensemble yields a bell-shaped distribution of potential energy which has a maximum around the average energy at temperature  $T$ . The conventional MC or MD simulations at constant temperature are expected to yield  $P_{\text{NVT}}(E; T)$ . An MC simulation based on the Metropolis algorithm (158) is performed with the following transition probability from a state  $x$  of potential energy  $E$  to a state  $x'$  of potential energy  $E'$ :

$$w(x \rightarrow x') = \min\left(1, \frac{W_B(E'; T)}{W_B(E; T)}\right) = \min(1, \exp(-\beta\Delta E)), \quad (7)$$

where

$$\Delta E = E' - E. \quad (8)$$

A MD simulation, on the other hand, is based on the following Newton equations of motion:

$$\dot{\mathbf{q}}_k = \frac{\mathbf{p}_k}{m_k}, \quad (9)$$

$$\dot{\mathbf{p}}_k = -\frac{\partial E}{\partial \mathbf{q}_k} = \mathbf{f}_k, \quad (10)$$

where  $f_k$  is the force acting on the  $k$ th atom ( $k = 1, \dots, N$ ). This set of equations actually yield the microcanonical ensemble, however, and we have to add a thermostat in order to obtain the canonical ensemble at temperature  $T$ . Here, we just follow Nosé's prescription (159, 160), and we have

$$\dot{q}_k = \frac{\dot{p}_k}{m_k}, \quad (11)$$

$$\dot{p}_k = -\frac{\partial E}{\partial q_k} - \frac{\dot{s}}{s} p_k = f_k - \frac{\dot{s}}{s} p_k, \quad (12)$$

$$\dot{s} = s \frac{P_s}{Q}, \quad (13)$$

$$\dot{P}_s = \sum_{k=1}^N \frac{\dot{p}_k^2}{m_k} - 3Nk_B T = 3Nk_B(T(t) - T), \quad (14)$$

where  $s$  is Nosé's scaling parameter,  $P_s$  is its conjugate momentum,  $Q$  is its mass, and the “instantaneous temperature”  $T(t)$  is defined by

$$T(t) = \frac{1}{3Nk_B} \sum_{k=1}^N \frac{\dot{p}_k(t)^2}{m_k}. \quad (15)$$

However, in practice, it is very difficult to obtain accurate canonical distributions of complex systems at low temperatures by conventional MC or MD simulation methods. This is because simulations at low temperatures tend to get trapped in one or a few of local-minimum-energy states. This difficulty is overcome by, for instance, the generalized-ensemble algorithms, which greatly enhance conformational sampling.

The REM is one of effective generalized-ensemble algorithms. The system for REM consists of  $M$  noninteracting copies (or, replicas) of the original system in the canonical ensemble at  $M$  different temperatures  $T_m$  ( $m = 1, \dots, M$ ). We arrange the replicas so that there is always exactly one replica at each temperature. Then there exists a one-to-one correspondence between replicas and temperatures; the label  $i$  ( $i = 1, \dots, M$ ) for replicas is a permutation of the label  $m$  ( $m = 1, \dots, M$ ) for temperatures, and vice versa:

$$\begin{cases} i = i(m) \equiv f(m), \\ m = m(i) \equiv f^{-1}(i), \end{cases} \quad (16)$$

where  $f(m)$  is a permutation function of  $m$  and  $f^{-1}(i)$  is its inverse.

Let  $X = \{x_1^{[i(1)]}, \dots, x_M^{[i(M)]}\} = \{x_{m(1)}^{[1]}, \dots, x_{m(M)}^{[M]}\}$  stand for a “state” in this generalized ensemble. Each “substate”  $x_m^{[i]}$  is specified by the coordinates  $q^{[i]}$  and momenta  $p^{[i]}$  of  $N$  atoms in replica  $i$  at temperature  $T_m$ :

$$x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m. \quad (17)$$

Because the replicas are noninteracting, the weight factor for the state  $X$  in this generalized ensemble is given by the product of Boltzmann factors for each replica (or at each temperature):

$$\begin{aligned} W_{\text{REM}}(X) &= \prod_{i=1}^M \exp \left\{ -\beta_{m(i)} H(q^{[i]}, p^{[i]}) \right\} \\ &= \prod_{m=1}^M \exp \left\{ -\beta_m H(q^{[i(m)]}, p^{[i(m)]}) \right\}, \\ &= \exp \left\{ -\sum_{i=1}^M \beta_{m(i)} H(q^{[i]}, p^{[i]}) \right\} \\ &= \exp \left\{ -\sum_{m=1}^M \beta_m H(q^{[i(m)]}, p^{[i(m)]}) \right\}, \end{aligned} \quad (18)$$

where  $i(m)$  and  $m(i)$  are the permutation functions in Eq. 16.

We now consider exchanging a pair of replicas in this ensemble. Suppose we exchange replicas  $i$  and  $j$  which are at temperatures  $T_m$  and  $T_n$ , respectively:

$$X = \left\{ \dots, x_m^{[i]}, \dots, x_n^{[j]}, \dots \right\} \rightarrow X' = \left\{ \dots, x_m^{[j]}, \dots, x_n^{[i]}, \dots \right\}. \quad (19)$$

Here,  $i$ ,  $j$ ,  $m$ , and  $n$  are related by the permutation functions in Eq. 16, and the exchange of replicas introduces a new permutation function  $f'$ :

$$\begin{cases} i = f(m) \rightarrow j = f'(m), \\ j = f(n) \rightarrow i = f'(n). \end{cases} \quad (20)$$

The exchange of replicas can be written in more detail as

$$\begin{cases} x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m \rightarrow x_m^{[j]'} \equiv (q^{[j]}, p^{[j]})_m, \\ x_n^{[j]} \equiv (q^{[j]}, p^{[j]})_n \rightarrow x_n^{[i]'} \equiv (q^{[i]}, p^{[i]})_n, \end{cases} \quad (21)$$

where the definitions for  $p^{[i']}$  and  $p^{[j']}$  will be given below. We remark that this process is equivalent to exchanging a pair of temperatures  $T_m$  and  $T_n$  for the corresponding replicas  $i$  and  $j$  as follows:

$$\begin{cases} x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m \rightarrow x_n^{[i]'} \equiv (q^{[i]}, p^{[i]})_n, \\ x_n^{[j]} \equiv (q^{[j]}, p^{[j]})_n \rightarrow x_m^{[j]'} \equiv (q^{[j]}, p^{[j]})_m. \end{cases} \quad (22)$$

In the original implementation of the REM (92–94), Monte Carlo algorithm was used, and only the coordinates  $q$  (and the

potential energy function  $E(q)$ ) had to be taken into account. In molecular dynamics algorithm, on the other hand, we also have to deal with the momenta  $p$ . We proposed the following momentum assignment in Eq. 21 (and in Eq. 22) (101):

$$\begin{cases} p^{[i]\prime} \equiv \sqrt{\frac{T_n}{T_m}} p^{[i]}, \\ p^{[j]\prime} \equiv \sqrt{\frac{T_m}{T_n}} p^{[j]}, \end{cases} \quad (23)$$

which we believe is the simplest and the most natural. This assignment means that we just rescale uniformly the velocities of all the atoms in the replicas by the square root of the ratio of the two temperatures so that the temperature condition in Eq. 4 may be satisfied immediately after replica exchange is accepted.

The transition probability of this replica-exchange process is given by the usual Metropolis criterion:

$$w(X \rightarrow X') \equiv w(x_m^{[i]} | x_n^{[j]}) = \min\left(1, \frac{W_{\text{REM}}(X')}{W_{\text{REM}}(X)}\right) = \min(1, \exp(-\Delta)), \quad (24)$$

where in the second expression (i.e.,  $w(x_m^{[i]} | x_n^{[j]})$ ) we explicitly wrote the pair of replicas (and temperatures) to be exchanged. From Eqs. 1, 2, 18, and 23, we have

$$\begin{aligned} & \frac{W_{\text{REM}}(X')}{W_{\text{REM}}(X)} \\ &= \exp\left\{-\beta_m\left[K(p^{[j]\prime}) + E(q^{[j]})\right] - \beta_n\left[K(p^{[i]\prime}) + E(q^{[i]})\right]\right. \\ &\quad \left.+ \beta_m\left[K(p^{[i]}) + E(q^{[i]})\right] + \beta_n\left[K(p^{[j]}) + E(q^{[j]})\right]\right\}, \\ &= \exp\left\{-\beta_m \frac{T_m}{T_n} K(p^{[j]}) - \beta_n \frac{T_n}{T_m} K(p^{[i]}) + \beta_m K(p^{[i]}) + \beta_n K(p^{[j]})\right. \\ &\quad \left.- \beta_m\left[E(q^{[j]}) - E(q^{[i]})\right] - \beta_n\left[E(q^{[i]}) - E(q^{[j]})\right]\right\}. \end{aligned} \quad (25)$$

Because the kinetic energy terms in this equation all cancel out,  $\Delta$  in Eq. 24 becomes

$$\Delta = \beta_m\left(E(q^{[j]}) - E(q^{[i]})\right) - \beta_n\left(E(q^{[j]}) - E(q^{[i]})\right), \quad (26)$$

$$= (\beta_m - \beta_n)\left(E(q^{[j]}) - E(q^{[i]})\right). \quad (27)$$

Here,  $i$ ,  $j$ ,  $m$ , and  $n$  are related by the permutation functions in Eq. 16 before the replica exchange:

$$\begin{cases} i = f(m), \\ j = f(n). \end{cases} \quad (28)$$

Note that after introducing the momentum rescaling in Eq. 23, we have the same Metropolis criterion for replica exchanges, i.e., Eqs. 24 and 27, for both MC and MD versions.

Without loss of generality, we can assume  $T_1 < T_2 < \dots < T_M$ . The lowest temperature  $T_1$  should be sufficiently low so that the simulation can explore the global-minimum-energy region, and the highest temperature  $T_M$  should be sufficiently high so that no trapping in an energy-local-minimum state occurs. A simulation of the REM is then realized by alternately performing the following two steps:

1. Each replica in canonical ensemble of the fixed temperature is simulated *simultaneously* and *independently* for a certain MC or MD steps.
2. A pair of replicas at neighboring temperatures, say  $x_m^{[i]}$  and  $x_{m+1}^{[j]}$ , are exchanged with the probability  $w(x_m^{[i]}|x_{m+1}^{[j]})$  in Eq. 24.

Note that in Step 2 we exchange only pairs of replicas corresponding to neighboring temperatures, because the acceptance ratio of the exchange process decreases exponentially with the difference of the two  $\beta$ 's (see Eqs. 27 and 24). Note also that whenever a replica exchange is accepted in Step 2, the permutation functions in Eq. 16 are updated. A random walk in “temperature space” is realized for each replica, which in turn induces a random walk in potential energy space. This alleviates the problem of getting trapped in states of energy local minima.

The REM simulation is particularly suitable for parallel computers. Because one can minimize the amount of information exchanged among nodes, it is best to assign each replica to each node (exchanging pairs of temperature values among nodes is much faster than exchanging coordinates and momenta). This means that we keep track of the permutation function  $m(i; t) = f^{-1}(i; t)$  in Eq. 16 as a function of MC or MD step  $t$  during the simulation. After parallel canonical MC or MD simulations for a certain steps (Step 1),  $M/2$  pairs of replicas corresponding to neighboring temperatures are simultaneously exchanged (Step 2), and the pairing is alternated between the two possible choices, i.e.,  $(T_1, T_2)$ ,  $(T_3, T_4)$ , ... and  $(T_2, T_3)$ ,  $(T_4, T_5)$ , ... .

After a long production run of a replica-exchange simulation, the canonical expectation value of a physical quantity  $A$  at temperature  $T_m$  ( $m = 1, \dots, M$ ) can be calculated by the usual arithmetic mean:

$$\langle A \rangle_{T_m} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)), \quad (29)$$

where  $x_m(k)$  ( $k = 1, \dots, n_m$ ) are the configurations obtained at temperature  $T_m$  and  $n_m$  is the total number of measurements made at  $T = T_m$ . The expectation value at any intermediate temperature  $T (= 1/k_B\beta)$  can also be obtained as follows:

$$\langle A \rangle_T = \frac{\sum_E A(E) P_{\text{NVT}}(E; T)}{\sum_E P_{\text{NVT}}(E; T)} = \frac{\sum_E A(E) n(E) \exp(-\beta E)}{\sum_E n(E) \exp(-\beta E)}. \quad (30)$$

The summation instead of integration is used in Eq. 30, because we often discretize the potential energy  $E$  with step size  $\epsilon(E = E_i; i = 1, 2, \dots)$ . Here, the explicit form of the physical quantity  $A$  should be known as a function of potential energy  $E$ . For instance,  $A(E) = E$  gives the average potential energy  $\langle E \rangle_T$  as a function of temperature, and  $A(E) = \beta^2(E - \langle E \rangle_T)^2$  gives specific heat.

The density of states  $n(E)$  in Eq. 30 is given by the multiple-histogram reweighting techniques (19, 20) as follows. Let  $N_m(E)$  and  $n_m$  be respectively the potential-energy histogram and the total number of samples obtained at temperature  $T_m = 1/k_B\beta_m$  ( $m = 1, \dots, M$ ). The best estimate of the density of states is then given by (19, 20)

$$n(E) = \frac{\sum_{m=1}^M g_m^{-1} N_m(E)}{\sum_{m=1}^M g_m^{-1} n_m \exp(f_m - \beta_m E)}, \quad (31)$$

where we have for each  $m$  ( $= 1, \dots, M$ )

$$\exp(-f_m) = \sum_E n(E) \exp(-\beta_m E). \quad (32)$$

Here,  $g_m = 1 + 2\tau_m$ , and  $\tau_m$  is the integrated autocorrelation time at temperature  $T_m$ . For many systems, the quantity  $g_m$  can safely be set to be a constant in the reweighting formulae (20), and hereafter we set  $g_m = 1$ .

Note that Eqs. 31 and 32 are solved self-consistently by iteration (19, 20) to obtain the density of states  $n(E)$  and the dimensionless Helmholtz free energy  $f_m$ . Namely, we can set all the  $f_m$  ( $m = 1, \dots, M$ ) to, e.g., zero initially. We then use Eq. 31 to obtain  $n(E)$ , which is substituted into Eq. 32 to obtain next values of  $f_m$ , and so on.

Moreover, the ensemble averages of any physical quantity  $A$  (including those that cannot be expressed as functions of potential energy) at any temperature  $T (= 1/k_B\beta)$  can now be obtained from the “trajectory” of configurations of the production run. Namely, we first obtain  $f_m$  ( $m = 1, \dots, M$ ) by solving Eqs. 31 and 32 self-consistently, and then we have (109) (see also (161))

$$\langle A \rangle_T = \frac{\sum_{m=1}^M \sum_{k=1}^{n_m} A(x_m(k)) \frac{1}{\sum_{\ell=1}^M n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))]} \exp[-\beta E(x_m(k))]}{\sum_{m=1}^M \sum_{k=1}^{n_m} \frac{1}{\sum_{\ell=1}^M n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))]} \exp[-\beta E(x_m(k))]}, \quad (33)$$

where  $x_m(k)$  ( $k = 1, \dots, n_m$ ) are the configurations obtained at temperature  $T_m$ .

Eqs. 30 and 31 or any other equations which involve summations of exponential functions often encounter with numerical difficulties such as overflows. These can be overcome by using, for instance, the following equation (23, 162): For  $C = A + B$  (with  $A > 0$  and  $B > 0$ ), we have

$$\begin{aligned} \ln C &= \ln \left[ \max(A, B) \left( 1 + \frac{\min(A, B)}{\max(A, B)} \right) \right], \\ &= \max(\ln A, \ln B) + \ln \{1 + \exp[\min(\ln A, \ln B) - \max(\ln A, \ln B)]\}. \end{aligned} \quad (34)$$

We now give more details about the momentum rescaling in Eq. 23 (163). Actually, Eq. 23 is only valid for the Langevin dynamics (164), Andersen thermostat (165), and Gaussian constraint method (166–168). The former two thermostats are based on the weight factor of Eq. 3 with Eqs. 1 and 2, and the Gaussian constraint method is based on the following weight factor:

$$W(q, p) = \delta \left( \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m_k} - \frac{gk_B T}{2} \right) \exp[-\beta E(q)]. \quad (35)$$

For Nosé's method (159, 160), which gives the equations of motion in Eqs. 11–14, the Nosé Hamiltonian is given by

$$H_{\text{Nose}} = \sum_{k=1}^N \frac{\tilde{\mathbf{p}}_k^2}{2m_k s^2} + E(q) + \frac{P_s^2}{2Q} + gk_B T \log s. \quad (36)$$

Here,  $g$  ( $= 3N$ ) is the number of degrees of freedom,  $s$  is a position variable of the thermostat,  $P_s$  is a momentum conjugate to  $s$ , and  $\tilde{\mathbf{p}}_k$  is a virtual momentum, which is related to the real momenta  $\mathbf{p}_k$  as  $\mathbf{p}_k = \tilde{\mathbf{p}}_k/s$ . The weight factor for the Nosé's method is then given by

$$W(q, \tilde{p}, s, P_s) = \delta(H_{\text{Nose}} - \mathcal{E}), \quad (37)$$

where  $\mathcal{E}$  is the initial value of  $H_{\text{Nose}}$ . Namely, in the Nosé's method, the entire system including the thermostat is in the microcanonical ensemble. Note that the mass  $Q$  of the thermostat can have different values in each replica in REMD simulations. The rescaling method for the Nosé thermostat is given by Eq. 23 and

$$P_s^{[i]\prime} = \sqrt{\frac{T_n Q_n}{T_m Q_m}} P_s^{[i]}, \quad P_s^{[j]\prime} = \sqrt{\frac{T_m Q_m}{T_n Q_n}} P_s^{[j]}, \quad (38)$$

$$\begin{aligned} s^{[i]\prime} &= s^{[i]} \exp \left[ \frac{1}{\beta k_B} \left( \frac{E(q^{[i]}) - \mathcal{E}_m}{T_m} - \frac{E(q^{[i]}) - \mathcal{E}_n}{T_n} \right) \right], \\ s^{[j]\prime} &= s^{[j]} \exp \left[ \frac{1}{\beta k_B} \left( \frac{E(q^{[j]}) - \mathcal{E}_n}{T_n} - \frac{E(q^{[j]}) - \mathcal{E}_m}{T_m} \right) \right], \end{aligned} \quad (39)$$

where  $\mathcal{E}_m$  and  $\mathcal{E}_n$  are the initial values of  $H_{\text{Nose}}$  in the simulations with  $T_m$  and  $T_n$ , respectively, before the replica exchange. Note that the real momenta have to be used in the rescaling method in Eq. 23, not the virtual momenta.

For the Nosé–Hoover thermostat (169), the states are specified by the following weight factor:

$$W(q, p, \zeta) = \exp \left[ -\beta \left( \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m_k} + E(q) + \frac{Q}{2} \zeta^2 \right) \right], \quad (40)$$

where  $\zeta$  is a velocity of the thermostat and  $Q$  is its mass parameter. The rescaling method for the Nosé–Hoover thermostat is given by Eq. 23 and

$$\zeta^{[i]\prime} = \sqrt{\frac{T_n Q_m}{T_m Q_n}} \zeta^{[i]}, \quad \zeta^{[j]\prime} = \sqrt{\frac{T_m Q_n}{T_n Q_m}} \zeta^{[j]}, \quad (41)$$

where  $Q_m$  and  $Q_n$  are the mass parameters in the replicas at temperature values  $T_m$  and  $T_n$ , respectively, before the replica exchange.

The rescaling method for the Nosé–Hoover thermostat can be generalized to the Nosé–Hoover chains (170) in a similar way. The weight factor for the Nosé–Hoover chains is given by

$$W(q, p, \zeta_1, \dots, \zeta_{\mathcal{L}}) = \exp \left[ -\beta \left( \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m_k} + E(q) + \sum_{\ell=1}^{\mathcal{L}} \frac{Q_{\ell}}{2} \zeta_{\ell}^2 \right) \right], \quad (42)$$

where  $\mathcal{L}$  is the number of thermostats,  $\zeta_{\ell}$  ( $\ell = 1, \dots, \mathcal{L}$ ) is the velocity of the  $\ell$ th thermostat, and  $Q_{\ell}$  ( $\ell = 1, \dots, \mathcal{L}$ ) is a mass parameter corresponding to the  $\ell$ th thermostat. A rescaling method for REMD with the Nosé–Hoover chains is given by Eq. 23 and the following:

$$\zeta_{\ell}^{[i]\prime} = \sqrt{\frac{T_n Q_{m,\ell}}{T_m Q_{n,\ell}}} \zeta_{\ell}^{[i]}, \quad \zeta_{\ell}^{[j]\prime} = \sqrt{\frac{T_m Q_{n,\ell}}{T_n Q_{m,\ell}}} \zeta_{\ell}^{[j]}, \quad (\ell = 1, \dots, \mathcal{L}), \quad (43)$$

where  $Q_{m,\ell}$  and  $Q_{n,\ell}$  are the mass parameters in the replicas at temperature values  $T_m$  and  $T_n$ , respectively, which correspond to the  $\ell$ th thermostat.

In the Nosé-Poincaré thermostat (171), the states are specified by  $x \equiv (q, \tilde{p}, s, P_s)$ , and the weight factor is given by

$$W(q, \tilde{p}, s, P_s) \propto \delta[s(H_{\text{Nose}} - \mathcal{E})], \quad (44)$$

where  $H_{\text{Nose}}$  is the Nosé Hamiltonian in Eq. 36 and  $\mathcal{E}$  is its initial value. A rescaling method of the Nosé-Poincaré thermostat is the same as in the Nosé's thermostat and given by Eqs. 23, 38, and 39 above.

## 2.2. Simulated Tempering

We now introduce another generalized-ensemble algorithm, the *simulated tempering* (ST) method (80, 81). In this method temperature itself becomes a dynamical variable, and both the configuration and the temperature are updated during the simulation with a weight:

$$W_{\text{ST}}(E; T) = \exp(-\beta E + \alpha(T)), \quad (45)$$

where the function  $\alpha(T)$  is chosen so that the probability distribution of temperature is flat:

$$\begin{aligned} P_{\text{ST}}(T) &= \int dE n(E) W_{\text{ST}}(E; T) \\ &= \int dE n(E) \exp(-\beta E + \alpha(T)) = \text{constant}. \end{aligned} \quad (46)$$

Hence, in simulated tempering, *temperature* is sampled uniformly. A free random walk in temperature space is realized, which in turn induces a random walk in potential energy space and allows the simulation to escape from states of energy local minima.

In the numerical work we discretize the temperature in  $M$  different values,  $T_m$  ( $m = 1, \dots, M$ ). Without loss of generality we can order the temperature so that  $T_1 < T_2 < \dots < T_M$ . The probability weight factor in Eq. 45 is now written as

$$W_{\text{ST}}(E; T_m) = \exp(-\beta_m E + \alpha_m), \quad (47)$$

where  $\alpha_m = \alpha(T_m)$  ( $m = 1, \dots, M$ ). Note that from Eqs. 46 and 47, we have

$$\exp(-\alpha_m) \propto \int dE n(E) \exp(-\beta_m E). \quad (48)$$

The parameters  $\alpha_m$  are therefore “dimensionless” Helmholtz free energy at temperature  $T_m$  (i.e., the inverse temperature  $\beta_m$  multiplied by the Helmholtz free energy).

Once the parameters  $\alpha_m$  are determined and the initial configuration and the initial temperature  $T_m$  are chosen, a simulated tempering simulation is realized by alternately performing the following two steps (80, 81):

1. A canonical MC or MD simulation at the fixed temperature  $T_m$  is carried out for a certain steps.

2. The temperature  $T_m$  is updated to the neighboring values  $T_m \pm 1$  with the configuration fixed. The transition probability of this temperature-updating process is given by the following Metropolis criterion (see Eq. 47):

$$w(T_m \rightarrow T_{m\pm 1}) = \min\left(1, \frac{W_{\text{ST}}(E; T_{m\pm 1})}{W_{\text{ST}}(E; T_m)}\right) = \min(1, \exp(-\Delta)), \quad (49)$$

where

$$\Delta = (\beta_{m\pm 1} - \beta_m)E - (\alpha_{m\pm 1} - \alpha_m). \quad (50)$$

Note that in Step 2 we update the temperature only to the neighboring temperatures in order to secure sufficiently large acceptance ratio of temperature updates.

We remark that when MD simulations are performed in Step 1 above, we also have to deal with the momenta  $p$ , and the kinetic energy term should be included in the weight factor. When temperature  $T_{m_0\pm 1}$  is accepted for  $T$ -update in Step 2, we rescale the momenta in the same way as in REMD (101, 155, 157):

$$\mathbf{p}'_k = \sqrt{\frac{T_{m_0\pm 1}}{T_{m_0}}} \mathbf{p}_k. \quad (51)$$

The kinetic energy terms then cancel out in Eq. 50, and we can use the same  $\Delta$  in the Metropolis criterion in Step 2 for both MC and MD simulations. Similar momentum scaling formulae given above should also be introduced for various other thermostats (163).

The simulated tempering parameters  $\alpha_m = \alpha(T_m)$  ( $m = 1, \dots, M$ ) can be determined by iterations of short trial simulations (see, e.g., Refs. (52, 82, 84) for details). This process can be nontrivial and very tedious for complex systems.

After the optimal simulated tempering weight factor is determined, one performs a long simulated tempering run once. The canonical expectation value of a physical quantity  $A$  at temperature  $T_m$  ( $m = 1, \dots, M$ ) can be calculated by the usual arithmetic mean from Eq. 29. The expectation value at any intermediate temperature can also be obtained from Eq. 30, where the density of states is given by the multiple-histogram reweighting techniques (19, 20). Namely, let  $N_m(E)$  and  $n_m$  be respectively the potential-energy histogram and the total number of samples obtained at temperature  $T_m = 1/k_B\beta_m$  ( $m = 1, \dots, M$ ). The best estimate of the density of states is then given by solving Eqs. 31 and 32 self-consistently.

Moreover, the ensemble averages of any physical quantity  $A$  (including those that cannot be expressed as functions of potential energy) at any temperature  $T (= 1/k_B\beta)$  can now be obtained from Eq. 33.

### 2.3. Multicanonical Algorithm

The third generalized-ensemble algorithm that we present is the *MUCA* (21, 22). In the multicanonical ensemble, each state is weighted by a non-Boltzmann weight factor  $W_{\text{MUCA}}(E)$  (which we refer to as the *multicanonical weight factor*) so that a uniform potential energy distribution  $P_{\text{MUCA}}(E)$  is obtained:

$$P_{\text{MUCA}}(E) \propto n(E) W_{\text{MUCA}}(E) \equiv \text{constant}. \quad (52)$$

The flat distribution implies that a free one-dimensional random walk in the potential energy space is realized in this ensemble. This allows the simulation to escape from any local-minimum-energy states and to sample the configurational space much more widely than the conventional canonical MC or MD methods.

The definition in Eq. 52 implies that the multicanonical weight factor is inversely proportional to the density of states, and we can write it as follows:

$$W_{\text{MUCA}}(E) \equiv \exp[-\beta_a E_{\text{MUCA}}(E; T_a)] = \frac{1}{n(E)}, \quad (53)$$

where we have chosen an arbitrary reference temperature,  $T_a = 1/k_B \beta_a$ , and the “multicanonical potential energy” is defined by

$$E_{\text{MUCA}}(E; T_a) \equiv k_B T_a \ln n(E) = T_a S(E). \quad (54)$$

Here,  $S(E)$  is the entropy in the microcanonical ensemble. Because the density of states of the system is usually unknown, the multicanonical weight factor has to be determined numerically by iterations of short preliminary runs (21, 22).

A multicanonical MC simulation is performed, for instance, with the usual Metropolis criterion (158): The transition probability of state  $x$  with potential energy  $E$  to state  $x'$  with potential energy  $E'$  is given by

$$\begin{aligned} w(x \rightarrow x') &= \min\left(1, \frac{W_{\text{MUCA}}(E')}{W_{\text{MUCA}}(E)}\right) = \min\left(1, \frac{n(E)}{n(E')}\right) \\ &= \min(1, \exp(-\beta_a \Delta E_{\text{MUCA}})), \end{aligned} \quad (55)$$

where

$$\Delta E_{\text{MUCA}} = E_{\text{MUCA}}(E'; T_a) - E_{\text{MUCA}}(E; T_a). \quad (56)$$

The MD algorithm in the multicanonical ensemble also naturally follows from Eq. 53, in which the regular constant temperature MD simulation (with  $T = T_a$ ) is performed by replacing  $E$  by  $E_{\text{MUCA}}$  in Eq. 12 (50, 54):

$$\begin{aligned} \dot{\mathbf{p}}_k &= -\frac{\partial E_{\text{MUCA}}(E; T_a)}{\partial \mathbf{q}_k} - \frac{\dot{s}}{s} \mathbf{p}_k \\ &= \frac{\partial E_{\text{MUCA}}(E; T_a)}{\partial E} f_k - \frac{\dot{s}}{s} \mathbf{p}_k. \end{aligned} \quad (57)$$

If the exact multicanonical weight factor  $W_{\text{MUCA}}(E)$  is known, one can calculate the ensemble averages of any physical quantity  $A$  at any temperature  $T (= 1/k_B\beta)$  from Eq. 30, where the density of states is given by (see Eq. 53)

$$n(E) = \frac{1}{W_{\text{MUCA}}(E)}. \quad (58)$$

In general, the multicanonical weight factor  $W_{\text{MUCA}}(E)$ , or the density of states  $n(E)$ , is not *a priori* known, and one needs its estimator for a numerical simulation. This estimator is usually obtained from iterations of short trial multicanonical simulations. The details of this process are described, for instance, in Refs. (36, 45). However, the iterative process can be nontrivial and very tedious for complex systems.

In practice, it is impossible to obtain the ideal multicanonical weight factor with completely uniform potential energy distribution. The question is when to stop the iteration for the weight factor determination. Our criterion for a satisfactory weight factor is that as long as we do get a random walk in potential energy space, the probability distribution  $P_{\text{MUCA}}(E)$  does not have to be completely flat with a tolerance of, say, an order of magnitude deviation. In such a case, we usually perform with this weight factor a multicanonical simulation with high statistics (production run) in order to get even better estimate of the density of states. Let  $N_{\text{MUCA}}(E)$  be the histogram of potential energy distribution  $P_{\text{MUCA}}(E)$  obtained by this production run. The best estimate of the density of states can then be given by the single-histogram reweighting techniques (17, 18) as follows (see the proportionality relation in Eq. 52):

$$n(E) = \frac{N_{\text{MUCA}}(E)}{W_{\text{MUCA}}(E)}. \quad (59)$$

By substituting this quantity into Eq. 30, one can calculate ensemble averages of physical quantity  $A(E)$  as a function of temperature. Moreover, the ensemble averages of any physical quantity  $A$  (including those that cannot be expressed as functions of potential energy) at any temperature  $T (= 1/k_B\beta)$  can also be obtained as long as one stores the “trajectory” of configurations from the production run. Namely, we have (109)

$$\langle A \rangle_T = \frac{\sum_{k=1}^{n_s} A(x_k) W_{\text{MUCA}}^{-1}(E(x_k)) \exp[-\beta E(x_k)]}{\sum_{k=1}^{n_s} W_{\text{MUCA}}^{-1}(E(x_k)) \exp[-\beta E(x_k)]}, \quad (60)$$

where  $x_k$  is the configuration at the  $k$ th MC (or MD) step and  $n_s$  is the total number of configurations stored. Note that when  $A$  is a function of  $E$ , Eq. 60 reduces to Eq. 30 where the density of states is given by Eq. 59.

Some remarks are in order. The major advantage of REM over other generalized-ensemble methods such as simulated tempering (80, 81) and MUCA (21, 22) lies in the fact that the weight factor is a priori known (see Eq. 18), while in simulated tempering and MUCA, the determination of the weight factors can be very tedious and time-consuming. In REM, however, the number of required replicas increases greatly ( $\propto \sqrt{N}$ ) as the system size  $N$  increases (92), while only one replica is used in simulated tempering and MUCA. This demands a lot of computer power for complex systems. Moreover, so long as optimal weight factors can be obtained, simulated tempering and MUCA are more efficient in sampling than the REM (15, 77, 87, 110).

## **2.4. Replica-Exchange Simulated Tempering and Replica-Exchange Multicanonical Algorithm**

The REST(86) and the REMUCA (105, 109, 110) overcome both the difficulties of ST and MUCA (the weight factor determinations are nontrivial) and REM (many replicas, or a lot of computation time, are required).

In REST (86), we first perform a short REM simulation (with  $M$  replicas) to determine the simulated tempering weight factor and then perform with this weight factor a regular ST simulation with high statistics. The first step is accomplished by the multiple-histogram reweighting techniques (19, 20), which give the dimensionless Helmholtz free energy. Let  $N_m(E)$  and  $n_m$  be respectively the potential-energy histogram and the total number of samples obtained at temperature  $T_m (= 1/k_B\beta_m)$  of the REM run. The dimensionless Helmholtz free energy  $f_m$  is then given by solving Eqs. 31 and 32 self-consistently by iteration.

Once the estimate of the dimensionless Helmholtz free energy  $f_m$  are obtained, the simulated tempering weight factor can be directly determined by using Eq. 47 where we set  $a_m = f_m$  (compare Eq. (48) with Eq. 32). A long simulated tempering run is then performed with this weight factor. Let  $N_m(E)$  and  $n_m$  be respectively the potential-energy histogram and the total number of samples obtained at temperature  $T_m (= 1/k_B\beta_m)$  from this simulated tempering run. The multiple-histogram reweighting techniques of Eqs. 31 and 32 can be used again to obtain the best estimate of the density of states  $n(E)$ . The expectation value of a physical quantity  $A$  at any temperature  $T (= 1/k_B\beta)$  is then calculated from Eq. 30.

We now present the REMUCA (105, 109, 110). In REMUCA, just as in REST, we first perform a short REM simulation (with  $M$  replicas) to determine the multicanonical weight factor and then perform with this weight factor a regular multicanonical simulation with high statistics. The first step is accomplished by the multiple-histogram reweighting techniques (19, 20), which give the density of states. Let  $N_m(E)$  and  $n_m$  be respectively the potential-energy histogram and the total number of samples

obtained at temperature  $T_m$  ( $= 1/k_B\beta_m$ ) of the REM run. The density of states  $n(E)$  is then given by solving Eqs. 31 and 32 self-consistently by iteration.

Once the estimate of the density of states is obtained, the multicanonical weight factor can be directly determined from Eq. 53 (see also Eq. 54). Actually, the density of states  $n(E)$  and the multicanonical potential energy,  $E_{\text{MUCA}}(E; T_0)$ , thus determined are only reliable in the following range:

$$E_1 \leq E \leq E_M, \quad (61)$$

where

$$\begin{cases} E_1 = \langle E \rangle_{T_1}, \\ E_M = \langle E \rangle_{T_M}, \end{cases} \quad (62)$$

and  $T_1$  and  $T_M$  are respectively the lowest and the highest temperatures used in the REM run. Outside this range, we extrapolate the multicanonical potential energy linearly (105):

$$\begin{aligned} \mathcal{E}_{\text{MUCA}}(E) \equiv & \\ & \begin{cases} \frac{\partial E_{\text{MUCA}}(E; T_0)}{\partial E} \Big|_{E=E_1} (E - E_1) + E_{\text{MUCA}}(E_1; T_0), & \text{for } E < E_1, \\ E_{\text{MUCA}}(E; T_0), & \text{for } E_1 \leq E \leq E_M, \\ \frac{\partial E_{\text{MUCA}}(E; T_0)}{\partial E} \Big|_{E=E_M} (E - E_M) + E_{\text{MUCA}}(E_M; T_0), & \text{for } E > E_M. \end{cases} \end{aligned} \quad (63)$$

For Monte Carlo method, the weight factor for REMUCA is given by substituting Eq. 63 into Eq. 53 (105, 109):

$$\begin{aligned} W_{\text{MUCA}}(E) &= \exp[-\beta_0 \mathcal{E}_{\text{MUCA}}(E)] \\ &= \begin{cases} \exp(-\beta_1 E), & \text{for } E < E_1, \\ \frac{1}{n(E)}, & \text{for } E_1 \leq E \leq E_M, \\ \exp(-\beta_M E), & \text{for } E > E_M. \end{cases} \end{aligned} \quad (64)$$

The multicanonical MC and MD runs are then performed respectively with the Metropolis criterion of Eq. 55 and with the modified Newton equation in Eq. 57, in which  $\mathcal{E}_{\text{MUCA}}(E)$  in Eq. 63 is substituted into  $E_{\text{MUCA}}(E; T_0)$ . We expect to obtain a flat potential energy distribution in the range of Eq. 61. Finally, the results are analyzed by the single-histogram reweighting techniques as described in Eq. 59 (and Eq. 30).

The formulations of REST and REMUCA are simple and straightforward, but the numerical improvement is great, because the weight factor determination for ST and MUCA becomes very difficult by the usual iterative processes for complex systems.

## 2.5. Simulated Tempering Replica-Exchange Method and Multicanonical Replica-Exchange Method

In the previous subsection we presented REST and REMUCA, which use a short REM run for the determination of the simulated tempering weight factor and the multicanonical weight factor, respectively. Here, we present two modifications of REM and refer to the new methods as the *STREM* (87) and *MUCAREM* (105, 109, 110). In STREM the production run is a REM simulation with a few replicas that perform ST simulations with different temperature ranges. Likewise, in MUCAREM, the production run is a REM simulation with a few replicas in multicanonical ensembles, i.e., different replicas perform MUCA simulations with different energy ranges.

While ST and MUCA simulations are usually based on local updates, a replica-exchange process can be considered to be a global update, and global updates enhance the conformational sampling further.

## 3. Multidimensional Extensions of the Three Generalized Ensemble Algorithms

### 3.1. General Formulations

We now give the general formulations for the multidimensional generalized-ensemble algorithms (154–156). Let us consider a generalized potential energy function  $E_{\lambda}(x)$ , which depends on  $L$  parameters  $\lambda = (\lambda^{(1)}, \dots, \lambda^{(L)})$ , of a system in state  $x$ . Although  $E_{\lambda}(x)$  can be any function of  $\lambda$ , we consider the following specific generalized potential energy function for simplicity:

$$E_{\lambda}(x) = E_0(x) + \sum_{\ell=1}^L \lambda^{(\ell)} V_{\ell}(x). \quad (65)$$

Here, there are  $L + 1$  energy terms,  $E_0(x)$  and  $V_{\ell}(x)$  ( $\ell = 1, \dots, L$ ), and  $\lambda^{(\ell)}$  are the corresponding coupling constants for  $V_{\ell}(x)$ .

After integrating out the momentum degrees of freedom, the partition function of the system at fixed temperature  $T$  and parameters  $\lambda$  is given by

$$\begin{aligned} Z(T, \lambda) &= \int dx \exp(-\beta E_{\lambda}(x)) \\ &= \int dE_0 dV_1 \cdots dV_L n(E_0, V_1, \dots, V_L) \exp(-\beta E_{\lambda}), \end{aligned} \quad (66)$$

where  $n(E_0, V_1, \dots, V_L)$  is the multidimensional density of states:

$$\begin{aligned} n(E_0, V_1, \dots, V_L) &= \int dx \delta(E_0(x) - E_0) \delta(V_1(x) \\ &\quad - V_1) \cdots \delta(V_L(x) - V_L). \end{aligned} \quad (67)$$

Here, the integration is replaced by a summation when  $x$  is discrete.

The expression in Eq. 65 is often used in simulations. For instance, in simulations of spin systems,  $E_0(x)$  and  $V_1(x)$  (here,  $L = 1$  and  $x = \{S_1, S_2, \dots\}$  stand for spins) can be respectively considered as the zero-field term and the magnetization term coupled with the external field  $\lambda^{(1)}$ . (For Ising model,  $E_0 = -J\sum_{i,j} S_i S_j$ ,  $V_1 = -\sum_i S_i$ , and  $\lambda^{(1)} = h$ , i.e., external magnetic field.) In umbrella sampling (30) in molecular simulations,  $E_0(x)$  and  $V_\ell(x)$  can be taken as the original potential energy and the (biasing) umbrella potential energy, respectively, with the coupling parameter  $\lambda^{(\ell)}$  (here,  $x = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$  where  $\mathbf{q}_k$  is the coordinate vector of the  $k$ th particle and  $N$  is the total number of particles). For the molecular simulations in the isobaric-isothermal ensemble,  $E_0(x)$  and  $V_1(x)$  (here,  $L = 1$ ) correspond respectively to the potential energy  $U$  and the volume  $\mathcal{V}$  coupled with the pressure  $\mathcal{P}$ . (Namely, we have  $x = \{\mathbf{q}_1, \dots, \mathbf{q}_N, \mathcal{V}\}$ ,  $E_0 = U$ ,  $V_1 = \mathcal{V}$ , and  $\lambda^{(1)} = \mathcal{P}$ , i.e.,  $E_\lambda$  is the enthalpy without the kinetic energy contributions). For simulations in the grand canonical ensemble with  $N$  particles, we have  $x = \{\mathbf{q}_1, \dots, \mathbf{q}_N, N\}$ , and  $E_0(x)$  and  $V_1(x)$  (here,  $L = 1$ ) correspond respectively to the potential energy  $U$  and the total number of particles  $N$  coupled with the chemical potential  $\mu$ . (Namely, we have  $E_0 = U$ ,  $V_1 = N$ , and  $\lambda^{(1)} = -\mu$ .)

Moreover, going beyond the well-known ensembles discussed above, we can introduce any physical quantity of interest (or its function) as the additional potential energy term  $V_\ell$ . For instance,  $V_\ell$  can be an overlap with a reference configuration in spin glass systems, an end-to-end distance, a radius of gyration in molecular systems, etc. In such a case, we have to carefully choose the range of  $\lambda^{(\ell)}$  values so that the new energy term  $\lambda^{(\ell)} V_\ell$  will have roughly the same order of magnitude as the original energy term  $E_0$ . We want to perform a simulation where a random walk not only in the  $E_0$  space but also in the  $V_\ell$  space is realized. As shown below, this can be done by performing a multidimensional REM, ST, or MUCA simulation.

We first describe the *multidimensional replica-exchange method* (MREM) (103). The crucial observation that led to this algorithm is: As long as we have  $M$  *non-interacting* replicas of the original system, the Hamiltonian  $H(q, p)$  of the system does not have to be identical among the replicas and it can depend on a parameter with different parameter values for different replicas. The system for the multidimensional REM consists of  $M$  non-interacting replicas of the original system in the “canonical ensemble” with  $M = M_0 \times M_1 \times \dots \times M_L$  different parameter sets  $\Lambda_m$  ( $m = 1, \dots, M$ ), where  $\Lambda_m \equiv (T_{m_0}, \boldsymbol{\lambda}_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$  with  $m_0 = 1, \dots, M_0$ ,  $m_\ell = 1, \dots, M_\ell$  ( $\ell = 1, \dots, L$ ). Because the replicas are non-interacting, the weight factor is given by the product of Boltzmann-like factors for each replica:

$$W_{\text{MREM}} \equiv \prod_{m_0=1}^{M_0} \prod_{m_1=1}^{M_1} \cdots \prod_{m_L=1}^{M_L} \exp(-\beta_{m_0} E_{\lambda_{m_0}}). \quad (68)$$

Without loss of generality we can order the parameters so that  $T_1 < T_2 < \cdots < T_{M_0}$  and  $\lambda_1^{(\ell)} < \lambda_2^{(\ell)} < \cdots < \lambda_{M_\ell}^{(\ell)}$  (for each  $\ell = 1, \dots, L$ ). The multidimensional REM is realized by alternately performing the following two steps:

1. For each replica, a “canonical” MC or MD simulation at the fixed parameter set is carried out simultaneously and independently for a certain steps.
2. We exchange a pair of replicas  $i$  and  $j$  which are at the parameter sets  $\Lambda_m$  and  $\Lambda_{m+1}$ , respectively. The transition probability for this replica-exchange process is given by

$$w(\Lambda_m \leftrightarrow \Lambda_{m+1}) = \min(1, \exp(-\Delta)), \quad (69)$$

where we have

$$\Delta = (\beta_{m_0} - \beta_{m_0+1}) \left( E_{\lambda_m}(q^{[j]}) - E_{\lambda_m}(q^{[i]}) \right), \quad (70)$$

for  $T$ -exchange, and

$$\Delta = \beta_{m_0} \left[ \left( E_{\lambda_{m_\ell}}(q^{[j]}) - E_{\lambda_{m_\ell}}(q^{[i]}) \right) - \left( E_{\lambda_{m_\ell+1}}(q^{[j]}) - E_{\lambda_{m_\ell+1}}(q^{[i]}) \right) \right], \quad (71)$$

for  $\lambda^{(\ell)}$ -exchange (for one of  $\ell = 1, \dots, L$ ). Here,  $q^{[i]}$  and  $q^{[j]}$  stand for configuration variables for replicas  $i$  and  $j$ , respectively, before the replica exchange.

We now consider the *multidimensional simulated tempering* (MST) which realizes a random walk both in temperature  $T$  and in parameters  $\lambda$  (154–156). The entire parameter set  $\Lambda = (T, \lambda) \equiv (T, \lambda^{(1)}, \dots, \lambda^{(L)})$  becomes dynamical variables, and both the configuration and the parameter set are updated during the simulation with a weight factor:

$$W_{\text{MST}}(\Lambda) \equiv \exp(-\beta E_\lambda + f(\Lambda)), \quad (72)$$

where the function  $f(\Lambda) = f(T, \lambda)$  is chosen so that the probability distribution of  $\Lambda$  is flat:

$$P_{\text{MST}}(\Lambda) \propto \int dE_0 dV_1 \cdots dV_L n(E_0, V_1, \dots, V_L) \exp(-\beta E_\lambda + f(\Lambda)) \equiv \text{constant}. \quad (73)$$

This means that  $f(\Lambda)$  is the dimensionless (“Helmholtz”) free energy:

$$\exp(-f(\Lambda)) = \int dE_0 dV_1 \cdots dV_L n(E_0, V_1, \dots, V_L) \exp(-\beta E_\lambda). \quad (74)$$

In the numerical work we discretize the parameter set  $\Lambda$  in  $M (= M_0 \times M_1 \times \cdots \times M_L)$  different values:  $\Lambda_m \equiv (T_{m_0}, \boldsymbol{\lambda}_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ , where  $m_0 = 1, \dots, M_0, m_\ell = 1, \dots, M_\ell$  ( $\ell = 1, \dots, L$ ). Without loss of generality we can order the parameters so that  $T_1 < T_2 < \cdots < T_{M_0}$  and  $\lambda_1^{(\ell)} < \lambda_2^{(\ell)} < \cdots < \lambda_{M_\ell}^{(\ell)}$  (for each  $\ell = 1, \dots, L$ ). The free energy  $f(\Lambda_m)$  is now written as  $f_{m_0, m_1, \dots, m_L} = f(T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ .

Once the initial configuration and the initial parameter set are chosen, the multidimensional ST is realized by alternately performing the following two steps:

1. A “canonical” MC or MD simulation at the fixed parameter set  $\Lambda_m = (T_{m_0}, \boldsymbol{\lambda}_m) = (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$  is carried out for a certain steps with the weight factor  $\exp(-\beta_{m_0} E_{\boldsymbol{\lambda}_m})$  (for fixed  $\Lambda_m$ ,  $f(\Lambda_m)$  in Eq. 72 is a constant and does not contribute).
2. We update the parameter set  $\Lambda_m$  to a new parameter set  $\Lambda_{m \pm 1}$  in which one of the parameters in  $\Lambda_m$  is changed to a neighboring value with the configuration and the other parameters fixed. The transition probability of this parameter updating process is given by the following Metropolis criterion:

$$w(\Lambda_m \rightarrow \Lambda_{m \pm 1}) = \min\left(1, \frac{W_{\text{MST}}(\Lambda_{m \pm 1})}{W_{\text{MST}}(\Lambda_m)}\right) = \min(1, \exp(-\Delta)). \quad (75)$$

Here, there are two possibilities for  $\Lambda_{m \pm 1}$ , namely,  $T$ -update and  $\lambda^{(\ell)}$ -update. For  $T$ -update, we have  $\Lambda_{m \pm 1} = (T_{m_0 \pm 1}, \boldsymbol{\lambda}_m)$  with

$$\Delta = (\beta_{m_0 \pm 1} - \beta_{m_0}) E_{\boldsymbol{\lambda}_m} - (f_{m_0 \pm 1, m_1, \dots, m_L} - f_{m_0, m_1, \dots, m_L}). \quad (76)$$

For  $\lambda^{(\ell)}$ -update (for one of  $\ell = 1, \dots, L$ ), we have  $\Lambda_{m \pm 1} = (T_{m_0}, \boldsymbol{\lambda}_{m_\ell \pm 1})$  with

$$\Delta = \beta_{m_0} (E_{\boldsymbol{\lambda}_{m_\ell \pm 1}} - E_{\boldsymbol{\lambda}_{m_\ell}}) - (f_{m_0, \dots, m_\ell \pm 1, \dots} - f_{m_0, \dots, m_\ell, \dots}), \quad (77)$$

where  $\boldsymbol{\lambda}_{m_\ell \pm 1} = (\dots, \lambda_{m_{\ell-1}}^{(\ell-1)}, \lambda_{m_\ell \pm 1}^{(\ell)}, \lambda_{m_{\ell+1}}^{(\ell+1)}, \dots)$  and  $\boldsymbol{\lambda}_{m_\ell} = (\dots, \lambda_{m_{\ell-1}}^{(\ell-1)}, \lambda_{m_\ell}^{(\ell)}, \lambda_{m_{\ell+1}}^{(\ell+1)}, \dots)$ .

We now describe the *multidimensional multicanonical algorithm* (MMUCA) which realizes a random walk in the  $(L + 1)$ -dimensional space of  $E_0(x)$  and  $V_\ell(x)$  ( $\ell = 1, \dots, L$ ).

In the multidimensional MUCA ensemble, each state is weighted by the MUCA weight factor  $W_{\text{MMUCA}}(E_0, V_1, \dots, V_L)$  so that a uniform energy distribution of  $E_0, V_1, \dots$ , and  $V_L$  may be obtained:

$$P_{\text{MMUCA}}(E_0, V_1, \dots, V_L) \propto n(E_0, V_1, \dots, V_L) W_{\text{MMUCA}}(E_0, V_1, \dots, V_L) \equiv \text{constant}, \quad (78)$$

where  $n(E_0, V_1, \dots, V_L)$  is the multidimensional density of states. From this equation, we obtain

$$\begin{aligned} W_{\text{MMUCA}}(E_0, V_1, \dots, V_L) &\equiv \exp(-\beta_a E_{\text{MMUCA}}(E_0, V_1, \dots, V_L)) \\ &\propto \frac{1}{n(E_0, V_1, \dots, V_L)}, \end{aligned} \quad (79)$$

where we have introduced an arbitrary reference temperature,  $T_a = 1/k_B\beta_a$ , and wrote the weight factor in the Boltzmann-like form. Here, the “multicanonical potential energy” is defined by

$$E_{\text{MMUCA}}(E_0, V_1, \dots, V_L; T_a) \equiv k_B T_a \ln n(E_0, V_1, \dots, V_L). \quad (80)$$

The multidimensional MUCA MC simulation can be performed with the following Metropolis transition probability from state  $x$  with energy  $E_\lambda = E_0 + \sum_{\ell=1}^L \lambda^{(\ell)} V_\ell$  to state  $x'$  with energy  $E_{\lambda'} = E_0' + \sum_{\ell=1}^L \lambda^{(\ell)} V_\ell'$ :

$$\begin{aligned} w(x \rightarrow x') &= \min\left(1, \frac{W_{\text{MMUCA}}(E_0', V_1', \dots, V_L')}{W_{\text{MMUCA}}(E_0, V_1, \dots, V_L)}\right) \\ &= \min\left(1, \frac{n(E_0, V_1, \dots, V_L)}{n(E_0', V_1', \dots, V_L')}\right). \end{aligned} \quad (81)$$

An MD algorithm in the multidimensional MUCA ensemble also naturally follows from Eq. 79, in which a regular constant temperature MD simulation (with  $T = T_a$ ) is performed by replacing the total potential energy  $E_\lambda$  by the multicanonical potential energy  $E_{\text{MMUCA}}$  in Eq. 12:

$$\dot{\mathbf{p}}_k = -\frac{\partial E_{\text{MMUCA}}(E_0, V_1, \dots, V_L; T_a)}{\partial \mathbf{q}_k} - \frac{\dot{s}}{s} \mathbf{p}_k. \quad (82)$$

We remark that the random walk in  $E_0$  and in  $V_\ell$  for the MUCA simulation corresponds to that in  $\beta$  and in  $\beta\lambda^{(\ell)}$  for the ST simulation:

$$\begin{cases} E_0 \longleftrightarrow \beta, \\ V_\ell \longleftrightarrow \beta\lambda^{(\ell)}, \quad (\ell = 1, \dots, L). \end{cases} \quad (83)$$

They are in conjugate relation.

### 3.2. Weight Factor Determinations for Multidimensional ST and MUCA

Among the three multidimensional generalized-ensemble algorithms described above, only MREM can be performed without much preparation because the weight factor for MREM is just a product of regular Boltzmann-like factors. On the other hand, we do not know the MST and MMUCA weight factors a priori and need to estimate them. As a simple method for these weight factor determinations, we can generalize the REST and REMUCA presented in the previous subsections to multidimensions.

Suppose we have made a single run of a short MREM simulation with  $M (= M_0 \times M_1 \times \dots \times M_L)$  replicas that correspond to  $M$

different parameter sets  $\Lambda_m$  ( $m = 1, \dots, M$ ). Let  $N_{m_0, m_1, \dots, m_L}$  ( $E_0, V_1, \dots, V_L$ ) and  $n_{m_0, m_1, \dots, m_L}$  be respectively the  $(L+1)$ -dimensional potential-energy histogram and the total number of samples obtained for the  $m$ th parameter set  $\Lambda_m = (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ . The generalized WHAM equations are then given by

$$\begin{aligned} n(E_0, V_1, \dots, V_L) \\ = \frac{\sum_{m_0, m_1, \dots, m_L} N_{m_0, m_1, \dots, m_L}(E_0, V_1, \dots, V_L)}{\sum_{m_0, m_1, \dots, m_L} n_{m_0, m_1, \dots, m_L} \exp(f_{m_0, m_1, \dots, m_L} - \beta_{m_0} E_{\lambda_m})}, \end{aligned} \quad (84)$$

and

$$\exp(-f_{m_0, m_1, \dots, m_L}) = \sum_{E_0, V_1, \dots, V_L} n(E_0, V_1, \dots, V_L) \exp(-\beta_{m_0} E_{\lambda_m}). \quad (85)$$

The density of states  $n(E_0, V_1, \dots, V_L)$  (which is inversely proportional to the MMUCA weight factor) and the dimensionless free energy  $f_{m_0, m_1, \dots, m_L}$  (which is the MST parameter) are obtained by solving Eqs. 84 and 85 self-consistently by iteration.

### 3.3. Expectation Values of Physical Quantities

We now present the equations to calculate ensemble averages of physical quantities with any temperature  $T$  and any parameter  $\lambda$  values.

After a long production run of MREM and MST simulations, the canonical expectation value of a physical quantity  $A$  with the parameter values  $\Lambda_m$  ( $m = 1, \dots, M$ ), where  $\Lambda_m \equiv (T_{m_0}, \lambda_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$  with  $m_0 = 1, \dots, M_0$ ,  $m_\ell = 1, \dots, M_\ell$  ( $\ell = 1, \dots, L$ ), and  $M (= M_0 \times M_1 \times \dots \times M_L)$ , can be calculated by the usual arithmetic mean:

$$\langle A \rangle_{T_{m_0}, \lambda_m} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)), \quad (86)$$

where  $x_m(k)$  ( $k = 1, \dots, n_m$ ) are the configurations obtained with the parameter values  $\Lambda_m$  ( $m = 1, \dots, M$ ), and  $n_m$  is the total number of measurements made with these parameter values. The expectation values of  $A$  at any intermediate  $T$  ( $= 1/k_B\beta$ ) and any  $\lambda$  can also be obtained from

$$\langle A \rangle_{T, \lambda} = \frac{\sum_{E_0, V_1, \dots, V_L} A(E_0, V_1, \dots, V_L) n(E_0, V_1, \dots, V_L) \exp(-\beta E_{\lambda})}{\sum_{E_0, V_1, \dots, V_L} n(E_0, V_1, \dots, V_L) \exp(-\beta E_{\lambda})}, \quad (87)$$

where the density of states  $n(E_0, V_1, \dots, V_L)$  is obtained from the multiple-histogram reweighting techniques. Namely, from the MREM or MST simulation, we first obtain the histogram

$N_{m_0, m_1, \dots, m_L}(E_0, V_1, \dots, V_L)$  and the total number of samples  $n_{m_0, m_1, \dots, m_L}$  in Eq. 84. The density of states  $n(E_0, V_1, \dots, V_L)$  and the dimensionless free energy  $f_{m_0, m_1, \dots, m_L}$  are then obtained by solving Eqs. 84 and 85 self-consistently by iteration. Substituting the obtained density of states  $n(E_0, V_1, \dots, V_L)$  into Eq. 87, one can calculate the ensemble average of the physical quantity  $A$  at any  $T$  and any  $\lambda$ .

Moreover, the ensemble average of the physical quantity  $A$  (including those that cannot be expressed as functions of  $E_0$  and  $V_\ell (\ell = 1, \dots, L)$ ) can be obtained from the “trajectory” of configurations of the production run (155). Namely, we first obtain  $f_{m_0, m_1, \dots, m_L}$  for each  $(m_0 = 1, \dots, M_0, m_1 = 1, \dots, M_1, \dots, m_L = 1, \dots, M_L)$  by solving Eqs. 84 and 85 self-consistently, and then we have

$$\langle A \rangle_{T, \lambda} = \frac{\sum_{m_0=1}^{M_0} \dots \sum_{m_L=1}^{M_L} \sum_{x_m} A(x_m) \frac{\exp(-\beta E_\lambda(x_m))}{\sum_{n_0=1}^{M_0} \dots \sum_{n_L=1}^{M_L} n_{n_0, \dots, n_L} \exp(f_{n_0, \dots, n_L} - \beta_{n_0} E_{\lambda, n}(x_m))}}{\sum_{m_0=1}^{M_0} \dots \sum_{m_L=1}^{M_L} \sum_{x_m} \frac{\exp(-\beta E_\lambda(x_m))}{\sum_{n_0=1}^{M_0} \dots \sum_{n_L=1}^{M_L} n_{n_0, \dots, n_L} \exp(f_{n_0, \dots, n_L} - \beta_{n_0} E_{\lambda, n}(x_m))}}, \quad (88)$$

where  $x_m$  are the configurations obtained at  $\Lambda_m = (T_{m_0}, \lambda_m) = (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ . Here, the trajectories  $x_m$  are stored for each  $\Lambda_m$  separately.

For the MMUCA simulation with the weight factor  $W_{\text{MMUCA}}(E_0, \dots, V_L)$ , the expectation values of  $A$  at any  $T (= 1/k_B\beta)$  and any  $\lambda$  can also be obtained from Eq. 87 by the single-histogram reweighting techniques as follows. Let  $N_{\text{MMUCA}}(E_0, V_1, \dots, V_L)$  be the histogram of the distribution of  $E_0, V_1, \dots, V_L$ ,  $P_{\text{MMUCA}}(E_0, V_1, \dots, V_L)$ , obtained by the production run. The best estimate of the density of states  $n(E_0, V_1, \dots, V_L)$  is then given by

$$n(E_0, V_1, \dots, V_L) = \frac{N_{\text{MMUCA}}(E_0, V_1, \dots, V_L)}{W_{\text{MMUCA}}(E_0, \dots, V_L)}. \quad (89)$$

Moreover, the ensemble average of the physical quantity  $A$  (including those that cannot be expressed as a function of  $E_0$  and  $V_\ell (\ell = 1, \dots, L)$ ) can be obtained as long as one stores the “trajectory” of configurations  $x_k$  from the production run. We have

$$\langle A \rangle_{T, \lambda} = \frac{\sum_{k=1}^{n_s} A(x_k) W_{\text{MMUCA}}^{-1}(E_0(x_k), \dots, V_L(x_k)) \exp(-\beta E_\lambda(x_k))}{\sum_{k=1}^{n_s} W_{\text{MMUCA}}^{-1}(E_0(x_k), \dots, V_L(x_k)) \exp(-\beta E_\lambda(x_k))}. \quad (90)$$

Here,  $x_k$  is the configuration at the  $k$ th MC (or MD) step and  $n_s$  is the total number of configurations stored.

### 3.4. Multidimensional Generalized-Ensemble Algorithms for the Isobaric–Isothermal Ensemble

As examples of the multidimensional formulations in the previous subsections, we present the generalized-ensemble algorithms for the isobaric-isothermal ensemble (or, the NPT ensemble) (157). Let us consider a physical system that consists of  $N$  atoms and that is in a box of a finite volume  $\mathcal{V}$ . The states of the system are specified by coordinates  $\mathbf{q} \equiv \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N\}$  and momenta  $\mathbf{p} \equiv \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$  of the atoms and volume  $\mathcal{V}$  of the box. The potential energy  $E(\mathbf{q}, \mathcal{V})$  for the system is a function of  $\mathbf{q}$  and  $\mathcal{V}$ .

In the isobaric-isothermal ensemble (159, 160, 165, 172), the probability distribution  $P_{\text{NPT}}(E, \mathcal{V}; T, \mathcal{P})$  for potential energy  $E$  and volume  $\mathcal{V}$  at temperature  $T$  and pressure  $\mathcal{P}$  is given by

$$P_{\text{NPT}}(E, \mathcal{V}; T, \mathcal{P}) \propto n(E, \mathcal{V}) W_{\text{NPT}}(E, \mathcal{V}; T, \mathcal{P}) = n(E, \mathcal{V}) e^{-\beta \mathcal{H}}. \quad (91)$$

Here, the density of states  $n(E, \mathcal{V})$  is given as a function of both  $E$  and  $\mathcal{V}$ , and  $\mathcal{H}$  is the “enthalpy” (without the kinetic energy contributions):

$$\mathcal{H} = E + \mathcal{P}\mathcal{V}. \quad (92)$$

This weight factor produces an isobaric-isothermal ensemble at constant temperature ( $T$ ) and constant pressure ( $\mathcal{P}$ ). Note that this is a special case of the general formulations in Eq. 7.65 with  $L = 1$ ,  $E_0 = E$ ,  $V_1 = \mathcal{V}$ , and  $\lambda^{(1)} = \mathcal{P}$ .

In order to perform the isobaric-isothermal MC simulation (172), we perform Metropolis sampling on the scaled coordinates  $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \dots, \boldsymbol{\sigma}_N\}$  where  $\boldsymbol{\sigma}_k = \mathcal{V}^{-1/3} \mathbf{q}_k$  ( $k = 1, 2, \dots, N$ ) ( $\mathbf{q}_k$  are the real coordinates) and the volume  $\mathcal{V}$  (here, the particles are placed in a cubic box of a side of size  $\mathcal{V}^{-1/3}$ ). The trial moves from state  $x$  with the scaled coordinate  $\boldsymbol{\sigma}$  with volume  $\mathcal{V}$  to state  $x'$  with the scaled coordinate  $\boldsymbol{\sigma}'$  and volume  $\mathcal{V}'$  are generated by uniform random numbers. The enthalpy is accordingly changed from  $\mathcal{H}(E(\boldsymbol{\sigma}, \mathcal{V}), \mathcal{V})$  to  $\mathcal{H}'(E(\boldsymbol{\sigma}', \mathcal{V}'), \mathcal{V}')$  by these trial moves. The trial moves will be accepted with the following Metropolis criterion:

$$w(x \rightarrow x') = \min(1, \exp[-\beta\{\mathcal{H}' - \mathcal{H} - Nk_B T \ln(\mathcal{V}'/\mathcal{V})\}]), \quad (93)$$

where  $N$  is the total number of atoms in the system.

As for the MD method in this ensemble, we just present the Nosé-Andersen algorithm (159, 160, 165). The equations of motion in Eqs. 11–14 are now generalized as follows:

$$\dot{\mathbf{q}}_k = \frac{\mathbf{p}_k}{m_k} + \frac{\dot{\mathcal{V}}}{3\mathcal{V}} \mathbf{q}_k, \quad (94)$$

$$\dot{\mathbf{p}}_k = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_k} - \left( \frac{\dot{s}}{s} + \frac{\dot{\mathcal{V}}}{3\mathcal{V}} \right) \mathbf{p}_k \quad (95)$$

$$= \mathbf{f}_k - \left( \frac{\dot{s}}{s} + \frac{\dot{\mathcal{V}}}{3\mathcal{V}} \right) \mathbf{p}_k, \quad (96)$$

$$\dot{s} = s \frac{P_s}{Q}, \quad (97)$$

$$\dot{P}_s = \sum_{i=1}^N \frac{\dot{\mathbf{p}}_i^2}{m_i} - 3Nk_B T = 3Nk_B(T(t) - T), \quad (98)$$

$$\dot{\mathcal{V}} = s \frac{P_V}{M}, \quad (99)$$

$$\dot{P}_V = s \left[ \frac{1}{3\mathcal{V}} \left( \sum_{i=1}^N \frac{\dot{\mathbf{p}}_i^2}{m_i} - \sum_{i=1}^N \mathbf{q}_i \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{q}_i} \right) - \frac{\partial \mathcal{H}}{\partial \mathcal{V}} \right] \quad (100)$$

$$= s(\mathcal{P}(t) - \mathcal{P}), \quad (101)$$

where  $M$  is the artificial mass associated with the volume,  $P_V$  is the conjugate momentum for the volume, and the “instantaneous pressure”  $\mathcal{P}(t)$  is defined by

$$\mathcal{P}(t) = \frac{1}{3\mathcal{V}} \left( \sum_{i=1}^N \frac{\mathbf{p}_i(t)^2}{m_i} + \sum_{i=1}^N \mathbf{q}_i(t) \cdot \mathbf{f}_i(t) \right) - \frac{\partial E}{\partial \mathcal{V}}(t). \quad (102)$$

In REM simulations for the NPT ensemble, we prepare a system that consists of  $M_T \times M_P$  noninteracting replicas of the original system, where  $M_T$  and  $M_P$  are the number of temperature and pressure values used in the simulation, respectively. The replicas are specified by labels  $i(i = 1, 2, \dots, M_T \times M_P)$ , temperature by  $m_0$  ( $m_0 = 1, 2, \dots, M_T$ ), and pressure by  $m_1$  ( $m_1 = 1, 2, \dots, M_P$ ).

To perform REM simulations, we carry out the following two steps alternately: (1) perform a usual constant NPT MC or MD simulation in each replica at assigned temperature and pressure and (2) try to exchange the replicas. If the temperature (specified by  $m_0$  and  $n_0$ ) or pressure (specified by  $m_1$  and  $n_1$ ) between the replicas is exchanged in Step 2, the transition probability from  $X \equiv \{\dots, (\sigma^{[i]}, \mathcal{V}^{[i]}; T_{m_0}, \mathcal{P}_{m_1}), \dots, (\sigma^{[j]}, \mathcal{V}^{[j]}; T_{n_0}, \mathcal{P}_{n_1}), \dots\}$  to  $X' \equiv \{\dots, (\sigma^{[i]}, \mathcal{V}^{[i]}; T_{n_0}, \mathcal{P}_{n_1}), \dots, (\sigma^{[j]}, \mathcal{V}^{[j]}; T_{m_0}, \mathcal{P}_{m_1}), \dots\}$  at the trial is given by (11, 120)

$$w_{\text{REM}}(X \rightarrow X') = \min[1, \exp(-\Delta_{\text{REM}})], \quad (103)$$

where

$$\begin{aligned} \Delta_{\text{REM}} = & (\beta_{m_0} - \beta_{n_0}) \left[ E(\sigma^{[j]}, \mathcal{V}^{[j]}) - E(\sigma^{[i]}, \mathcal{V}^{[i]}) \right] \\ & + (\beta_{m_0} \mathcal{P}_{m_1} - \beta_{n_0} \mathcal{P}_{n_1}) (\mathcal{V}^{[j]} - \mathcal{V}^{[i]}). \end{aligned} \quad (104)$$

In ST simulations for the NPT ensemble, we introduce a function  $f(T, \mathcal{P})$  and use a weight factor  $W_{\text{ST}}(E, \mathcal{V}; T, \mathcal{P}) \equiv \exp[-\beta(E + \mathcal{P}\mathcal{V}) + f(T, \mathcal{P})]$  so that the distribution function  $P_{\text{ST}}(T, \mathcal{P})$  of  $T$  and  $\mathcal{P}$  may be uniform:

$$P_{\text{ST}}(T, \mathcal{P}) \propto \int_0^\infty d\mathcal{V} \int_V dq W_{\text{ST}}[E(q, \mathcal{V}), \mathcal{V}; T, \mathcal{P}] = \text{constant}. \quad (105)$$

From Eq. 105, it is found that  $f(T, \mathcal{P})$  is formally given by

$$f(T, \mathcal{P}) = -\ln \left\{ \int_0^\infty d\mathcal{V} \int_V dq \exp[-\beta(E(q, \mathcal{V}) + \mathcal{P}\mathcal{V})] \right\}, \quad (106)$$

and this function is the dimensionless Gibbs free energy except for a constant.

To perform ST simulations, we again discretize temperature and pressure into  $M_0 \times M_1$  set of values  $(T_{m_0}, \mathcal{P}_{m_1})(m_0 = 1, \dots, M_0, m_1 = 1, \dots, M_1)$ . We carry out the following two steps alternately: (1) perform a usual constant NPT MC or MD simulation and (2) try to update the temperature or pressure. In Step 2 the transition probability from the state  $X \equiv \{\sigma, \mathcal{V}; T_{m_0}, \mathcal{P}_{m_1}\}$  to the state  $X' \equiv \{\sigma, \mathcal{V}; T_{n_0}, \mathcal{P}_{n_1}\}$  is given by

$$w_{\text{ST}}(X \rightarrow X') = \min[1, \exp(-\Delta_{\text{ST}})], \quad (107)$$

where

$$\begin{aligned} \Delta_{\text{ST}} = & (\beta_{n_0} - \beta_{m_0})E(\sigma, \mathcal{V}) + (\beta_{n_0}\mathcal{P}_{n_1} - \beta_{m_0}\mathcal{P}_{m_1})\mathcal{V} \\ & - (f_{n_0, n_1} - f_{m_0, m_1}). \end{aligned} \quad (108)$$

We remark that when we perform MD simulations with REM and ST, the momenta should be rescaled if the replicas are exchanged for the temperature in REM and the temperature is updated in ST as shown above in the previous subsections.

From the production run of REM or ST simulations in the NPT ensemble, we can calculate isobaric-isothermal averages of a physical quantity  $A$  at  $(T_{m_0}, \mathcal{P}_{m_1})(m_0 = 1, \dots, M_0, m_1 = 1, \dots, M_1)$  by the usual arithmetic mean:

$$\langle A \rangle_{T_{m_0}, \mathcal{P}_{m_1}} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)), \quad (109)$$

where  $x_m(k)$  ( $k = 1, \dots, n_m$ ) are the configurations obtained with the parameter values  $(T_{m_0}, \mathcal{P}_{m_1})$  and  $n_m$  is the total number of measurements made with these parameter values. The expectation values of  $A$  at any intermediate temperature  $T (= 1/k_B\beta)$  and any intermediate pressure  $\mathcal{P}$  can also be obtained from

$$\langle A \rangle_{T,\mathcal{P}} = \frac{\sum_{E,\mathcal{V}} A(E, \mathcal{V}) n(E, \mathcal{V}) \exp(-\beta(E + \mathcal{P}\mathcal{V}))}{\sum_{E,\mathcal{V}} n(E, \mathcal{V}) \exp(-\beta(E + \mathcal{P}\mathcal{V}))}, \quad (110)$$

where the density of states  $n(E, \mathcal{V})$  is obtained from the multiple-histogram reweighting techniques. Namely, from the REM or ST simulation, we first obtain the histogram  $N_{m_0, m_1}(E, \mathcal{V})$  and the total number of samples  $n_{m_0, m_1}$ . The density of states  $n(E, \mathcal{V})$  and the dimensionless free energy  $f_{m_0, m_1}$  are then obtained by solving the following equations self-consistently by iteration (see Eqs. 84 and 85 above):

$$n(E, \mathcal{V}) = \frac{\sum_{m_0, m_1} N_{m_0, m_1}(E, \mathcal{V})}{\sum_{m_0, m_1} n_{m_0, m_1} \exp(f_{m_0, m_1} - \beta_{m_0}(E + \mathcal{P}_{m_1}\mathcal{V}))}, \quad (111)$$

and

$$\exp(-f_{m_0, m_1}) = \sum_{E, \mathcal{V}} n(E, \mathcal{V}) \exp(-\beta_{m_0}(E + \mathcal{P}_{m_1}\mathcal{V})). \quad (112)$$

Substituting the obtained density of states  $n(E, \mathcal{V})$  into Eq. 110, one can calculate the ensemble average of the physical quantity  $A$  at any  $T$  and any  $\mathcal{P}$ .

We now introduce the MUCA into the isobaric-isothermal ensemble and refer to this generalized-ensemble algorithm as the MUBATH (72–75). The molecular simulations in this generalized ensemble perform random walks both in the potential energy space and in the volume space.

In the MUBATH ensemble, each state is sampled by the MUBATH weight factor  $W_{\text{mbt}}(E, \mathcal{V}) \equiv \exp\{-\beta_a \mathcal{H}_{\text{mbt}}(E, \mathcal{V})\}$  ( $\mathcal{H}_{\text{mbt}}$  is referred to as the multibaric-multithermal enthalpy) so that a uniform distribution in both potential energy  $E$  and volume  $\mathcal{V}$  is obtained (72):

$$\begin{aligned} P_{\text{mbt}}(E, \mathcal{V}) &\propto n(E, \mathcal{V}) W_{\text{mbt}}(E, \mathcal{V}) \\ &= n(E, \mathcal{V}) \exp\{-\beta_a \mathcal{H}_{\text{mbt}}(E, \mathcal{V})\} \equiv \text{constant}, \end{aligned} \quad (113)$$

where we have chosen an arbitrary reference temperature,  $T_a = 1/k_B \beta_a$ .

The MUBATH MC simulation can be performed by replacing  $\mathcal{H}$  by  $\mathcal{H}_{\text{mbt}}$  in Eq. 93:

$$w(x \rightarrow x') = \min(1, \exp[-\beta_a \{\mathcal{H}'_{\text{mbt}} - \mathcal{H}_{\text{mbt}} - Nk_B T_a \ln(\mathcal{V}'/\mathcal{V})\}]). \quad (114)$$

In order to perform the MUBATH MD simulation, we just solve the above equations of motion (Eqs. 94–101) for the regular isobaric-isothermal ensemble (with arbitrary reference temperature  $T = T_a$ ), where the enthalpy  $\mathcal{H}$  is replaced by the multibaric-multithermal enthalpy  $\mathcal{H}_{\text{mbt}}$  in Eqs. 95 and 100 (74).

In order to calculate the isobaric-isothermal-ensemble averages, we employ the single-histogram reweighting techniques (17, 18). The expectation value of a physical quantity  $A$  at any  $T$  and any  $\mathcal{P}$  is obtained by substituting the following density of states into Eq. 110:

$$n(E, \mathcal{V}) = \frac{N_{\text{mbt}}(E, \mathcal{V})}{W_{\text{mbt}}(E, \mathcal{V})}, \quad (115)$$

where  $N_{\text{mbt}}(E, \mathcal{V})$  is the histogram of the probability distribution  $P_{\text{mbt}}(E, \mathcal{V})$  of potential energy and volume that was obtained by the MUBATH production run.

## 4. Examples of Simulation Results

We tested the effectiveness of the generalized-ensemble algorithms by using a system of a 17-residue fragment of ribonuclease  $T_1$  (in Refs. (110, 173)) (154–156). It is known by experiments that this peptide fragment forms  $\alpha$ -helical conformations (173). We have performed a two-dimensional REM simulation and a two-dimensional ST simulation. In these simulations, we used the following energy function:

$$E_\lambda = E_0 + \lambda E_{\text{SOL}}, \quad (116)$$

where we set  $L = 1$ ,  $V_1 = E_{\text{SOL}}$ , and  $\lambda^{(1)} = \lambda$  in Eq. 65. Here,  $E_0$  is the potential energy of the solute and  $E_{\text{SOL}}$  is the solvation free energy. The parameters in the conformational energy as well as the molecular geometry were taken from ECEPP/2 (174–176).

The solvation term  $E_{\text{SOL}}$  is given by the sum of terms that are proportional to the solvent-accessible surface area of heavy atoms of the solute (177). For the calculations of solvent-accessible surface area, we used the computer code NSOL (178).

The computer code KONF90 (7, 8) was modified in order to accommodate the generalized-ensemble algorithms. The simulations were started from randomly generated conformations. We prepared eight temperatures ( $M_0 = 8$ ) which are distributed exponentially between  $T_1 = 300$  K and  $T_{M_0} = 700$  K (i.e., 300.00, 338.60, 382.17, 431.36, 486.85, 549.49, 620.20, and 700.00 K) and four equally spaced  $\lambda$  values ( $M_1 = 4$ ) ranging from 0 to 1 (i.e.,  $\lambda_1 = 0$ ,  $\lambda_2 = 1/3$ ,  $\lambda_3 = 2/3$ , and  $\lambda_4 = 1$ ) in the two-dimensional REM simulation and the two-dimensional ST simulation. Simulations with  $\lambda = 0$  (i.e.,  $E_\lambda = E_0$ ) and with  $\lambda = 1$  (i.e.,  $E_\lambda = E_0 + E_{\text{SOL}}$ ) correspond to those in gas phase and in aqueous solution, respectively.

We first present the results of the two-dimensional REM simulation. We used 32 replicas with the eight temperature values and the four  $\lambda$  values given above. Before taking the data, we made the two-dimensional REM simulation of 100,000 MC sweeps with each replica for thermalization. We then performed the

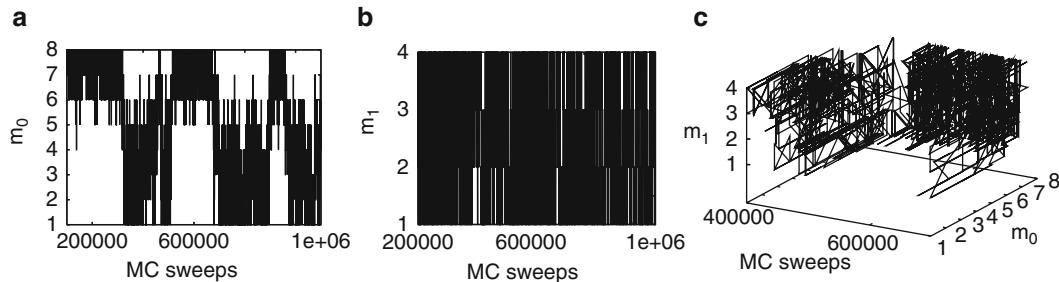


Fig. 1. Time series of the labels of  $T_{m_0}$ ,  $m_0$ , (a) and  $\lambda_{m_1}$ ,  $m_1$ , (b) as functions of MC sweeps, and that of both  $m_0$  and  $m_1$  for the region from 400,000 MC sweeps to 700,000 MC sweeps (c). The results were from one of the replicas (Replica 1). In (a) and (b), MC sweeps start at 100,000 and end at 1,100,000 because the first 100,000 sweeps have been removed from the consideration for thermalization purpose.

two-dimensional REM simulation of 1,000,000 MC sweeps for each replica to determine the weight factor for the two-dimensional ST simulation. At every 20 MC sweeps, either  $T$ -exchange or  $\lambda$ -exchange was tried (the choice of  $T$  or  $\lambda$  was made randomly). In each case, either set of pairs of replicas  $((1,2), \dots, (M-1,M))$  or  $((2,3), \dots, (M,1))$  was also chosen randomly, where  $M$  is  $M_0$  and  $M_1$  for  $T$ -exchange and  $\lambda$ -exchange, respectively.

In Fig. 1 we show the time series of labels of  $T_{m_0}$  (i.e.,  $m_0$ ) and  $\lambda_{m_1}$  (i.e.,  $m_1$ ) for one of the replicas. The replica realized a random walk not only in temperature space but also in  $\lambda$  space. The behavior of  $T$  and  $\lambda$  for other replicas was also similar (see Ref. (156)). From Fig. 1, one finds that the  $\lambda$ -random walk is more frequent than the  $T$ -random walk.

We also show the time series of temperature  $T$ , total energy  $E_{\text{TOT}}$ , conformational energy  $E_C$ , solvation free energy  $E_{\text{SOL}}$ , and end-to-end distance  $D$  for the same replica in Fig. 2. From Figs. 2 (a) and 2(e), we find that at lower temperatures the end-to-end distance is about 8 Å, which is the length of a fully  $\alpha$ -helical conformation, and that at higher temperatures it fluctuates much for a range from 7 Å to 14 Å. It suggests that  $\alpha$ -helix structures exist at low temperatures and random-coil structures occur at high temperatures. There are transitions from/to  $\alpha$ -helix structures to/from random coils during the simulation. It indicates that the REM simulation avoided getting trapped in local-minimum-energy states and sampled a wide conformational space.

The canonical probability distributions of  $E_{\text{TOT}}$  and  $E_{\text{SOL}}$  at the 32 conditions obtained from the two-dimensional REM simulation are shown in Fig. 3. For an optimal performance of the REM simulation, there should be enough overlaps between all pairs of neighboring distributions, which will lead to sufficiently uniform and large acceptance ratios of replica exchanges. There are indeed ample overlaps between the neighboring distributions in Fig. 3.

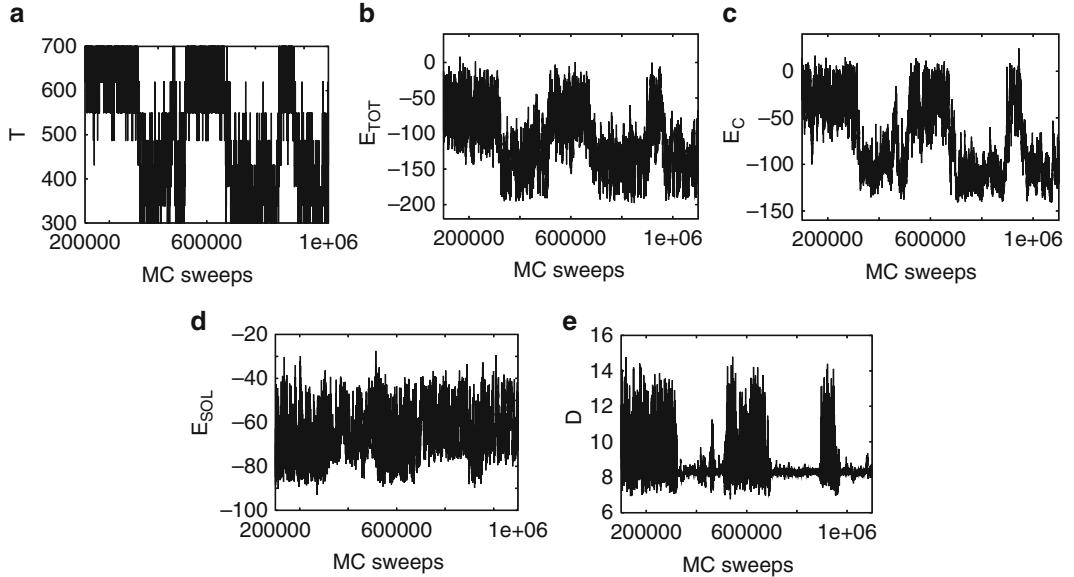


Fig. 2. Time series of the temperature  $T$  (a), total energy  $E_{\text{TOT}}$  (b), conformational energy  $E_C$  (c), solvation free energy  $E_{\text{SOL}}$  (d), and end-to-end distance  $D$  (e) for the same replica as in Fig. 1. The temperature is in K, the energy is in kcal/mol, and the end-to-end distance is in Å.

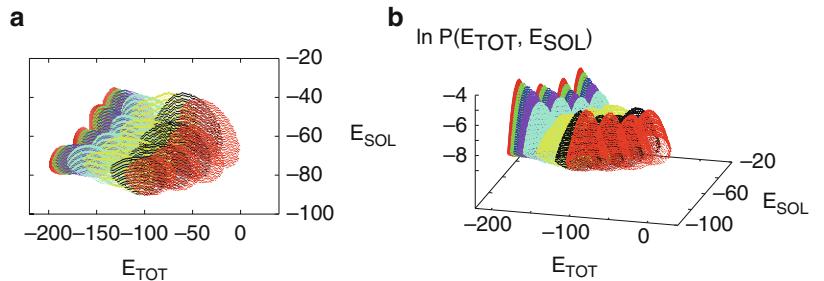


Fig. 3. Contour curves and histograms of distributions of the total energy  $E_{\text{TOT}}$  and the solvation free energy  $E_{\text{SOL}}$  ((a) and (b)) from the two-dimensional REM simulation.

We now use the results of the two-dimensional REM simulation to determine the weight factors for the two-dimensional ST simulation by the multiple-histogram reweighting techniques. Namely, by solving the generalized WHAM equations in Eqs. 84 and 85 with the obtained histograms at the 32 conditions (see Fig. 3), we obtained 32 values of the ST parameters  $f_{m_0, m_1}$  ( $m_0 = 1, \dots, 8$ ;  $m_1 = 1, \dots, 4$ ).

After obtaining the ST weight factor,  $W_{\text{ST}} = \exp(-\beta_{m_0}(E_C + \lambda_{m_1}E_{\text{SOL}}) + f_{m_0, m_1})$ , we carried out the two-dimensional ST simulation of 1,000,000 MC sweeps for data collection after 100,000 MC sweeps for thermalization. At every 20 MC sweeps, either  $T_{m_0}$  or  $\lambda_{m_1}$  was respectively updated to  $T_{m_0 \pm 1}$  or  $\lambda_{m_1 \pm 1}$  (the choice of  $T$  or  $\lambda$  update and the choice of  $\pm 1$  were made randomly).

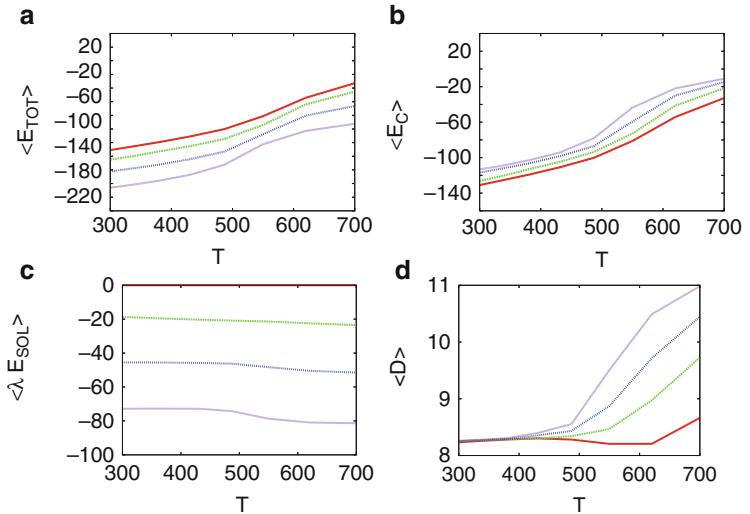


Fig. 4. The average total energy (a), average conformational energy (b), average of  $\lambda \times E_{\text{SOL}}$  (c), and average end-to-end distance (d) with all the  $\lambda$  values as functions of temperature. The lines colored in red, green, blue, and purple are for  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ , and  $\lambda_4$ , respectively. They are in order from above to below in (a) and (c) and from below to above in (b) and (d).

We show the average total energy, average conformational energy, average  $\lambda \times E_{\text{SOL}}$ , and average end-to-end distance in Fig. 4. The results are in good agreement with those of the REM simulation (data not shown).

We found that the results of the two-dimensional ST simulation are in complete agreement with those of the two-dimensional REM simulation for the average quantities. The only difference between the two simulations is the number of replicas. In the present simulation, while the REM simulation used 32 replicas, the ST simulation used only one replica. Hence, we can save much computer power with ST.

A second example of our multidimensional generalized-ensemble simulations is a pressure ST (PST) simulation in the isobaric-isothermal ensemble (157). This simulation performs a random walk in one-dimensional pressure space. The system that we simulated is ubiquitin in explicit water. This system has been studied by high-pressure NMR experiments and known to undergo high-pressure denaturations (179, 180). Ubiquitin has 76 amino acids and it was placed in a cubic box of 6,232 water molecules. Temperature was fixed to be 300 K throughout the simulations, and we prepared 100 values of pressure ranging from 1 bar to 10,000 bar. Temperature and pressure were controlled by Hoover-Langevin method (181), and particle mesh Ewald method (182, 183) was employed for electrostatic interactions. The time step was 2.0 fsec. The force field CHARMM22 (184) with CMAP

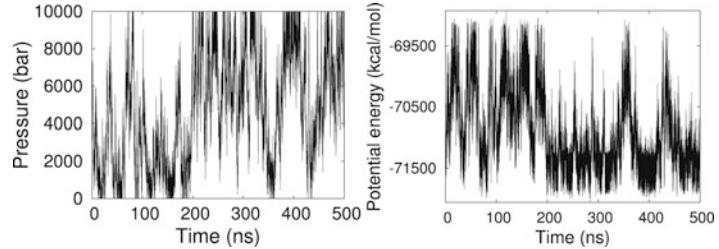


Fig. 5. Time series of pressure (*left*) and potential energy (*right*) during the PST production run.

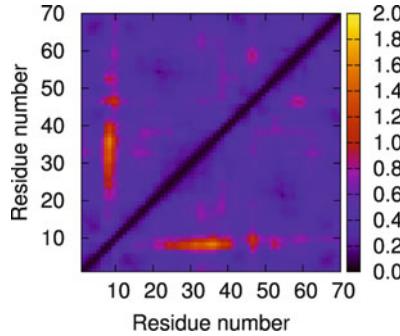


Fig. 6. Fluctuations of distance between pairs of  $C^\alpha$  atoms that was calculated from the PST production run.

(185, 186) and TIP3P water model (184, 187) were used, and the program package NAMD version 2.7b3 (188) was modified to incorporate the PST algorithm.

We first performed 100 independent conventional isobaric-isothermal simulations of 4 nsec with  $T = 300$  K (i.e.,  $M_0 = 1$ ) and 100 values of pressure (i.e.,  $M_1 = 100$ ). Using the obtained histogram  $N_{m_0, m_1}(E, V)$  of potential energy and volume distribution, we obtained the ST parameters  $f_{m_0, m_1}$  by solving the WHAM equations in Eqs. 111 and 112. We then performed the PST production of 500 nsec and repeated it 10 times with different seeds for random numbers (so, the total simulation time for the production run is 5.0  $\mu$ s).

In Fig. 5, we show the time series of pressure and potential energy during the PST production run.

In the figure, we see a random walk in pressure between 1 bar and 10,000 bar. A random walk in potential energy is also observed, and it is anti-correlated with that of pressure, as it should be.

We calculated the fluctuations  $\sqrt{\langle d^2 \rangle - \langle d \rangle^2}$  of the distance  $d$  between pairs of  $C^\alpha$  atoms. The results are shown in Fig. 6.

We see that large fluctuations are observed between residues around 7–10 and around 20–40, which are in accord with the experimental results (179, 180).

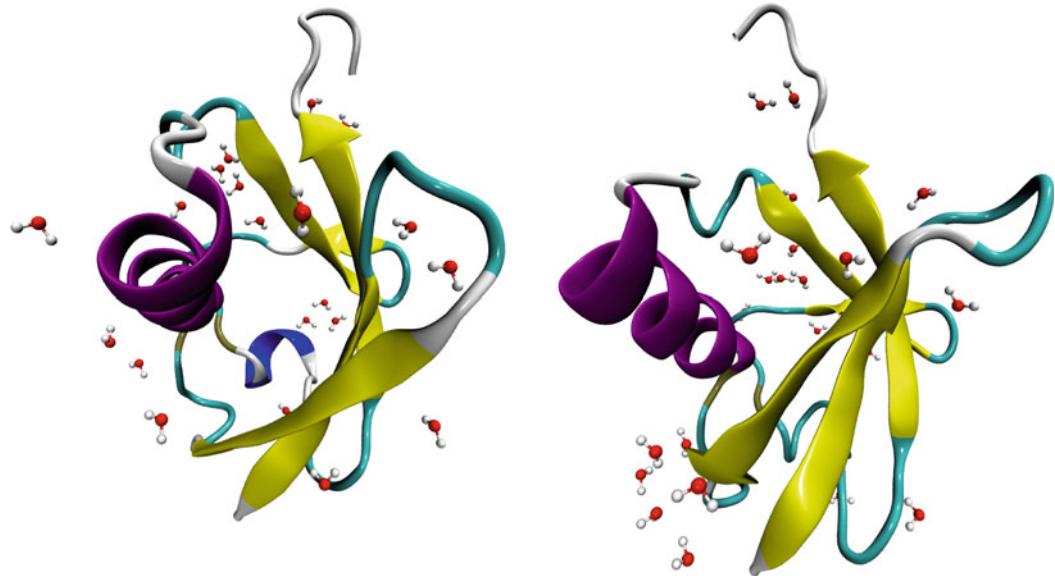


Fig. 7. Snapshots of ubiquitin during the PST production run at low pressure (*left*) and at high pressure (*right*).

The fluctuating distance corresponds to that between the turn region of the  $\beta$ -hairpin and the end of the  $\alpha$ -helix as depicted in Fig. 7. While at low pressure this distance is small, at high pressure it is larger and water comes into the created open region.

---

## 5. Conclusions

In this article we first introduced three well-known generalized-ensemble algorithms, namely, REM, ST, and MUCA, which can greatly enhance conformational sampling of biomolecular systems. We then presented various extensions of these algorithms. Examples are the general formulations of the multidimensional REM, ST, and MUCA. We generalized the original potential energy function  $E_0$  by adding any physical quantities  $V_\ell$  of interest as a new energy term with a coupling constant  $\lambda^{(\ell)}$  ( $\ell = 1, \dots, L$ ). The simulations in multidimensional REM and multidimensional ST algorithms realize a random walk in temperature and  $\lambda^{(\ell)}$  ( $\ell = 1, \dots, L$ ) spaces. On the other hand, the simulation in multidimensional MUCA algorithms realizes a random walk in  $E_0, V_1, \dots, V_L$  spaces.

While the multidimensional REM simulation can be easily performed because no weight factor determination is necessary, the required number of replicas can be quite large and computationally demanding. We thus prefer to use the multidimensional ST or MUCA, where only a single replica is simulated, instead of REM. However, it is very difficult to obtain optimal weight factors

for the multidimensional ST and MUCA. Here, we have proposed a powerful method to determine these weight factors. Namely, we first perform a short multidimensional REM simulation and use the multiple-histogram reweighting techniques to determine the weight factors for multidimensional ST and MUCA simulations.

The multidimensional generalized-ensemble algorithms that were presented in the present article will be very useful for Monte Carlo and molecular dynamics simulations of complex systems such as spin glass, polymer, and biomolecular systems.

## Acknowledgments

Some of the results were obtained by the computations on the supercomputers at the Institute for Molecular Science, Okazaki, and the Institute for Solid State Physics, University of Tokyo, Japan. This work was supported, in part, by Grants-in-Aid for Scientific Research on Innovative Areas (“Fluctuations and Biological Functions”) and for the Next-Generation Super Computing Project, Nanoscience Program, and Computational Materials Science Initiative from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

## References

- Kirkpatrick S, Gelatt CD Jr, Vecchi MP (1983) Optimization by simulated annealing. *Science* 220:671–680
- Nilges M, Clore GM, Gronenborn AM (1988) Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett* 229:317–324
- Brünger AT (1988) Crystallographic refinement by simulated annealing. Application to a 2.8 Å resolution structure of aspartate aminotransferase. *J Mol Biol* 203:803–816
- Wilson SR, Cui W, Moskowitz JW, Schmidt KE (1988) Conformational analysis of flexible molecules—location of the global minimum energy conformation by the simulated annealing method. *Tetrahedron Lett* 29:4373–4376
- Kawai H, Kikuchi T, Okamoto Y (1989) A prediction of tertiary structures of peptide by the Monte Carlo simulated annealing method. *Protein Eng* 3:85–94
- Wilson C, Doniach S (1989) A computer model to dynamically simulate protein folding: studies with crambin. *Proteins* 6:193–209
- Kawai H, Okamoto Y, Fukugita M, Nakazawa T, Kikuchi T (1991) Prediction of  $\alpha$ -helix folding of isolated C-peptide of ribonuclease A by Monte Carlo simulated annealing. *Chem Lett* 1991:213–216
- Okamoto Y, Fukugita M, Nakazawa T, Kawai H (1991)  $\alpha$ -helix folding by Monte Carlo simulated annealing in isolated C-peptide of ribonuclease A. *Protein Eng* 4:639–647
- Hansmann UHE, Okamoto Y (1999) Generalized-ensemble approach for protein folding simulations. In: Stauffer D (ed) Annual Reviews of Computational Physics VI. World Scientific, Singapore, pp 129–157
- Mitsutake A, Sugita Y, Okamoto Y (2001) Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers* 60:96–123
- Sugita Y, Okamoto Y (2002) Free-energy calculations in protein folding by generalized-ensemble algorithms. In: Schlick T, Gan HH (eds) Lecture notes in computational science and engineering. Springer, Berlin, pp 304–332. e-print: cond-mat/0102296

12. Okamoto Y (2004) Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J Mol Graphics Mod* 22:425–439. e-print: cond-mat/0308360
13. Kokubo H, Okamoto Y (2006) Replica-exchange methods and predictions of helix configurations of membrane proteins. *Mol Sim* 32:791–801
14. Itoh SG, Okumura H, Okamoto Y (2007) Generalized-ensemble algorithms for molecular dynamics simulations. *Mol Sim* 33:47–56
15. Sugita Y, Mitsutake A, Okamoto Y (2008) Generalized-ensemble algorithms for protein folding simulations. In: Janke W (ed) Lecture notes in physics. Rugged free energy landscapes: common computational approaches in spin glasses, structural glasses and biological macromolecules. Springer, Berlin, pp 369–407. e-print: arXiv:0707.3382v1 [cond-mat.stat-mech]
16. Okamoto Y (2009) Generalized-ensemble algorithms for studying protein folding. In: Kuwajima K, Goto Y, Hirata F, Kataoka M, Terazima M (eds) Water and Biomolecules. Springer, Berlin, pp 61–95
17. Ferrenberg AM, Swendsen RH (1988) New Monte Carlo technique for studying phase transitions. *Phys Rev Lett* 61:2635–2638
18. Ferrenberg AM, Swendsen RH (1989) New Monte Carlo technique for studying phase transitions errata. *Phys Rev Lett* 63:1658
19. Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo data analysis. *Phys Rev Lett* 63:1195–1198
20. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J Comput Chem* 13:1011–1021
21. Berg BA, Neuhaus T (1991) Multicanonical algorithms for 1st order phase transitions. *Phys Lett* B267:249–253
22. Berg BA, Neuhaus T (1992) Multicanonical ensemble: a new approach to simulate first-order phase transitions. *Phys Rev Lett* 68:9–12
23. Berg BA (2004) Introduction to Monte Carlo simulations and their statistical analysis. World Scientific, Singapore
24. Janke W (1998) Multicanonical Monte Carlo simulations. *Phys A* 254:164–178
25. Lee J (1993) New Monte Carlo algorithm: entropic sampling. *Phys Rev Lett* 71:211–214
26. Lee J (1993) New Monte Carlo algorithm: entropic sampling errata. *Phys Rev Lett* 71:2353
27. Hao WH, Scheraga HA (1994) Monte Carlo simulation of a first-order transition for protein folding. *J Phys Chem* 98:4940–4948
28. Mezei M (1987) Adaptive umbrella sampling—self-consistent determination of the non-Boltzmann bias. *J Comput Phys* 68:237–248
29. Bartels C, Karplus M (1998) Probability distributions for complex systems: adaptive umbrella sampling of the potential energy. *J Phys Chem B* 102:865–880
30. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J Comput Phys* 23:187–199
31. Wang F, Landau DP (2001) Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys Rev Lett* 86:2050–2053
32. Wang F, Landau DP (2001) Determining the density of states for classical statistical models: a random walk algorithm to produce a flat histogram. *Phys Rev E* 64:056101
33. Yan Q, Faller R, de Pablo JJ (2002) Density-of-states Monte Carlo method for simulation of fluids. *J Chem Phys* 116:8745–8749
34. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99:12562–12566
35. Trebst S, Huse DA, Troyer M (2004) Optimizing the ensemble for equilibration in broad-histogram Monte Carlo simulations. *Phys Rev E* 70:046701
36. Berg BA, Celik T (1992) New approach to spin-glass simulations. *Phys Rev Lett* 69:2292–2295
37. Berg BA, Hansmann UHE, Neuhaus T (1993) Simulation of an ensemble with varying magnetic field: a numerical determination of the order-order interface tension in the D=2 Ising model. *Phys Rev B* 47:497–500
38. Janke W, Kappler S (1995) *Phys Rev Lett* 74:212–215
39. Berg BA, Janke W (1998) *Phys Rev Lett* 80:4771–4774
40. Hatano N, Gubernatis JE (2000) A multicanonical Monte Carlo study of the 3D +/- J spin glass. *Prog Theor Phys (Suppl)* 138:442–447
41. Berg BA, Billoire A, Janke W (2000) Spin-glass overlap barriers in three and four dimensions. *Phys Rev B* 61:12143–12150
42. Berg BA, Muguruma C, Okamoto Y (2007) Residual entropy of ordinary ice from multicanonical simulations. *Phys Rev B* 75:092202

43. Hansmann UHE, Okamoto Y (1993) Prediction of peptide conformation by multicanonical algorithm—new approach to the multiple-minima problem. *J Comput Chem* 14:1333–1338
44. Hansmann UHE, Okamoto Y (1994) Comparative study of multicanonical and simulated annealing algorithms in the protein folding problem. *Physica A* 212:415–437
45. Okamoto Y, Hansmann UHE (1995) Thermodynamics of helix-coil transitions studied by multicanonical algorithms. *J Phys Chem* 99:11276–11287
46. Wilding NB (1995) Critical-point and coexistence-curve properties of the Lennard–Jones fluid: a finite-size scaling study. *Phys Rev E* 52:602–611
47. Kolinski A, Galazka W, Skolnick J (1996) On the origin of the cooperativity of protein folding: implications from model simulations. *Proteins* 26:271–287
48. Urakami N, Takasu M (1996) Multicanonical Monte Carlo simulation of a polymer with stickers. *J Phys Soc Jpn* 65:2694–2699
49. Kumar S, Payne P, Vásquez M (1996) Method for free-energy calculations using iterative techniques. *J Comput Chem* 17:1269–1275
50. Hansmann UHE, Okamoto Y, Eisenmenger F (1996) Molecular dynamics, Langevin and hybrid Monte Carlo simulations in a multicanonical ensemble. *Chem Phys Lett* 259:321–330
51. Hansmann UHE, Okamoto Y (1996) Monte Carlo simulations in generalized ensemble: multicanonical algorithm versus simulated tempering. *Phys Rev E* 54:5863–5865
52. Hansmann UHE, Okamoto Y (1997) Numerical comparisons of three recently proposed algorithms in the protein folding problem. *J Comput Chem* 18:920–933
53. Noguchi H, Yoshikawa K (1997) First-order phase transition in a stiff polymer chain. *Chem Phys Lett* 278:184–188
54. Nakajima N, Nakamura H, Kidera A (1997) Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. *J Phys Chem B* 101:817–824
55. Bartels C, Karplus M (1997) Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations. *J Comput Chem* 18:1450–1462
56. Higo J, Nakajima N, Shirai H, Kidera A, Nakamura H (1997) Two-component multicanonical Monte Carlo method for effective conformation sampling. *J Comput Chem* 18:2086–2092
57. Iba Y, Chikenji G, Kikuchi M (1998) Simulation of lattice polymers with multi-self-overlap ensemble. *J Phys Soc Jpn* 67:3327–3330
58. Mitsutake A, Hansmann UHE, Okamoto Y (1998) Temperature dependence of distributions of conformations of a small peptide. *J Mol Graphics Mod* 16:226–238; 262–263
59. Hansmann UHE, Okamoto Y (1999) Effects of side-chain charges on alpha-helix stability in C-peptide of ribonuclease A studied by multicanonical algorithm. *J Phys Chem B* 103:1595–1604
60. Shimizu H, Uehara K, Yamamoto K, Hiwatari Y (1999) Structural phase transition of di-block polyampholyte. *Mol Sim* 22:285–301
61. Ono S, Nakajima N, Higo J, Nakamura H (1999) The multicanonical weighted histogram analysis method for the free-energy landscape along structural transition paths. *Chem Phys Lett* 312:247–254
62. Mitsutake A, Okamoto Y (2000) Helix-coil transitions of amino-acid homo-oligomers in aqueous solution studied by multicanonical simulations. *J Chem Phys* 112:10638–10647
63. Sayano K, Kono H, Gromiha MM, Sarai A (2000) Multicanonical Monte Carlo calculation of the free-energy map of the base-amino acid interaction. *J Comput Chem* 21:954–962
64. Yasar F, Celik T, Berg BA, Meirovitch H (2000) Multicanonical procedure for continuum peptide models. *J Comput Chem* 21:1251–1261
65. Mitsutake A, Kinoshita M, Okamoto Y, Hirata F (2000) Multicanonical algorithm combined with the RISM theory for simulating peptides in aqueous solution. *Chem Phys Lett* 329:295–303
66. Cheung MS, Garcia AE, Onuchic JN (2002) Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc Natl Acad Sci USA* 99:685–690
67. Kamiya N, Higo J, Nakamura H (2002) Conformational transition states of a beta-hairpin peptide between the ordered and disordered conformations in explicit water. *Protein Sci* 11:2297–2307
68. Jang SM, Pak Y, Shin SM (2002) Multicanonical ensemble with Nose–Hoover molecular dynamics simulation. *J Chem Phys* 116:4782–4786
69. Terada T, Matsuo Y, Kidera A (2003) A method for evaluating multicanonical potential function without iterative refinement: application to conformational sampling of a

- globular protein in water. *J Chem Phys* 118:4306–4311
70. Berg BA, Noguchi H, Okamoto Y (2003) Multioverlap simulations for transitions between reference configurations. *Phys Rev E* 68:036126
  71. Bachmann M, Janke W (2003) Multicanonical chain-growth algorithm. *Phys Rev Lett* 91:208105
  72. Okumura H, Okamoto Y (2004) Monte Carlo simulations in multibaric-multithermal ensemble. *Chem Phys Lett* 383:391–396
  73. Okumura H, Okamoto Y (2004) Monte Carlo simulations in generalized isobaric-isothermal ensembles. *Phys Rev E* 70:026702
  74. Okumura H, Okamoto Y (2004) Molecular dynamics simulations in the multibaric-multithermal ensemble. *Chem Phys Lett* 391:248–253
  75. Okumura H, Okamoto Y (2006) Multibaric-multithermal ensemble molecular dynamics simulations. *J Comput Chem* 27:379–395
  76. Itoh SG, Okamoto Y (2004) Multi-overlap molecular dynamics methods for biomolecular systems. *Chem Phys Lett* 400:308–313
  77. Sugita Y, Okamoto Y (2005) Molecular mechanism for stabilizing a short helical peptide studied by generalized-ensemble simulations with explicit solvent. *Biophys J* 88:3180–3190
  78. Itoh SG, Okamoto Y (2007) Effective sampling in the configurational space of a small peptide by the multicanonical-multioverlap algorithm. *Phys Rev E* 76:026705
  79. Munakata T, Oyama S (1996) Adaptation and linear-response theory. *Phys Rev E* 54:4394–4398
  80. Lyubartsev AP, Martinovski AA, Shevkunov SV, Vorontsov-Velyaminov PN (1992) New approach to Monte Carlo calculation of the free energy—method of expanded ensemble. *J Chem Phys* 96:1776–1783
  81. Marinari E, Parisi G (1992) Simulated tempering—a new Monte Carlo scheme. *Europhys Lett* 19:451–458
  82. Marinari E, Parisi G, Ruiz-Lorenzo JJ (1997) Numerical simulations of spin glass systems. In: Young AP (ed) *Spin glasses and random fields*. World Scientific, Singapore, pp 59–98
  83. Escobedo FA, de Pablo JJ (1995) Monte Carlo simulation of the chemical potential of polymers in an expanded ensemble. *J Chem Phys* 103:2703–2710
  84. Irbäck A, Potthast F (1995) Studies of an off-lattice model for protein folding—sequence dependence and improved sampling at finite temperature. *J Chem Phys* 103:10298–10305
  85. Irbäck A, Sandelin E (1999) Monte Carlo study of the phase structure of compact polymer chains. *J Chem Phys* 110:12256–12262
  86. Mitsutake A, Okamoto Y (2000) Replica-exchange simulated tempering method for simulations of frustrated systems. *Chem Phys Lett* 332:131–138
  87. Mitsutake A, Okamoto Y (2004) Replica-exchange extensions of simulated tempering method. *J Chem Phys* 121:2491–2504
  88. Park S, Pande V (2007) Choosing weights for simulated tempering. *Phys Rev E* 76:016703
  89. Zheng L, Chen M, Yang W (2009) Simultaneous escaping of explicit and hidden free energy barriers: application of the orthogonal space random walk strategy in generalized ensemble based conformational sampling. *J Chem Phys* 130:234105
  90. Zhang C, Ma J (2010) Enhanced sampling and applications in protein folding in explicit solvent. *J Chem Phys* 132:244101
  91. Kim J, Straub JE (2010) Generalized simulated tempering for exploring strong phase transitions. *J Chem Phys* 133:154101
  92. Hukushima K, Nemoto K (1996) Exchange Monte Carlo method and application to spin glass simulations. *J Phys Soc Jpn* 65:1604–1608
  93. Hukushima K, Takayama H, Nemoto K (1996) Application of an extended ensemble method to spin glasses. *Int J Mod Phys C* 7:337–344
  94. Geyer CJ (1991) Markov chain Monte Carlo maximum likelihood. In: Keramidas EM (ed) *Computing science and statistics: proceedings 23rd symposium on the interface*. Interface Foundation, Fairfax Station, pp 156–163
  95. Swendsen RH, Wang J-S (1986) Replica Monte Carlo simulation of spin glasses. *Phys Rev Lett* 57:2607–2609
  96. Kimura K, Taki K (1991) Time-homogeneous parallel annealing algorithm. In: Vichnevetsky R, Miller, JJH (eds) *IMACS 91 Proceedings of the 13th World Congress on Computation and Applied Mathematics*, vol 2. pp 827–828
  97. Frantz DD, Freeman DL, Doll JD (1990) Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking—applications to atomic clusters. *J Chem Phys* 93:2769–2784
  98. Tesi MC, van Rensburg EJJ, Orlandini E, Whittington SG (1996) Monte Carlo study of the interacting self-avoiding walk model in three dimensions. *J Stat Phys* 82:155–181

99. Iba Y (2001) Extended ensemble Monte Carlo. *Int J Mod Phys C* 12:623–656
100. Hansmann UHE (1997) Parallel tempering algorithm for conformational studies of biological molecules. *Chem Phys Lett* 281:140–150
101. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151
102. Wu MG, Deem MW (1999) Efficient Monte Carlo methods for cyclic peptides. *Mol Phys* 97:559–580
103. Sugita Y, Kitao A, Okamoto Y (2000) Multi-dimensional replica-exchange method for free-energy calculations. *J Chem Phys* 113:6042–6051
104. Woods CJ, Essex JW, King MA (2003) The development of replica-exchange-based free-energy methods. *J Phys Chem B* 107:13703–13710
105. Sugita Y, Okamoto Y (2000) Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape. *Chem Phys Lett* 329:261–270
106. Gront D, Kolinski A, Skolnick J (2000) Comparison of three Monte Carlo conformational search strategies for a proteinlike homopolymer model: folding thermodynamics and identification of low-energy structures. *J Chem Phys* 113:5065–5071
107. Verkhivker GM, Rejto PA, Bouzida D, Arthurs S, Colson AB, Freer ST, Gehlhaar DK, Larson V, Luty BA, Marrone T, Rose PW (2001) Parallel simulated tempering dynamics of ligand-protein binding with ensembles of protein conformations. *Chem Phys Lett* 337:181–189
108. Fukunishi F, Watanabe O, Takada S (2002) On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. *J Chem Phys* 116:9058–9067
109. Mitsutake A, Sugita Y, Okamoto Y (2003) Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. I. Formulation and benchmark test. *J Chem Phys* 118:6664–6675
110. Mitsutake A, Sugita Y, Okamoto Y (2003) Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. II. Application to a more complex system. *J Chem Phys* 118:6676–6688
111. Sikorski A, Romiszowski P (2003) Thermodynamical properties of simple models of protein-like heteropolymers. *Biopolymers* 69:391–398
112. Lin CY, Hu CK, Hansmann UHE (2003) Parallel tempering simulations of HP-36. *Proteins* 52:436–445
113. La Penna G, Mitsutake A, Masuya M, Okamoto Y (2003) Molecular dynamics of C-peptide of ribonuclease A studied by replica-exchange Monte Carlo method and diffusion theory. *Chem Phys Lett* 380:609–619
114. Kokubo H, Okamoto Y (2004) Prediction of membrane protein structures by replica-exchange Monte Carlo simulations: case of two helices. *J Chem Phys* 120:10837
115. Kokubo H, Okamoto Y (2009) Analysis of helix-helix interactions of bacteriorhodopsin by replica-exhcange simulations. *Biophys J* 96:765–776
116. Falcioni M, Deem DW (1999) A biased Monte Carlo scheme for zeolite structure solution. *J Chem Phys* 110:1754–1766
117. Yan Q, de Pablo JJ (1999) Hyper-parallel tempering Monte Carlo: application to the Lennard-Jones fluid and the restricted primitive model. *J Chem Phys* 111:9509–9516
118. Nishikawa T, Ohtsuka H, Sugita Y, Mikami M, Okamoto Y (2000) Replica-exchange Monte Carlo method for Ar fluid. *Prog Theor Phys (Suppl)* 138:270–271
119. Kofke DA (2002) On the acceptance probability of replica-exchange Monte Carlo trials. *J Chem Phys* 117:6911–6914
120. Okabe T, Kawata M, Okamoto Y, Mikami M (2001) Replica-exchange Monte Carlo method for the isobaric-isothermal ensemble. *Chem Phys Lett* 335:435–439
121. Ishikawa Y, Sugita Y, Nishikawa T, Okamoto Y (2001) Ab initio replica-exchange Monte Carlo method for cluster studies. *Chem Phys Lett* 333:199–206
122. Garcia AE, Sanbonmatsu KY (2001) Exploring the energy landscape of a beta hairpin in explicit solvent. *Proteins* 42:345–354
123. Zhou RH, Berne BJ, Germain R (2001) The free energy landscape for beta hairpin folding in explicit water. *Proc Natl Acad Sci USA* 98:14931–14936
124. Garcia AE, Sanbonmatsu KY (2002)  $\alpha$ -Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci USA* 99:2782–2787
125. Zhou RH, Berne BJ (2002) *Proc Natl Acad Sci USA* 99:12777–12782

126. Feig M, MacKerell AD, Brooks CL III (2003) Force field influence on the observation of pi-helical protein structures in molecular dynamics simulations. *J Phys Chem B* 107:2831–2836
127. Rhee YM, Pande VS (2003) Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophys J* 84:775–786
128. Paschek D, Garcia AE (2004) Reversible temperature and pressure denaturation of a protein fragment: a replica exchange molecular dynamics simulation study. *Phys Rev Lett* 93:238105
129. Paschek D, Gnanakaran S, Garcia AE (2005) Simulations of the pressure and temperature unfolding of an  $\alpha$ -helical peptide. *Proc Natl Acad Sci USA* 102:6765–6770
130. Pitera JW, Swope W (2003) Understanding folding and design: replica-exchange simulations of “Trp-cage” fly miniproteins. *Proc Natl Acad Sci USA* 100:7587–7592
131. Ohkubo YZ, Brooks CL III (2003) Exploring Flory’s isolated-pair hypothesis: Statistical mechanics of helix-coil transitions in polyalanine and the C-peptide from RNase A. *Proc Natl Acad Sci USA* 100:13916–13921
132. Fenwick MK, Escobedo FA (2003) Hybrid Monte Carlo with multidimensional replica exchanges: conformational equilibria of the hypervariable regions of a llama V-HH antibody domain. *Biopolymers* 68:160–177
133. Xu HF, Berne BJ (2000) Multicanonical jump walking annealing: an efficient method for geometric optimization. *J Chem Phys* 112:2701–2708
134. Faller R, Yan Q, de Pablo JJ (2002) Multicanonical parallel tempering. *J Chem Phys* 116:5419–5423
135. Fenwick MK, Escobedo FA (2003) Expanded ensemble and replica exchange methods for simulation of protein-like systems. *J Chem Phys* 119:11998–12010
136. Murata K, Sugita Y, Okamoto Y (2004) Free energy calculations for DNA base stacking by replica-exchange umbrella sampling. *Chem Phys Lett* 385:1–7
137. Felts AK, Harano Y, Gallicchio E, Levy RM (2004) Free energy surfaces of  $\beta$ -hairpin and  $\alpha$ -helical peptides generated by replica exchange molecular dynamics with the AGBNP implicit solvent model. *Proteins* 56:310–321
138. Mitsutake A, Kinoshita M, Okamoto Y, Hirata F (2004) Combination of the replica-exchange Monte Carlo method and the reference interaction site model theory for simulating a peptide molecule in aqueous solution. *J Phys Chem B* 108:19002–19012
139. Baumketner A, Shea JE (2005) Free energy landscapes for amyloidogenic tetrapeptides dimerization. *Biophys J* 89:1493–1503
140. Yoda T, Sugita Y, Okamoto Y (2007) Cooperative folding mechanism of a beta-hairpin peptide studied by a multicanonical replica-exchange molecular dynamics simulation. *Proteins* 66:846–859
141. Roitberg AE, Okur A, Simmerling C (2007) Coupling of replica exchange simulations to a non-Boltzmann structure reservoir. *J Phys Chem B* 111:2415–2418
142. Rosta E, Buchete N-Y, Hummer G (2009) Thermostat artifacts in replica exchange molecular dynamics simulations. *J Chem Theory Comput* 5:1393–1399
143. Yoda T, Sugita Y, Okamoto Y (2010) Hydrophobic core formation and dehydration in protein folding studied by generalized-ensemble simulations. *Biophys J* 99:1637–1644
144. De Simone A, Derreumaux P (2010) Low molecular weight oligomers of amyloid peptides display  $\beta$ -barrel conformations: a replica exchange molecular dynamics study in explicit solvent. *J Chem Phys* 132:165103
145. Hukushima K (1999) Domain-wall free energy of spin-glass models: numerical method and boundary conditions. *Phys Rev E* 60:3606–3614
146. Whitfield TW, Bu L, Straub JE (2002) Generalized parallel sampling. *Physica A* 305:157–171
147. Kwak W, Hansmann UHE (2005) Efficient sampling of protein structures by model hopping. *Phys Rev Lett* 95:138102
148. Bunker A, Dünweg B (2000) Parallel excluded volume tempering for polymer melts. *Phys Rev E* 63:016701
149. Liu P, Kim B, Friesner RA, Bern BJ (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc Natl Acad Sci USA* 102:13749–13754
150. Affentranger R, Tavernelli I, Di Iorio EE (2006) A novel Hamiltonian replica exchange MD protocol to enhance protein conformational space sampling. *J Chem Theory Comput* 2:217–228
151. Lou H, Cukier RI (2006) Molecular dynamics of apo-adenylate kinase: a distance replica exchange method for the free energy of conformational fluctuations. *J Phys Chem B* 110:24121–24137

152. Mu Y (2009) Dissociation aided and side chain sampling enhanced Hamiltonian replica exchange. *J Chem Phys* 130:164107
153. Itoh SG, Okumura H, Okamoto Y (2010) Replica-exchange method in van der Waals radius space: overcoming steric restrictions for biomolecules. *J Chem Phys* 132:134105
154. Mitsutake A, Okamoto Y (2009) From multidimensional replica-exchange method to multidimensional multicanonical algorithm and simulated tempering. *Phys Rev E* 79:047701
155. Mitsutake A, Okamoto Y (2009) Multidimensional generalized-ensemble algorithms for complex systems. *J Chem Phys* 130:214105
156. Mitsutake A (2009) Simulated-tempering replica-exchange method for the multidimensional version. *J Chem Phys* 131:094105
157. Mori Y, Okamoto Y (2010) Generalized-ensemble algorithms for the isobaric-isothermal ensemble. *J Phys Soc Jpn* 79:074003
158. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092
159. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* 52:255–268
160. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* 81:511–519
161. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:124105
162. Berg BA (2003) Multicanonical simulations step by step. *Comp Phys Commun* 153:397–406
163. Mori Y, Okamoto Y (2010) Replica-exchange molecular dynamics simulations for various constant temperature algorithms. *J Phys Soc Jpn* 79:074001
164. Allen MP, Tildesley DJ (1987) Computer Simulation of Liquids. Oxford, New York, p 259
165. Andersen HG (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72:2384–2393
166. Hoover WG, Ladd AJC, Moran B (1982) High strain rate plastic flow studied via non-equilibrium molecular dynamics. *Phys Rev Lett* 48:1818–1820
167. Evans DJ (1983) Computer experiment for non-linear thermodynamics of couette flow. *J Chem Phys* 78:3297–3302
168. Evans DJ, Morriss GP (1983) The isothermal isobaric molecular dynamics ensemble. *Phys Lett A* 98:433–436
169. Hoover WG (1985) Canonical dynamics—equilibrium phase space distributions. *Phys Rev A* 31:1695–1697
170. Martyna GJ, Klein ML, Tuckerman M (1992) Nosé–Hoover chains—the canonical ensemble via continuous dynamics. *J Chem Phys* 97:2635–2643
171. Bond SD, Leimkuhler BJ, Laird BB (1999) The Nosé–Poincaré method for constant temperature molecular dynamics. *J Comput Phys* 151:114–134
172. McDonald IR (1972) NpT-ensemble Monte Carlo calculations for binary liquid mixtures. *Mol Phys* 23:41–58
173. Myers JK, Pace CN, Scholtz JM (1997) A direct comparison of helix propensity in proteins and peptides. *Proc Natl Acad Sci USA* 94:2833–2837
174. Momany FA, McGuire RF, Burgess AW, Scheraga HA (1975) Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J Phys Chem* 79:2361–2381
175. Némethy G, Pottle MS, Scheraga HA (1983) Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. *J Phys Chem* 87:1883–1887
176. Sippl MJ, Némethy G, Scheraga HA (1984) Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O–H...O=C hydrogen bonds from packing configurations. *J Phys Chem* 88:6231–6233
177. Ooi T, Oobatake M, Némethy G, Scheraga HA (1987) Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc Natl Acad Sci USA* 84:3086–3090
178. Masuya M, unpublished; see <http://biocomputing.cc/nsol/>.
179. Kitahara R, Akasaka K (2003) Close identity of a pressure-stabilized intermediate with a kinetic intermediate in protein folding. *Proc Natl Acad Sci USA* 100:3167–3172
180. Kitahara R, Yokoyama S, Akasaka K (2005) NMR snapshots of a fluctuating protein structure: ubiquitin at 30 bar–3 kbar. *J Mol Biol* 347:277–285

181. Quigley D, Probert MJ (2004) Landau–de Gennes dynamics in constant pressure extended systems. *J Chem Phys* 120:11432–11441
182. Darden T, York D, Pedersen L (1993) Particle mesh Ewald—an  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
183. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593
184. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkner M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
185. MacKerell AD Jr, Feig M, Brooks CL III (2004) Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 126:698–699
186. MacKerell AD Jr, Feig M, Brooks CL III (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25:1400–1415
187. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
188. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802

# Chapter 8

## Force Fields for Classical Molecular Dynamics

Luca Monticelli and D. Peter Tieleman

### Abstract

In this chapter we review the basic features and the principles underlying molecular mechanics force fields commonly used in molecular modeling of biological macromolecules. We start by summarizing the historical background and then describe classical pairwise additive potential energy functions. We introduce the problem of the calculation of nonbonded interactions, of particular importance for charged macromolecules. Different parameterization philosophies are then presented, followed by a section on force field validation. We conclude with a brief overview on future perspectives for the development of classical force fields.

**Key words:** Force field, Molecular mechanics, Molecular dynamics, Computer simulation

---

### 1. Introduction: Simple Molecular Mechanics Force Fields

Computer simulations are a powerful tool to study the structure and dynamics of biological macromolecules. In principle, all details of molecular structures and interactions can be predicted from first principles using quantum mechanics. Unfortunately, many of the problems we are interested in cannot be tackled with quantum mechanics due to the extremely high computational cost. The development of simplifications is therefore necessary in the investigation of structure and dynamics of biological macromolecules.

In force field methods (also known as molecular mechanics methods) the electronic degrees of freedom of the molecules are ignored, and only motions of the nuclei are calculated. The fundamental assumption underlying all molecular mechanics methodologies is the Born–Oppenheimer approximation, which allows separating nuclear and electronic motions (which happen on different time scales) and therefore to write the energy of a system as a function of nuclear coordinates only. Two additional assumptions are also involved in all common force fields: additivity and

transferability. Additivity means that the (potential) energy of any system can be written as a sum of different potentials with a simple physical interpretation (bond deformations, electrostatics, dispersion forces, etc.). Transferability means that potential energy functions developed on a small set of molecules can be applied to a much wider range of molecules with similar chemical groups. The correctness of molecular mechanics force fields depends on the validity of these assumptions.

Most classical force fields rely on five terms with a simple physical interpretation: there are potential energy terms associated with deformation of bond and angle geometry (stretching/compression of bonds, bending of angles), terms associated with the rotation about certain dihedral angles (torsions), and the so-called nonbonded terms, describing the electrostatic interactions and terms describing the dispersion interactions and repulsion when atoms overlap (van der Waals forces). More complex force fields include additional terms accounting for atomic polarizability (see next chapter) and more complex coupling terms such as cross coupling between bonds and angle.

A force field consists of the equations chosen to model the potential energy and their associated parameters. It is important to keep in mind that molecular mechanics force fields are empirical: the separation of the potential energy in terms with a simple physical interpretation is not strictly correct, as there is no unique way to “translate” quantum mechanical effects into classical mechanics equations. As a consequence, there is no unique solution for the optimal set of functions and parameters. In other words, a force field is supposed to reproduce a certain potential energy surface (PES) derived from quantum mechanics, but different force fields might reproduce the same PES by summing up contributions that are different. For this reason, it is in general not correct to mix energy functions and parameters from different force fields to represent a certain molecule.

Despite all this, current force fields for biological macromolecules look very similar to each other in their functional forms and often in the parameters. The way energy contributions are separated into different terms is basically the same as devised by the pioneers of the field in the 1960s (see next section). The main reason for the observed “convergence” of different force fields towards the same functional form seems to be convenience. The separation of energy terms allows an easy interpretation of how the different components affect the ability to reproduce molecular structure and dynamics. Also, separability of energy terms is the basis for multiple time step algorithms for molecular dynamics simulations, in which bonded and nonbonded components are updated with different frequencies, allowing for faster simulations. Finally, bond stretching and bending parameters for biomolecules turn out to be fairly independent of the rest of the force field. Therefore, adapting them from different force fields is sometimes acceptable.

In the following sections we briefly survey the historical background on molecular mechanics force fields, with particular reference to the case of proteins. Then we describe in some detail each term of classical force fields, and the different methodologies that have been adopted for the calculation of nonbonded interactions. We touch upon the different parameterization philosophies and the problem of force field validation, and we conclude with an outlook on future developments.

---

## 2. Historical Background on Force Fields

While the first molecular mechanics calculations date back to the 1940s (see ref. 1), significant progress in the field had to wait until the end of the 1960s, when the first computers became available. Pioneers in the development of classical force fields for proteins and organic molecules were Shneior Lifson (Weizmann Institute, Israel), Harold Scheraga (Cornell University, USA), and Norman Allinger (Wayne State and University of Georgia, USA). Most of the programs and force fields in use today for organic and biological macromolecules are based on the work of these pioneers and their coworkers (1, 2).

The first force fields were parameterized to reproduce structural properties and vibrational spectra (3). After the first atomistic simulations of a protein (4), the first coarse-grained model (5), and the first QM-MM simulations (6) (all completed in the mid-1970s), more systematic efforts to develop consistent force fields were made. A complete review of available force fields is beyond the scope of the present work. Good reviews on this topic can be found in refs. (1, 7, 8). We briefly summarize the features and the evolution of some of the force fields commonly used for biological macromolecules.

Among the programs and force fields that were developed in the early 1980s, widely used were Amber, CHARMM, and OPLS. In the initial Amber force field (9), polar hydrogen atoms were explicitly represented but the ones bound to carbon atoms were treated as “united atoms,” and most parameters were developed and tested primarily on gas-phase simulations. Partial charges were derived from quantum mechanics calculations via fitting partial atomic charges to the electrostatic potential (and are generally referred to as ESP charges) (10). The united atom (UA) approach proved to be insufficient in the treatment of aromatic rings and in the conformational equilibria of five-membered aliphatic rings, as well as in the prediction of vibrational spectra. In the 1990s, the need to produce potentials suitable for condensed phase simulations induced the Amber developers to include liquid phase properties (densities, heats of vaporization) in the further developments of

this force field. A major update of the AMBER force field was published in 1995 and is still widely used (11). As a further development, Kollman and coworkers developed a program (named antechamber) that automatically generates force field parameters for an arbitrary molecule consistent with the rest of the Amber force field (12).

The CHARMM program was also developed originally in the early 1980s and included its own force field (13). Like for Amber, the initial CHARMM force field represented explicitly only polar hydrogens and used a united atom representation for the nonpolar hydrogens. Parameters were developed and tested mainly on gas-phase simulations, but this parameterization also used more sophisticated fits to quantum mechanics calculations, typically including hydrogen bonded complexes between water and different molecular fragments. As for Amber, in the 1990s the development of CHARMM started to take into account liquid phase properties directly in the parameterization. Although long-range electrostatic interactions are usually taken into account using Ewald summation techniques (see below), the CHARMM force field still enforces neutral charge groups (groups of adjacent atoms must have zero net charge). CHARMM has seen several major updates, reviewed in ref. (14).

The development of the OPLS force field was also started in the early 1980s. In contrast with Amber and CHARMM, OPLS aimed from the beginning at reproducing liquid state properties like densities and heats of vaporization. This was achieved by deriving nonbonded interactions by fitting thermodynamic properties of simple liquids. Like the previous two force fields, OPLS also used a united atoms representation initially and later moved to an all-atom representation (15–17). Bonded interactions were taken directly from the 1984 Amber force field initially (UA version), and the 1986 Amber force field later (AA). Like CHARMM, OPLS also uses neutral charge groups.

A number of other force fields are currently available for simulations of biological macromolecules. Among them, GROMOS (18, 19) is widely used, while ENCAD (20) is quite popular in the field of protein folding. The latest versions of the GROMOS force field (53A5 and 53A6) stand out as the first general purpose biomolecular force field parameterized explicitly to reproduce free energies of solvation. The choice of free energies of solvation as a target is motivated by their great importance in a number of biological processes, including the self-assembly of biological macromolecules, molecular transport across membranes and protein folding. One limitation of this force field is the use of two different sets of partial atomic charges to be used in polar and apolar environments. As a consequence, it is not clear which set of charges is more appropriate for a molecule in a heterogeneous environment. Indeed, this may be an example where the classical force field approach breaks down.

Most major force fields contain parameters for different polar and apolar solvents, common organic molecules, proteins, carbohydrates, nucleic acids, and some common phospholipids. However, a large fraction of efforts in force field development during the past few decades has been devoted to the improvement of protein force fields, with less work carried out on the other classes of biological macromolecules. Only a few specialized force fields are in common use today for sugars, nucleic acids, and lipids. Advantages and disadvantages of these “specialized” force fields are described in the chapters dedicated to each of these classes of molecules (Chaps. 15–18).

### 3. Classical Pairwise Additive Force Fields

As mentioned above, the expression for the potential energy of a molecular system that is used most frequently for simple organic molecules and biological macromolecules is the following:

$$\begin{aligned}
 V(r) = & \sum_{\text{bonds}} \frac{k_d}{2} (d - d_0)^2 + \sum_{\text{angles}} \frac{k_\theta}{2} (\theta - \theta_0)^2 \\
 & + \sum_{\text{dihedrals}} \frac{k_\phi}{2} (1 + \cos(n\phi - \phi_0)) + \sum_{\text{impropers}} \frac{k_\psi}{2} (\psi - \psi_0)^2 \\
 & + \sum_{\substack{\text{non-bonded} \\ \text{pairs}(i,j)}} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{\substack{\text{non-bonded} \\ \text{pairs}(i,j)}} \frac{q_i \times q_j}{4\epsilon_D r_{ij}}
 \end{aligned} \tag{1}$$

The bond stretching term is usually represented with a harmonic potential (the first term in Eq. 1).  $k_d$  represents the force constant of the bond and is generally very high, indicating that it takes a large amount of energy to stretch or compress a chemical bond significantly.  $d_0$  represents the reference bond length, which is the value of the bond length when all other terms in the potential energy function are zero (for example, in a harmonic oscillator in vacuum). If other energy terms are not zero (as it normally happens in large molecules), they will also influence the bond length, which can very well be different from  $d_0$  at equilibrium.

The harmonic potential is a good approximation to the energy of a system only for small deviations from the reference bond length. For larger deviations, the Morse potential provides a more accurate representation of the potential energy for bond stretching:

$$V(d) = D_e \{1 - \exp[-\alpha(d - d_0)]\}^2. \tag{2}$$

In this case,  $D_e$  is the depth of the potential energy minimum,  $\alpha$  is related to the force constant, and  $d_0$  is the reference bond length.

To improve the description of the bond stretching/compression energy profile using additional terms beyond the harmonic term can be used:

$$V(d) = \sum_{\text{bonds}} \frac{k_b}{2} (d - d_0)^2 \left[ 1 - k'(d - d_0) - k''(d - d_0)^2 - k'''(d - d_0)^3 \dots \right]. \quad (3)$$

Since large deviations from the reference bond length are rare in simulations of biological macromolecules, potential energy functions of this kind are rarely used.

The second term in Eq. 1 represents the potential energy change upon deformation of angles. Values of force constants are typically lower than for bond stretching, indicating that it takes less energy for a bond angle to deviate from its reference value. Like for bonds, the accuracy of the force field can be increased by incorporating higher order contributions. This is sometimes useful for reproducing energies of highly strained molecules, but in general higher order terms are not included in common force fields.

The third term in Eq. 1 is often referred to as the torsional term and represents the potential energy of the molecular system as a function of the rotation about each dihedral angle. The energies involved in this case are significantly lower than for bond stretching and angle bending. Torsional terms include atoms that are separated by three bonds, and are therefore in position 1–4 relative to each other. While the interactions in bonds and angles are usually considered as devoid of any van der Waals and electrostatic contribution, interactions involving atoms in position 1–4 usually include a van der Waals and Coulomb component, often scaled by some factor. There is no unique way of determining the balance between the nonbonded and torsional component in the potential energy profile observed upon rotation of a dihedral angle. Usually force field developers try to reproduce such potential energy profiles (obtained from quantum mechanics calculations, see below) using a sum of all three components (van der Waals, electrostatics, and torsional).

In the potential energy function:

$$V(\phi) = \sum_{\text{dihedrals}} \frac{k_\phi}{2} (1 + \cos(n\phi - \phi_0)). \quad (4)$$

$k_\phi$  is related to the barrier height,  $n$  to the number of minima in the energy function (usually referred to as multiplicity) and  $\phi_0$  is named the phase factor, determining the position of the minima. There often is more than one term for each dihedral in a molecule, yielding a truncated cosine series. An alternative way to express the torsional energy is the following:

$$V(\phi) = \sum_{n=\theta}^N C_n \cos(\phi)^n, \quad (5)$$

with a variable number of terms, usually between 4 and 6. Equation 4 is too symmetric, with the same well depth and barrier height for each well and barrier, and normally combines with 1–4 nonbonded interactions to produce the desired torsion profile. Equation 5 is more flexible and sometimes implies 1–4 interactions that are set to zero. Both are valid approaches to describe the energy landscape of dihedral rotations, but they are not entirely equivalent. Without 1–4 nonbonded interactions there is no coupling of dihedral energies as function of dihedral angle with bonds and angle terms because these modify the 1–4 distances in addition to the dihedral angle term.

The term  $V(\psi) = \sum_{\text{impropers}} \frac{k_\psi}{2} (\psi - \psi_0)^2$  is usually introduced in order to preserve planarity of groups with flat geometry, and sometimes the chirality of certain groups. In the first case, the term provides a penalty function for bending out-of-plane. A good example of this is the peptide (amide) bond in proteins, or aromatic rings in some amino acid side chains.

As mentioned in the introduction, the separation of energy terms between bond stretching, angle bending, etc., is not justified on the basis of the quantum nature of intra- and intermolecular interactions, but it is a very convenient approximation and it provides useful insight into the determinants of molecular structure and dynamics. One of the limitations of this approximation is found when one tries to accurately reproduce vibrational spectra. In real systems, equilibrium bond lengths are coupled to bond angles and torsional energies. So-called cross-terms can be introduced in force fields to take into account the interdependence of bonds and angles, bonds and torsions, or angle and torsions. These are generally sufficient to reproduce accurately both structural properties, but sometimes not vibrational frequencies, which might depend on the coupling between more than two terms. One simple case of cross term is the Urey-Bradley potential, in which the variation in energy upon angle bending is described as a function of the distance between atoms in position 1–3:

$$V(r_{j,3}) = \sum_{\text{angles}} \frac{k_{r_{1,3}}}{2} (r_{1,3} - r_{1,3}^0)^2 \quad (6)$$

Cross-terms are rarely used in large-scale simulations of biological macromolecules.

The fifth and sixth term in Eq. 1 represent the so-called nonbonded terms, which are calculated for pairs of atoms separated by three or more bonds and between atoms in different molecules. The van der Waals component of the potential:

$$V(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (7)$$

is known as Lennard-Jones 12-6 potential. It contains an attractive and a repulsive term. The physical origin of the attractive term lies in the dispersion forces generated between instantaneous dipoles, which arise from fluctuations in electronic charge distributions in all molecules. The quantum mechanical origin of dispersion forces was first explained by London (7). A simple Drude model can be used to explain the dispersion interaction in classical terms (7). The repulsive term reflects the observation that, below a certain distance (typically around 0.3 nm), atoms repel each other. Physically, this effect is due to the Pauli exclusion principle. A realistic functional form for the repulsive potential is:  $V(r) = \exp(-2kr/a_0)$ , where  $a_0$  is the Bohr radius. In practice, the calculation of exponential functions is avoided as they are computationally expensive. The form with a power 12 repulsive term is a reasonable approximation of the exponential function in the interesting range of distances, and is very fast to calculate.

Electric charge in molecules is not distributed evenly, because different atoms have different power to attract electrons (i.e., different electronegativity). Some electrostatic properties of molecules can be calculated from first principles: for example, the electrostatic potential can be calculated directly from the electron density, which is defined by the wave function. One possible way to represent the electrostatic potential of a molecule is by the sum of electrostatic potentials generated by charges placed on atomic nuclei, i.e., partial atomic charges. This is a convenient representation, as the electrostatic potential can easily be calculated as:

$$V(r) = \sum_{\substack{\text{non-bonded} \\ \text{pairs } i,j}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (8)$$

where  $r_{ij}$  is the distance between nuclei  $i$  and  $j$ . Atom-centered charges have the advantage that Coulomb forces are easy to calculate and act directly on the nuclei, which is very convenient if one wants to use those forces in molecular mechanics simulations. Therefore, a lot of effort has been put into developing methods to determine partial charges that reproduce electrostatic properties of molecules, and in particular the electrostatic potential obtained from quantum mechanics calculations (7).

Unfortunately it is not always possible to reproduce the electrostatic potential (or other electrostatic properties of molecules) with atom-centered partial charges. For example, an oxygen molecule has no net charge and no dipole moment but it does have a significant quadrupole moment. This quadrupole moment is important for several molecular properties but cannot be reproduced using only atom-centered charges. Also, atom-centered partial charges imply an isotropic charge distribution around the nucleus, which is often a significant approximation. There are

several ways to represent the asymmetric charge distribution in molecules, e.g., via distributed multipoles. We do not treat this subject here, as these methods are rarely used in simulations on biological macromolecules (8).

We should point out that, contrary to the electrostatic potential, the partial charge distribution in a molecule is not an observable quantity. Although electrostatic properties of molecules are a consequence of its electronic charge distribution, it is not possible to determine partial atomic charges unambiguously from first principles. Remarkably, the electrostatic potential of a molecule can be reproduced equally well by different distributions of atom-centered partial charges. Also, the charge distribution within a molecule and therefore its electrostatic potential depends on the molecular conformation and on the chemical environment. Thus not only is there no unique solution to the problem of determining partial atomic charges, but the solution should also depend on the conformation of the molecule and its environment. This problem can be tackled by the so-called polarizable force fields (see next chapter).

#### **4. Calculation of Nonbonded Interactions: Methods, Challenges and Limitations**

The calculation of nonbonded interactions in biomolecular simulations is a key issue. Electrostatic interactions are very long-ranged, as they decay as  $1/\text{distance}$ . The dispersion contribution decays much faster (sixth power of the distance), but they are all of the same sign (attractive) and their effect is also long-ranged.

Numerically, these problems occur elsewhere in science. A  $1/\text{distance}$  dependence in particular also describes gravity. Over time, several approaches have been adopted in molecular dynamics simulations to calculate long-range interactions. They fall into a number of different categories.

In the simplest case, interactions are simply truncated after a certain distance, typically 0.8–1.4 nm. This makes the properties of the model strongly cutoff dependent and causes significant artifacts at the cutoff distance. In water, for instance, water structure shows a marked dependence on the cutoff, with spurious correlation lengths that vary with cutoff length (21). In dilute ionic solutions the structure strongly depends on the cutoff, with spurious peaks in the radial distribution function at the cutoff distance and other artifacts (22). The underlying reason for this is that cutting off the interaction at a specific distance causes a discontinuity in the potential and therefore the force and its derivative.

By smoothing the potential, forces and derivatives of the forces at the cutoff, some artifacts can be lessened but this involves unphysical potentials. The general approach is to use a “shift” or

“switch” function that modifies the normal nonbonded potential so that near the cutoff potential, the force as derivative of the potential, and preferably the second derivative of the potential go to zero instead of a finite value. There are several approaches that differ in whether they act on the derivative of the force as well as the force itself, the distance over which they modify the physical potential function, and the mathematical shape of the function. All of these solutions are somewhat unphysical because they modify the standard  $1/r$  or  $1/r^6$  interactions that have a solid physical basis. Shift or switch functions on the Coulomb interactions are almost never (and probably should not be) used anymore, as better solutions exist. Shift or switch functions on the Lennard-Jones potential are much less problematic as at typical cutoff distances the value of Lennard-Jones interactions is substantially smaller than for Coulomb interactions. Examples of equations are given in ref. (23).

A more elegant method to avoid sudden cutoff effects is to assume that at larger distances a central atom sees essentially a homogeneous environment and to describe this environment analytically (24). This leads to so-called reaction field methods for electrostatic interactions, where a long-distance component based on some assumption of the medium replaces the pair potential after a cutoff. Like a shift or switch function, this is also a modification of the “normal” Coulomb potential. In this case, however, there is a good physical basis as in solution after some distance it is reasonable to assume a charged atom no longer sees the individual charges but rather the averaged dielectric properties of the medium. There are simple analytical equations for dielectric media or media with a salt concentration (24). A major advantage of this method is that it is computationally cheap and easy to parallelize on large numbers of processors. One of the largest simulations to date on one of the largest computers in the world made use of this to simulate millions of atoms with good scaling over tens of thousands of processor cores (25). The approximation breaks down in nonhomogeneous media such as a large protein or a membrane environment, because it relies on the assumption that the bulk phase far away from an atom is the same everywhere. In a membrane this is obviously not true, as the interior of a membrane has a dielectric constant of 1–2, while inside large biomolecules the dielectric constant is also much lower than 80. However, the impact of this in practice may not be very large and will depend on specific systems. For example, in simulations of peptides (26) and a protein crystal (27) the reaction field method maintained the experimental structures similar to Particle Mesh Ewald (PME) (see below), and much better than with cutoffs. In a lipid bilayer, simple cutoffs did not result in an appropriate structure but both reaction field and PME gave stable bilayers, with slightly different structural parameters (28). In more highly charged DNA systems the RF method did not result in stable structures, while PME does give stable structures (29, 30).

A third approach assumes the simulation cell is replicated indefinitely and uses Fourier-based methods to solve the Poisson equation for this replicated system. There are several algorithms for this, of which PME is the most popular (31, 32). The basis of PME is an interpolation of the reciprocal-space Ewald sum. All of the interactions in the periodically replicated system are summed, and therefore include all long-range electrostatic interactions. This method does not cut off the potential but changes the simulation system into a truly periodic system, which in some cases may cause artifacts. These can occur in a number of ways, but the most problematic case is free energy calculations where the amount of charge in the system changes (33, 34). This is a hard problem where the results depend critically on the choice of algorithms and boundary conditions, although recent progress has clarified the assumptions in the main approaches and provided correction factors for several assumptions (34). In certain cases with highly charged proteins in small simulation cells the artificial periodicity of PME stabilizes structures unnaturally, although these cases seem to require rather pathetic conditions (35). A third class of cases involves detailed free energy calculations of ion permeation through low dielectric regions such as bilayers. Here one can show that the ion when in the low dielectric region interacts with its periodic images, although this effect turns out to be not very strong (36). Despite these drawbacks, in general PME is widely used with considerable success. Highly charged systems such as polynucleotides only give stable trajectories with PME or similar methods, protein simulations appear to be stable and consistent with experiment, and membrane simulations are well-behaved.

Lastly, there are other methods to solve the 1/distance problem, including multipole expansion methods. Multipole expansion methods have been investigated in the past but were considered too inefficient for “small” systems (37). However, as systems of over a 100,000 atoms are getting very common and parallelization of algorithms is a key concern, this method may be worth looking at again. Another alternative is to consider nonperiodic systems, such as spherical boundary conditions. This requires careful consideration of the boundary of the all-atom solvent with the outside of the sphere (33).

Today PME is normally considered the recommended choice. The choice is not arbitrary; a particular force field really should be developed for use with one of these methods, as the accuracy of the force field is not maintained when switching between different methods to calculate electrostatic interactions. As an example of this for the rather sensitive case of a lipid bilayer, Anezo et al. compared the effect of modifying electrostatic interaction algorithms and several other simulation parameters while maintaining the same force field on a dipalmitoylphosphatidylcholine bilayer and found significant differences in bilayer structure (28).

As force fields in general move towards full treatment of long-range interactions, parameters will become less dependent on the details of the methods (38). In practice, most commonly used force fields are either parameterized for use with PME or with a reaction field (GROMOS). The latter seems to perform similar with PME for proteins in solution, which is consistent with previous tests. Further use of both methods will no doubt help to further identify properties that may depend on the choice of electrostatics treatment.

The same problem arises in the treatment of long-range Lennard-Jones interactions. Although these decay faster (as  $1/r^6$ ), they are always attractive and have a significant effect on, e.g., pressure. They can be corrected by assuming a homogeneous “Lennard-Jones density” beyond a cutoff (39). This approximation is strictly valid only for noble gases but it proved to be useful for isotropic systems in general (38). The situation is more complicated for nonisotropic systems, including proteins in water and any system with a large interface. Several other methods exist (40, 41), analogous to the options for Coulomb interactions, but these are not in common use yet and require testing and probably reparameterization of at least parts of most current force fields.

---

## 5. Parameterization Philosophy

As outlined above, classical molecular dynamics seeks to approximate the underlying “real” energy surface of a system of molecules by simple mathematical functions from which forces between atoms can easily be calculated. There is no unique choice of such functions, and both the functions themselves and the parameters in them can be chosen freely.

A cornerstone of classical force fields is the philosophy that similar chemical groups in different molecules interact in the same way. This is both a practical and a philosophical assumption. Practically, it means a force field can be developed based on reproducing the energy surfaces for a set of small molecules that are typically well-characterized and contain the functional groups that occur in all biomolecules. This reduces the amount of work required to develop a force field and gives a force field a very broad applicability. In addition, in practice it would be impossible to develop a force field for the protein lysozyme based on properties of lysozyme alone, as there would not be enough experimental data to determine all parameters in a force field. Philosophically, it guards against the danger of over-fitting. If a model only reproduces the information that was used to create it, it has no predictive value. However, the building block approach allows modeling all biomolecules, including in complex mixed systems that combine proteins, nucleic

acids, carbohydrates, and lipids, from a common basis that does not use specific knowledge of the system of interest.

Although the functional form of the potential and its parameters can be freely chosen, the functional form of the potential has become more or less standard as described above. Several extensions to more complex functions that were originally primarily meant to reproduce gas-phase properties of small molecules have essentially disappeared (e.g., cross-terms between angles and bonds), largely because quantum mechanics is a better approach to such problems with modern computers. Extensions that lead to nonadditive force fields are described in the next section. Assuming we accept the potential form of Eq. 1, how do we develop parameters?

There are many possible criteria. The bonded terms are relatively straightforward. Historically, bonds and angles were taken from small-molecule crystallography and vibrational spectroscopy. Because these terms do not couple strongly to other terms and do not contribute too much to the internal conformational freedom of more complex molecules, their values tend to be not too critical. These days it is more convenient (and usually more accurate) to use quantum mechanics methods on small model compounds to get reasonable values for bonds and angles as well as for force constants.

Dihedral terms appear also relatively straightforward. Historically these are a weak part of most force fields because experimental information on the exact shape of the dihedral potential is impossible to get. For hydrocarbons such as butane for instance, spectroscopy can give relative populations of *trans* and gauche conformations, which translate into a free energy difference between three angles, but not enough information about the rest of the potential. Again, quantum mechanics is the method of choice for these interactions. A complication of dihedrals, however, is the coupling between dihedral terms and 1,4 interactions. The 1,4 interactions are scaled nonbonded interactions although they are often considered part of the dihedral interactions (a bonded term). When parameterizing dihedrals for a molecule in which the Lennard-Jones parameters and charges have not been finalized yet, the dihedral parameters will have to change with changes in these parameters to reproduce the underlying quantum mechanical PES.

Parameterizing the nonbonded interactions is much more complex. As described above, the value of those interactions depends on how the long-range part is treated, which is a choice that has to be made before parameterizing a force field. Neither partial charges nor Lennard-Jones parameters have a direct connection to quantum mechanics or experiment. In practice, force fields define a number of atom types, e.g., a carbon atom that is part of a carbonyl group, a carbon atom that is part of a linear aliphatic chain, and a carbon atom that is surrounded by four other nonhydrogen atoms. Each of the different carbon types identified has its

own Lennard-Jones parameters and possible charges, which are assumed to be constant among the same chemical groups in different molecules. The parameterization problem is then to optimize a limited number of “atom types” instead of every possible molecule.

The choice of model molecules is important, as they should represent all local chemical environments encountered in the biological molecules of interest. For small molecules liquid properties including density and heat of evaporation are easy to calculate. More recently free energies of solvation in different solvents and free energies of transfer between solvents have become readily available by simulation. Reproducing the partitioning of parts of molecules between different environments, with different degrees of polarity, hydrogen bonding capability and other properties, is a key requirement for accurately simulating protein folding and melting, peptide aggregation, and interactions of peptides and proteins with membranes.

In this model compound approach, parameters are developed for a particular force field, with a particular set of simulation parameters that define the treatment of Lennard-Jones and Coulomb interactions, macroscopic boundary conditions including pressure and temperature, and details of a particular potential function. Thus in theory parameters from different force fields cannot easily or reliably be mixed. In practice the accuracy of mixed force fields in reproducing thermodynamic properties such as partitioning or dynamic properties like the behavior of a membrane protein in a membrane do not seem overly sensitive to mixing force fields with similar design philosophies, although control simulations are essential. Some discussion of this issue in the context of membrane simulations is found in ref. (42).

---

## 6. Validation of Force Fields

The validation of force fields and simulation results ultimately relies on detailed comparison with experimental data. Force fields are parameterized on a certain set of properties, but to be useful they should be able to accurately reproduce or predict quantities that can be accurately measured but were not used in the parameterization. In general, any experimental data that can be directly compared to the same properties calculated from simulations is useful, although a large number of caveats exist. These include the use of derived (interpreted) experimental data that was obtained through incorrect assumptions from primary experimental data, technical problems with simulation software, and insufficient convergence of simulation results. Van Gunsteren and coworkers have outlined a large number of scenarios in two interesting papers that describe these issues in more detail (43, 44). In practice, force fields can be

tested on a growing number of experimental properties as computers have become faster and experiments are reaching ever higher resolution in time and space. The previously mentioned free energies of solvation are one example of important properties that can be simulated. Others include melting temperatures of proteins and nucleic acids, phase transition temperatures of lipid bilayers, and folding/unfolding equilibria of small proteins and peptides. For very large or complex systems, direct validation can be impossible. In such cases, successful tests of smaller similar systems should give enough confidence in the larger systems to seriously consider their results.

---

## 7. Outlook and Perspectives

Methods based on classical force fields and the classical force fields themselves now have a nearly 50-year long history. Simulations have progressed from initial applications to Lennard-Jones liquids on a time scale of picoseconds to current state of the art applications of very complex biological models with hundreds of thousands of atoms on a time scale of micro and soon milliseconds. It is clear that there will be a growing number of applications and problems that can be studied effectively with classical methods.

Although they entail approximations, this is true for every physics-based model. The key question is whether the level of approximation is appropriate for the problem of interest. Within the assumption of a classical force field, current computer and software capabilities offer significant room for improvement in parameters used. Increased computer power also has removed some of the practical rationale for key assumptions in the force fields, for example the simple potential forms. Although there are good philosophical reasons for using simple potentials, technically it is now completely feasible to use arbitrary tabulated potentials with tabulated derivatives. Experiments also continue to evolve and yield new experimental properties that can be used to develop and validate force fields, including previously unaccessible properties such as NMR relaxation times on time scales that have only recently become routine in classical simulations. Thus it continues to be worthwhile to test and improve classical methods.

At the same time, much effort is now focused on both polarizable (more detailed) and coarse-grained (less detailed) models, as well as on models that combine quantum mechanics in some part of the system of interest with classical mechanics in the rest of the system. Together, these methods add to an ever-growing toolkit to study a variety of interesting problems that are united by a common basis in the need for an accurate description of interactions at the atomistic level.

## Acknowledgments

D.P.T. is an Alberta Innovates Health Solutions Scientist and the Alberta Innovates Technology Futures Strategic Chair in (Bio) Molecular Simulation. Work in his group supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

## References

- Schlick T (2010) Molecular modeling and simulation: an interdisciplinary guide, vol 21, 2nd edn. Springer, New York
- Levitt M (2001) The birth of computational structural biology. *Nat Struct Biol* 8:392–393
- Lifson S, Warshel A (1968) Consistent force field for calculations of conformations, vibrational spectra and enthalpies of cycloalkane and *n*-alkane molecules. *J Chem Phys* 49:5116
- McCammon JA, Gelin BR, Karplus M (1977) Dynamics of folded proteins. *Nature* 267:585–590
- Levitt M, Warshel A (1975) Computer simulation of protein folding. *Nature* 253:694–698
- Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227–249
- Leach AR (2001) Molecular modelling: principles and applications, 2nd edn. Prentice Hall, Englewood Cliffs
- Ponder JW, Case DA (2003) Force fields for protein simulations. *Adv Protein Chem* 66:27–85
- Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J Am Chem Soc* 106:765–784
- Besler BH, Merz KM, Kollman PA (1990) Atomic charges derived from semiempirical methods. *J Comput Chem* 11:431–439
- Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J Am Chem Soc* 117:5179–5197
- Wang JM, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25:247–260
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimisation, and dynamics calculations. *J Comput Chem* 4:187–217
- Mackerell AD Jr (2004) Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem* 25:1584–1604
- Jorgensen WL, Maxwell DS, TiradoRives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
- Jorgensen WL, Madura JD, Swenson CJ (1984) Optimized intermolecular potential functions for liquid hydrocarbons. *J Am Chem Soc* 106:6638–6646
- Jorgensen WL (1981) Transferable intermolecular potential functions for water, alcohols and ethers. Application to liquid water. *J Am Chem Soc* 103:335–340
- Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF (1999) The GROMOS biomolecular simulation program package. *J Phys Chem A* 103:3596–3607
- Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676
- Levitt M, Hirshberg M, Sharon R, Daggett V (1995) Potential-energy function and parameters for simulations of the molecular-dynamics of proteins and nucleic-acids in solution. *Comput Phys Commun* 91:215–231
- Hess B (2002) Determining the shear viscosity of model liquids from molecular dynamics simulations. *J Chem Phys* 116:209–217
- Brooks CL, Pettitt BM, Karplus M (1985) Structural and energetic effects of truncating long ranged interactions in ionic and polar fluids. *J Chem Phys* 83:5897–5908
- Apol E, Apostolov R, Berendsen HJC, Van Buuren A, Bjelkmar P, Van Drunen R, Feenstra A, Groenhof G, Kasson P, Larsson P, Meulenhoff P, Murtola T, Pall S, Pronk S, Schulz R, Shirts MR, Sijbers A, Tielemans DP, Hess B, van der Spoel D,

- Lindahl E (2010) GROMACS user manual version 4.5.4, <http://www.gromacs.org>
24. Tironi IG, Sperb R, Smith PE, Vangunsteren WF (1995) A generalized reaction field method for molecular-dynamics simulations. *J Chem Phys* 102:5451–5459
  25. Schulz R, Lindner B, Petridis L, Smith JC (2009) Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer. *J Chem Theory Comput* 5:2798–2808
  26. Monticelli L, Simoes C, Belvisi L, Colombo G (2006) Assessing the influence of electrostatic schemes on molecular dynamics simulations of secondary structure forming peptides. *J Phys Condens Matter* 18:S329–S345
  27. Walser R, Hunenberger PH, van Gunsteren WF (2001) Comparison of different schemes to treat long-range electrostatic interactions in molecular dynamics simulations of a protein crystal. *Proteins* 43:509–519
  28. Anezo C, Vries AHD, Holtje H-D, Tieleman DP, Marrink S-J (2003) Methodological issues in lipid bilayer simulations. *J Phys Chem B* 107:9424–9433
  29. Krautler V, Hunenberger PH (2008) Explicit-solvent molecular dynamics simulations of a DNA tetradecanucleotide duplex: lattice-sum versus reaction-field electrostatics. *Mol Simulat* 34:491–499
  30. Cheatham TE, Miller JL, Fox T, Darden TA, Kollman PA (1995) Molecular-dynamics simulations on solvated biomolecular systems—the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J Am Chem Soc* 117:4193–4194
  31. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
  32. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald potential. *J Chem Phys* 103:8577–8592
  33. Warshel A, Sharma PK, Kato M, Parson WW (2006) Modeling electrostatic effects in proteins. *Biochim Biophys Acta* 1764:1647–1676
  34. Kastenholz MA, Hunenberger PH (2006) Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids. *J Chem Phys* 124(22):224501
  35. Kastenholz MA, Hunenberger PH (2004) Influence of artificial periodicity and ionic strength in molecular dynamics simulations of charged biomolecules employing lattice-sum methods. *J Phys Chem B* 108:774–788
  36. Allen TW, Andersen OS, Roux B (2004) Energies of ion conduction through the gramicidin channel. *Proc Natl Acad Sci USA* 101:117–122
  37. Esselink K (1995) A comparison of algorithms for long-range interactions. *Comput Phys Commun* 87:375–395
  38. Shirts MR, Mobley DL, Chodera JD, Pande VS (2007) Accurate and efficient corrections for missing dispersion interactions in molecular Simulations. *J Phys Chem B* 111:13052–13063
  39. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Oxford University Press, New York
  40. Venable RM, Chen LE, Pastor RW (2009) Comparison of the extended isotropic periodic sum and particle mesh ewald methods for simulations of lipid bilayers and monolayers. *J Phys Chem B* 113:5855–5862
  41. Veld PJ, Ismail AE, Grest GS (2007) Application of Ewald summations to long-range dispersion forces. *J Chem Phys* 127(14):144711
  42. Tieleman DP, MacCallum JL, Ash WL, Kandt C, Xu ZT, Monticelli L (2006) Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid-protein interactions, side chain transfer free energies and model proteins. *J Phys Condens Matter* 18: S1221–S1234
  43. van Gunsteren WF, Mark AE (1998) Validation of molecular dynamics simulation. *J Chem Phys* 108:6109–6116
  44. van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, Gee P, Geerke DP, Glattli A, Hunenberger PH, Kastenholz MA, Ostenbrink C, Schenk M, Trzesniak D, van der Vegt NFA, Yu HB (2006) Biomolecular modeling: goals, problems, perspectives. *Angew Chem Int Ed* 45:4064–4092

# Chapter 9

## Polarizable Force Fields

Hanne S. Antila and Emppu Salonen

### Abstract

This chapter provides an overview of the most common methods for including an explicit description of electronic polarization in molecular mechanics force fields: the induced point dipole, shell, and fluctuating charge models. The importance of including polarization effects in biomolecular simulations is discussed, and some of the most important achievements in the development of polarizable biomolecular force fields to date are highlighted.

**Key words:** Biomolecular simulation, Polarizable force fields, Electronic polarization, Molecular mechanics

---

### 1. Introduction

The aim of this chapter is to introduce the reader to the basic methodologies for including electronic polarization in present-day molecular mechanics (MM) force fields for biomolecules. We will restrict our treatment to empirical atomistic models of relatively low computational cost due to their popularity in current force field development. Topics such as higher induced multipole moments, hyperpolarizability and polarizable continuum solvent models, as well as polarizable coarse-grained models, are left for a more specialized treatment elsewhere. We will highlight some of the most important achievements on the development and application of polarizable force fields in biomolecular simulation to date. A comprehensive review of these issues is, however, beyond the scope of this chapter. For further insight on polarizable force fields, the interested reader is referred to various review articles on the topic (1–8).

We will start by introducing the physical mechanisms of polarization in molecular systems (section 1.1) and emphasizing the importance of an explicit description of polarization effects

in biomolecular simulations (section 1.2). Section 2 covers the three most widely used methods for including electronic polarization in MM force fields, namely, the induced point dipole (IPD) (section 2.1), shell (section 2.2) and fluctuating charge (FQ) (section 2.3) models. We will also briefly cover other methodologies used to model electronic polarization in atomistic simulations (section 2.4). Section 3 provides a brief overview of the state-of-the-art of biomolecular simulations employing polarizable force fields. The chapter is concluded with an outlook on the future of polarizable force fields in classical molecular simulations (section 4).

The treatment of the subjects in this chapter is mainly technical in nature. We assume that the reader is familiar with the general features of atomistic force fields and handling of electrostatic interactions, introduced in the chapters by Monticelli and Tieleman, and Sagui et al., respectively.

### 1.1. Polarization

Electric polarization means the redistribution of charge in space due to an electric field. For molecular systems, the polarizing electric field can be simply due to the other molecules in a condensed-phase system or a macroscopic field applied over the entire system. Classically, polarization of molecules is divided into three main contributions (see Fig. 1) (4):

1. *Orientational polarization*, which is the rigid rotation of molecules due to their coupling to an electric field.
2. *Geometric or atomic polarization*, which results from the displacement of atoms in the molecules with respect to each other. In the language of MM force fields, these conformational changes can be due to bond stretching, valence angle bending, and rotations around dihedral angles.

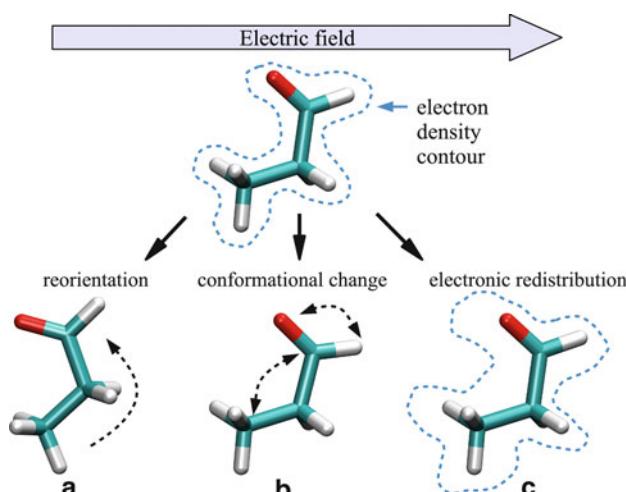


Fig. 1. Effective division of molecular polarization into three components: (a) orientational; (b) geometric (atomic); and (c) electronic polarization.

3. *Electronic polarization*, which refers to changes in the electronic structure of the molecules, that is, redistribution of the electron clouds around the nuclei.

This division of polarization into distinct components is just an effective one, as in reality the different components of polarization are interrelated. Displacements of the atomic nuclei in a molecule will always result in electronic polarization up to some degree. However, we shall retain this division into distinct types of molecular polarization for the sake of discussion regarding force fields. Only such force fields which feature explicit models of electronic polarization are called *polarizable force fields*.

Taking electronic polarization into account results in *adaptivity* of the molecular models to their environment. As we shall see in section 2, the terms related to electronic polarization in the total potential energy function of the system cannot be treated simply as a sum of independent pairwise interactions. This *non-additivity* is an important property of polarizable force fields and it comes with a significant increase in the computational time required to carry out the force and energy calculations.

## **1.2. Importance of Polarization Effects**

Considerations of taking electronic polarizability explicitly into account in biomolecular simulations extend at least as far as the 1970s and the work of Warshel and co-workers (9, 10). The choice of excluding explicit treatment of electronic polarization in force fields has not been due to physical motivation, but rather because of practical reasons: the simplicity of the potential energy function and the reduced computational cost of additive force field simulations. A common method to include electronic polarization effects – in an *implicit* manner – in non-polarizable force fields has been to fit atomic partial charges to quantum-mechanically calculated electrostatic potentials. By using a theoretical framework which is known to overestimate molecular dipole moments, such as the Hartree-Fock method, the resulting partial charges are enhanced from their more realistic gas-phase values. This is then intended to reflect electronic polarization in condensed phases (3). The current non-polarizable force fields can well reproduce structural, dynamic, and thermodynamic properties of bulk organic compounds and various biological supramolecular structures, such as lipid bilayers. On the other hand, in general, they are not very accurate in reproducing the potential energy surfaces associated with molecular conformational changes (11–13). In the case of non-polar molecules, this issue can be remedied in part with the use of (fixed) atom-based electric multipoles instead of point charges only. Polar molecules also require an explicit description of intramolecular polarization in order to reproduce gas-phase *ab initio* potential energy surfaces in a satisfactory manner.

Polarization is an important contribution to the total potential energy of condensed-phase systems. It accounts to about 10–20% of the interaction energy between molecules at the van der Waals distance (effective distance of closest approach). For polar molecules electronic polarization is a significant factor for the free energy of solvation even in solvents of low relative permittivity (14, 15). In this context it should be noted that force field parameters optimized with fixed partial charges commonly “embed” the molecular polarization energies in the other potential energy terms. An interesting exception to this practice is the SPC/E water model (16). The experimental potential energy that was used as a target for parameterizing the model was corrected by the polarization energy, resulting in more accurate values of density and diffusion coefficient in comparison to the preceding SPC water model.

Biomolecular systems are inherently heterogeneous with respect to their chemical composition and physical properties. Including an explicit description of electronic polarization in force fields is very likely to improve the transferability of molecular models from one environment to another (11, 14). This is particularly important for the description of molecules in the vicinity of ions and for systems comprised of subregions of different dielectric properties. Biologically relevant examples of such systems are lipid bilayers, which are important components of cell membranes and serve as the native environments of membrane proteins. Lipid bilayers are characterized by large variations in the local relative permittivities. At the center of the bilayer the local permittivity is very low,  $\epsilon \sim 1 - 2$ , but strongly increases towards the lipid-water interface, where the permittivity component in the bilayer plane has been estimated to reach values as high as  $\epsilon_{\parallel} \sim 100 - 1,000$  (17–19). With increasing distance from the bilayer center, the local permittivity gradually decreases to that of bulk water,  $\epsilon \sim 80$ . The ability of force fields to properly reproduce the electrostatic potential difference over a lipid bilayer membrane and the partitioning of the amino acid residues of membrane proteins, as well as small solutes (e.g. drugs) into the membrane, is an important factor for the reliability of the empirical simulations (18–20).

Other examples of the importance of polarization effects in the context of biological macromolecules include folding and stabilization of local structures of proteins (21, 22), ligand-protein interactions (23, 24), and nucleic acids in the presence of divalent cations (6, 25). Water, as a highly polar compound and the environment of biomolecules, is an especially important molecule in force field development. Various popular non-polarizable models of water (e.g., TIP3P, TIP4P, SPC, SPC/E, and their further modifications) have structural and electrostatic properties which are intermediate between those in the gas and liquid phases. The parameters employed in a force field affect such important properties of water

as surface tension, relative permittivity, hydrogen bond energies, and hydration energies of solvents. Various polarizable models of water produce a good description of both the gas and liquid phases (26–31). The adaptivity of the polarizable force fields is reflected in realistic values of molecular dipole moments in the gas (1.85 D) and liquid ( $\sim 2.5 - 2.9$  D) phases. These models can be expected to pave the way toward more realistic simulation studies of interfacial systems.

## 2. Methods for Modelling Polarization

In this section we will overview the three most commonly used methodologies for including electronic polarization in MM force fields: the induced point dipole model (IDP) (section 2.1), the shell model (section 2.2), and the fluctuating charge (FQ) model (section 2.3). The polarizable versions of the most widely used biomolecular force fields, such as CHARMM, OPLS, AMBER, and GROMOS, as well as entirely new polarizable force fields (e.g. AMOEBA (32, 33)), employ one of the methods covered here.

A popular approach for describing the redistribution of molecular charge (e.g. electronic polarization) is to use the electrostatic multipole expansion. The components of this expansion are the induced dipoles, quadrupoles, octupoles, and so forth. The first two models featured in this section, the IPD and shell models, feature induced dipoles only. This choice has been motivated by the order of magnitudes of the multipole expansion terms, the dipole term being the most important one (4). In general, the energy associated with electronic polarization in the induced dipole approximation is a sum of three contributions,

$$U_{\text{tot}} = U_{\text{pol}} + U_{\text{dip-stat}} + U_{\text{dip-dip}} \quad (1)$$

where  $U_{\text{pol}}$  is the energy required to create the induced dipoles (polarization energy),  $U_{\text{dip-stat}}$  is the interaction energy between the dipoles and the static charge distribution (e.g. due to fixed partial charges), and  $U_{\text{dip-dip}}$  is the interaction energy between the induced dipoles.

The third method featured in this section, the FQ model, produces all orders in the induced charge moments. As we will see in section 2.3, the total energy expression of this model is quite different from that of Eq. 1.

### 2.1. Induced Point Dipole Model

#### 2.1.1. General Description

In the IPD model electronic polarization is described by polarizable point dipoles, which are assigned to some pre-selected sites in molecules. These sites are typically the atomic interaction centers (e.g., the nuclei). The total molecular polarization is a result of the

interaction of these dipoles with each other and their environment. In the case of a single, isolated point dipole and assuming linear response, the induced dipole moment  $\boldsymbol{\mu}$  is

$$\boldsymbol{\mu} = \boldsymbol{\alpha} \mathbf{E}(\mathbf{r}), \quad (2)$$

where  $\boldsymbol{\alpha}$  is the  $3 \times 3$  polarizability tensor of the dipole and  $\mathbf{E}(\mathbf{r})$  is the electric field strength at its position  $\mathbf{r}$ .

In a system consisting of  $N$  inducible dipoles, the dipole moment  $i$  is given by the combined effect of the electric field and induction due to the other dipoles,

$$\boldsymbol{\mu}_i = \boldsymbol{\alpha}_i \mathbf{E}(\mathbf{r}_i) = \boldsymbol{\alpha}_i \left[ \mathbf{E}^0(\mathbf{r}_i) - \sum_{j \neq i}^N \mathbf{T}_{ij} \boldsymbol{\mu}_j \right]. \quad (3)$$

Here  $\mathbf{E}^0(\mathbf{r}_i)$  is the electric field due to the static charge distribution in the system. In present-day force fields, this corresponds to the fixed atomic partial charges and a possible external electric field applied over the system. The dipole–dipole interactions in Eq. 3 are given by the tensor (35)

$$\mathbf{T}_{ij} = \frac{1}{r_{ij}^3} \mathbf{I} - \frac{3}{r_{ij}^5} \begin{bmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{bmatrix}, \quad (4)$$

where  $\mathbf{I}$  is the unit tensor, and  $x$ ,  $y$ , and  $z$  are the cartesian coordinates of the vector  $\mathbf{r}_{ij}$  connecting the interacting dipoles  $i$  and  $j$  ( $r_{ij} \equiv |\mathbf{r}_{ij}|$ ).

The coupling of the dipoles through the interaction tensor  $\mathbf{T}_{ij}$  in Eq. 3 means that the instantaneous values of the dipole moments are interdependent. A change in one of the dipole moments will induce a change in all the other ones. Calculating the energy of the system thus requires a self-consistent solution with respect to the induced dipole moments. To this end, Eq. 3 can be rewritten in the form of a matrix equation (35),

$$\tilde{\mathbf{A}} \tilde{\boldsymbol{\mu}} = \tilde{\mathbf{E}}, \quad (5)$$

where  $\tilde{\mathbf{A}}$  is a  $3N \times 3N$  matrix containing the inverses of the dipole polarizability tensors along the diagonal.  $\tilde{\mathbf{E}}$  and  $\tilde{\boldsymbol{\mu}}$  are  $3N$ -component vectors where the dipole moments  $\boldsymbol{\mu}_i$  and electric field strength at each polarizable site  $\mathbf{E}_i^0$  are given in the order corresponding to that of the polarizabilities in  $\tilde{\mathbf{A}}$ . That is,

$$\begin{bmatrix} \boldsymbol{\alpha}_1^{-1} & \mathbf{T}_{12} & \dots & \mathbf{T}_{1N} \\ \mathbf{T}_{21} & \boldsymbol{\alpha}_2^{-1} & \dots & \mathbf{T}_{2N} \\ \vdots & \ddots & \vdots & \\ \mathbf{T}_{N1} & \mathbf{T}_{N2} & \dots & \boldsymbol{\alpha}_N^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_N \end{bmatrix} = \begin{bmatrix} \mathbf{E}_1^0 \\ \mathbf{E}_2^0 \\ \vdots \\ \mathbf{E}_N^0 \end{bmatrix}. \quad (6)$$

Inverting  $\tilde{\mathbf{A}}$  provides the solution for the dipole moments,

$$\tilde{\boldsymbol{\mu}} = \tilde{\mathbf{B}} \tilde{\mathbf{E}}, \quad (7)$$

with

$$\tilde{\mathbf{B}} = \tilde{\mathbf{A}}^{-1} = (\tilde{\boldsymbol{\alpha}}^{-1} + \tilde{\mathbf{T}})^{-1}. \quad (8)$$

This exact method of solving the induced dipole moments through matrix inversion is an  $O(N^3)$  operation. Considering that a typical biomolecular simulation contains  $10^4 - 10^6$  atoms, it is clear that for actual dynamic simulations (such as molecular dynamics, MD), the matrix inversion method is unfeasible. Thus, one needs to resort to approximative techniques, such as iterative schemes or those based on the use of extended Lagrangians (2). In the latter approach, the dipole moments are treated as dynamic quantities following their own “equations of motion”, and their magnitudes are computed only once during an MD timestep.

### 2.1.2. Polarization and Interaction Energy

The total interaction energy for the IPD model, written according to Eq. 1, is

$$U_{\text{tot}} = \frac{1}{2} \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{E}(\mathbf{r}_i) - \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{E}^0(\mathbf{r}_i) + \frac{1}{2} \sum_{i=1}^N \boldsymbol{\mu}_i \mathbf{T}_{ij} \boldsymbol{\mu}_j, \quad (9)$$

$$= \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \left[ \frac{1}{2} \mathbf{E}(\mathbf{r}_i) - \mathbf{E}^0(\mathbf{r}_i) + \frac{1}{2} \sum_{i=1}^N \mathbf{T}_{ij} \boldsymbol{\mu}_j \right], \quad (10)$$

where the first term is the polarization energy  $U_{\text{pol}}$  (that is, the energy required to create the dipoles), the second one the interaction of the induced dipoles with the static charge distribution (e.g. fixed atomic partial charges)  $U_{\text{dip-stat}}$ , and the last term the mutual interaction of the dipoles  $U_{\text{dip-dip}}$ . From Eq. 3 we see that

$$\mathbf{E}(\mathbf{r}_i) = \mathbf{E}^0(\mathbf{r}_i) - \sum_{i=1}^N \mathbf{T}_{ij} \boldsymbol{\mu}_j, \quad (11)$$

and the total interaction energy can be expressed in a simple form:

$$U_{\text{tot}} = -\frac{1}{2} \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{E}^0(\mathbf{r}_i), \quad (12)$$

which depends only on the dipole moments and the electric field due to the static charge distribution.

If the atomic polarizabilities are isotropic (e.g., the polarization tensor is diagonal and  $\alpha_{xx} = \alpha_{yy} = \alpha_{zz}$ ), the induced dipole moments are parallel to the electric field,  $\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_i$ . Note that the polarizability  $\alpha_i$  is now a scalar quantity. Inserting this relation into the first term of Eq. 10 provides an alternative expression for the polarization energy,

$$U_{\text{pol}} = \frac{1}{2} \sum_{i=1}^N \frac{\mu_i^2}{\alpha_i}. \quad (13)$$

### 2.1.3. Molecular Polarizability

In the parametrization and validation of a force field, it is commonly required to calculate polarizabilities for a set of selected molecules. For a molecule consisting of  $N$  point dipole polarizabilities, the total polarizability is obtained by contracting the tensor  $\tilde{\mathbf{B}}$  (see Eq. 8 above) to a  $3 \times 3$  matrix  $\boldsymbol{\alpha}_{\text{mol}}$ ,

$$\boldsymbol{\mu}_{\text{mol}} = \left[ \sum_{i=1}^N \sum_{j=1}^N \mathbf{B}_{ij} \right] \mathbf{E} = \boldsymbol{\alpha}_{\text{mol}} \mathbf{E}. \quad (14)$$

The three eigenvalues of  $\boldsymbol{\alpha}_{\text{mol}}$  correspond to the diagonal components of molecular polarizability,  $\alpha_{\text{mol}}^{xx}$ ,  $\alpha_{\text{mol}}^{yy}$ ,  $\alpha_{\text{mol}}^{zz}$ . Experimental values of molecular polarizabilities are typically obtained from measurements of gas-phase refraction indices (34), yielding only the mean molecular polarizabilities,

$$\bar{\alpha}_{\text{mol}} \equiv \frac{1}{3} (\alpha_{\text{mol}}^{xx} + \alpha_{\text{mol}}^{yy} + \alpha_{\text{mol}}^{zz}). \quad (15)$$

Further information on the individual components of molecular polarizabilities can be obtained from *ab initio* calculations, and experimental Kerr effect and light scattering depolarization data (35).

### 2.1.4. Polarization Catastrophe

Due to the point-like character of the induced dipoles, the IPD model has a specific problem related to dipole–dipole interactions at small distances. For simplicity, let us consider a system of two induced point dipoles  $i$  and  $j$ , separated by the distance  $r$ , with scalar polarizabilities  $\alpha_i$  and  $\alpha_j$ , respectively. The polarizability of the system along the line joining the dipoles  $\alpha_{\parallel}$ , and perpendicular to it,  $\alpha_{\perp}$ , can be calculated by using Eq. 14. The resulting polarizabilities are (36)

$$\alpha_{\parallel} = (\alpha_i + \alpha_j + 4\alpha_i\alpha_j/r^3)/(1 - 4\alpha_i\alpha_j/r^6) \quad (16)$$

and

$$\alpha_{\perp} = (\alpha_i + \alpha_j - 2\alpha_i\alpha_j/r^3)/(1 - \alpha_i\alpha_j/r^6). \quad (17)$$

From the denominator of Eq. 16, it can be seen that when the distance between the dipoles approaches  $(4\alpha_i\alpha_j)^{1/6}$ , the polarizability  $\alpha_{\parallel}$  goes to infinity. This effect is called the *polarization catastrophe*. It results from the unphysically strong interaction of point dipoles aligned in the same direction (the so-called head-to-tail interaction).

There are two obvious ways of avoiding the polarization catastrophe. The first option is to modify the polarizability tensor so that at short separations the polarizability of the point dipoles is damped to some finite value (37). The other approach is to modify the interaction tensor, so that at short separations the dipole–dipole interactions are made to correspond to those between finite-sized dipoles or charge distributions, instead of point-like dipoles (35, 36).

### 2.1.5. Thole's Model for Damped Interactions

A damping scheme which has been adapted to several present-day polarizable force fields is the modified interaction model by Thole (36). The model employs atom interaction centers as polarizable sites with scalar polarizabilities  $\alpha_i$ . Anisotropic molecular polarizabilities then originate from the mutual interactions of the dipoles according to Eq. 3. Thole's model avoids the polarization catastrophe by replacing the usual dipole interaction tensor, Eq. 4, by one that realistically screens the dipole–dipole interactions at short separations. By introducing a reduced distance specific for each pair of interacting dipoles,

$$u \equiv r/(\alpha_i \alpha_j)^{1/6}, \quad (18)$$

the interaction tensor is rewritten as

$$\begin{aligned} (T_{ij})_{pq} &= \delta_{pq} r^{-3} - 3r_p r_q r^{-5}, \\ &= (\alpha_i \alpha_j)^{-1/2} (\delta_{pq} u^{-3} - 3u_p u_q u^{-5}), \\ &= (\alpha_i \alpha_j)^{-1/2} t_{pq}(u), \end{aligned} \quad (19)$$

where the indices  $p$  and  $q$  designate cartesian coordinates, and  $\delta_{pq}$  is the Kronecker delta. The *shape function*  $t$  of the interaction can be interpreted as the interaction tensor between two dipoles of unit polarizabilities (36).

In general, the tensor  $T_{ij}$  assumes that the interacting dipoles are each made of two point charges that are displaced by an infinitesimal distance from each other. The total interaction is then obtained from a sum of four interactions between the point charges. Thole's model replaces the other point charge in such interactions with some well-behaved, spherically symmetric charge distribution  $\rho(u)$  around the atom nucleus. By doing so, the dipole–dipole interactions become screened, and the polarization catastrophe is avoided. After considering various possible charge distribution functions for  $t$ , Thole concluded that a linear and an exponential form were the most suitable ones (36). As an example, the exponential charge distribution is

$$\rho(u) = (\alpha^3 / 8\pi) e^{-au}, \quad (20)$$

where  $a$  is a characteristic damping parameter of the model. The corresponding form of the shape function is

$$\begin{aligned} t_{pq} &= \delta_{pq}/u^3 [1 - (a^2 u^2/2 + au + 1)e^{-au}] \\ &\quad - 3u_p r_q/r^5 [1 - (a^3 r^3/6 + a^2 r^2/2 + ar + 1)e^{-au}]. \end{aligned} \quad (21)$$

Thole's original atom-based parametrizations were obtained by fitting to a set of 15 experimental molecular polarizabilities (36). The results were then tested by calculating polarizabilities for molecules not included in the learning set, with the conclusion that only a single value of polarizability per element was required

for accurate results. That is, polarizability in the Thole model is independent of the local chemical environment of the atom in question. Along with the parameter  $\alpha$  for each charge distribution model, Thole determined the respective polarizability values for H, C, N, and O. Van Duijnen and Swart later provided an improved set of parameters by using an extended learning set of 52 molecules (38), including molecules containing F, S, Cl, Br, and I. Swart and co-workers have also presented Thole model parameters specifically optimized for amino acid residues using polarizabilities obtained with time-dependent density functional theory calculations (39). Most recently, Wang et al. presented parametrizations of the Thole model, with different charge distribution models, based on a set of 420 molecules (40, 41).

#### *2.1.6. Practical Application of the IPD Model*

The appeal of the IPD model lies mainly in the transferability of its element-based polarizabilities. Once these have been determined, it is straightforward to add the description of electronic polarization into any molecule in the force field. On the other hand, introduction of the IPD model to most of the current MD simulation codes will require the implementation of entirely new types of interactions according to Eq. 10. Obtaining the self-consistent solution of the induced dipole moments will also require a suitably efficient algorithm (42). The long-range dipole–dipole interactions will need to be taken into account with a method similar to the ones commonly used to handle point charges (see the chapter on electrostatics by Sagui et al.). Such algorithms have been developed (43–46) but not yet implemented in the most popular MD simulation packages.

As opposed to the two other models for electronic polarization introduced in this chapter, the IPD model does not require any shorter timesteps than the current non-polarizable force fields. The computational overhead related to the IPD model comes from calculating the interactions of the induced dipoles, obtaining the self-consistent solution for the dipole moments, and handling of the long-range electrostatic interactions. Depending on the choice of methodology, the computational cost of a polarizable force field based on the IPD model is about 2–4 times that of a corresponding non-polarizable one (2, 46).

### **2.2. Shell Model**

#### *2.2.1. General Description*

In the shell model a polarizable site is described by two charged particles, a heavy *core particle* and a massless or very light *shell particle*. The particles are connected to each other by a harmonic spring with a force constant  $k$ . The shell charge  $-q_s$  ( $q_s \geq 0$ ) and the core charge  $+q_s$  are kept fixed in the simulations, and polarization in the model is obtained by the displacement of the core and shell particles with respect to each other. This produces a finite-length dipole with a dipole moment

$$\boldsymbol{\mu} = -q_s \Delta \mathbf{r}, \quad (22)$$

where  $\Delta\mathbf{r}$  is the displacement vector from the core to the shell particle. Self-interactions, e.g., the interactions between the core and shell particles of the same polarizable site, are not taken into account.

In molecular force fields the core particles are typically assigned to the atomic interaction centers (nuclei), and atom  $i$  thus carries a total charge of  $\tilde{Q}_i = Q_i + q_{s,i}$ , where  $Q_i$  is the standard atomic partial charge. In MD simulations, the core particles follow the standard equations of motion of the associated nuclei. The shell particles, in turn, can be interpreted to describe the redistributions of the valence electrons, and in dynamic simulations, they are treated either *adiabatically* (see section 2.2.3) or *dynamically* (see section 2.2.4).

The shell model is often also called the *Drude oscillator model* or *charge-on-spring model*. Historically, the former name has been used for simulations of the liquid state, while the term shell model has been mainly used in the context of simulations of solid state materials, such as ionic crystals. However, all three names have become interchangeable in the literature of biomolecular simulation.

### 2.2.2. Shell Model Polarization and Interaction Energy

The energy associated with the core–shell pairs is

$$U_{\text{tot}} = \frac{1}{2} \sum_{i=1}^N k_i (\Delta r_i)^2 + \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N-1} \sum_{j>i}^N \left[ \frac{\tilde{Q}_i \tilde{Q}_j}{|\mathbf{r}_{ij}|} + \frac{\tilde{Q}_i q_{s,j}}{|\mathbf{r}_{ij} + \Delta\mathbf{r}_j|} + \frac{q_{s,i} \tilde{Q}_j}{|\mathbf{r}_{ij} - \Delta\mathbf{r}_i|} + \frac{q_{s,i} q_{s,j}}{|\mathbf{r}_{ij} - \Delta\mathbf{r}_i + \Delta\mathbf{r}_j|} \right], \quad (23)$$

where  $r_{ij}$  is the core–core distance between the interacting atoms. The first term is the polarization energy,  $U_{\text{pol}}$ , associated with separating the core and shell particles by a distance  $\Delta r_i$ . This is simply the mechanical energy of the associated harmonic spring. The following terms in parentheses are the core–core, core–shell, and shell–shell Coulomb interactions, respectively.

### 2.2.3. Adiabatic Shell Model

In the adiabatic case, the shell particles are massless, and each particle is set at its minimum energy position (zero net force) at every simulation step. The electric field  $\mathbf{E}$  exerts a force  $\mathbf{F}_E = -q_s \mathbf{E}(\mathbf{r} + \Delta\mathbf{r})$  on the shell particle. At mechanical equilibrium the restoring force of the core–shell harmonic spring is equal in magnitude to the force due to the electric field

$$\mathbf{F}_E + \mathbf{F}_{\text{spring}} = 0. \quad (24)$$

With  $\mathbf{F}_{\text{spring}} = -k\Delta\mathbf{r}$ , one obtains the relation

$$\Delta\mathbf{r} = -\frac{q_s \mathbf{E}}{k}. \quad (25)$$

Since the displacement of a shell particle is affected by the electric fields of the other shell particles, a self-consistent solution for all the displacements is required (cf. section 2.1.1). Failing to maintain the self-consistency results in drag forces on the shell particles and simulation instability (47). The solution can be obtained through the inversion of a matrix containing the coupled shell particle displacements, given by Eq. 25, or by iteration. The former is exact and needs to be done only once, while the latter is much faster but requires, in addition to reasonable initial values of the dipole moments, sufficient convergence of the iterations for stable simulations.

Comparing Eqs. 22 and 25 shows that polarization produced by the shell model is isotropic and produces dipole moments according to the linear response

$$\Delta\boldsymbol{\mu} = \frac{q_s^2}{k} \mathbf{E}. \quad (26)$$

Using Eq. 26 and the general expression  $\boldsymbol{\mu} = \alpha \mathbf{E}$ , the relation between the parameters  $q_s$ ,  $k$  and polarizability  $\alpha$  is

$$\alpha = \frac{q_s^2}{k}. \quad (27)$$

Note that in order to obtain a certain value of  $\alpha$ , it is possible to use any physically reasonable combination of  $q_s$  and  $k$ . For example, the SWM4-NDP water model by Lamoureux et al. (29) has  $q_s = -1.71636e$  and  $k = 4.184 \times 10^5$  kJ/nm<sup>2</sup> on a single polarizable site set at the oxygen atom. These parameter choices result in  $\alpha/4\pi\epsilon_0 = 0.97825 \times 10^{-3}$  nm<sup>3</sup>. This is significantly lower than the experimental gas-phase value of  $\alpha/4\pi\epsilon_0 = 1.45 \times 10^{-3}$  nm<sup>3</sup> (34). The SWM4-NDP water model and other molecule parametrizations in the same CHARMM framework are based on reproducing the dielectric constants of pure liquids, while treating the molecular polarizabilities as free parameters (7). For this force field, empirically determined scaling factors for the gas-phase polarizabilities range from 0.7 to 1.0, depending on the molecule type. This approach has been justified by QM calculations of pure liquids and small solutes therein (48, 49) which have yielded lower values of molecular polarizabilities in solution than in the gas phase.

#### 2.2.4. Dynamic Shell Model

In the *dynamic case*, parts of the atomic masses, typically of the order of 0.1–1 amu, are assigned to the respective shell particles, which then follow their own equations of motion. This is a type of an extended Lagrangian approach similar to the ones mentioned in 2.1.1. To prevent the shell particle thermalization, which would decrease the accuracy of the simulations and even result in simulation instabilities (47), the shell particles are coupled to a low-temperature ( $\sim 1$  K) thermostat of their own. The advantage of this scheme is in avoiding the time-consuming calculations for self-consistent solutions of the

shell particle displacements. However, the low masses of the shell particles necessitate the use of short timesteps for integrating the equations of motion ( $\sim 0.5\text{--}1$  fs) in order to guarantee the accuracy and stability of the simulations.

In the adiabatic case the polarizability of a shell model site was obtained by assuming mechanical equilibrium for the core–shell pair (cf. 2.2.3). In order to derive a similar expression for the dynamic case, we first recall that the polarizability of a shell model site is isotropic. Assuming that variations in the total electric field are small enough at length scales comparable to the sizes of the core–shell dipoles, the related polarization energy is given by Eq. 13. Equating this with the general expression for the shell model polarization energy (the first term in Eq. 23) gives

$$k\Delta r^2 = \frac{\mu^2}{\alpha}. \quad (28)$$

Inserting the expression for the dipole moment in Eq. 22, we recover the expression for the polarizability of a shell model site in the adiabatic case, Eq. 27.

#### 2.2.5. Steric and Dispersion Interactions

The interaction center for the steric and dispersion interactions, modelled commonly with the 12–6 Lennard-Jones potential, is usually set at the shell particle. The physical reasoning here is that interatomic repulsion (exchange) and dispersion interactions can be mainly associated with the valence electrons. The repulsive interactions, modelled by the  $r^{-12}$  part of the Lennard-Jones potential, prevent the polarizable atoms from coming arbitrarily close to each other. In most cases this is sufficient to avoid the polarization catastrophe (see 2.1.4). However, as demonstrated by Yu et al. (50), in the case of highly polarizable atoms or polyvalent ions, additional means are required to ensure the stability of the simulation. One option is to use an additional anharmonic interaction at high enough core–shell separations (50),

$$U_{\text{hyp}} = k_{\text{hyp}}(\Delta r - \Delta r_{\text{cut}})^4, \quad (29)$$

where  $k_{\text{hyp}}$  is a force constant and  $\Delta r_{\text{cut}}$  is a cutoff distance beyond which the additional restoring force is applied. Another method to ensure the stability of the simulations is to assign a small repulsive potential also to the core particle. A third option is to use some kind of a damping scheme, like the one introduced in section 2.1.5. That is, to replace point charge interactions at short separations with those between point charges and smooth charge distributions. This approach is used in the polarizable CHARMM force field based on the shell model (50, 51).

#### 2.2.6. Practical Application of the Shell Model

Electrostatic energy in the shell model is solely based on interactions between point charges, and no special interactions (aside from the harmonic springs connecting the core–shell pairs) need to be

introduced into the existing simulation codes. Consequently, the shell model has already been implemented in many commonly used MD packages, such as GROMACS, CHARMM, and GROMOS.

On the other hand, for each polarizable site, the method introduces twice the amount of charged particles. This translates to four times as many Coulomb interactions as in a non-polarizable force field (as interactions between the core and shell particles of the same polarizable site are not taken into account). Moreover, in fully dynamic shell model simulations, a small timestep ( $\sim 0.5\text{--}1$  fs) needs to be used due to the light shell particles. Part of this extra computational cost could be overcome by multi-timestep methods treating the motion of the shell particles and atomic nuclei at separate timescales.

As opposed to the IPD model, which uses highly transferable element-based parameters, the shell model parameters  $k_i$  and  $q_i$  for each polarizable site are not necessarily transferable between different molecular moieties. For each new molecule included in the force field library, the polarization parameters need to be carefully optimized in order to reproduce, for instance, the correct molecular polarizabilities and electrostatic properties. To this end, the developers of the polarizable CHARMM force field have published comprehensive procedures for parametrizing molecules for the shell model (51–53).

### **2.3. Fluctuating Charge Model**

#### *2.3.1. General Description*

The third commonly used model for describing polarization in MM force fields is the FQ model. This model is also often called the *electronegativity equalization model* (see 2.3.2 below). The FQ model does not employ induced dipoles like the IPD or shell models. Instead, the values of atomic partial charges, which are fixed in non-polarizable force fields, are made dynamic quantities. In this way the model produces molecular polarizability to all orders in the charge moments. In addition, the FQ model allows intra- and intermolecular charge transfer, as opposed to the IPD and shell models.

The instantaneous values of the partial charges are found at each timestep by minimizing the electrostatic energy of the system. For simplicity, let us first consider the case of a single atom. The energy required to create a charge  $Q$  on a neutral atom can be expressed as a truncated second-order Taylor series,

$$U(Q) = U^0 + Q \left( \frac{\partial U}{\partial Q} \right)_0 + \frac{1}{2} Q^2 \left( \frac{\partial^2 U}{\partial Q^2} \right)_0, \quad (30)$$

where  $U^0 = U(0)$  is a constant. The partial derivatives of energy in the series have clear physical interpretations. Taking a neutral atom as the reference state and assuming that the truncated Taylor series of Eq. 30 is valid in the range  $Q = \pm e$ , the ionization potential (IP) of the atom is  $IP = U(+e) - U(0)$ . In turn, the electron affinity

(EA) of the atom is  $EA = U(-e) - U(0)$ . The first-order term coefficient in Eq. 30 can then be expressed as

$$e \left( \frac{\partial U}{\partial Q} \right)_0 = (IP + EA)/2 \equiv \chi^0, \quad (31)$$

which is the definition of (absolute) electronegativity by Mulliken (54). In the same fashion, the second-order term coefficient is

$$\frac{1}{2} e^2 \left( \frac{\partial^2 U}{\partial Q^2} \right)_0 = (IP - EA)/2 \equiv \eta^0, \quad (32)$$

which is the atomic hardness, e.g., resistance to the flow of electrons (55). The atom-specific values of electronegativity  $\chi^0$  and atomic hardness  $\eta^0$  can either be determined from experimental data or *ab initio* calculations (56, 57). However, since these quantities are used in empirical force fields, they are often treated simply as adjustable parameters. The aim is that the optimized parameters would be specific to small functional groups commonly found in biomolecules and hence transferable between different molecules in the force field (58). This kind of an approach is similar to the fitting of the  $r^{-6}$  dispersion term coefficients in Lennard-Jones potentials: while in theory the Lennard-Jones coefficients are related to atomic polarizabilities, in most MM force fields, they are taken as empirical parameters in order to reproduce the desired condensed-phase properties.

For a system of  $N$  atoms, the total electrostatic energy expression, replacing the usual summation over Coulomb interactions in non-polarizable force fields ( $\propto Q_i Q_j / r_{ij}$ ), is

$$U_{\text{tot}} = \sum_i (U_i^0 + \chi_i^0 Q_i) + \sum_{i,j} \eta_{ij}(r_{ij}) Q_i Q_j, \quad (33)$$

where the coupling terms  $\eta_{ij}$  depend on the distance between the interacting atoms  $i$  and  $j$ , and  $\eta_{ii}(0) = \eta_i^0$ . Note that in order to simplify the expression for the total energy, the charges are now given in units of  $e$  (cf. Eqs. 31 and 32).

### 2.3.2. Obtaining the Partial Charges $Q_i$

The instantaneous atomic partial charges are obtained by minimizing Eq. 33 with respect to the charges, with the boundary condition that the net charge of the system (or parts thereof, see 2.3.4),  $Q_{\text{tot}} = \sum_i Q_i$ , is conserved. The boundary condition can be imposed by using a Lagrange multiplier (2),

$$U_{\text{tot}} = U_{\text{tot}} - \lambda \left( Q_{\text{tot}} - \sum_i Q_i \right). \quad (34)$$

Minimizing this expression with respect to the charges leads to

$$\left( \frac{\partial U_{\text{tot}}}{\partial Q_i} \right) - \lambda = 0 \quad \forall i \quad (35)$$

Since the multiplier is the same for each partial derivative with respect to the atomic charges, one obtains

$$\left( \frac{\partial U_{\text{tot}}}{\partial Q_i} \right) = \left( \frac{\partial U_{\text{tot}}}{\partial Q_j} \right) \quad \forall i, j \quad (36)$$

Recalling the definition of the Mulliken electronegativity, Eq. 31, we notice that minimizing the electrostatic energy leads equalization of the atomic electronegativities. Hence, the FQ model is often also called the method of electronegativity equalization. It should be noted that the electronegativities in molecules are not the same as for isolated atoms, but instead

$$\chi_i = \left( \frac{\partial U_{\text{tot}}}{\partial Q_i} \right) = \chi_i^0 + 2\eta_{ii}Q_i + \sum_{j \neq i} \eta_{ij}Q_j. \quad (37)$$

This set of coupled linear equations requires a self-consistent solution. As with the two previous models, the self-consistency can be obtained through matrix inversion, iteration, or by using an extended Lagrangian where the partial charges follow their own “equations of motion” with fictitious masses  $m_Q$ . As in the case of the shell model (cf. see section 2.2.4), thermostating of the charges to a very low temperature ( $\sim 1$  K) provides an acceptable separation of the charge and nuclear degrees of freedom and stable simulation trajectories (59).

### 2.3.3. Coupling Terms $\eta_{ij}$

The physical meaning of the coupling terms  $\eta_{ij}$  in Eq. 33 is that they represent Coulomb interactions between the atoms. At long enough separations,  $\eta_{ij}$  should thus give the usual expression for Coulombic interactions between point charges,  $\eta_{ij} = (8\pi\epsilon_0 r_{ij})^{-1}$ . (Note the extra factor of 1/2 which comes from avoiding double counting of the interaction energies in Eq. 36). At short separations, where the atom-centered charge distributions overlap, a physically plausible approach is to replace the point charge interactions with those between charge distributions  $\rho(\mathbf{r})$  (2),

$$\eta_{ij}(\mathbf{r}) = \frac{1}{8\pi\epsilon_0} \int \frac{\rho_i(\mathbf{r}_i)\rho_j(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{r}|} d\mathbf{r}_i d\mathbf{r}_j. \quad (38)$$

The charge distributions used can be either spherical or elongated. One option is to use atom-centered Slater-type orbitals (56),

$$\phi_{n\zeta} = Ar^{n-1} e^{-\zeta r}, \quad (39)$$

where  $A$  is a normalization constant,  $n$  the principal quantum number, and  $\zeta$  a screening exponent. A computationally more efficient alternative to the Coulomb interaction integrals in Eq. 38 is to use some kind of an empirical combining rule for  $\eta_{ij}$ .

For instance, Nalewajski et al. (60) have proposed a simple form based on the atom type hardness values

$$\eta_{ij}(\eta_i, \eta_j, r_{ij}) = \frac{\frac{1}{2}(\eta_i + \eta_j)}{8\pi\epsilon_0 \left(1 + \frac{1}{4}(\eta_i + \eta_j)^2 r_{ij}^2\right)^{1/2}}. \quad (40)$$

Note that for sufficient large interatomic distances  $r_{ij}$  ( $\sim 0.25$  nm (59)), the screened interaction term of Eq. 40 finds the Coulomb form  $\propto 1/r_{ij}$ . The values of  $\eta_{ij}$  can be fitted to many-body energies of small molecular clusters and molecular dipole moments, electrostatic potential, and polarizabilities. In reference to the latter, the relation between the molecular dipole polarizability and atomic hardness tensors is (59)

$$\boldsymbol{\alpha} = \Delta\mathbf{r}\boldsymbol{\eta}^{-1}\Delta\mathbf{r}^T, \quad (41)$$

where  $\Delta\mathbf{r}$  represents the positions of atoms with respect to the center of geometry.

#### 2.3.4. Restrictions on the Flow of Charge

Equation (33) gives the total electrostatic energy for a system of  $N$  atoms. This expression allows completely free flow of charge throughout the system. While charge transfer effects can be important in quantum-mechanical description of molecular systems (see the chapter by Groenhof on QM/MM methods), in the FQ model, a completely unrestricted electronegativity equalization leads to unrealistically high transfer of charge over long distances (2, 61). This results in overestimation of polarizabilities and molecular dipole moments and is especially problematic for macromolecules such as proteins and DNA. Thus, in practical applications of the FQ model, the redistribution of charge is allowed only within each molecule, or even just certain parts of molecules. The boundary conditions used for the electronegativity equalization, Eq. 36, are the conservation of the net charge of each molecule (or molecule fragment). As an example, the FQ model of the DMPC phospholipid by Davis et al. (18) allows electronegativity equalization only within small fragments of 5–7 heavy atoms, corresponding to small compounds used for parametrizing the force field.

The flow of charge can also be regulated by having atom-pair-specific hardness values, as is done in the bond-charge increment (BCI) (58) and atom–atom charge transfer (AACT) (61) models. In this way charge transfer can even be restricted only to atoms covalently bonded to each other.

#### 2.3.5. Practical Application of the FQ Model

The FQ model has some practical advantages over the two previous models covered in this chapter. It does not entail any new types of interactions (cf. the IPD model) nor does it increase the number of charge–charge interactions (cf. the shell model). If the partial charge degrees of freedom are treated with an extended Lagrangian

scheme, there is only a small computational overhead associated with the dynamics and potential energy minimization with respect to the charges (of the order of 10%) (18, 26, 59). However, the continuously changing values of atomic partial charges necessitate the use of short timesteps in MD simulations. On the one hand, the fictitious charge masses  $m_Q$  in the extended Lagrangian approach need to be small enough so that the charges adapt promptly to changes in their environment. On the other hand, the charge masses should be large enough to allow the use of a reasonably large simulation timestep. In practice, atomistic simulations employing the FQ model are limited to timesteps of  $\sim$ 0.5–1 fs (26, 59), though the simulation times could be reduced with multi-timestep algorithms.

In addition to the computational overhead and practical issues related to restricting the flow of charge, there are certain well-known physical issues related to the FQ model. Since the atomic polarization response by the changing values of the partial charges is isotropic, there is no out-of-plane polarization for planar or linear molecular moieties. This leads to problems, e.g., when simulating charged groups above aromatic rings. Due to the same reason, the FQ model yields too low energies for bifurcated hydrogen bonds (11, 62). These issues can be fixed by combining the FQ polarization model with another one, as is done for instance in the hybrid IPD/FQ model by Stern and co-workers (11).

#### **2.4. Other Methods for Modelling Electronic Polarization**

In addition to the polarization models introduced in sections 2.1–2.3, we briefly mention here a few other methods for including polarization effects in molecular simulations. Naturally, the most accurate way to model electronic polarization in atoms and molecules is by quantum-mechanical methods. However, the high computational cost inhibits the applicability of the full QM approach, and hence, one has to resort to more approximative methods. One option is to use *hybrid QM/MM methods* introduced in the chapter by Groenhof. In QM/MM, only a small subregion of the system, such as one comprised by a ligand and the active site of a protein (63), is treated quantum-mechanically. The rest of the system is modelled using a classical MM force field. The accuracy of the electronic polarizability in the QM part will depend on the chosen level of theory. For the rest of the system, polarization can be included either implicitly or explicitly, e.g., by using one of the methods introduced above. Another option for including QM effects into the computationally efficient force field simulations is to base the molecular interaction terms on electron-structure calculations, as is done in the *X-Pol force field* by Gao and co-workers (64, 65, 66). In this method, electronic polarization is included in the force field by calculating the instantaneous values of atomic partial charges from the molecular wave functions.

In the *continuum dielectric approach*, a molecule is depicted at the atomic level and the solvent or environment as a continuum medium characterized by a dielectric constant (3, 6, 67). Examples of this type of an approach are the studies by Schnieders et al. (68, 69), where the atomistic AMOEBA force field was combined with a continuum solvent. A further simplification of the continuum dielectric approach was recently presented by Tan et al. (70, 71). In their model both the solute and solvent are treated at the continuum level, the only difference between them being their dielectric constant values, while polarization is included as a density of induced dipole moments. The *Langevin dipole (LD) solvation model* is an intermediate between the continuum and discrete approaches (9). In the LD model the use of the dielectric constant is avoided, but a detailed description of solvent molecules is also omitted. The solvent is approximated by polarizable point dipoles fixed on a cubic grid, and the dipoles reorient according to a Langevin-type equation.

### 3. Biomolecular Simulations Based on Polarizable Force Fields

Although atomistic simulations employing polarizable force fields have thus far mainly focused on water, ions, and simple organic liquids (see the review articles (2, 6, 7)), there has been a steady progress in the methodological development toward comprehensive biomolecular force fields. We provide here some highlights of biomolecular simulation studies employing the polarizable force field methodologies presented in this chapter.

#### 3.1. Peptides and Proteins

##### 3.1.1. N-Methylacetamide

N-Methylacetamide (NMA) is a small molecule analog of the peptide backbone of proteins which has been used as a starting point for parametrizing comprehensive polarizable protein force fields. Cieplak et al. showed that taking electronic polarization explicitly into account with the IPD model produces NMA-water hydrogen bonding free energies in a better agreement with *ab initio* calculations than a non-polarizable model (72). The FQ model of NMA by Rick and Berne accurately reproduced the changes in gas-phase charge distributions arising from conformational changes and functional group substitution (73). Calculations carried out with a dielectric continuum solvent also reproduced the aqueous-phase charge distributions and free energies of solvation.

Several parametrizations of NMA based on the shell model (51) were recently developed for the polarizable CHARMM force field by MacKerell and co-workers. The best agreement with experimental data was obtained with the parameter set which most accurately reproduced the *ab initio* gas-phase values for the dipole polarizability

tensor elements. In contrast to the additive CHARMM force field, the polarizable models of NMA accurately reproduced its bulk liquid properties, including the anomalously high relative permittivity.

Xie et al. presented parametrizations of NMA, other amides, and liquid alkanes for an IPD-model-based variant of CHARMM22 called PIPF-CHARMM (74). In addition to providing a good description of bulk liquid properties, detailed parametrizations of the conformational energetics of NMA and alanine dipeptide were concluded to be an encouraging starting point for a complete force field for proteins.

### 3.1.2. AMBER ff02rl

Wang and co-workers optimized peptide backbone dihedrals (75) for the AMBER *ff02* force field (72), which employs the IPD model for electronic polarization. Alanine dipeptide was used as a model system of the parameter fitting for the new AMBER *ff02.rl* force field. The dipeptide as well as a longer Ace-(Ala)<sub>7</sub>-Nme peptide were simulated in water. Conformational distributions of the peptides were concluded to be in agreement with experiments.

### 3.1.3. OPLS/PFF

Since the late 1990s, Friesner, Berne, and co-workers have worked on constructing a comprehensive force field for proteins (5). The first studies focused on parametrizing FQ (58) and hybrid FQ/IPD (11) models for small peptides. Transferability of the parametrizations to other peptides was demonstrated, and calculations of gas-phase conformational energies of these molecules showed clear improvement with respect to the non-polarizable OPLS-AA force field.

The first comprehensive version of the polarizable protein force field, named OPLS/PFF, was presented in 2002 (76). The parametrization of the nonbonded interactions were primarily based on gas-phase *ab initio* calculations. For the most part, the bonded energy terms were adapted from the OPLS-AA force field, but key torsional terms for the amino acid residues were reparametrized. The force field was tested with energy minimization and short MD studies in vacuum. The results showed accuracy comparable to OPLS-AA, but the authors noted that a proper validation of the force field required more realistic simulations of proteins in solvation. The force field was subsequently employed in explicit solvent simulations of bovine pancreatic trypsin inhibitor (BPTI) (45). OPLS/PFF was demonstrated to retain the native structure of the protein as accurately as OPLS-AA. However, a more detailed analysis showed noticeable differences between the polarizable and non-polarizable force field simulations, especially in hydrogen bonding patterns and structural and dynamic properties of water around the protein (77).

The second version of OPLS/PFF (67) included more extensive parametrizations of small organic molecules and an accompanying continuum solvent model, which was fit to reproduce free energy of solvation data for a large set of molecules. The test simulations

included structure optimizations of proteins and protein-ligand complexes, with reasonable agreement to experimental structure data. The importance of accounting for electronic polarizability in  $pK_a$  shift calculations was demonstrated in the case of carboxylic acid residues of the turkey ovomucoid third domain protein (78).

### **3.1.4. FQ CHARMM Force Field**

In 2004, Patel et al. presented an FQ polarizable force field for proteins based on the additive CHARMM22 force field (59, 79). Protein structures obtained with gas-phase energy minimizations as well as short dynamic simulations in vacuum were shown to be comparable to those produced by OPLS/PFF (76). More importantly, condensed-phase MD simulations of six small proteins in solvation resulted in root-mean-square deviations of 1–2.5 Å from the native structures, in line with the accuracy of CHARMM22. The protein force field was later used by the Patel group in studies of a membrane ion channel (19, 80) (see below).

## **3.2. Lipids**

As discussed in section 1.2, lipid bilayers represent systems whose simulation is likely to benefit a great deal from polarizable models. Although there are several well-established additive force fields for lipids, the first polarizable models were introduced only very recently. Harder et al. employed the shell model CHARMM force field to study DPPC lipid monolayers at an air–water interface. They showed that only by using a polarizable model, a good agreement with the experimentally determined interfacial dipole potential could be obtained (20). In a similar fashion, Vorobyov and Allen studied the effect of explicit electronic polarization on the interfacial potential of a DPPC lipid bilayer membrane (81). In this case, the lipid model was based on the CHARMM27 additive force field, with only the lipid hydrocarbon tails chosen as shell model sites. In reference to the non-polarizable force field, a reduced membrane dipole potential was obtained, in line with experiments. The polarizable DPPC model was also used by the same group to study the partitioning of arginine side chain analogs in a DPPC membrane (82), and the effect of membrane deformation on the permeation energetics of a charged arginine analog (83).

Davis et al. have recently simulated lipid bilayer membranes by using FQ models based on CHARMM27 (18, 19). The polarizable models of the DMPC and DPPC lipids produced structural properties of the membranes similar to those obtained with CHARMM27. A noticeable difference between the force fields, though, was a more pronounced permeation of water into the membrane in the case of the polarizable model. This was attributed to the weaker dipole moments of water molecules deeper in the membrane, resulting in a decreased free energy cost of permeating the membrane. Interestingly, in contrast to the study by Vorobyov and Allen (81), in this case, the use of a polarizable force field resulted in a small increase of the membrane dipole potential.

The FQ force field, named CHEQ by the authors, was further used to study the stability of the ion channel gramicidin A in a DMPC bilayer (19) and K<sup>+</sup> ion permeation through the channel (80).

### 3.3. Nucleic Acids

Baucom et al. used MD simulations to study a DNA decamer crystal with different versions of the AMBER force field (84), including the polarizable *ff02* parametrization. Although all the AMBER variants tested performed reasonably well on a 25 ns timescale, the polarizable models of DNA and water provided the best agreement with experimental data. In a follow-up study (25), the authors showed that AMBER *ff02* reproduced sequence-specific effects in the same DNA decamer better than the non-polarizable AMBER *ff99* force field. With a simulation set-up corresponding to the unit cell environment, the ideal DNA B form was shown to converge toward the appropriate crystallographic structure. The AMBER *ff02* force field was also used by Vladimirov et al. (85) to calculate solvent reorganization energies of electron transfer in A-DNA, B-DNA, and rhodamine 6G-DNA complexes, showing a reduction of ~30% from the energies obtained with the AMBER *ff99* force field.

In 2005, Anisimov et al. carried out the first polarizable CHARMM force field nucleic acid simulations (52). Base-base and model compound-water interactions of the shell model parametrization were generally shown to be comparable in accuracy to those of the non-polarizable CHARMM27 force field. However, a 1- ns simulation study of a DNA octamer in an aqueous environment with sodium counterions showed noticeable differences to the CHARMM27 force field, and the structure steadily deviated from that of the DNA B form. An improved parametrization of the polarizable nucleic acid bases was recently introduced (86). The representation of electrostatics in the new model was significantly improved with respect to the CHARMM27 force field. The heats of sublimation of the bases were also better reproduced than with the additive model, though at a small expense of accuracy in molecular volumes and gas-phase binding enthalpies of the bases.

---

## 4. Outlook

Considering the importance of polarization effects, it is actually quite surprising to note for how long non-polarizable force fields have been the most popular tools in the computational study of biomolecular systems. While non-polarizable force fields employing fixed partial charges do feature both orientational and geometric polarization (cf. 1.1), an energetic contribution as important as electronic polarization cannot be overlooked in any force field that aims to produce realistic descriptions of nonbonded

interactions in molecular systems. Moreover, the adaptivity of polarizable force fields is likely to improve both the general accuracy and transferability of the next-generation force fields.

For polarizable force fields to become widely used tools in the future, the associated special interactions and algorithms need to be implemented to the molecular simulation software in an efficient way. Replacing the electrostatic interactions solely based on fixed point charges by one of the non-additive descriptions of electronic polarization further entails a complete reoptimization of the existing non-polarizable force field parameters. Since it has taken nearly four decades for the non-polarizable force fields to reach the present state of accuracy, this is not a trivial task. On the other hand, entirely new MM force fields featuring explicit descriptions of electronic polarization from the first version onward are being developed. The AMOEBA force field (32, 33) is an example of this new generation of force fields. Although AMOEBA is based on a similar total potential energy expression as the current non-polarizable force fields, it further features a multipole representation of the fixed charges in molecules and the IPD model for electronic polarization.

As opposed to the implementation and parametrization of polarizable force fields, their computational cost is hardly an issue anymore. Novel algorithms for treating the degrees of freedom associated with electronic polarization make such simulations only about 2–4 times slower than the ones carried out with non-polarizable force fields. More efficient implementations, decreasing the computational overhead as low as a factor of 1.2–1.3, have also been realized (45, 87). Depending on the application, a trade-off between a reduction in the simulation timescale and the potential gain in accuracy associated with polarizable models seems quite appealing. With the advent of GPU-based computing, which is especially suited for the intensive calculations related to the methods introduced in this chapter, we anticipate a significant increase in the performance of polarizable force fields in the near future.

---

## Acknowledgements

We would like to thank Jaakko Uusitalo and Praveen Nedumpully for their criticism and suggestions on the manuscript. Financial support from the Academy of Finland is gratefully acknowledged.

## References

1. Halgren, T. A., and Damm, W. (2001) Polarizable force fields. *Curr. Opin. Struct. Biol.* 11, 236–242
2. Rick, S. W., and Stuart, S. J. (2002) Potentials and algorithms for incorporating polarizability in computer simulations. *Rev. Comp. Chem.* 18, 89–146
3. Ponder, J. W., and Case, D. A. (2003) Force fields for protein simulations. *Adv. Protein. Chem.* 66, 27–85
4. Yu, H., and van Gunsteren, W. F. (2005) Accounting for polarization in molecular simulation. *Comput. Phys. Commun.* 172, 69–85
5. Friesner, R. A. (2006) Modeling polarization in proteins and protein-ligand complexes: methods and preliminary results. *Adv. Protein. Chem.* 72, 79–104
6. Cieplak, P., Dupradeau, F., Duan, Y., and Wang, J. (2009) Polarization effects in molecular mechanical force fields. *J. Phys.: Condens. Matter* 21, 333102
7. Lopes, P. E. M., Roux, B., and MacKerell, Jr., A. D. (2009) Molecular modeling and dynamics studies with explicit inclusion of electronic-polarizability: Theory and applications. *Theor. Chem. Acc.* 124, 11–28
8. Illingworth, C. J., and Domene, C. (2009) Many-body effects and simulations of potassium channels. *Proc. Roy. Soc. A* 465, 1701–1716
9. Warshel, A., and Levitt, M. (1976) Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103, 227–249
10. Warshel, A. (1979) Calculations of chemical processes in solutions. *J. Phys. Chem.* 83, 1640–1650.
11. Stern, H. A., Kaminski, G. A., Banks, J. L., Zhou, R., Berne, B. J., and Friesner, R. A. (1999) Fluctuating charge, polarizable dipole, and combined models: parametrization from *ab initio* quantum chemistry. *J. Phys. Chem. B* 103, 4730–4737
12. Kaminsky, J., and Jensen, F. (2007) Force field modeling of amino acid conformational energies. *J. Chem. Theor. Comput.* 3, 1774–1788
13. Rasmussen, T. D., Ren, P., Ponder, J. W., and Jensen, F. (2007) Force field modeling of conformational energies: importance of multipole moments and intramolecular polarization. *Int. J. Quantum. Chem.* 107, 1390–1395
14. Geerke, D. P., and van Gunsteren, W. F. (2007) Calculation of the free energy of polarization: quantifying the effect of explicitly treating electronic polarization on the transferability of force-field parameters. *J. Phys. Chem. B* 111, 6425–6436
15. Jorgensen, W. L., McDonald, N. A., Selmi, M., and Rablen, P. R. (1995) Importance of polarization for dipolar solutes in low-dielectric media: 1,2-dichloroethane and water in cyclohexane. *J. Am. Chem. Soc.* 117, 11809–11810
16. Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The missing term in effective pair potentials. *J. Phys. Chem.* 91, 6269–6271
17. Stern, H. A., and Feller, S. E. (2003) Stern HA, Feller SE (2003) Calculation of the dielectric permittivity profile for a nonuniform system: application to a lipid bilayer simulation. *J. Chem. Phys.* 118, 3401–3412
18. Davis, J. E., Rahaman, O., and Patel, S. (2009) Molecular dynamics simulations of a DMPC bilayer using nonadditive interaction models. *Biophys. J.* 96, 385–402
19. Davis, J. E., and Patel, S. (2009) Charge equilibration force fields for lipid environments: applications to fully hydrated DPPC bilayers and DMPC-embedded Gramicidin A. *J. Phys. Chem. B* 113, 9183–9196
20. Harder, E., MacKerell, Jr., A. D., and Roux, B. (2009) Many-body polarization effects and the membrane dipole potential. *J. Am. Chem. Soc.* 131, 2760–2761
21. Gresh, N., Guo, H., Salahub, D. R., Roques, B. P., and Kafafi, S. A. (1999) Critical role of anisotropy for the dimerization energies of two protein-protein recognition motifs: *cis*-N-methylacetamide versus a  $\beta$ -sheet conformer of alanine dipeptide. A joint *ab initio*, density functional theory, and molecular mechanics investigation. *J. Am. Chem. Soc.* 121, 7885–7894
22. Tong, Y., Ji, C. G., Mei, Y., and Zhang, J. Z. H. (2003) Simulation of NMR data reveals that proteins' local structures are stabilized by electronic polarization. *J. Am. Chem. Soc.* 131, 8636–8641
23. Caldwell, J. W., and Kollman, P. A. (1995) Cation- $\pi$  interactions: Nonadditive effects are critical in their accurate representation. *J. Am. Chem. Soc.* 117, 4177–4178
24. Jiao, D., Golubkov, P. A., Darden, T. A., and Ren, P. (2008) Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. USA* 105, 6290–6295

25. Babin, V., Baucom, J., Darden, T. A., and Sagui, C. (2006) Molecular dynamics simulations of DNA with polarizable force fields: convergence of an ideal B-DNA structure to the crystallographic structure. *J. Phys. Chem. B* **110**, 11571–11581
26. Rick, S. W., Stuart, S. J., and Berne, B. J. (1994) Dynamical fluctuating charge force fields: application to liquid water. *J. Chem. Phys.* **101**, 6141–6156
27. Ren, P., and Ponder, J. W. (2003) Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **107**, 5933–5947
28. Yu, H., and van Gunsteren, W. F. (2004) Charge-on-spring polarizable water models revisited: from water clusters to liquid water to ice. *J. Chem. Phys.* **121**, 9549–9564
29. Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B., and MacKerellJr., A. D. (2006) A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **418**, 245–249
30. Bauer, B. A., and Patel, S. (2009) Properties of water along the liquid-vapor coexistence curve via molecular dynamics simulations using the polarizable TIP4P-QDP-LJ water model. *J. Chem. Phys.* **131**, 084709
31. Kunz, A.-P. E., and van Gunsteren, W. F. (2009) Development of a nonlinear classical polarization model for liquid water and aqueous solutions: COS/D. *J. Phys. Chem. A* **113**, 11570–11579
32. Ponder, J. W., Wu, C., Ren, P., Pande, V. S., Chodera, J. D., Schnieders, M. J., Haque, I., Mobley, D. L., Lambrecht, D. S., DiStasio,Jr., R. A., Head-Gordon, M., Clark, G. N. I., Johnson, M. E., and Head-Gordon, T. (2010) Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **114**, 2549–2564
33. Ren, P., Wu, C., and Ponder, J.W. (2011) Polarizable atomic multipole-based molecular mechanics for organic molecules. *J. Chem. Theory Comput.* **7**, 3143–3161
34. CRC Handbook of Chemistry and Physics, 90th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, USA, 2009.
35. Applequist, J., Carl, J. R., and Fung, K. K. (1972) Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *J. Am. Chem. Soc.* **94**, 2952–2960
36. Thole, B. T. (1981) Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **59**, 341–350
37. Birge, R. R. (1980) Calculation of molecular polarizabilities using an anisotropic atom point dipole interaction model which includes the effect of electron repulsion. *J. Chem. Phys.* **72**, 5312–5319
38. van Duijnen, P. T., and Swart, M. (1998) Molecular and atomic polarizabilities: Thole's model revisited. *J. Phys. Chem. A* **102**, 2399–2407
39. Swart, M., Snijders, J. G., and van Duijnen, P. T. (2004) Polarizabilities of amino acid residues. *J. Comp. Meth. Sci. Eng.* **4**, 419–425
40. Wang, J., Cieplak, P., Li, J., Hou, T., Luo, R., and Duan, Y.(2011) Development of polarizable models for molecular mechanical calculations I: Parametrization of atomic polarizability. *J. Phys. Chem. B* **115**, 3091–3099
41. Wang, J., Cieplak, P., Li, J., Wang, J., Cai, Q., Hsieh, M., Lei, H., Luo, R., and Duan, Y. (2011) Development of polarizable models for molecular mechanical calculations II: induced dipole models significantly improve accuracy of intermolecular interaction energies. *J. Phys. Chem. B* **115**, 3100–3111
42. Xie, W., Pu, J., and Gao, J. (2009) A coupled polarization-matrix inversion and iteration approach for accelerating the dipole convergence in a polarizable potential function. *J. Phys. Chem. A* **113**, 2109–2116
43. Nymand, T. M., and Linse, P. (2000) Ewald summation and reaction field methods for potentials with atomic charges, dipoles, and polarizabilities. *J. Chem. Phys.* **112**, 6152–6160
44. Toukmaji, A., Sagui, C., Board, J., and Darden, T. (2000) Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **113**, 10913–10927
45. Harder, E., Kim, B., Friesner, R. A., and Berne, B. J. (2005) Efficient simulation method for polarizable protein force fields: application to the simulation of BPTI in liquid water. *J. Chem. Theor Comput.* **1**, 169–180
46. Wang, W., and Skeel, R. D. (2005) Fast evaluation of polarizable forces. *J. Chem. Phys.* **123**, 164107
47. Lamoureux, G., MacKerell, Jr., A. D., and Roux, B. (2003) A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **119**, 5185–5197
48. Morita, A., and Kato, S. (1999) An *ab initio* analysis chem. Theory Comput., of medium perturbation on molecular polarizabilities. *J. Chem. Phys.* **110**, 11987–11998
49. Schropp, B., and Tavan, P. (2008) The polarizability of point-polarizable water models: density functional theory/molecular mechanics results. *J. Phys. Chem. B* **112**, 6233–6240

50. Yu, H., Whitfield, T. W., Harder, E., Lamoureux, G., Vorobyov, I., Anisimov, V. M., MacKerell, Jr., A. D., and Roux, B. (2010) Simulating monovalent and divalent ions in aqueous solution using a Drude polarizable force field. *J. Chem. Theor. Comput.* **6**, 774–786
51. Harder, E., Anisimov, V. M., Whitfield, T., MacKerell, Jr., A. D., and Roux, B. (2008) Understanding the dielectric properties of liquid amides from a polarizable force field. *J. Phys. Chem. B* **112**, 3509–3521
52. Anisimov, V. A., Lamoureux, G., Vorobyov, I. V., Huang, N., Roux, B., and MacKerell, Jr., A. D. (2005) Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J. Chem. Theor. Comput.* **1**, 153–168
53. Harder, E., Anisimov, V. M., Vorobyov, I. V., Lopes, P. E. M., Noskov, S. Y., MacKerell, Jr., A. D., and Roux, B. (2006) Atomic level anisotropy in the electrostatic modeling of lone pairs for a polarizable force field based on the classical Drude oscillator. *J. Chem. Theor. Comput.* **2**, 1587–1597
54. Mulliken, R. S. (1934) A new electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities. *J. Chem. Phys.* **2**, 782–793
55. Parr, R. G., and Pearson, R. G. (1983) Absolute hardness: companion parameter to absolute electronegativity. *J. Am. Chem. Soc.* **105**, 7512–7516
56. Rappe, A. K., and Goddard III, W. A. (1991) Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **95**, 3358–3363
57. Chelli, R., Ciabatti, S., Cardini, G., Righini, R., and Procacci, P. (1999) Calculation of optical spectra in liquid methanol using molecular dynamics and the chemical potential equalization method. *J. Chem. Phys.* **111**, 4218–4229
58. Banks, J. L., Kaminski, G. A., Zhou, R., Mainz, D. T., Berne, B. J., and Friesner, R. A. (1999) Parametrizing a polarizable force field from ab initio data. I. The fluctuating point charge model. *J. Chem. Phys.* **110**, 741–754
59. Patel, S., and Brooks III, C. L. (2004) CHARMM fluctuating charge force field for proteins: I Parametrization and application to bulk organic liquid simulations. *J. Comput. Chem.* **25**, 1–15
60. Nalewajski, R. F., Korchowiec, J., and Zhou, Z. (1988) Molecular hardness and softness parameters and their use in chemistry. *Int. J. Quant. Chem. Quant. Chem. Symp.* **22**, 349–366
61. Chelli, R., Procacci, P., Righini, R., and Califano, S. (1999) Electrical response in chemical potential equalization schemes. *J. Chem. Phys.* **111**, 8569–8575
62. Liu, Y. P., Kim, K., Berne, B. J., Friesner, R. A., and Rick, S. W. (1998) Constructing *ab initio* force fields for molecular dynamics simulations. *J. Chem. Phys.* **108**, 4739–4755
63. Cho, A. E., Guallar, V., Berne, B. J., and Friesner, R. (2005) Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem.* **26**, 915–931
64. Gao, J. (1997) Toward a molecular orbital derived empirical potential for liquid simulations. *J. Phys. Chem. B* **101**, 657–663
65. Gao, J. (1998) A molecular-orbital derived polarization potential for liquid water. *J. Chem. Phys.* **109**, 2346–2354
66. Xie, W., and Gao, J. Design of a next generation force field: The X-Polpotential. *J. Chem. Theor. Comput.* **2007**, 1890–1900
67. Maple, J. R., Cao, Y., Damm, W., Halgren, T. A., Kaminski, G. A., Zhang, L. Y., and Friesner, R. A. (2005) A polarizable force field and continuum solvation methodology for modeling of protein-ligand interactions. *J. Chem. Theor. Comput.* **1**, 694–715
68. Schnieders, M. J., Baker, N. A., Ren, P., and Ponder, J. W. (2007) Polarizable atomic multipole solutes in a Poisson-Boltzmann continuum. *J. Chem. Phys.* **126**, 124114
69. Schnieders, M. J., and Ponder, J. W. (2007) Polarizable atomic multipole solutes in a generalized Kirkwood continuum. *J. Chem. Theor. Comput.* **3**, 2083–2097
70. Tan, T.-H., and Luo, R. (2007) Continuum treatment of electronic polarization effect. *J. Chem. Phys.* **126**, 094103
71. Tan, T.-H., Tan, C., and Luo, R. (2008) Continuum polarizable force field within the Poisson-Boltzmann framework. *J. Phys. Chem. B* **112**, 7675–7688
72. Cieplak, P., Caldwell, J., and Kollman, P. (2001) Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *J. Comp. Chem.* **22**, 1048–1057
73. Rick, S. W., and Berne, B. J. (1996) Dynamical fluctuating charge force fields: the aqueous solvation of amides. *J. Am. Chem. Soc.* **118**, 672–679

74. Xie, W., Pu, J., MacKerell, Jr., A. D., and Gao, J. (2007) Development of a polarizable intermolecular potential function (PIPF) for liquid amides and alkanes. *J. Chem. Theor Comput.* 3, 1878–1889
75. Wang, Z.-X., Zhang, W., Wu, C., Lei, H., Cieplak, P., and Duan, Y. (2006) Strike a balance: optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. *J. Comput. Chem.* 27, 781–790. See also the correction in Wang et al. (2006) *J. Comput. Chem.* 27, 994
76. Kaminski, G. A., Stern, H. A., Berne, B. J., Friesner, R. A., Cao, Y. X., Murphy, R. B., Zhou, R., and Halgren, T. A. (2002) Development of a polarizable force field for proteins via *ab initio* quantum chemistry: First generation model and gas phase tests. *J. Comput. Chem.* 23, 1515–1531
77. Kim, B., Young, T., Harder, E., Friesner, R. A., and Berne, B. J. (2005) Structure and dynamics of the solvation of bovine pancreatic trypsin inhibitor in explicit water: a comparative study of the effects of solvent and protein polarizability. *J. Phys. Chem. B* 109, 16529–16538
78. MacDermaid, C. M., and Kaminski, G. A. (2007) Electrostatic polarization is crucial for reproducing pKa shifts of carboxylic residues in turkey ovomucoid third domain. *J. Phys. Chem. B* 111, 9036–9044
79. Patel, S., MacKerell, Jr., A. D., and Brooks III, C. L. (2004) CHARMM fluctuating charge force field for proteins: II Protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* 25, 1504–1514
80. Patel, S., Davis, J. E., and Bauer, B. A. (2009) Exploring ion permeation energetics in gramicidin A using polarizable charge equilibration force fields. *J. Am. Chem. Soc.* 131, 13890–13891
81. Vorobyov, I., and Allen, T. W. (2010) The electrostatics of solvent and membrane interfaces and the role of electronic polarizability. *J. Chem. Phys.* 132, 185101
82. Vorobyov, I., Li, L., and Allen, T. W. (2008) Assessing atomistic and coarse-grained force fields for protein-lipid interactions: the formidable challenge of an ionizable side chain in a membrane. *J. Phys. Chem. B* 112, 9588–9602
83. Vorobyov, I., Li, L., and Allen, T. W. (2008) Electrostatics of deformable lipid membranes. *Biophys J* 98:2904–2913
84. Baucom, J., Transue, T., Fuentes-Cabrera, M., Krahin, J. M., Darden, T. A., and Sagui, C. (2004) Molecular dynamics simulations of the *d*(CCAACGTTGG)<sub>2</sub> decamer in crystal environment: comparison of atomic point-charge, extra-point, and polarizable force fields. *J. Chem. Phys.* 121, 6998–7008
85. Vladimirov, E., Ivanova, A., and Röösli, N. (2009) Solvent reorganization energies in A-DNA, B-DNA, and rhodamine 6G-DNA complexes from molecular dynamics simulations with a polarizable force field. *J. Phys. Chem. B* 113, 4425–4434
86. Baker, C. M., Anisimov, V. M., and MacKerell, Jr., A. D. (2011) Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *J. Phys. Chem. B* 115, 580–596
87. Jiang, W., Hardy, D. J., Phillips, J. C., MacKerell, Jr., A. D., Schulten, K., and Roux, B. (2011) High-performance scalable molecular dynamics simulations of a polarizable force field based on classical Drude oscillators in NAMD. *J. Phys. Chem. Lett.* 2, 87–92

# Chapter 10

## Electrostatics Interactions in Classical Simulations

G. Andrés Cisneros, Volodymyr Babin, and Celeste Sagui

### Abstract

Electrostatic interactions are crucial for both the accuracy and performance of atomistic biomolecular simulations. In this Chapter we review well-established methods and current developments for calculating such interactions with a focus on their efficiency and accuracy. Specifically, we review the classical Ewald summations, particle-particle particle-method particle-method Ewald algorithms, multigrid, fast multipole, and local methods. We also highlight some recent developments targeting more accurate, yet classical, representation of the molecular charge distribution.

**Key words:** Electrostatics, Ewald summation, Particle mesh Ewald, Particle-particle particle-mesh, PME, P3M, Multigrid, Fast multipole method, Molecular dynamics

---

### 1. Introduction

Electrostatic interactions in biomolecular simulations have traditionally been modeled using the atom-centered ‘partial charge’ approximation in which the full charge density of the system is replaced by point charges, and, possibly, higher moments, assigned to every atom. Despite the apparent simplicity of the model, the long range nature of the Coulomb interactions demands carefully constructed algorithms in order to take advantage of the existing and emerging computer architectures. In the past, long-range forces in macromolecular simulations were truncated with artificial cutoffs. Advancements in computer architectures coupled with algorithmic developments have made possible the inclusion of full long-range interactions. These can now be computed at a small cost, generally less than that of a typical cutoff calculation. In addition, the use of cutoffs has been shown to lead to severe artifacts in simulations of water (1–6), peptides (7, 8), proteins (9), DNA (10–13), and lipids (14–19). For instance, the truncation scheme in lipid membranes induces artificial structure at the scale of the truncation distance, and

this results in change of large-scale properties such as phase behavior and elasticity (15, 18, 19). The use of switching or shifting functions can alleviate the problem; however, artifacts still are present.

In this work, we review important algorithmic developments used in biomolecular simulations with explicit solvent. We will not consider the treatment of systems that have a finite size, such as quasi two-dimensional systems in slab geometry. Methods developed for this geometry include several adaptations of the Ewald summation (20–24), Lekner summation (25–31), and the MMM2D method (32, 33). Real-space methods like multigrid or fast multipoles can naturally deal with these types of geometries, and they will be briefly reviewed below. Likewise, the space limit precludes the review of implicit solvent methods; a good review is given in (34).

---

## 2. The Electrostatic Problem

The first element needed for an electrostatic description is the assignment of a charge distribution in space, and its mathematical formulation. While the positive nuclear charge may be considered discrete on the atomic scale, the representation of the continuous negative electronic charge distribution poses a problem. In classical descriptions, the electronic density is replaced by a discrete set of point charges, dipoles, and perhaps higher-order multipoles. The total electrostatic potential due to the charge distribution is then expressed as a sum of successive multipolar potentials. In principle, such a “distributed multipole” description can exactly describe the potential due to the true charge density, at points distant from the expansion centers where “penetration” effects are negligible (35). In practice this expansion is truncated, usually at the monopole level in most current force fields. The distributed monopoles are usually referred to as “partial” charges, since they need not have integral values. The long-range effects are most pronounced for the monopolar potential, which decays with distance  $r$  as  $1/r$ . Physically, the Coulomb electrostatic energy due to a set of charges represents a well-posed mathematical problem: by knowing all the partial charges  $q_i$  and their positions  $\mathbf{r}_i$ , one can compute the interactions, no matter how complicated the configurations of charges may be. In practice such computation is far from trivial, due to two main computational considerations: the finite system size and the large size of the molecules. Since all simulations have finite-size constraints, the infinite range of the interactions, as well as other quantities that characterize the properties of the system in the thermodynamic limit, are dealt with through boundary conditions. With respect to molecular size, for a system with  $N$  charges, the

electrostatic energy represents a sum over  $N(N - 1)/2$  pairs, i.e., it is an  $\mathcal{O}(N^2)$  calculation. Thus, a direct pair computation is too costly for typical biomolecular sizes.

Boundary conditions may be conveniently divided into periodic boundary conditions (PBC) and non-periodic boundary conditions (NPBC). Generally vacuum and implicit solvent simulations are coupled to NPBC, while explicit solvent simulations can be carried out with both PBC and NPBC (when real-space methods are used). Intermediate approaches, where the solute is surrounded by a thin layer of explicit waters, have also been treated with both NPBC and PBC. In these models, special care is needed for the boundary between the explicit and continuum descriptions, as it may become a source of artifacts. Although both NPBC and PBC have been criticized as introducing artifacts in the simulations, there is growing consensus on the soundness of PBC in the treatment of long-range forces. In the next sections we review current methods for dealing with electrostatics computation.

### 3. Ewald-Based Methods

The Ewald summation method was introduced in 1921 by Ewald (36) to compute the electrostatic energy of ionic crystals. These are described by a Coulomb series that is only conditionally convergent, i.e., its limit depends on how the series is summed. With the use of a Jacobi theta transform, the slowly, conditionally convergent Coulomb sum was converted into two sums that converge rapidly and absolutely: the *direct* and the *reciprocal* sums, with the conditionally convergent nature of the series hidden in a surface term (37). This term depends on the dipole moment of the unit cell  $U$ , the dielectric constant of the surrounding medium, and the order of summation, that is, on the asymptotic shape of the array of copies of  $U$ , and vanishes when the surrounding medium has infinite dielectric (“tin-foil” boundary conditions). Although the surface term is only rarely used, its use has been recommended for calculations of the dielectric constant of the sample (38). Ewald-based methods use the error function erf and the complementary error function erfc to write the Coulomb interaction as  $1/r = \text{erfc}(\beta r)/r + \text{erf}(\beta r)/r$ . The first term is short-range and gives rise to the direct sum, while the second term is long-range, and is handled in the reciprocal sum via Fourier transforms. Such splitting of the Coulombic interactions involves a smooth function of  $r$  and therefore avoids any form of discontinuity in the forces, or higher-order derivatives of  $1/r$ . Ewald-based methods include the original Ewald summation (36), the particle-particle particle-mesh method (39, 40), the particle-mesh Ewald algorithm (41, 42), and the fast Fourier Poisson (FFP) (43) method.

### 3.1. Ewald Summations

Under PBC, the system to be simulated comprises a unit simulation cell  $U$ , whose edges are given by the vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ , and  $\mathbf{a}_3$  (not necessarily orthogonal). The volume of  $U$  is given by  $V = \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3$ . A lattice vector is defined as  $\mathbf{n}_1 = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$  for all integer triples  $(n_1, n_2, n_3)$ , with  $n_1, n_2, n_3$  not all zero. A standard way (37, 44, 45) of treating Coulombic interactions under PBC is to model the system as a large but finite ensemble of copies of the unit cell  $U$ , which is then immersed in a dielectric medium. Each particle at position  $\mathbf{r}_i$  within the cell interacts with all the other particles in  $U$ , with their and its own periodic images, as well as with the reaction field induced in the surrounding dielectric medium. Infinite lattice sums are obtained in the limit of an infinite ensemble. The traditional Coulomb (monopolar) potential  $\phi_0(\mathbf{r}_i)$  acting on charge  $i$  is expressed as:

$$\phi_0(\mathbf{r}_i) = \sum'_{\mathbf{n}} \sum_{j=1}^N \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}|}, \quad (1)$$

where the prime indicates that terms with  $i = j$  and  $\mathbf{n} = 0$  are omitted.

The Coulomb infinite series conditionally converges slowly to a finite limit only if  $U$  is electrically neutral, when it is not, the standard practice is to add a uniform background charge density (the “neutralizing plasma”), that fixes the problem with the monopole term, without disturbing the higher-order terms. If one carries out a multipolar expansion of the electronic density, octupolar and higher-order terms in the Coulomb potential converge *absolutely*, but the monopolar and dipolar terms do not. For a neutral unit cell, monopolar and dipolar terms converge but only conditionally.

The Ewald method has also been extended to higher-order multipoles (46–49). Here we illustrate an extension of the Ewald method to a system of  $N$  atoms where electrostatics is represented simply by point partial charges,  $q_1, q_2, \dots, q_N$ , and point dipoles,  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N$  (some of which can be zero) at positions  $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$  within the unit cell, assuming charge neutrality:  $q_1 + q_2 + \dots + q_N = 0$ . As before, the edges of the unit cell are denoted by vectors  $\mathbf{a}_\alpha$ ,  $\alpha = 1, 2, 3$ , and the conjugate reciprocal vectors  $\mathbf{a}_\alpha^*$  are defined by the relations  $\mathbf{a}_\alpha^* \cdot \mathbf{a}_\beta = \delta_{\alpha\beta}$  (the Kronecker delta), for  $\alpha, \beta = 1, 2, 3$ . The point charge  $q_i$  at position  $\mathbf{r}_i$  has fractional coordinates  $s_{\alpha i}$ ,  $\alpha = 1, 2, 3$  defined by  $s_{\alpha i} = \mathbf{a}_\alpha^* \cdot \mathbf{r}_i$ . The charges and dipoles interact with each other, and with their periodic images. The electrostatic energy of the unit cell is then written as:

$$E(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{2} \sum'_{\mathbf{n}} \sum_{i,j} (q_i + \mathbf{p}_i \cdot \nabla_i) (q_j - \mathbf{p}_j \cdot \nabla_j) \left( \frac{1}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}|} \right) \quad (2)$$

where the outer sum is over the vectors  $\mathbf{n}_1 = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ , the prime indicating that terms with  $i = j$  and  $\mathbf{n} = 0$  are omitted.

For molecular systems, corrections are introduced to account for the “masked pairs,” which are atom pairs  $(i, j) \in M$ ,  $M$  the masked list, whose nonbond interactions should not be calculated, since they are accounted for by other terms in the potential. These masked pairs are easily omitted from the direct sum. However, since all pair interactions are unavoidably included in the Fourier treatment of the reciprocal sum, the reciprocal part of the masked pair contributions must be separately subtracted. Computationally, this is conveniently achieved by skipping these pairs in the direct sum and separately adding the sum over  $(i, j) \in M$  of the “adjusted” interaction energy, which is the difference between the direct sum interaction and the standard Coulomb interaction between the charge, dipole pairs. Similarly the self-energy, which is the reciprocal part of the interactions of the charge, dipole pairs with themselves, must be removed. This is achieved by calculating the adjusted interaction energy of the charge, dipole pair with an identical pair at distance  $r$  and taking the limit as  $r \rightarrow 0$  (the result is independent of the path taken). The electrostatic energy can then be written as a sum of four terms (47), the direct, the reciprocal, the adjusted, and the self terms:  $E = E_{dir} + E_{rec} + E_{adj} + E_{self}$ .

The direct sum energy is then given by

$$E_{dir} = \frac{1}{2} \sum_{\mathbf{n}}^* \sum_{i,j=1}^N (q_i + \mathbf{p}_i \cdot \nabla_i)(q_j - \mathbf{p}_j \cdot \nabla_i) \frac{\text{erfc}(\beta|\mathbf{r}_j - \mathbf{r}_i + \mathbf{n}|)}{|\mathbf{r}_j - \mathbf{r}_i + \mathbf{n}|} \quad (3)$$

where the asterisk over the sum denotes that the terms with  $\mathbf{n} = 0$  and either  $j = i$  or  $(i, j) \in M$  are omitted. The adjusted energy has a similar form to Eq. 3, with erfc replaced by -erf:

$$E_{adj} = -\frac{1}{2} \sum_{(i,j) \in M} (q_i + \mathbf{p}_i \cdot \nabla_i)(q_j - \mathbf{p}_j \cdot \nabla_i) \frac{\text{erf}(\beta|\mathbf{r}_j - \mathbf{r}_i + \mathbf{n}|)}{|\mathbf{r}_j - \mathbf{r}_i + \mathbf{n}|}. \quad (4)$$

The reciprocal energy  $E_{rec}$  has the same form as the traditional monopolar case, but it requires a generalization of the structure factor  $S(\mathbf{m})$  to include the dipolar interactions. With the reciprocal lattice vectors  $\mathbf{m}$  defined by  $\mathbf{m} = m_1 \mathbf{a}_1^* + m_2 \mathbf{a}_2^* + m_3 \mathbf{a}_3^*$  with  $m_1, m_2, m_3$  integers not all zero, the reciprocal electrostatic energy is given by

$$E_{rec} = \frac{1}{2\pi V} \sum_{\mathbf{m} \neq 0} \frac{\exp(-\frac{\pi^2 \mathbf{m}^2}{\beta^2})}{\mathbf{m}^2} S(\mathbf{m}) S(-\mathbf{m}). \quad (5)$$

Here,  $S(\mathbf{m})$  is the structure factor

$$S(\mathbf{m}) = \sum_{j=1}^N [q_j + 2\pi i \mathbf{p}_j \cdot \mathbf{m}] \exp(2\pi i \mathbf{m} \cdot \mathbf{r}_j), \quad (6)$$

where  $\exp(2\pi i \mathbf{m} \cdot \mathbf{r}_j) = \exp(2\pi i(m_1 s_{1j} + m_2 s_{2j} + m_3 s_{3j}))$  and  $s_{\alpha j}$ ,  $\alpha = 1, 2, 3$  are the fractional coordinates of site  $j$ , defined above.

Finally, the self-energy term is given by

$$E_{self} = -\frac{\beta}{\sqrt{\pi}} \sum_i^N \left( q_i^2 + \frac{2\beta^2}{3} |\mathbf{p}_i|^2 \right). \quad (7)$$

The electrostatic field and force on atom  $i$  at position  $\mathbf{r}_i$  are computed as the negative gradient of the electrostatic potential  $\phi(\mathbf{r}_i)$  and electrostatic energy  $E(\mathbf{r}_i)$ , respectively (and thus they have a direct, adjusted and reciprocal components):

$$\begin{aligned} \mathbf{E}(\mathbf{r}_i) &= -\nabla_i \phi(\mathbf{r}_i), \\ \mathbf{F}(\mathbf{r}_i) &= -\nabla_i E(\mathbf{r}_i) = -(q_i + \mathbf{p}_i \cdot \nabla_i) \nabla_i \phi(\mathbf{r}_i), \\ \boldsymbol{\tau}(\mathbf{r}_i) &= \mathbf{p}_i \times \mathbf{E}(\mathbf{r}_i). \end{aligned} \quad (8)$$

An important difference with the monopolar case, is that, for non-zero dipoles, the force is no longer parallel to the field, and a torque  $\boldsymbol{\tau}(\mathbf{r}_i)$  needs to be calculated. This torque, in turn, leads to extra terms in the atomic forces (47). The electric field has also a self term or “self field,” given by

$$\mathbf{E}_{self}(\mathbf{r}_i) = \frac{4\beta^3}{3\sqrt{\pi}} \mathbf{p}_i. \quad (9)$$

The stress tensor  $\Pi$ , necessary for constant pressure simulations using Ewald summations, is computed in the usual manner (50–52), where the direct (adjusted) part is given by combining the various components of the direct (adjusted) forces and distances. The computation of the reciprocal tensor, on the other hand, gives a new term compared to the monopolar contribution. The latter originates on the  $\mathbf{a}$ -dependence of the volume  $V$  and of the kernel  $\exp(-\pi^2 \mathbf{m}^2 / \beta^2) / \mathbf{m}^2$ . In the monopolar case the structure factor  $S(\mathbf{m})$  is independent of the vectors  $\mathbf{a}_\alpha$ ,  $\alpha = 1, 2, 3$ . The dipolar interaction, as Eq. 6 shows, introduces a dependence of  $S(\mathbf{m})$  on  $\mathbf{a}$ . The *additional* contribution is:

$$V\Pi_{rec\ \alpha\beta}^{(1)} = -\sum_{i=1}^N \mathbf{p}_{i\beta} \mathbf{E}_{rec\ \alpha}(\mathbf{r}_i), \quad (10)$$

where  $\mathbf{E}_{rec\ \alpha}(\mathbf{r}_i)$  is the  $\alpha$  component of the reciprocal electrostatic field at position  $\mathbf{r}_i$ . Of course, other nonbond interactions such as van der Waals interactions also contribute to the virial tensor, in addition to the bond, angle, and dihedral interaction contributions (53).

The equations above are valid for *total* dipoles  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N$ . Each total dipole  $\mathbf{p}_i$  can be written as the sum of a *permanent* dipole  $\mathbf{s}_i$  and an *induced* dipole  $\mathbf{d}_i$ ,  $\mathbf{p}_i = \mathbf{s}_i + \mathbf{d}_i$ . The induced polarization is a many-body problem, where the induced dipoles all depend on

each other. For the simplest case where the induced dipole  $\mathbf{d}_i$  is directly proportional to the total electric field at  $\mathbf{r}_i$ ,  $\mathbf{E}_i^T$ :

$$\mathbf{d}_i = \alpha_i \mathbf{E}_i^T = \alpha_i \left[ \mathbf{E}_i^0 + \sum_{j=1, j \neq i}^N \mathbf{T}_{ij} \cdot \mathbf{d}_j \right], \quad (11)$$

where  $\alpha_i$  is the polarizability of site  $i$ .  $\mathbf{E}_i^T$  is the total electric field at  $\mathbf{r}_i$ , which is the sum of an “external” field  $\mathbf{E}_i^0$  (due to charges, permanent dipoles, and any other applied field) at  $\mathbf{r}_i$ , and an interaction field that describes, through the dipole interaction tensor  $\mathbf{T}_{ij}$ , the net effect of all the other induced dipoles  $\mathbf{d}_j$  on dipole  $\mathbf{d}_i$ . Under this simple representation of the polarization, the total energy given in Eq. 2 is augmented by the term  $\sum_{i=1}^N (|\mathbf{d}_i|^2 / 2\alpha_i)$ .

Clearly all the equations above recover the traditional Ewald expression when the dipoles are set to zero. The erfc function in the Ewald direct sum originates in Gaussian screening functions centered at each point particle. The reciprocal sum is given by the Fourier series representation of the solution to Poisson’s equation in PBC, using the sum of the compensating Gaussians, again centered at each point particle, as a source distribution. The relative contributions and computational cost of the direct and reciprocal sum can be tuned by varying the width (the  $\beta$  parameter in the above expressions) of the Gaussian distributions. Narrow Gaussians (large  $\beta$ s) make the erfc in the direct sum decay fast, and a shorter cutoff is needed for the direct sum, while more terms are needed in the reciprocal space. Wide Gaussians have the opposite effect: the exponential in the reciprocal sum decays fast and fewer terms are needed, while the direct sum needs more terms. In particular, if the Gaussians are chosen to vanish (within a prescribed tolerance) at a standard cutoff distance independent of system size  $N$ , conventionally taken to coincide with the cutoff of the Lennard-Jones interactions, the direct sum is  $\mathcal{O}(N)$ , but the reciprocal sum becomes  $\mathcal{O}(N^2)$ . If, on the other hand, the Gaussians are chosen to vanish at a cutoff of half the cell size, the direct sum is over all minimum image pairs, and thus is  $\mathcal{O}(N^2)$ , while the reciprocal sum is  $\mathcal{O}(N)$ .

Possible artifacts of the Ewald summation methods have been explored by many authors in the past (an extensive listing is presented in (54)). The general conclusion is that the artifacts are negligible for high dielectric solvents, with typical simulation cell sizes. A recent example involves an octaalanine peptide with charged termini in explicit solvent, for which severe effects due to the use of Ewald sums were reported in the past. Villarreal and Montich (55) have shown that there are no periodicity-induced artifacts, and that incomplete sampling is likely to affect the results to a larger extent than any artifacts induced by the use of Ewald sums. These conclusions seem to be supported by other long-time simulations (54–57). The importance of correct sampling for a

system with explicit solvent and PBC can be illustrated with a very simple system: zwitterionic trialanine (58). The effect of boundary conditions on calculated free energies has also been intensely studied (59). In particular, ionic charging free energies provide a very simple system to study these effects, and have been shown to exhibit remarkably system size consistency (60) in Ewald calculations in high dielectric solvents. For low dielectric solvents, analytical, finite-size corrections have been derived (61–63). Size-consistency is a necessary but not sufficient condition for correct ionic charging free energies. For instance, these energies were computed both with Ewald in PBC and with spherical boundary conditions (64, 65). Size consistent results were obtained in both cases, but the results differed substantially. The difference between the two boundary conditions was traced to the electric potential drop across the vacuum-water interface in the spherical cluster. Using Gauss' law together with the spherical symmetry, a rigorous treatment of the charging process in cluster interiors is possible. The results for a variety of simple ions using this treatment are in quantitative agreement with those using Ewald summation, even when low dielectric solvents (with system size corrections) were used. From these results (64, 65), it is clear that ionic charging free energies are correct when using the Ewald summation and that the surface potential of water must be considered in spherical clusters.

### 3.2. Particle-Mesh Methods

The particle-mesh (PM) methods, in analogy to the Ewald method, split the electrostatic potential and energy into a direct and a reciprocal part, but the reciprocal part is computed on a mesh. With a fixed cutoff (say 8 or 9 Angstroms), the direct part scales as  $\mathcal{O}(N)$ , and the scaling of the reciprocal part determines the overall scaling of the method. PM methods compute the reciprocal potential by solving Poisson's equation of a smooth charge distribution on a grid. This can be done via fast Fourier transform (FFT) under PBC or via a discrete Poisson solver in real space under a wide variety of boundary conditions. FFT PM methods include the particle-particle particle-mesh method (39, 40) (P3M), the particle-mesh Ewald algorithm (41, 42) (PME), and the FFP (43) method.

All the FFT PM methods transform the reciprocal sum into a sum over coefficients of the discrete Fourier transform of a gridded charge distribution, whose calculation is accelerated to  $\mathcal{O}(N \log N)$  via three-dimensional FFTs. Although the methods can be used with different switching functions (39) (that give the relative weight of direct and reciprocal parts), the use of Gaussian screening functions as in the original Ewald method and an optimal Green's function (66) greatly improves the accuracy of PM methods. The methods differ in how they transform the continuous charge density due to the sum of compensating Gaussians onto a regular three-dimensional grid, and how they compensate for the loss of accuracy introduced in this process. These methods have been

carefully studied and compared in the literature (40, 67–71). Of these, the PME algorithm of Darden et al. (41, 42) is perhaps the most broadly used. The method approximates the complex exponentials in the reciprocal sum by an Euler spline interpolation, based on cardinal B-splines in the smooth PME (42). B-splines are factored into three one-dimensional B-splines which can be differentiated analytically via recursion relations. The forces are obtained analytically from the gradient of the energy in the smooth PME method at very low cost, requiring only two FFTs (versus four FFTs for equivalent accuracy in the P3M method). The actual steps in the algorithms for either P3M or PME are well-known and described in the literature, and so are the differences and similarities between the two methods (40, 68–71). PME for molecular simulations with explicit solvent has been implemented in most of the major biomolecular codes: AMBER (72), CHARMM (73), NAMD (74), GROMACS (75), DESMOND (76), etc.

Another variant of the particle mesh approach is the FFP method, proposed by York and Yang (43). This method directly samples the compensating Gaussians onto the grid, and avoids loss of accuracy from interpolation by use of a clever identity, later used in multigrid methods. Due to the cost of sampling the Gaussians, this method is not competitive for the modest accuracies appropriate for present force fields, but it appears to be more efficient than the above methods for high accuracy requirements.

Recent methodological advances include the staggered mesh Ewald (StME) algorithm by Cerutti et al. (77), that uses an old technique named “interlacing”(78, 79) to improve the accuracy of the smooth PME method by averaging the results of two separate reciprocal space calculations. This allows to replace the original mesh in the PME method with two separate coarser meshes, each less than one-third of the original size. The resulting StME can improve the accuracy of computed forces by 1.2 orders of magnitude and exceed the efficiency of the smooth PME with similarly optimized parameters. In addition, the method also offers advantages for parallel implementations. The success of interlacing is due to the fact that it suppresses the aliasing effects in the forces.

A promising combination of the SPME and multigrid ideas has been presented recently by Cerutti and Case (80). The method, termed “multi-level Ewald” (MLE), splits the convolution required by standard P3M algorithms into a series of convolutions over slab-shaped subregions of the original simulation cell. A convolution over the entire cell is performed using a much coarser mesh. The MLE approximation does not seem to introduce significant errors compared to the SPME method. It is anticipated that the division of the single convolution into many independent subproblems will be useful for extending the parallel scaling of molecular simulations.

The parallel scaling of the PME methods is usually limited by the scaling of the three-dimensional FFT that requires all-to-all

communication. For typical meshes on commodity networks, the number of messages to be sent and their associated latency is the primary determinant of FFT communication time (76). To improve the overall parallel performance, the latency-sensitive FFT computations are typically interleaved with the latency-tolerant (and dominant) cutoff limited calculations. The division of work can be either explicit (75, 81), in which case a number of processors can be dedicated to the reciprocal PME computations, or adaptive. The latter approach is used in the NAMD program (74) taking advantage of its message-driven architecture. It allows for great scalability of the resulting molecular dynamics code (82).

With respect to real-space PM methods, multigrid has become relatively popular because of its favorable scaling as  $\mathcal{O}(N)$  and because of its appeal for parallelization. Parallelization, however, is far from a settled issue and recent breakthroughs in the treatment of FFT allow the PME methods in codes such as NAMD to be highly parallelizable. At present, it is hard to predict how the FFT PM and the multigrid methods will compare in the next parallelization breakthrough. First introduced in the 1970s by Brandt (83–85), multigrid methods can solve linear and non-linear elliptic partial differential equations and integro-differential equations on grids with  $K$  points in  $\mathcal{O}(K)$  operations (86). Their appeal derives from the ease of handling any sort of boundary conditions; great functionality in implicit solvent methods, like Poisson-Boltzmann approaches (87–92); possible advantages in the implementation of highly accurate, continuous electrostatics; possible adaptability in efficient multiple-time-step integrators that separate fast and slow interactions (93); etc. In particular, multigrid or the related multi-level methods (88–92) have been very successful in the static calculation of electrostatic potentials for large, complex molecules in implicit solvent, specially with the use of finite elements. In this chapter, we restrict ourselves to multigrid as applied to molecular dynamics with explicit solvent.

Multigrid algorithms introduce successive auxiliary grids with larger mesh spacings in order to eliminate all different frequency components of the error in Fourier space (83–85). The resulting accelerated convergence on all length scales allows for the solution of elliptical partial differential equations on grids with  $K$  points in  $\mathcal{O}(K)$  operations (83). For condensed-matter systems, this translates into a truly  $\mathcal{O}(N)$  linearly scaling method. By comparison, other well-known relaxation methods such as the Gauss-Seidel, the alternating direction implicit method and the iterative SOR methods exhibit a convergence rate that is lower than that of multigrid methods. For instance, in (94) it is shown that an *optimal* SOR method scales as  $\mathcal{O}(N^{4/3})$ , because the cost per iteration also increases as the size of the system increases.

A first attempt to use multigrid in the context of biomolecular simulations consisted of an adaptive multigrid technique

implemented into the original fast multipole method (FMM) (95). This technique is competitive with the FMM only for low accuracy (comparable to FMM with 4th–6th order multipoles). Molecular dynamics was not implemented, but given what is known about lack of energy conservation for FMM with lower-order multipoles, one would expect that the method did not represent a solution for molecular dynamics. Sagui and Darden (96) introduced a multigrid method appropriate for molecular dynamics of biomolecules in explicit solvent. Along general lines, the method consists of: (i) a charge assignment scheme, which both interpolates and smoothly assigns the charges onto a grid; (ii) the solution of Poisson’s equation on the grid via multigrid methods; and (iii) the back interpolation of the “reciprocal” forces from the grid to the particle space, to be added to the direct and adjusted forces. For classical calculations, the charge assignment scheme (i.e., the spreading of the charges on the mesh) was found to be crucial for the performance of the method. To obtain reasonable accuracies, the charge density on the grid must be as smooth as possible. This means that the charges must be spread over a fairly large number of grid points, that can quickly become an important time-consuming part of the algorithm. Two methods were used for interpolating the charge: a lattice diffusion multigrid scheme (LDM) and a lattice Gaussian multigrid scheme (LGM). The LDM method uses a diffusion scheme with an adjustable diffusion constant and results in relative force errors of approximately  $0.5 \times 10^{-3}$ , which is not satisfactory for biomolecular simulations. The LGM on the other hand, uses Gaussians as interpolating functions and gives relative force errors of  $10^{-4}$  or less, and relative energy errors of approximately  $10^{-5}$  – at considerably more cost than the LDM method. Both methods showed excellent energy conservation in the nanosecond scale. The high accuracy of the LGM method was achieved by using, instead of the traditional partition in the Ewald method, the partition proposed in the FFP (43). In addition, since standard discretization schemes proved not accurate enough, a deferred defect correction scheme (83, 97, 98) for the LGM was applied.

Since then, several other variations of multigrid have attempted to speed up the method (99–101). The challenge has been to obtain a method that is efficient, while preserving accuracy, since relative force errors for reliable biomolecular MD trajectories need to be of the order of  $10^{-4}$ . Goedecker and Chauvin (100) proposed a modified multigrid V cycle scheme that employs wavelets and is more efficient than the standard V cycle. Instead of transferring the residue to the coarser grid levels, this scheme works with copies of the charge density on the different grid levels that were obtained from the underlying charge density on the finest grid by wavelet transformations. The authors show that halfway V cycles with the restriction and prolongation steps based on wavelet theory are the most efficient approach for the solution of the three-dimensional

Poisson's equation on the grid. Currently, it is not clear whether wavelets will improve efficiency when applied to charge and force interpolation.

Two recent studies (102, 103) have applied convolutions to accelerate the charge assignment and force interpolation steps in the LGM method (which simply used separability of three-dimensional Gaussians to pretabulate one-dimensional Gaussians). Both schemes effectively reduce the number of mesh point at which the charge spreading function is nonzero by proceeding in two stages. In the first stage, charges are spread onto the mesh, and in the second-stage, a convolution with a Gaussian is performed to further spread the charges. Shan et al. (103) used another Gaussian (with a smaller support) for the first stage of charge spreading, and this same Gaussian was used for the final interpolation to calculate the forces. Banerjee and Board (102), on the other hand, used a set of weights on the grid that then were further convolved with the second Gaussian. These same weights are used for the back interpolation to compute the forces, but in a different manner. On a single-processor, both convolution-based methods reduce the computational burden by a factor of about 2 when compared to the earlier LGM. Parallelization studies were not reported in (103). The convolutions implemented in (102) perform well in parallel environment, although the multigrid implementation itself performs poorly in communication constrained parallel environments and needs further improvement of parallel performance.

Among other novel applications of the multigrid method, there has been its implementation in dissipative particle dynamics (DPD) by Groot (104), and its application to higher-order multipoles by Sagui et al. (49). Groot incorporated electrostatics into DPD by solving for the electrostatic field locally on a grid with a multigrid method. The method captures the essential features of electrostatic interactions and satisfies screening of the potential near a charged surface and the Stillinger-Lovett moment conditions. The method is specially applicable to (coarse-grained) simulations of mixed surfactants near oil-water interfaces or to simulations of Coulombic polymer-surfactant interactions.

## 4. Other Methods

### 4.1. Fast Multipole Methods

FMMs treat Coulombic interactions exactly for particles within the same or neighboring subcells of the simulation cell, and evaluate the potential for more distant particles through multipolar expansions. Note that close range interactions are handled exactly, unlike those in the lattice sums. However, the splitting of the Coulombic interactions is *not* a smooth function of  $r$ , and thus, these methods

may suffer in comparison to lattice sum methods in their simulation behavior (105). The original FMM (106, 107) uses NPBC, but it has also been extended to PBC (40, 108–110).

Here, we briefly review the original FMM of Greengard and Rokhlin (106, 111–114). The method can be discussed in terms of a number of contributing developments. The main elements of the method are: the following (i) multipole expansions in terms of spherical coordinates; (ii) octal tree constructions, which consist of successive subdivisions of the initial cell into self-similar subcells in order to calculate the expansion of the far field created by the particles inside the cell; and (iii) translation operators for transformations of multipole expansions. The key of the tree methods is to compute interactions between well-separated clusters through multipole expansions with rigorous bounds on the error. For a given precision  $\epsilon$ , it is necessary to use  $p = -\log_{\sqrt{3}}(\epsilon)$  terms in the expansion (113) (i.e.,  $p$  is the multipolar order). The major hurdle in the original FMM scheme is the conversion of multipole to local translations in the “downward pass,” a transformation which requires approximately  $189 Np^4$  operations. More recent efforts include the combination of multipole expansions with exponential or plane wave expansions, which diagonalize the translation operators (113). The most recent adaptive FMM for the Laplace equation uses new compression techniques and diagonal forms for the translation operators to accelerate the convergence of the method (115). For relatively uniform charge distributions, the method scales as  $27Ns + 2Np^2 + 189 \frac{N}{s} p^2 + 20 \frac{N}{s} p^3$ , where  $s$  is the number of charges per box at the finest level.

The relative efficiency of FMM versus PM-based approaches for biomolecular simulations in explicit solvent probably is still an open question. A traditional study is that reported by Pollock and Glosli (40), although this is rather old now. The authors concluded that the P3M approach is significantly more efficient than FMM for any conceivable system size, despite the superior asymptotic scaling of the latter ( $O(N)$  versus  $O(N \log N)$ ). They also discuss other advantages of the P<sup>3</sup>M method, the most obvious one being the ease of coding. Also, non-cubic periodic cells can be easily implemented with the particle-mesh methods, but so easily with the FMM. Although parallel implementations of FMM are supposed to scale better than P3M or PME with the number of processors, we are not aware of recent comparisons. Skeel et al. (99) carried out a comparison between multigrid and FMM methods. Their results suggest that the multigrid methods are superior to FMMs. The FMM approach, on the other hand, has found a niche in solid state physics and scattering problems. The method may be better suited to problems involving highly nonuniform particle density, since the tree approach lends itself well to an adaptive approach (116), whereas in these cases, the grid size in the PM methods will grow

faster than linearly with the number of particles (however, local finer subgrids, especially in multigrid methods, can circumvect this problem). In general, FMM suffers from a lack of energy conservation unless very high accuracy is employed (96, 105, 117), whereas the PM methods have very good energy conservation properties. Efficient parallel implementations can be found in (118, 119). A comprehensive and insightful review of the method is given by Kurzak and Pettitt (120).

#### 4.2. Local Methods

The electrostatic problem described in the previous sections can be considered as the static limit of Maxwellian electrodynamics, where the magnetic field  $\mathbf{B}$  and the current  $\mathbf{J}$  produced by moving charges have vanished. Since the electric field  $\mathbf{E}$  obeys Gauss' law,  $\text{div } \mathbf{E} = 4\pi\rho$ , in the static limit  $\mathbf{E} = -\nabla\phi$  and the electrostatic potential is the solution of Poisson's equation  $\nabla^2\phi = -4\pi\rho$ . This elliptic partial differential equation demands a *global* solution in space, which is the origin of the high computational cost of electrostatics.

In full electrodynamics, however, electromagnetic waves propagate with finite speed of light  $c$  as  $\dot{\mathbf{B}} = -c \text{ curl } \mathbf{E}$  and  $\dot{\mathbf{E}} = c \text{ curl } \mathbf{B} - 4\pi\mathbf{J}$ . In contrast to the Poisson equation, solutions to these hyperbolic differential equations require only *local* operations. Since Maxwell's equations conserve Gauss' law if it is obeyed as an initial condition, Coulomb interactions could be generated by coupling charges directly to the time varying electromagnetic fields and solving for their dynamics. This approach could be used in a condensed-matter framework if values other than the speed of light are used for  $c$  in order to propagate the interactions. Maggs and Rossetto (121) showed rigorously that the Coulomb interaction is recovered in the thermodynamic limit *independently* of the value of  $c$ . They studied a constrained partition function of the form

$$\mathcal{Z}(\{\mathbf{r}_i\}) = \int \mathcal{D}\mathbf{E} \prod_{\mathbf{r}} \delta(\text{div } \mathbf{E}(\mathbf{r}) - 4\pi\rho(\{\mathbf{r}_i\})) e^{-\mathcal{U}/k_B T}, \quad (12)$$

where the energy and charge density are given by  $\mathcal{U} = \int \mathbf{E}^2/2 d^3\mathbf{r}$  and  $\rho(\mathbf{r}) = \sum_i e_i \delta(\mathbf{r} - \mathbf{r}_i)$ , and showed that a Gibbs distribution characterized by  $1/r$  interactions is generated when the electric field  $\mathbf{E} = -\nabla\phi + \mathbf{E}_{tr}$  is allowed to fluctuate under the constraint of Gauss' law. In the partition function the Coulombic part completely decouples from the unconstrained or transverse (free) electromagnetic field  $\mathbf{E}_{tr}$  (122), which is not present in conventional electrostatics. This result is nontrivial and shows that the Coulomb interaction is more than the  $c = \infty$  limit of Maxwell's equations; it is fundamentally a consequence of Gauss' law and does not require  $\text{curl } \mathbf{E} = 0$ . Thus, rather than quenching the curl degrees of freedom in the electric field to zero, one can perform an integration over them and obtain the Coulomb interaction even though the electrostatic energy  $U \geq \int (\nabla\phi)^2/2 d^3\mathbf{r}$ . Performing

this integration while maintaining Gauss' law requires only local operations and obviates the need for computing long-ranged interactions at every timestep.

This insight can be applied to molecular dynamics by propagating Maxwell equations and coupling the resulting fields to the motions of the particles (123), with Langevin thermostats on both particle and magnetic field degrees of freedom. The resulting equations have been integrated on a grid (123, 124), which verified the presence of an effective Coulombic interaction. The principal technical challenge lies in the efficient coupling of charges and fields. As in all mesh-based methods, the charges need to be interpolated from the continuum onto the electrostatic grid to generate the current  $\mathbf{J}$ , and the electric field must be extrapolated to compute the electrostatic force  $\mathbf{F}_{\text{el}}$ . Similar to the multigrid method, high accuracy requires smooth interpolation over a large volume, which becomes the dominant computational part of the algorithm. The current implementation (125) uses truncated Gaussians for the interpolation of the point charges to the electric grid with relative force errors greater than  $1 \times 10^{-3}$  which are considered large for atomistic biomolecular simulations. The method has similar advantages to multigrid methods:  $\mathcal{O}(N)$  scaling, ease of parallelization, flexibility for all types of boundary conditions, and the ability to handle inhomogeneous dielectric environments.

## 5. Better Representation of Electrostatics

More accurate representations of electrostatics can be achieved through the inclusion of higher-order multipoles, polarizability, and continuous electrostatics. Of course, the gain in accuracy has to be balanced with the increase in the cost of the simulations, and new algorithms are needed to make feasible the more physical representations of the electronic cloud. Here we review some of the recent developments.

Electrostatic interactions have traditionally been modeled using an atom-centered point-charge (“partial charge”) representation of the molecular charge density. The most popular methods for extracting charges from molecular wavefunctions are based on fitting atomic charges to the molecular electrostatic potential (MEP), computed with ab initio, density functional theory or semiempirical wavefunctions. The charge fitting procedure consists of minimizing the squared deviation between the Coulombic potential produced by the atomic charges and the MEP. These non-bond potentials are then expressed as a sum of spherically isotropic atom-atom potentials. Such representations are believed to be an important source of error in current force fields (35).

The fit to the MEP can be improved either by adding more charge sites (126) or by including higher-order multipoles at the atomic sites or bonds. Even with these improvements the fit to the MEP remains poor in regions near the atomic nuclei, where the charge densities overlap. As a consequence, the electrostatic interaction energy must be corrected for “penetration” effects at close range (127) (usually this error is absorbed into the exchange-repulsion term), and the optimal values of the point multipoles may be poorly determined (128, 129). Nevertheless the use of off-center charges and/or higher-order atomic point multipoles can significantly improve the treatment of electrostatics (35, 130–134), and numerous schemes have been developed to partition the electronic cloud into multipoles (131, 135–137).

In addition to fixed distributed multipoles, a better accuracy requires the inclusion of polarization and, when possible, of the charge density overlap. Polarization is the result of the change in the molecular charge density due to the presence of other molecules or to changes in the geometry of the molecule. These effects are known as inter-and intra-molecular polarization respectively (35). This effect can be introduced by the use of an explicit polarization energy contribution. Several methods have been developed including the Drude oscillator (138, 139), fluctuating charge (140, 141), and induced dipole model (142–144). This has given rise to several polarizable force fields (145–149).

The *distributed* multipole analysis first introduced by Stone assigns distributed multipole moments to several sites in the molecule (i.e., atoms and bond midpoints) and gives a more accurate representation of the electrostatic potential than one-center, molecular multipole expansion. The generalization of Ewald summation to atomic multipoles up to quadrupoles was given by Smith (46). Since then, a few groups (47–49, 143) have extended the Ewald method to take into account multipoles at the atomic and other point sites. However, the multipoles greatly increase the cost of calculations within the Ewald framework. For instance, an electrostatic representation including charges, dipoles, and quadrupoles costs approximately 100 times more than a representation with only charges, using the Ewald formalism, thus rendering multipolar representations in biomolecular simulations prohibitively expensive. In order to surmount this difficulty, PME-based methods have been introduced. A first approach for *large-scale* biomolecular simulations was introduced in 2000 by Toukmaji et al. (47), who developed—in addition to the classical Ewald treatment—a PME-based treatment of fixed and induced point dipoles. Both methods have been implemented into the sander molecular dynamics module of AMBER (versions 6–11), along with several schemes for solving the induced dipoles (iterative, Car-Parrinello). The PME based implementation is quite efficient; for a 1 fs timestep, it is only approximately 1.4 times more expensive

than a calculation including only charges. Since then, a polarizable empirical force field based on the classical Drude oscillator model (150) was implemented in CHARMM in 2005. Even these very simple representations of polarizability make a big difference for various biomolecular systems (151–161).

An efficient implementation of higher-order multipoles in a Cartesian tensor formalism was introduced in 2004 by Sagui et al. (49, 162). The long-range electrostatic interactions are divided in two sums according to the usual Ewald scheme: the *direct* sum, which evaluates the fast-varying, particle-particle interactions, considered up to a given cutoff in real space, and the *reciprocal* sum, which evaluates the smoothly varying, long-range part of the interaction. When implementing multipoles, one has to take care of additional physics that is not present in the usual treatment of charges. First, the higher-order multipoles produce additional contributions to the reciprocal virial, that arise from the dependence of the structure factor on the reciprocal lattice vector. Second, all the multipolar components that appear in the expressions of energy, forces, etc. are given in a global coordinate system. It is necessary therefore to transform the local multipole moments—generally defined in reference to the molecule—to a global framework before any calculation starts. This is achieved by defining “frames” (local orthogonal coordinate systems). Third, to carry out molecular dynamics, the torques produced by every multipole need to be converted into atomic forces.

In order to accelerate the evaluation of the Ewald-like sums, the direct part was implemented using a McMurchie-Davidson formalism (163), while the reciprocal part was implemented in three different ways: using an Ewald scheme, a PME-based approach, and a multigrid-based approach. The standard matrix implementation of multipole interactions up to hexadecapole-hexadecapole costs three orders of magnitude more than charge-charge interactions. Instead, due to the use of the SPME and the factorizability of B-splines, the multipolar SPME is very efficient. For instance, for the same grid density and spline order, calculating reciprocal sum interactions up to hexadecapole-hexadecapole is only twice as expensive as calculating charge-charge interactions. Therefore, by transferring more of the computation of the interactions to reciprocal space and using a small cutoff for the direct sum, it is possible to preserve the accuracy of the calculations at a moderate cost. In fact, a considerably accurate calculation of interactions up to hexadecapole-hexadecapole costs only a factor of 8.5 more (for relative force errors of  $\sim 5 \times 10^{-4}$ ) than a regular AMBER implementation with only charge-charge interactions. Furthermore, a “regular” cutoff of 8 Angstroms for the Coulomb summation (with the acceleration provided by the McMurchie-Davidson scheme) is approximately six times more expensive and has a force error two orders of magnitude larger

than the complete calculation for which most of the interaction is computed in reciprocal space via the PME method. This code has been adapted to achieve a fast implementation of the AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) force field of Ren and Ponder (143, 164–166). The force field includes fixed atomic multipoles up to quadrupole level as well as inducible atomic dipoles using a Thole damping model. A PME-based implementation of the multipolar code for AMOEBA has been released through AMBER 9–11.

Even if polarization is taken into account, the neglect of the charge density overlap when point charges or multipoles are used gives rise to the so-called penetration errors (35, 167, 168) in the Coulomb interactions at close distances. This effect may be compensated to an extent by the use of damping functions (169–172). Another possibility to avoid the penetration errors is by using continuous representations of the molecular charge density. Several methods have been introduced to include better descriptions of the electron density. The Hansen-Coppens formalism describes the static density of a molecular fragment by a superposition of aspherical pseudoatoms composed of nuclei-centered density units (173–175). Gavezzotti introduced a method to adapt electron densities from molecular orbital calculations to calculate intermolecular energies by direct numerical integration (176).

Recently, Cisneros et al. (168, 177, 178) introduced a force field based on density fitting, termed the Gaussian electrostatic model (GEM) (177). In this method, the molecular charge density obtained from quantum mechanical (QM) calculations is fitted to a linear combination of Gaussian functions centered at specific sites. The fitted densities of the fragments are employed to calculate each of the components of the intermolecular interaction separately, thus also including many-body effects such as non-additivity. The philosophy of GEM is to carry out the fitting in order to reproduce only the gas-phase ab initio QM intermolecular interaction results from the constrained space orbital variation (CSOV) energy decomposition scheme (179, 180).

The initial formalism in this approach involved the analytical fitting of QM densities to s-type functions (GEM-0) (177). This enabled the calculation of intermolecular interactions with average errors well below 0.2 kcal/mol for each component of the energy as well as the total intermolecular energy. The formalism was subsequently extended for the Coulomb and exchange-repulsion terms to enable the use of Gaussian fitting functions of arbitrary angular momentum (178). The use of higher angular momentum functions required the improvement of the fitting methods to reduce numerical instabilities. Normalized Hermite Gaussian basis sets were employed for the fit of the molecular densities. The use of Hermite Gaussian functions has the additional advantage of facilitating a point multipole decomposition determination at each expansion

site. Since the basis functions with higher angular momentum have directionality, a reference molecular frame formalism similar to that defined in (49) was employed for the rotation of the fitted expansion coefficients.

The use of molecular densities results in the need to compute a large number of two center integrals for the intermolecular interaction. A significant computational speedup can be achieved by using reciprocal space methods based on Ewald sums. In this way, the integrals are calculated in direct and reciprocal space depending on the exponent of the Gaussian charge densities.

Here we describe how the Ewald formalism can be extended to take into account the Gaussian distribution. As before, let  $U$  denote a unit cell whose edges are given by the vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ , and  $\mathbf{a}_3$ . An idealized infinite crystal can be generated by all periodic translations  $\mathbf{n} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$  for all integer triples  $(n_1, n_2, n_3)$ , with  $n_1, n_2, n_3$  not all zero. Now, consider a collection of  $N$  normalized spherical Gaussian charge densities  $\rho_1 \dots \rho_N$  (e.g., GEM-0) centered at  $\{\mathbf{R}_1 \dots \mathbf{R}_N\} \in U$  with exponents  $\alpha_i$ , i.e.  $\rho_i(\mathbf{r}) = (\alpha_i/\pi)^{3/2} \exp(-\alpha_i(\mathbf{r} - \mathbf{R}_i)^2)$ , and let  $q_1 + \dots + q_N = 0$ . Note that  $N$  need not be limited only to atomic positions, that is, GEM-0 includes sites on the oxygen lone pairs and the bisector line between the two hydrogens (177). The Coulomb energy of the central unit cell within a large spherical crystal, due to the interactions of the Gaussian charge distributions with each other and all periodic images within the crystal can be calculated using Ewald methods.

In particular, to determine the reciprocal part in the Ewald sum, it is necessary to grid the Gaussian densities. However, this can become intractable for Gaussian functions with large exponents (compact Gaussians). In the initial implementation the charge densities were classified into compact or diffuse Hermite Gaussians based on a given Ewald exponent  $\beta$ . Therefore, if the exponent of a given Hermite was above the cutoff, it was considered compact, and diffuse ( $\alpha_i < \beta$ ) otherwise. With this, the contributions involving diffuse Hermites can be calculated in reciprocal space exclusively (178).

This was later improved by the realization that the Ewald exponent,  $\beta$  may be different for each pair  $ij$  (181). Thus,  $\beta$  is chosen to be infinite for  $ij$  pairs where at least one of the Gaussians is diffuse. In this way, all pairs that involve diffuse Hermites are evaluated in reciprocal space. That is, given  $\beta > 0$  a Gaussian distribution  $q_i \rho_i$  is classified as compact ( $i \in c$ ) if  $\alpha_i \geq 2\theta$  and diffuse ( $i \in d$ ) otherwise. Subsequently, for  $i, j \in \{c\}$ , select  $\beta$  so that  $1/\theta = 1/\alpha_i + 1/\alpha_j + 1/\beta$ , otherwise  $\beta = \infty$ . With this, the Ewald expression becomes

$$\begin{aligned}
E = & \frac{1}{2} \sum_{\mathbf{n}}' \sum_{(i,j) \in c \times c} q_i q_j \left\{ \frac{\operatorname{erfc}(\theta^{1/2} |\mathbf{R}_{ij} - \mathbf{n}|) - \operatorname{erfc}(\mu_{ij}^{1/2} |\mathbf{R}_{ij} - \mathbf{n}|)}{|\mathbf{R}_{ij} - \mathbf{n}|} \right\} \\
& + \frac{1}{2\pi V} \sum_{\mathbf{m} \neq 0} \sum_{(i,j) \in c \times c}^N q_i q_j \frac{\exp(-\pi^2 \mathbf{m}^2 / \theta)}{\mathbf{m}^2} \exp(-2\pi i \mathbf{m} \cdot \mathbf{R}_{ij}) \\
& + \frac{1}{2\pi V} \sum_{\mathbf{m} \neq 0} \sum_{(i,j) \notin c \times c}^N q_i q_j \frac{\exp(-\pi^2 \mathbf{m}^2 / \mu_{ij})}{\mathbf{m}^2} \exp(-2\pi i \mathbf{m} \cdot \mathbf{R}_{ij}) \\
& - \frac{\pi}{2V} \sum_{i,j=1}^N q_i q_j \left( \frac{1}{\theta} + \frac{1}{\alpha_i} + \frac{1}{\alpha_j} \right) - \sum_{i=1}^N q_i^2 \left( \frac{\theta}{\pi} \right)^{1/2} \\
& - \sum_{i=1}^N q_i^2 \left( \frac{\alpha_i}{\pi} \right)^{1/2} + \frac{2\pi \mathbf{D}^2}{3V} + \epsilon(K),
\end{aligned} \tag{13}$$

where the first term corresponds to the direct part of the Ewald sum, the second and third terms to the reciprocal part,  $\mathbf{R}_{ij} = \mathbf{R}_i - \mathbf{R}_j$ , the term involving the unit cell dipole  $\mathbf{D} = q_1 \mathbf{R}_1 + \dots + q_N \mathbf{R}_N$  is the surface term,  $\epsilon(K)$  denotes a quantity that converges to 0 as  $K \rightarrow \infty$ ,  $\mathbf{m}$  denotes the reciprocal lattice vectors, and  $1/\mu_{ij} = 1/\alpha_i + 1/\alpha_j$ .

The above equation applies only to spherical ( $l = 0$ ) Gaussians. This expression can be further generalized for the case when the auxiliary and optimized bases used for the fitting include Gaussians with  $l > 0$  (178). In this case, the direct space contributions can be efficiently evaluated by using the McMurchie-Davidson (MD) recursion (163). This recursion was used to calculate the required erfc and higher derivatives for the multipoles (49). This approach was also employed for the Hermite Gaussians (178), where it was shown that the MD recursion is applicable to other types of operators besides  $1/r$ . For the reciprocal sums, two methods were implemented: SPME (42) and FFP (43).

These methods were applied to simulate a series of water boxes of 64–1,024 molecules under PBC (178, 181). These calculations showed a significant computational speed improvement when FFP and PME were used compared to full Ewald sums. Alternative methods for fitting have been also explored. Numerical fitting of molecular properties, in particular electrostatic potential, have given good results. The numerical fitting has resulted in two improvements. The first is the reduction of numerical instabilities in the calculation of the coefficients. The second is a reduction in the number of fitting sites needed for accurate reproduction of the MEP (182). Overall, the use of numerical fits produces results similar to the analytical fits with errors around 0.2 kcal/mol. More recently the numerical fitting has been improved by introducing

spherical grids coupled to a Gaussian type weighting function for discarding points near the core (183). The reduction in the number of fitting sites results in significant computational speedups. For example, an extremely accurate calculation with GEM using a three site fitting model for water with interactions up to d-type Hermite Gaussians was carried out in a box of 4,096 waters in PBC using the A1 auxiliary basis set (184) with coefficients fitted to the electrostatic potential. The resulting energy/force calculation costs (in one processor) only one order of magnitude more than the calculation using a TIP3P water model (182).

---

## 6. Perspective

At present, the performance of traditional classical codes, where electrostatics is represented by point charges, is driven by the successive generations of computer architectures with parallelization issues dominating the game. At the frontier, very large systems and milli-second time scales are now feasible (82, 185). It is to be expected that simulations on such scales will become more commonplace with improvements in hardware prompting questions regarding the validity of the force fields on these scales. Interesting work remains to be done to characterize the effect of point-charge electrostatics in extremely long molecular dynamics simulations. There is a good chance that undesired artifacts will be discovered as more and more systems are studied. Future studies will likely need more accurate representations of non-bonded interactions. This, in turn, will require the development of new parallel implementations to allow for the efficient evaluation of long-range electrostatic interactions with more accurate representations of the molecular charge density.

---

## Acknowledgements

This research was supported by NSF under grant FRG-0804549, and by WayneState University.

## References

1. Alper HE, Levy RM (1989) Computer-simulations of the dielectric-properties of water—studies of the simple point-charge and transferable intermolecular potential models. *J Chem Phys* 91:1242–1251
2. Alper HE, Bassolino D, Stouch TR (1993) Computer-simulation of a phospholipid monolayer-water system—the influence of long-range forces on water-structure and dynamics. *J Chem Phys* 98:9798–9807

3. Alper HE, Bassolino-Klimas D, Stouch TR (1993) The limiting behavior of water hydrating a phospholipid monolayer—a computer simulation study. *J Chem Phys* 99:5547–5559
4. Feller SE, Pastor RW, Rojnuckarin A, Bogusz A, Brooks BR (1996) Effect of electrostatic force truncation on interfacial and transport properties of water. *J Phys Chem* 100:17011–17020
5. Mark P, Nilsson L (2002) Structure and dynamics of liquid water with different long-range interaction truncation and temperature control methods in molecular dynamics simulations. *J Comp Chem* 23:1211–1219
6. Yonetani Y (2006) Liquid water simulation: a critical examination of cutoff length. *J Chem Phys* 124:204501
7. Smith PE, Pettitt BM (1991) Peptides in ionic solutions: a comparison of the Ewald and switching function techniques. *J Chem Phys* 95:8430–8441
8. Schreiber H, Steinhauser O (1992) Cutoff size does strongly influence molecular-dynamics results on solvated polypeptides. *Biochemistry* 31:5856–5860
9. York DM, Darden TA, Pedersen LG (1993) The effect of long-range electrostatic interactions in simulations of macromolecular crystals—a comparison of the Ewald and truncated list methods. *J Chem Phys* 99:8345–8348
10. York DM, Yang WT, Lee H, Darden T, Pedersen LG (1995) Toward the accurate modeling of DNA—the importance of long-range electrostatics. *J Am Chem Soc* 117:5001–5002
11. Cheatham III TE, Kollman PA (1997) Molecular dynamics simulations highlight structural differences among DNA:DNA, RNA:RNA, and DNA:RNA hybrid duplexes. *J Am Chem Soc* 119:4805–4825
12. Auffinger P, Westhof E (1998) Molecular dynamics simulations of nucleic acids. In: Encyclopedia of computational chemistry. Wiley, New York
13. Norberg J, Nilsson L (2000) On the truncation of long-range electrostatic interactions in DNA. *Biophys J* 79:1537–1553
14. Feller SE, Pastor RW, Rojnuckarin A, Bogusz S, Brooks BR (1996) Effect of electrostatic force truncation on interfacial and transport properties of water. *J Phys Chem* 100:17011–17020
15. Patra M, Karttunen M, Hyvonen MT, Falck E, Vattulainen I (2004) Lipid bilayers driven to a wrong lane in molecular dynamics simulations by subtle changes in long-range electrostatic interactions. *J Phys Chem B* 108:4485–4494
16. Anézo C, de Vries AH, Höltje HD, Tieleman DP (2003) Methodological issues in lipid bilayer simulations. *J Phys Chem B* 107:9424–9433
17. Khelashvili GA, Scott HL (2004) Combined Monte Carlo and molecular dynamics simulation of hydrated 18:0 sphingomyelin-cholesterol lipid bilayers. *J Chem Phys* 120:9841–9847
18. Patra M, Hyvonen MT, Falck E, Sabouri-Ghomi M, Vattulainen I, Karttunen M (2007) Long-range interactions and parallel scalability in molecular simulations. *Comp Phys Comm* 176:14–22
19. Karttunen M, Rottler J, Vattulainen I, Sagui C (2008) Computational modeling of membrane bilayers. Elsevier, New York
20. Heyes DM, Barber M, Clarke JHR (1977) Molecular dynamics computer simulation of surface properties of crystalline potassium chloride. *J Chem Soc Faraday Trans 2* 72:1485–1496
21. Hautman J, Klein ML (1992) An Ewald summation method for planar surfaces and interfaces. *Mol Phys* 75:379–395
22. de Leeuw SW, Perram JW (1979) Electrostatic lattice sums for semi-infinite lattices. *Mol Sim* 37:1313–1322
23. Nijboer BRA, de Wette FW (1957) On the calculation of lattice sums. *Physica* 23:309–321
24. Yeh IC, Berkowitz ML (1999) Ewald summation for systems with slab geometry. *J Chem Phys* 111:3155–3162
25. Lekner J (1989) Summation of dipolar fields in simulated liquid-vapor interfaces. *Physica A* 157:826–838
26. Lekner J (1991) Summation of Coulomb fields in computer-simulated disordered systems. *Physica A* 176:485–498
27. Lekner J (1998) Coulomb forces and potentials in systems with an orthorhombic unit cell. *Mol Sim* 20:357
28. Sperb R (1994) Extension and simple proof of Lekner's summation formula for coulomb forces. *Mol Sim* 13:189–193
29. Sperb R (1998) An alternative to Ewald sums part I: identities for sums. *Mol Sim* 20:179–200
30. Sperb R (1999) An alternative to Ewald sums, part 2: the Coulomb potential in a periodic system. *Mol Sim* 22:199–212
31. Strebel R, Sperb R (2001) An alternative to Ewald sums. part 3: implementation and results. *Mol Sim* 27:61–74
32. Arnold A, Holm C (2002) A novel method for calculating electrostatic interactions in 2D

- periodic slab geometries. *Chem Phys Lett* 354:324–330
33. Arnold A, Holm C (2002) MMM2D: a fast and accurate summation method for electrostatic interactions in 2D slab geometries. *Comput Phys Comm* 148:327–348
34. Baker NA (2005) Improving implicit solvent simulations: a Poisson-centric view. *Curr Opin Struct Bio* 15:137–143
35. Stone AJ (1996) *The theory of intermolecular forces*. Clarendon Press, Oxford
36. Ewald P (1921) Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann Phys* 64:253–287
37. DeLeeuw SW, Perram JW, Smith ER (1980) Simulation of electrostatic systems in periodic boundary conditions I: lattice sums and dielectric constants. *Proc R Soc Lond A* 373:27–56
38. Boresch S, Steinhauser O (1997) Presumed versus real artifacts of the Ewald summation technique: the importance of dielectric boundary conditions. *Ber Bunsenges Phys Chem* 101:1019–1029
39. Hockney RW, Eastwood JW (1981) *Computer simulation using particles*. McGraw-Hill, New York
40. Pollock E, Glosli J (1996) Comments on PPPM, FMM, and the Ewald method for large periodic Coulombic systems. *Comp Phys Comm* 95:93–110
41. Darden TA, York DM, Pedersen LG (1993) Particle mesh Ewald: an  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
42. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593
43. York D, Yang W (1994) The fast Fourier Poisson (FFP) method for calculationg Ewald sums. *J Chem Phys* 101:3298–3300
44. Smith ER (1994) Calculating the pressure in simulations using periodic boundary conditions. *J Stat Phys* 77:449–472
45. Caillol JP (1994) Comments on the numerical simulations of electrolytes in periodic boundary conditions. *J Chem Phys* 101:6080–6090
46. Smith W (1982) Point multipoles in the Ewald summation. *CCP5 Inform Q* 4:13–25
47. Toukmaji A, Sagui C, Board JA, Darden T (2000) Efficient PME-based approach to fixed and induced dipolar interactions. *J Chem Phys* 113:10913–10927
48. Aguado A, Madden P (2003) Ewald summation of electrostatic multipole interactions up to the quadrupolar level. *J Chem Phys* 119:7471–7483
49. Sagui C, Pedersen LG, Darden TA (2004) Towards an accurate representation of electrostatics in classical force fields: efficient implementation of multipolar interactions in biomolecular simulations. *J Chem Phys* 120:73–87
50. Smith W (1987) Coping with the pressure: how to calculate the virial. *CCP5 Information Quarterly* 26:43–50
51. Alejandre J, Tildesley DJ, Chapela GA (1995) Molecular dynamics simulation of the orthobaric densities and surface tension of water. *J Chem Phys* 102:4574–4583
52. Nose S, Klein M (1983) Constant pressure molecular dynamics for molecular systems. *Mol Phys* 50:1055–76
53. Brown D, Clarke JHR (1991) A loose coupling, constant-pressure, molecular dynamics algorithm for use in the modelling of polymer materials. *Comp Phys Comm* 62:360–369
54. Reif MM, Kraeutler V, Kastenholz MA, Daura X, Huenenberger PH (2009) Molecular dynamics simulations of a reversibly folding beta-Heptapeptide in methanol: influence of the treatment of long-range electrostatic interactions. *J Phys Chem B* 113:3112–3128
55. Villarreal MA, Montich GG (2005) On the Ewald artifacts in computer simulations. The test-case of the octalanine peptide with charged termini. *J Biomol Struct Dyn* 23:135–142
56. Monticelli L, Colombo G (2004) The influence of simulation conditions in molecular dynamics investigations of model beta-sheet peptides. *Theo Chem Acc* 112:145–157
57. Monticelli L, Simões C, Belvisi L, Colombo G (2006) Assessing the influence of electrostatic schemes on molecular dynamics simulations of secondary structure forming peptides. *Journal of Physics: condensed Matter* 18: S329–S345
58. Babin V, Roland C, Darden TA, Sagui C (2006) The free energy landscape of small peptides as obtained from metadynamics with umbrella sampling corrections. *J Chem Phys* 125:204909
59. Levy RM, Gallicchio E (1998) Computer simulations with explicit solvent: recent progress in the thermodynamic decomposition of free energies, and in modeling electrostatic effects. *Ann Rev Phys Chem* 49:531–567

60. Hummer G, Pratt LR, Garcia AE (1996) On the free energy of ionic hydration. *J Phys Chem* 100:1206–1215
61. Figueiredo F, Buono GSD, Levy RM (1997) On finite-size corrections to the free energy of ionic hydration. *J Phys Chem B* 101:5622–5623
62. Hummer G, Pratt LR, Garcia AE (1997) Ion sizes and finite-size corrections for ionic-solvation free energies. *J Chem Phys* 107:9275–9277
63. Sakane S, Ashbaugh HS, Wood RH (1998) Continuum corrections to the polarization and thermodynamic properties of Ewald sum simulations for ions and ion pairs at infinite dilution. *J Phys Chem B* 102:5673–5682
64. Darden T, Pearlman D, Pedersen L (1998) Ionic charging free energies: spherical versus periodic boundary conditions. *J Chem Phys* 109:10921–10935
65. Herce D, Darden T, Sagui C (2003) Calculation of ionic charging free energies in simulation systems with atomic charges, dipoles and quadrupoles. *J Chem Phys* 119:7621–7632
66. Ferrell R, Bertschinger E (1994) Particle-mesh methods on the connection machine. *J Mod Phys* 5:933–956
67. Toukmaji A, Board JA (1996) Ewald sum techniques in perspective: a survey. *Comput Phys Comm* 95:78–92
68. Darden T, Toukmaji A, Pedersen L (1997) Long-range electrostatic effects in biomolecular simulations. *J Chim Phys* 94:1346–1364
69. Deserno M, Holm C (1998) How to mesh up Ewald sums i: a theoretical and numerical comparison of various particle mesh routines. *J Chem Phys* 109:7678–7693
70. Sagui C, Darden TA (1999) Molecular dynamics simulations of biomolecules: long-range electrostatic Effects. *Annu Rev Biophys Biomol Struct* 28:155–179
71. Sagui C, Darden TA (1999) P3M and PME: a comparison of the two methods. In: Pratt LR, Hummer G (eds) Simulation and theory of electrostatic interactions in solution. AIP, Melville, NY
72. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods R (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
73. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J Comp Chem* 4:187–217
74. Phillips J, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comp Chem* 26:1781–1802
75. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theo Comp* 4:435–447
76. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kollossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: SC '06: proceedings of the 2006 ACM/IEEE conference on supercomputing. ACM Press
77. Cerutti DS, Duke RE, Darden TA, Lybrand TP (2009) Staggered mesh Ewald: an extension of the smooth particle-mesh Ewald method adding great versatility. *J Chem Theory Comput* 5:2322–2338
78. Chen L, Langdon AB, Birdsall CK (1974) Reduction of grid effects in simulation plasmas. *J Comp Phys* 14:200–222
79. Hockney RW, Eastwood JW (1988) *Computer simulation using particles*. Adam Hilger, Bristol
80. Cerutti DS, Case DA (2010) Multi-level ewald: a hybrid multigrid/fast Fourier transform approach to the electrostatic particle-mesh problem. *J Chem Theory Comput* 6:443–458
81. Duke RE (2006) Amber 9 manual: amber. scripps.edu/doc9/amber9.pdf. AMBER 9, D. A. Case et al., University of California, San Francisco
82. Schulten K, Phillips JC, Kale LV, Bhatele A (2008) In: Bader D (ed) Petascale computing: algorithms and applications. Chapman & Hall/CRC Press, New York
83. Brandt A (1977) Multi-level adaptive solutions to boundary value problems. *Math Comput* 31:333–390
84. Brandt A (1994) SIAM J Num Anal 31:1695
85. Brandt A (2001) Multiscale scientific computation: review 2001. In: Barth TJ, Chan TF, Haimes R (eds) Multiscale and multiresolution methods: theory and applications. Springer Verlag, Heidelberg
86. Beck TL (2000) Real-space mesh techniques in density-functional theory. *Rev Mod Phys* 72:1041–1080
87. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. *Science* 268:1144–1149

88. Holst M, Saied F (1993) Multigrid solution of the Poisson-Boltzmann equation. *J Comp Chem* 14:105–113
89. Holst M, Kozack RE, Saied F, Subramaniam S (1994) Treatment of electrostatic effects in proteins: multigrid-based Newton iterative method for solution of the full Poisson-Boltzmann equation. *Proteins Struct Fun Gen* 18:231–241
90. Holst M, Saied F (1995) Numerical solution of the nonlinear Poisson-Boltzmann equation: developing more robust and efficient methods. *J Comp Chem* 16:337–364
91. Ripoll DR, Vorobjev YN, Liwo A, Vila JA, Scheraga HA (1996) Coupling between folding and ionization equilibria: effects of pH on the conformational preferences of polypeptides. *J Mol Biol* 264:770–783
92. Vorobjev YN, Scheraga HA (1997) A fast adaptive multigrid boundary element method for macromolecular electrostatic computations in a solvent. *J Comp Chem* 18:569–583
93. Qian X, Schlick T (2002) Efficient multiple-time-step integrators with distance-based force splitting for particle-mesh-Ewald molecular dynamics simulations. *J Chem Phys* 116:5971–5983
94. Press W, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in FORTRAN: the art of scientific computing. Cambridge University Press, Cambridge
95. Zaslavsky L, Schlick T (1998) An adaptive multigrid technique for evaluating long-range forces in biomolecular simulations. *Appl Math Comput* 97:237–250
96. Sagui C, Darden TA (2001) Multigrid methods for classical molecular dynamics simulations of biomolecules. *J Chem Phys* 114:6578–6591
97. Hackbusch W (1985) Multigrid methods and applications. Springer, Berlin
98. Briggs EL, Sullivan DJ, Bernholc J (1996) A real-space multigrid-based approach to large-scale electronic structure calculations. *Phys Rev B* 54:14362–14375
99. Skeel RD, Tezcan I, Hardy DJ (2002) Multiple grid methods for classical molecular dynamics. *J Comp Chem* 23:673–684
100. Goedecker S, Chauvin C (2003) Combining multigrid and wavelet ideas to construct more efficient multiscale algorithms. *J Theo Comp Chem* 2:483–495
101. Izaguirre JA, Hampton SS, Matthey T (2005) Parallel multigrid summation for the N-body problem. *J Parall Distrib Comp* 65:949–962
102. Banerjee S, J A Board J (2005) Efficient charge assignment and back interpolation in multigrid methods for molecular dynamics. *J Comp Chem* 26:957–967
103. Shan Y, Klepeis JL, Eastwood MP, Dror RO, Shaw DE (2005) Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J Chem Phys* 122:054101
104. Groot RD (2003) Electrostatic interactions in dissipative particle dynamics—simulation of polyelectrolytes and anionic surfactants. *J Chem Phys* 118:11265–11277
105. Bishop T, Skeel R, Schulten K (1997) Difficulties with multiple stepping and fast multipole algorithm in molecular dynamics. *J Comp Chem* 18:1785–91
106. Greengard L, Rokhlin V (1987) A fast algorithm for particle simulations. *J Comp Phys* 73:325–348
107. Board JA, Causey JW, Leathrum JF, Windemuth A, Schulten K (1992) Accelerated molecular dynamics simulation with the parallel fast multipole algorithm. *Chem Phys Lett* 198:89–94
108. Schmidt KE, Lee MA (1991) Implementing the fast multipole method in three dimensions. *J Stat Phys* 63:1223–1235
109. Lambert CG, Darden TA, Board JA (1996) A multipole-based method for efficient calculation of forces and potentials in macroscopic periodic assemblies of particles. *J Comp Phys* 126:274–285
110. Figueirido F, Levy R, Zhou R, Berne B (1997) Large scale simulation of macromolecules in solution: combining the periodic fast multipole method with multiple time step integrators. *J Chem Phys* 106:9835–9849
111. Greengard LF (1988) The rapid evaluation of potential fields in particle systems. The MIT Press, Cambridge
112. Greengard LF (1994) Fast algorithms for classical physics. *Science* 265:909–914
113. Greengard L, Rokhlin V (1997) A new version of the fast multipole method for the Laplace equation in three dimensions. *Acta Numerica* 6:229–270
114. Greengard L, Rokhlin V (1997) A fast algorithm for particle simulation. *J Comput Phys* 135:280–292
115. Cheng H, Greengard L, Rokhlin V (1999) A fast adaptive multipole algorithm in three dimensions. *J Chem Phys* 155:468–498
116. Fenley M, Olson W, Chua K, Boschitsch A (1994) Fast adaptive multipole method for computation of electrostatic energy in simulations of polyelectrolyte DNA. *J Comp Chem* 17:976

117. Zhou R, Berne B (1995) A new molecular dynamics method combining the reference system propagator algorithm with a fast multipole method for simulating proteins and other complex systems. *J Chem Phys* 103:9444–9459
118. Ogata S, Campbell T, Kalia R, Nakano A, Vashishta P, Vemparala S (2003) Scalable and portable implementation of the fast multipole method on parallel computers. *Comp Phys Comm* 153:445–461
119. Kurzak J, Pettitt BM (2005) Massively parallel implementation of a fast multipole method for distributed memory machines. *J Parall Distrib Comp* 65:870–881
120. Kurzak J, Pettitt BM (2006) Fast multipole methods for particle dynamics. *Molecular Simulation* 32:775–790
121. Maggs AC, Rossetto V (2002) Paper1. *Phys Rev Lett* 88:196402
122. Alastuey A, Appel W (2000) *Physica A* 276:508
123. Rottler J, Maggs AC (2004) Local molecular dynamics with coulombic interactions. *Phys Rev Lett* 93:170201
124. Pasichnyk I, Duenweg B (2004) bla. *J Phys Cond Mat* 16:S3999
125. Rottler J (2007) Local electrostatics algorithm for classical molecular dynamics simulations. *J Chem Phys* 127
126. Dixon R, Kollman P (1997) Advancing beyond the atom-centered model in additive and non-additive molecular mechanics. *J Comp Chem* 18:1632–1646
127. Wheatley R, Mitchell J (1994) Gaussian multipoles in practice: electrostatic energies for Intermolecular potentials. *J Comp Chem* 15:1187–1198
128. Bayly C, Cieplak P, Cornell W, Kollman P (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - the RESP model. *J Phys Chem* 97:10269–10280
129. Franci MM, Chirlian LA (1999) The pluses and minuses of mapping atomic charges to Electrostatic potentials. In: Lipkowitz K, Boyd DB (eds) *Reviews in computational chemistry*. VCH Publishers, New York
130. Price S (1999) In: Lipkowitz K, Boyd DB (eds) *Reviews in computational chemistry*, VCH Publishers, New York
131. Popelier P (2000) Atoms in molecules: an introduction. Prentice Hall, Harlow
132. Kosov DS, Popelier PLA (2000) Atomic partitioning of molecular electrostatic potentials. *J Phys Chem A* 104:7339–7345
133. Popelier PLA, Joubert L, Kosov DS (2001) Convergence of the electrostatic interaction based on topological atoms. *J Phys Chem A* 105:8254–8261
134. Popelier PLA, Kosov DS (2001) *J Chem Phys* 114:6539–6547
135. Proft FD, Alsenov CV, Peeters A, Langenaecker W, Geerlings P (2002) Atomic charges, dipole moments and Fukai functions using the Hirshfeld partitioning of the electron density. *J Comp Chem* 23:1198–1209
136. Bader R (1990) Atoms in molecules: a quantum theory. Clarendon Press, Oxford
137. Sagui C, Pomorski P, Darden TA, Roland C (2004) Ab initio calculation of electrostatic multipoles with Wannier functions for large-scale biomolecular simulations. *J Chem Phys* 120:4530–4544
138. Yu H, Hansson T, van Gunsteren WF (2003) Development of a simple, self-consistent polarizable model for liquid water. *J Chem Phys* 118:221–234
139. Lamoureux G, Alexander D MacKerell J, Roux B (2003) A simple polarizable model of water based on classical Drude oscillators. *J Chem Phys* 119:5185–5197
140. Rick SW, Stuart SJ, Berne BJ (1994) Dynamical fluctuating charge force fields: Application to liquid water. *J Chem Phys* 101:6141–6156
141. Ribeiro MCC, Almeida LCJ (1999) Fluctuating charge model for polyatomic ionic systems: a test case with diatomic anions. *J Chem Phys* 110:11445–11448
142. Caldwell J, Dang LX, Kollman PA (1990) Implementation of nonadditive intermolecular potentials by use of molecular dynamics: development of a water-water potential and water-ion cluster interactions. *J Am Chem Soc* 112:9144–9147
143. Ren P, Ponder JW (2003) Polarizable atomic multipole water model for molecular mechanics simulation. *J Phys Chem B* 107:5933–5947
144. Eling D, Darden TA, Woods RJ (2007) Gaussian induced dipole polarization model. *J Comp Chem* 28:1261–1274
145. Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, Robert A DiStasio Jr, Head-Gordon M, Clark GNI, Johnson ME, Head-Gordon T (2010) Current status of the AMOEBA polarizable force field. *J Phys Chem B* 114:2549–2564
146. Case DA, Cheatham III TE, Darden TA, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comp Chem* 26:1668–1688

147. Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, Krauss M, Garmer D, Basch H, Cohen D (1996) An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J Chem Phys* 105:1968–1986
148. Gagliardi L, Lindh R, Karlström G (2004) Local properties of quantum chemical systems: the LoProp approach. *J Chem Phys* 121:4494–4500
149. Gresh N, Claverie P, Pullman A (1982) Computation of intermolecular interactions: expansion of a charge energy contribution in the framework of an additive procedure - application to hydrogen bonded systems. *Int J Quantum Chem* 22:199
150. Anisimov VM, Lamoureux G, Vorobyov IV, Huang N, Roux B, MacKerell AD (2005) Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J Chem Theory Comput* 1:153–168
151. Jungwirth P, Tobias D (2002) Ions at the air/water interface. *J Phys Chem B* 106:6361–6373
152. Baucom J, Transue T, Fuentes-Cabrera M, Krahn J, Darden T, Sagui C (2004) Molecular dynamics simulations of the d (CCAACGTTGG)<sub>2</sub> decamer in crystal environment: comparison of atomic point-charge, extra-point and polarizable force fields. *J Chem Phys* 121:6998–7008
153. Vrbka L, Mucha M, Minofar B, Jungwirth P, Brown E, Tobias D (2004) Propensity of soft ions for the air/water interface. *Curr Opin Coll Interface Sci* 9:67–73
154. Herce D, Perera L, Darden T, Sagui C (2005) Surface solvation for an ion in a water cluster. *J Chem Phys* 122:024513
155. Babin V, Baucom J, Darden TA, Sagui C (2006) Molecular dynamics simulations of DNA with polarizable force fields: convergence of an ideal B-DNA structure to the crystallographic structure. *J Phys Chem B* 110:11571–11581
156. Harder E, Anisimov VM, Whitfield TW, MacKerell AD, Roux B (2008) Understanding the dielectric properties of liquid amides from a polarizable force field. *J Phys Chem B* 112:3509–3521
157. Vladimirov E, Ivanova A, Roesch N (2008) Effect of solvent polarization on the reorganization energy of electron transfer from molecular dynamics simulations. *J Chem Phys* 129
158. Harder E, MacKerell AD, Roux B (2009) Many-body polarization effects and the membrane dipole potential. *J Am Chem Soc* 131:2760+
159. Vladimirov E, Ivanova A, Roesch N (2009) Solvent reorganization energies in A-DNA, B-DNA, and rhodamine 6G-DNA complexes from molecular dynamics simulations with a polarizable force field. *J Phys Chem B* 113:4425–4434
160. Cieplak P, Dupradeau FY, Duan Y, Wang J (2009) Polarization effects in molecular mechanical force fields. *J Phys Condens Matter* 21: 333102
161. Yan T, Wang Y, Knox C (2010) On the structure of ionic liquids: comparisons between electronically polarizable and nonpolarizable models I. *J Phys Chem B* 114:6905–6921
162. Sagui C, Roland C, Pedersen LG, Darden TA (2005) New distributed multipole methods for accurate electrostatics in large-scale biomolecular simulations. In: Leimkuhler B, Chipot C, Elber R, Laaksonen A, Mark A, Schlick T, Schuette C, Skeel R (eds) *New algorithms for macromolecular simulations*. Springer, Berlin
163. McMurchie L, Davidson E (1978) One- and two-electron integrals over cartesian Gaussian functions. *J Comput Phys* 26:218–231
164. Ren P, Ponder JW (2002) A consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J Comput Chem* 23:1497–1506
165. Ponder JW, Case DA (2003) Force fields for protein simulation. *Adv Protein Chem* 66:27
166. Ren P, Ponder JW (2004) Temperature and pressure dependence of the AMOEBA water model. *J Phys Chem B* 108:13427–13437
167. Qian W, Krimm S (2005) Limitations of the molecular multipole expansion treatment of electrostatic interactions for C-H...O and O-H...O hydrogen bonds and application of a general charge density approach. *J Phys Chem A* 109:5608–5618
168. Cisneros GA, Piquemal JP, Darden TA (2005) Intermolecular electrostatic energies using density fitting. *J Chem Phys* 123:044109
169. Piquemal JP, Gresh N, Giessner-Prettre C (2003) Improved formulas for the calculation of the electrostatic contribution to the intermolecular interaction energy from multipolar expansion of the electronic distribution. *J Phys Chem A* 107:10353–10359
170. Cisneros GA, Tholander SNI, Parisel O, Darden TA, Elking D, Perera L, Piquemal JP (2008) Simple formulas for improved point-charge electrostatics in classical force fields and hybrid quantum mechanical/

- molecular mechanical embedding. *Int J Quantum Chem* 108:1905–1912
171. Freitag MA, Gordon MS, Jensen JH, Stevens WJ (2000) Evaluation of charge penetration between distributed multipolar expansions. *J Chem Phys* 112:7300–7306
  172. Wang B, Truhlar DG (2010) Including charge penetration effects in molecular modeling. *J Chem Theo Comp* 6:3330
  173. Volkov A, Koritsanszky T, Coppens P (2004) Combination of the exact potential and multipole methods (EP/MM) for evaluation of intermolecular electrostatic interaction energies with pseudoatom representation of molecular electron densities. *Chem Phys Lett* 391:170–175
  174. Volkov A, Li X, Koritsanszky T, Coppens P (2004) *Ab initio* quality electrostatic atomic and molecular properties including intermolecular energies from a transferable theoretical pseudoatom databank. *J Phys Chem A* 108:4283–4300
  175. Volkov A, Coppens P (2004) Calculation of electrostatic interaction energies in molecular dimers from atomic multipole moments obtained by different methods of electron density partitioning. *J Comp Chem* 25:921–934
  176. Gavezzotti A (2002) Calculation of intermolecular interaction energies by direct numerical integration over electron densities I. Electrostatic and polarization energies in molecular crystals. *J Phys Chem B* 106:4145–4154
  177. Piquemal JP, Cisneros GA, Reinhardt P, Gresh N, Darden TA (2006) Towards a force field based on density fitting. *J Chem Phys* 124:104101
  178. Cisneros GA, Piquemal JP, Darden TA (2006) Generalization of the Gaussian electrostatic model: extension to arbitrary angular momentum, distributed multipoles, and speedup with reciprocal space methods. *J Chem Phys* 125:184101
  179. Bagus PS, Hermann K, Bauschlicher CW Jr (1984) A new analysis of charge transfer and polarization for ligand–metal bonding: model studies for Al<sub>4</sub>CO and Al<sub>4</sub>NH<sub>3</sub>. *J Chem Phys* 80:4378–4386
  180. Piquemal JP, Marquez A, Parisel O, Giessner-Prettre C (2005) A CSOV study of the difference between HF and DFT intermolecular interaction energy values: the importance of the charge transfer contribution. *J Comp Chem* 26:1052–1062
  181. Darden TA (2007) In: Shmueli U (ed) Dual bases in crystallographic computing in international tables of chrystallography. Kluwer Academic Publishers, Dordrecht
  182. Cisneros GA, Elking DM, Piquemal JP, Darden TA (2007) Numerical fitting of molecular properties to Hermite Gaussians. *J Phys Chem A* 111:12049–12056
  183. Elking DM, Cisneros GA, Piquemal JP, Darden TA, Pedersen LG (2010) Gaussian multipole model (GMM). *J Chem Theo Comp* 6:190–202
  184. Godbout N, Andzelm J (1998) DGauss Version 2.0, 2.1, 2.3, 4.0: the file that contains the A1 and P1 auxiliary basis sets can be obtained from the CCL WWW site at <http://ccl.osc.edu/cca/data/basis-sets/DGauss/basis.v3.html>. Computational Chemistry List, Ltd., Ohio
  185. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346

# Chapter 11

## An Introduction to Best Practices in Free Energy Calculations

Michael R. Shirts and David L. Mobley

### Abstract

Free energy calculations are extremely useful for investigating small-molecule biophysical properties such as protein-ligand binding affinities and partition coefficients. However, these calculations are also notoriously difficult to implement correctly. In this chapter, we review standard methods for computing free energy via simulation, discussing current best practices and examining potential pitfalls for computational researchers performing them for the first time. We include a variety of examples and tips for how to set up and conduct these calculations, including applications to relative binding affinities and small-molecule solvation free energies.

**Key words:** Free energy calculation, Alchemical methods, Thermodynamic integration, Bennett acceptance ratio, WHAM, MBAR, Solvation free energy, Binding free energy

---

### 1. Introduction

The phrase *free energy calculation* describes a large family of simulation procedures to calculate free energy differences between physical systems. The free energy is a fundamental quantity describing the stability of a system because the free energy of a system is minimized if and only if the system is at equilibrium with its environment. A system held at constant temperature and volume minimizes its Helmholtz free energy; a system held at constant temperature and pressure minimizes its Gibbs free energy.

Calculating the free energy difference between two different states is extremely useful in simulations of biological systems. If we can calculate the free energy difference between two arbitrary molecular systems we can determine small-molecule transfer free energies and partition coefficients and thus predict the concentration of the molecule in each phase. If we can calculate the free

energy of binding of a small molecule to a receptor, we know the dissociation constant of the receptor.

There has been a significant increase in enthusiasm about free energy calculations over the last decade. This new interest is primarily due to three major advancements. First, methodological innovations make the calculations easier and more robust. Second, the methods have been implemented in most molecular simulation packages, making them available to many users. Finally, steady increases in computer power make it possible to apply free energy methods to a broader range of systems. But free energy calculations are still difficult to carry out and there is significant confusion about appropriate techniques. This methodological review is designed to help researchers who have some familiarity with molecular simulations and knowledge of statistical mechanics, but are looking for more guidance on performing free energy calculations.

There are several reasons free energy calculations are among the most difficult types of biomolecular simulations. Most simulation packages require extensive manual adjustments to input files to carry out free energy calculations. Such calculations also can be extremely sensitive to certain choices of parameters that are unimportant for simulations of a single thermodynamic state. Additionally, the vast number of methodologies available can lead to a bewildering number of choices that a researcher must make to perform free energy calculations.

Standard computational methods for calculating free energies use molecular simulations to generate independent samples from the equilibrium distribution of the molecular system. Then, the information from these samples is analyzed using statistical methods to obtain an estimate of the free energy difference. Because of the statistical nature of the simulations and the analysis, free energy calculations give *estimated* free energies, and repeating the calculation from different starting configurations or different random seeds will give different free energy estimates. To emphasize: free energy results from simulations are *not* exact results; they are statistical estimates obtained from sampling thermodynamic probability distributions. Consequently, error analysis must always be performed to identify the statistical noise in the calculation, and no free energy calculation should ever be published without a statistically robust estimate of the uncertainty in the calculation.

Free energy methods provide an estimate of the correct free energy of some change *given a particular set of parameters and physical assumptions*. This does necessarily correspond to the value that would be observed for that change experimentally. The goal of good free energy methods is to converge to the unique free energy for that model. This is the “correct” free energy for a given model of the system, for example, a given choice of force field. Only after this free energy is determined accurately can parameters of the model be improved, though substantial care must be taken to avoid overfitting. This chapter will not address finding or developing the best

molecular model for a particular problem but instead will focus on the problem of computing free energies of a given model.

We begin with a brief theoretical overview of the methods that are useful for calculating free energies in biophysical applications. We then outline the steps that must be performed for typical free energy calculations, including setup, running the simulation, and data analysis and conclude with specific examples of these calculations for free energies of solvation and relative ligand binding affinities.

### 1.1. Theoretical Principles

Here, we focus on standard approaches to gather statistical samples from simulations using classical molecular mechanics potential energy functions. All of the standard approaches are variations of the same statistical sampling procedure. The main differences between approaches are in the types of energy data collected from simulation samples and the analysis performed with these data. In this discussion, we will assume standard classical molecular mechanics models, including harmonic bond and angle terms, periodic dihedral terms, and non-bonded terms consisting of point charges and Lennard-Jones repulsion/dispersion terms. We will not address free energies with polarizable models or with mixed QM/MM simulations (1), since these are not well developed enough to be of interest in most applications yet.

Free energy differences between states are directly related to the ratio of probabilities of those states. Rigorously, the free energy difference between two thermodynamic states in a constant volume ensemble is:

$$\Delta A_{ij} = -k_B T \ln \frac{Q_j}{Q_i} = -k_B T \ln \frac{\int_{V_j} e^{-\frac{U_j(\vec{q})}{k_B T}} d\vec{q}}{\int_{V_i} e^{-\frac{U_i(\vec{q})}{k_B T}} d\vec{q}}$$

where  $\Delta A_{ij}$  is the Helmholtz free energy difference between state  $j$  and state  $i$ ,  $k_B$  the Boltzmann constant,  $Q$  the canonical partition function,  $T$  is the temperature in Kelvin,  $U_i$  and  $U_j$  are the potential energies as a function of the coordinates and momenta  $\vec{q}$  for two states, and  $V_i$  and  $V_j$  are the *phase-space volumes* of  $\vec{q}$  over which we sample. The phase space volume is the total set of coordinates and momenta where the system has nonzero probability of being found. For ease of notation, we will also use  $(k_B T)^{-1} = \beta$  in this chapter. When the end states have different masses, the contributions of the changing masses to the kinetic energy must also be included in the integral. Generally, masses do not change in most standard simulation setups, and we use the potential energy  $U$  instead of the more general Hamiltonian  $H$  for clarity. From this basic definition, we point out several key facts:

- We are always calculating free energy *differences*, not absolute free energies. All of the quantities that are of interest in biological or thermodynamic experiments are free energy differences between two thermodynamic states, so we must always

specify two states. Even “absolute” free energies of binding are still free energy differences between two states: the ligand restricted to the binding site, and the ligand free to explore all other configurations.

- The temperature  $T$  does not have a subscript, because we are considering free energy differences at the same temperature. One can in theory compute a free energy difference between two states at different temperatures,  $\Delta A_{ij} = -k_B T_j \ln Q_j + k_B T_i \ln Q_i$ . However, this free energy is no longer a ratio of partition functions and it should never be necessary to calculate a free energy this way in systems of biological interest. When a researcher thinks of temperature dependence of the free energy, he or she is usually thinking about the effect of temperature on the free energy difference between two states, both of which are at the same temperature.
- There are two different phase-space volumes,  $V_i$  and  $V_j$ , in the integrals. These two volumes are usually the same, but they are not required to be. This difference can cause severe errors that are hard to detect afterward, because the standard free energy method we discuss here presupposes  $V_i = V_j$ . For example, suppose we are interested in computing the difference in the free energy of solvation between a hard sphere solute with radius  $\sigma$  (state  $i$ ) and a solute with a standard Lennard-Jones repulsion/dispersion potential with radius  $\sigma$  (state  $j$ ). In this case, the phase-space volume  $V_i$  does not include any configurations with solvent molecules closer to the solute than  $\sigma$ , but  $V_j$  does. The standard free energy methods presented here will likely fail when applied to systems like these.

We can easily modify the above discussion to deal with the Gibbs free energy instead. If we replace  $U_i$  and  $U_j$  with  $U_i + PV_i$  and  $U_j + PV_j$ , respectively, and integrate over all container volumes  $V$  (not to be confused with the phase-space volumes discussed above) in addition to integrating over the coordinates  $\vec{q}$ , then we will get the Gibbs free energy  $G$  (instead of  $A$ ) and the isobaric-isothermal partition function  $\Xi$  (instead of  $Q$ ). However, all the derivations presented in this chapter can be extended directly to that case. At constant 1 atm pressure, with near incompressible liquids like water, the change in free energy related to changes in average volume will be relatively small. For example, a change in average volume corresponding to the elimination of a 1-nm sphere would result in a  $PV$  work contribution to free energy of 0.032 kJ/mol or 0.008 kcal/mol, which is generally smaller than the error in all but the most precise experiments. This is only an approximation, as it ignores fluctuations in molecular size, but illustrates that we can generally neglect the  $PV$  component of the free energy and perform calculations at NVT if we are careful to

make sure that the simulation is actually at the average volume for the state. To make clear the discussion is of the NVT case, we will use the Helmholtz free energy difference  $\Delta A$ . Again, any simulation method that includes proper isobaric-isothermal sampling of volumes can simply insert  $U + P_{\text{ref}}V$  in place of  $U$ , and all the subsequent derivations will hold. If masses change, then the Hamiltonian  $H$  can be substituted for the potential energy  $U$ , and all subsequent derivations will also hold.

In the subsequent sections, we review analysis methods such as the Zwanzig relationship, thermodynamic integration, the Bennett acceptance ratio (BAR), the weighted histogram method (WHAM), and the multistate Bennett acceptance ratio (MBAR). “Free energy perturbation” is a common term for these methods that directly compute the free energy difference as a function of changing molecular structure. “Perturbation” usually refers to an approximate theory that can be written as a series expansion. Free energy perturbation, however, is exact. The term “perturbation” here refers to the changes in the *chemical identity* because simulations frequently involve changes in chemical identity, such as an amine to an alcohol, or a methyl group to a chlorine, or the disappearance of atoms. To avoid confusion, we will use the term “alchemical” to refer to this class of methods rather than “perturbation.”

## 1.2. Zwanzig Relationship

The most historically well-known method for calculating free energy differences from simulations is the Zwanzig relationship (2). This method is sometimes called *free energy perturbation* (another reason for avoiding this term!) or *exponential averaging*. The free energy between two potentials  $U_0(\vec{q})$  and  $U_1(\vec{q})$  over a coordinate and momentum space  $\vec{q}$  can be calculated as:

$$\Delta A = -\beta^{-1} \ln \left\langle e^{-\beta(U_1(\vec{q}) - U_0(\vec{q}))} \right\rangle_0 = -\beta^{-1} \ln \left\langle e^{-\beta\Delta U(\vec{q})} \right\rangle_0 \quad (1)$$

We will refer to this method as EXP for exponential averaging. Although the equation is exact if the phase volume for both states is the same, many studies have demonstrated that except in the case of rather small changes, EXP converges very poorly as a function of the number of samples collected. Free energy differences that appear to have converged may only indicate very poor overlap between the two states (3, 4). Except for very specific cases, where the difference between potential energy distributions is known to always be very small for all  $\vec{q}$ , on the order of 1–2  $k_B T$ , EXP should not be used. However, if these differences are shown to be small and it is much easier to simulate one end point state of the calculation than the other, then EXP can be very useful.

For example, if a researcher is interested in the change in free energy due to a slightly different, more expensive potential, then a simulation can be performed with the cheaper potential, and EXP can be used to evaluate the free energy difference between the

cheaper and the more expensive potential. This process is very different, and much more accurate, than simply approximating the free energy difference as the difference of the average energy of two potentials.

EXP is significantly more accurate when the final, “target” state phase space  $V_j$  is a subset of the initial phase space  $V_i$  (4–8) which is actually simulated. The insertion of a small rigid molecule into a dense fluid is an example of this feature of EXP since the set of accessible configurations of the molecule solvated in the dense fluid is a subset of the configurations in which the molecules have no interactions with the solvent. Although there are specific instances where this can be exploited, it has somewhat limited application; in many cases, neither phase space is strictly a subset of the other. Even if such a relationship between phase spaces  $V_i$  and  $V_j$  exists, it can be difficult to identify a priori which one is better to simulate from, and whichever state it is it may be significantly less convenient to sample from.

### 1.3. Multiple Intermediates

In most instances where the states of interest have very little phase-space overlap, the transformation can be broken into a series of intermediate states that do have good overlap with each other. Here, consider  $K - 1$  free energy calculations spanning a series of  $K$  states that *do* have phase-space overlap, where  $k = 1$  and  $k = K$  are our states of interest. Mathematically,

$$\Delta A_{1,K} = \sum_{i=1}^{K-1} \Delta A_{i,i+1}$$

A separate free energy calculation is then performed for each of the individual  $\Delta A$ 's. Since we usually care only about the free energy differences between the end states, the precise form of the intermediates is unimportant. This leaves us free to choose intermediate states that have significant overlap with one another, which means we can choose completely unphysical states if they lead to less total error for the transformation. Statistical uncertainty grows quickly when the amount of phase space becomes small, so the total uncertainty decreases quickly as a function of the number of intermediates. Common examples of nonphysical intermediates include an atom with charges turned off, an atom with van der Waals parameters that are partway between a carbon and a nitrogen, or a “softened” atomic site that solvent molecules can penetrate.

It is both conceptually useful and mathematically convenient to think of these intermediate states as lying along a continuous pathway connecting the initial and final states. The parameterized distance along this path connecting the initial and final states is traditionally called  $\lambda$ , with  $\lambda = 0$  corresponding to the initial state and  $\lambda = 1$  corresponding to the final state. Because these pathways often connect two molecules, we call them alchemical states.

We can then think of the potential describing the system as a function of both the  $\lambda$  parameter and the coordinates  $\vec{q}$ , writing this as  $U(\lambda, \vec{q})$ . We must then perform simulations of  $U(\lambda, \vec{q})$  at a series of  $\lambda$  values, generating samples that will allow us to estimate each of the  $\Delta A_{i,i+1}$  free energy differences.

#### 1.4. Thermodynamic Integration

By taking the derivative of the free energy with respect to the variable  $\lambda$  describing the distance along the series of intermediate alchemical states, we find that:

$$\begin{aligned} dA/d\lambda &= \frac{d}{d\lambda} \int e^{-\beta U(\lambda, \vec{q})} d\vec{q} = \left\langle \frac{dU(\lambda, \vec{q})}{d\lambda} \right\rangle_\lambda, \\ \Delta A &= \int_0^1 \left\langle \frac{dU(\lambda, \vec{q})}{d\lambda} \right\rangle_\lambda d\lambda, \end{aligned} \quad (2)$$

Computing free energies using this formula is called *thermodynamic integration*, abbreviated as TI in this chapter, and is often done using standard numerical integration techniques. Averaging over  $\left\langle \frac{dU}{d\lambda} \right\rangle$  requires fewer uncorrelated samples to reach a given level of relative error than averaging  $e^{-\beta \Delta U(x)}$ , as the range is usually narrower with a more Gaussian distribution. However, this is only true as long as  $\left\langle \frac{dU}{d\lambda} \right\rangle$  is well behaved, an important condition we will address later. Because we can only simulate a limited number of intermediates, we must use some type of numerical integration of the integral. By definition, numerical integration introduces bias, which must be minimized sufficiently that it is well beneath the level of statistical noise.

Various numerical integration schemes are possible, but the trapezoid rule provides a simple, flexible, and robust scheme. All types of numerical integration can be written as:

$$\Delta A \approx \sum_{k=1}^K w_k \left\langle \frac{dU(\lambda, \vec{q})}{d\lambda} \right\rangle_k$$

where the weights  $w_k$  correspond to a particular choice of numerical integration. Researchers have tried a large number of different integration schemes (9–11). Many other integration routines require specific choices of  $\lambda$  to minimize bias, which makes them unsuitable when the intermediates have widely-varying levels of uncertainty. For starting researchers, we therefore recommend a simple trapezoid rule scheme, as it allows for maximal flexibility in which values of  $\lambda$  are simulated. If we are using the trapezoid rule, with equal spacing between points, this becomes  $w_1 = w_K = \frac{1}{2(K-1)}$ , while  $w_k = \frac{1}{K-1}$  for  $k \neq 1, K$ . For the trapezoid rule with uneven spacing, then  $w_1 = \frac{\lambda_2 - \lambda_1}{2}$ ,  $w_K = \frac{\lambda_K - \lambda_{K-1}}{2}$ , and  $w_k = \frac{\lambda_{k+1} - \lambda_{k-1}}{2}$  for  $k \neq 1, K$ . Although the trapezoid rule is very robust, some improvements can be made by using a fit of the data to a polynomial (11) or to some other functional form (10). Since fits to higher

order polynomials can have numerical instabilities for some underlying functions, and alternate functional forms might only be appropriate with some types of transformations, some expertise and experience are required to perform such numerical integration modifications.

TI can be extremely simple to apply for some paths, but most paths require derivatives with respect to  $\lambda$  to be calculated directly in the code. If the pathway is chosen such that  $U(\lambda, \vec{q}) = (1 - \lambda)U_0(\vec{q}) + \lambda U_1(\vec{q})$ , then  $\frac{dU}{d\lambda} = U_1(\vec{q}) - U_0(\vec{q})$ , which can be easily calculated in post-processing by evaluating the same configuration at the initial and final states. If the pathway is not linear in the potential, then the derivative must be calculated analytically in the code. Unfortunately, many problems of interest require using pathways that are not linear, as we discuss below. If the code does compute  $\frac{dU}{d\lambda}$ , then TI is perhaps the simplest method to use, as it involves a very little post-processing, and the analysis requires only simple averages and sums. As long as care is taken to make sure that enough intermediates are used to reduce bias in the integration well below the statistical noise, then TI gives very robust free energy results. In general, the more curvature  $\langle \frac{dU}{d\lambda} \rangle$  has, the more intermediates will be required. In biophysical cases with standard potentials,  $\langle \frac{dU}{d\lambda} \rangle$  is a differentiable function, and any substantial apparent kinks in the curve indicate that more  $\lambda$  values are required near that kink to reduce error due to numerical integration.

### 1.5. Bennett Acceptance Ratio

Differences in the potential energy in both directions between two states can be used to give a significantly improved estimate of the free energy over that obtained by EXP. The difference between the potential energy of the same configuration  $\vec{q}$  for two different states along the pathway is  $\Delta U_{ij}(\vec{q})$ . There is a direct relationship between the distributions of potential energy differences  $\Delta U_{ji}(\vec{q})$  of states sampled from  $i$  and  $\Delta U_{ij}(\vec{q})$  sampled from the state  $j$  (12). Because of this relationship, there is a very robust, statistically optimal way to use information from both states  $i$  and  $j$  together to obtain an improved estimate of the free energy difference between two states. This was first derived by Bennett and is hence generally known as the Bennett acceptance ratio. Bennett's original derivation started with a relationship for the free energy difference

$$\Delta A_{ij} = -\ln k_B T \frac{Q_j}{Q_i} = k_B T \ln \frac{\langle \alpha(\vec{q}) \exp[-\beta \Delta U_{ij}(\vec{q})] \rangle_j}{\langle \alpha(\vec{q}) \exp[-\beta \Delta U_{ji}(\vec{q})] \rangle_i} \quad (3)$$

which is true for any function  $\alpha(\vec{q}) > 0$  for all  $\vec{q}$ . Bennett then used variational calculus to find the choice of  $\alpha(\vec{q})$  minimizing the

variance of the free energy (13), resulting in an implicit function of  $\Delta A$  easily solvable numerically:

$$\sum_{i=1}^{n_i} \frac{1}{1 + \exp(\ln(n_i/n_j) + \beta\Delta U_{ij} - \beta\Delta A))} - \sum_{i=1}^{n_j} \frac{1}{1 + \exp(\ln(n_j/n_i) - \beta\Delta U_{ji} + \beta\Delta A))} = 0 \quad (4)$$

where  $n_i$  and  $n_j$  are the number of samples from each state. We will refer to this method as BAR. A separate derivation shows that the same formula provides a maximum likelihood estimate of the free energy, given the samples from the two states (14). Both derivations give the same robust estimate for the variance and uncertainty of the free energy. Studies have demonstrated both the theoretical and practical superiority of BAR over EXP in molecular simulations (3, 4), and BAR can be shown to converge to EXP in the limit that all samples are from a single state (13, 14). Significantly less overlap between the configurational space of each state is required in order to converge results than in the case of EXP, though some overlap must still exist.

It is difficult to compare TI and BAR on a theoretical basis, because the two approaches use different information. However, practical experience indicates BAR generally performs more efficiently. More precisely, given a fixed length of each simulation, fewer intermediate states are required for BAR than for TI to give equivalent level of statistical precision. TI can be as efficient as BAR under conditions where the integrand is very smooth (3, 15), such as charging or small changes in bonded or nonbonded parameters. In other cases, such as the pathways required to completely eliminate large numbers of atomic sites, BAR is much more efficient than TI or EXP for free energies of larger molecular changes (3, 4, 16). Additionally, no analytical computation of  $dU/d\lambda$  is required, which must be built into the simulation code.

### **1.6. Weighted Histogram Analysis Method**

The *weighted histogram analysis method* (WHAM) provides a way to use information from all of the intermediate  $\lambda$  values in computing free energy differences between states. Most free energy calculations require simulations at a number of different intermediates, and we would prefer to use as much thermodynamic information as possible from all of these simulations simultaneously to save computational cycles. Multiple histogram weighting techniques were first introduced by Ferrenberg and Swendsen (17) in order to capture all of the thermodynamic information from all sampled states in computations of free energies and other observables. WHAM is a histogram reweighing technique introduced in 1992 by Kumar and collaborators for alchemical simulations (18). WHAM is the lowest uncertainty method for calculating free energies using samples collected from discrete states. However, it

introduces biases for continuous distributions, such the energies of atomistic simulations, because all variables must be discretized into bins. Other variations of WHAM based on maximum likelihood (19) and Bayesian methods (20) have also been developed. Beginners should generally not write their own WHAM implementation, because solving the nonlinear equations correctly can be very challenging. The GROMACS and CHARMM molecular mechanics packages include WHAM-based free energy calculations (21–23), and several other stand-alone WHAM implementations are available, so new development of WHAM tools is generally not necessary, other than for pedagogical reasons.

One can reduce the WHAM equations to a simpler form by shrinking the width of the histograms to zero (18, 22), yielding a set of iterative equations which estimate the free energies from  $K$  states simultaneously:

$$A_i = -\beta^{-1} \ln \sum_{k=1}^K \sum_{n=1}^{N_k} \frac{\exp[-\beta U_i(\vec{q}_{kn})]}{\sum_{k'=1}^K N_{k'} \exp[\beta A_{k'} - \beta U_{k'}(\vec{q}_{kn})]} \quad (5)$$

where  $i$  runs from 1 to  $K$ , the  $A_i$  are the free energies of each state,  $\vec{q}_{kn}$  is the  $n$ th sample from the  $k$ th state, and the  $U_i$  are the potentials of these  $K$  states. Although this looks like a formula for absolute free energies, not a formula for free energy differences, the equations are only unique up to an additive constant, so we must fix one of free energies, usually the initial state. This means we are effectively calculating free energy differences from that reference state. The derivation of this approximation is somewhat suspicious for finite numbers of samples, as the derivation involves finding the weighting factors that minimize the variance in the occupancy of the bins, which becomes undefined as the bin width and therefore the average number of samples per bin goes to zero!

### 1.7. Multistate Bennett Acceptance Ratio (MBAR)

A multistate extension of BAR called the multistate Bennett acceptance ratio, or MBAR (24), was recently introduced to overcome the binning issues with WHAM. In this approach, a series of  $K \times K$  weighting functions  $\alpha_{ij}(\vec{q})$  are derived to minimize the uncertainties in free energy differences between all  $K$  states considered simultaneously. The lowest variance estimator is exactly the WHAM equation in the limit of zero-width histograms Eq. 5. WHAM can be therefore be interpreted as a histogram-based approximation to MBAR. This MBAR derivation additionally gives the statistical uncertainty of the calculated free energies, which is not available in WHAM. MBAR has no histograms and hence no histogram bias, and is guaranteed to have lower bias and variance than WHAM. However, in many cases, the bins for WHAM are small enough so that the difference in free energies between the two methods is negligible compared to the statistical precision required. If WHAM is implemented directly in the code, it may not be worth the

additional gain to switch to MBAR, as the statistical uncertainty can be obtained by alternate methods such as the bootstrap method which we will be described below.

### **1.8. Nonequilibrium Methods**

Nonequilibrium simulations can also be used to compute free energies. In a physical or alchemical process where thermodynamic variables change over some interval of time, some amount of work  $W$  is required to make this change. If this is done infinitely slowly, the process is reversible, and  $W$  will be the free energy difference between the end states. However, if the change is performed in a finite amount of time, this process will not be reversible, and hence, the work will not equal to the free energy. Jarzynski noticed (92) that the free energy of the transformation can be written as an average over the nonequilibrium trajectories that started from an equilibrium ensemble:

$$\Delta G = \beta^{-1} \ln \langle e^{-\beta W} \rangle_0. \quad (6)$$

If the switching is instantaneous, then Eq. 6 is identical to EXP because the instantaneous work is simply the change in potential energy  $\Delta U_{ij}$ .

A version of BAR (though not MBAR) can be constructed with the non-equilibrium work (12, 14). A number of studies have compared nonequilibrium pathways to the equilibrium pathways. It appears that under most circumstances, equilibrium simulations are about the same or slightly more efficient than free energies calculated from ensembles of non-equilibrium simulations (25, 26). It is thus not yet clear the extent to which free energy calculations using Jarzynski's relationship will be useful in ligand binding calculations in the future, because of the extra complications of running many separate trials. This is an area of intense research, partly because this formalism has proven useful in treating nonequilibrium experiments as well as simulations and partly because there still are some tantalizing possibilities for substantially increasing the efficiency in free energy calculations. However, we do not recommend that beginners use these methods, as they add an extra degree of complication to both the simulation and the analysis. Further information on how to implement such nonequilibrium calculations can be found in other reviews (27).

## **2. Methods**

### **2.1. Outline of Free Energy Calculations**

Calculating a free energy requires a molecular simulation package that (1) generates samples from the equilibrium distribution of the states of interest, as well as from any intermediate states that might

be required, and (2) extracts basic energetic information from those states. These are the only essential ingredients.

Several key ingredients in a simulation package can help make free energy calculations much more convenient. The key features of a code that make it easy to calculate free energies efficiently are (1) the ability to simulate low variance intermediate states, (2) easy, efficient calculation of the required energetic information (either  $\Delta U_{ij}$  or  $\frac{dU}{dx}$ ); and (3) some degree of automation of the analysis of this information. Many free energy calculations can be performed with any molecular mechanics or Monte Carlo code, though calculations sufficiently efficient to study large and complicated systems require code specifically set up to support good free energy practices.

In what follows, we discuss how to conduct free energy calculations, striving to avoid specifics about particular codes and tools. It is impossible to give full descriptions of proper steps for all simulation packages, as the free energy capabilities of virtually all simulation packages are evolving rapidly. Most of the common packages used for biological or other condensed matter simulations (AMBER, CHARMM, GROMACS, GROMOS, DL\_POLY, LAMMPS, Desmond) have at least one of the standard free energy functionalities built in, but certainly not all programs have all or even most free energy functionalities. With collaborators, we have been developing a web site, <http://www.alchemistry.org>, which is intended to provide more in-depth information with code-specific instructions and cover points too detailed to include in this present format, as well as offering example input and result files.

All methods for computing free energy differences presented in this chapter consist of the following steps:

1. Construct a thermodynamic cycle that allows easy calculation of the free energy of interest, and determine the end states for each calculation required by the thermodynamic cycle.
2. Choose a sequence of intermediate states connecting the two end states for each free energy calculation.
3. Perform equilibrium simulations of the states of interest and any required intermediate states to collect uncorrelated, independent samples.
4. Extract the information of interest required for the desired free energy method from the sampled configurations.
5. Analyze the information from the simulations to obtain a statistical estimate for the free energy, including an estimate of statistical error.

### 2.1.1. Construct Thermodynamic Cycles and Choose End States

In some cases, the free energy difference of interest can be calculated without any intermediate states, and a researcher only needs to identify the two actual physical states that are the end points of the free energy of interest. This is relatively rare. So, it is important to note again that since the free energy is a state function, any series of transformations connecting the two end points gives the correct free energy. In many cases, it will be significantly more effective to use less direct paths that are more efficient.

Relative binding free energies can be calculated simply using thermodynamic cycles, providing a good example of their utility. These could be computed by performing two separate free energies of binding and subtracting them (see Fig. 1). However, free energies of binding can require extremely long simulation times because they require removing the entire ligand from the environment of the protein and the solvent, a process that can require either prohibitively large amounts of computer power or fairly involved constraining potentials to improve convergence. However, as can be seen in Fig. 1, this same free energy difference can be written as the difference of two different nonphysical processes, the changing of molecule *A* to molecule *B* while bound, and the changing of molecule *A* to *B* while unbound:

$$\Delta A_{\text{bind}} = \Delta A_{\text{bind}}^B - \Delta A_{\text{bind}}^A = \Delta A_{A \rightarrow B}^{\text{bound}} - \Delta A_{A \rightarrow B}^{\text{unbound}}.$$

Since the unbound protein is the same in both cases, no simulation needs to be performed of the unbound protein.

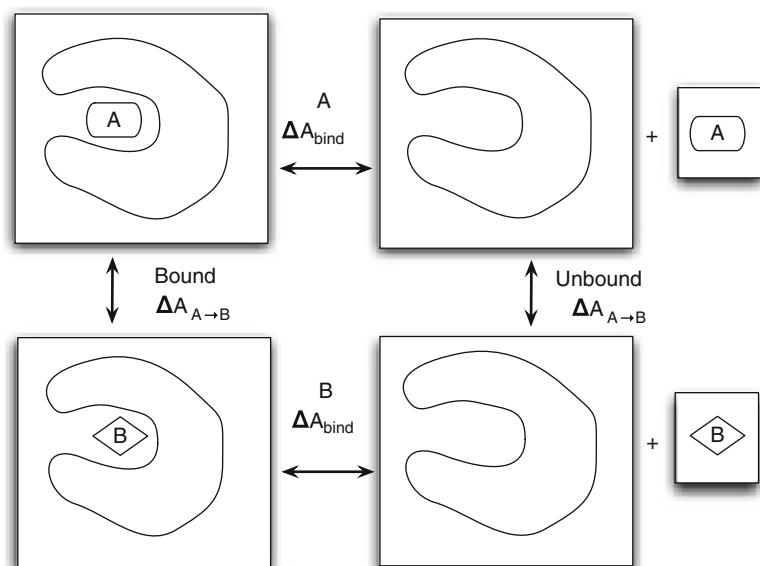


Fig. 1. The thermodynamic cycle for the relative binding affinities of ligands *A* and *B* to a host molecule.

The next step is to determine which simulations correspond to the end states of the free energy differences of interest. This must be done carefully. An example for computing a solvation free energy of a small-molecule solute can demonstrate how important it is to analyze the states carefully. For a solvation calculation at constant temperature and volume, the initial state of the calculation consists of the solute and some quantity of solvent in a specified constant volume. The final state consists of the same small-molecule solute in vacuum in the same volume as in the initial state, plus the same number of solvent molecules as the initial state, also in the same volume as the initial state. In the case of constant pressure simulations, which is usually of the most interest, all states should be at the same pressure, meaning the final box without the solute must shrink somewhat. Although it is not possible to have a simulation in vacuum at finite pressure, the pressure of the ideal gas state will cancel out.

A typical beginner's error is to use a final end state with all energetic terms of the solute turned off, which is not correct; in the vapor phase, the intramolecular interactions of the solute should remain turned on. Only the *intermolecular* interactions should be turned off. If the final state has all energetic terms turned off, then an additional calculation must be carried out in which these terms are turned back on in vacuum.

There is one exception to the requirement to have simulations performed at both end states. If the two end states are very close in phase space, such that the variance in energy differences  $U_i(\vec{q}) - U_j(\vec{q})$  is less than  $1-2 k_B T$ , then simulations can be done with just a single end state, using EXP. For example, if the free energy difference of interest is simply the difference due to differences between two very similar parameterizations of the same molecule, then EXP will likely be sufficient, as long as the calculation is not carried out so near a phase boundary that the potential perturbation carries the simulation over that boundary.

### 2.1.2. Choose a Series of Intermediate States

If the end states of the transformation of interest do not have significant overlap in phase space, a series of intermediate states is required. The judicious choice of these intermediates is one of the most complicated aspects of free energy calculations.

It is important to clarify some of the terms used in free energy calculations. When performing *equilibrium* simulations of intermediate states along a pathway, any distinction between "forward" and "backward" is arbitrary. If one state contains an atom in state A that is not present in state B, then interpreting A as the initial state and B as the final state means that this atom is disappearing or being destroyed or annihilated, whereas treating state B as the initial state means that the same atom is being created or introduced into the system. The choice of words to describe this change is entirely semantic. We will generally refer to either of these changes as *decoupling*, where only intermolecular interactions are turned

off, or *annihilation*, which refers to turning off all interactions with the system, both intermolecular and intramolecular, rather than creation or coupling.

The simplest (though frequently not the best) choice for most transformations between two potential functions  $U_0$  and  $U_1$  is the linear pathway. For example,

$$U(\lambda, \vec{q}) = (1 - \lambda) U_0(\vec{q}) + \lambda U_1(\vec{q}) + U_{\text{unaffected}}(\vec{q}) \quad (7)$$

where  $U_{\text{unaffected}}(\vec{q})$  is the potential due to interactions which do not change as a function of intermediate state. For annihilation, it will be the solvent-solvent interactions; for decoupling, it will be the solvent-solvent *and* solute-solute interactions.

A significant problem with this approach is that equal spacing in  $\lambda$  does not actually lead to equal phase-space overlap between states. If a Lennard-Jones function is used to for atomic exclusion and dispersion interactions, as is typical for biomolecular interactions, then when  $\lambda = 0.1$ , nearly at one end state, the excluded volume for a OPLS-AA united methane sphere (i.e., the volume with energy above 2–3  $k_B T$ ) will still be 60–70% of the original volume.

More severely, this choice of parameterization with a  $r^{-12}$  leads to a singularity in  $\frac{dU}{d\lambda}$  at  $r = 0$ , which then cannot be integrated numerically. Some studies try to approximate this difference by extrapolation, but this is extremely unreliable and error-prone. Therefore, a linear pathway in energy should not be used to annihilate or decouple atoms. By using a power of  $\lambda \geq 4$  instead of a strictly linear parameterization (such as  $U(\lambda) = (1 - \lambda)^4 U_0 + \lambda^4 U_1$ ), then the integral of  $dU / d\lambda$  will converge. However, it will converge rather slowly in number of samples and can cause numerical instabilities (28, 29). For any nonzero  $\lambda$ , whatever the power, there will be small “fence posts,” particles with a small impenetrable core (29). One possible way to avoid issues with these “fence posts” has been to shrink the entire molecular structure. However, this can create problems with nonbonded interactions as the molecular framework shrinks, causing instabilities in integration in molecular dynamics (29–31), and is generally not practical for large numbers of bonds.

Fortunately, there are now standard ways to handle the decoupling of atomic sites in an efficient way, the “soft-core-potential” approach (32, 33). In this approach, the infinity at  $r = 0$  of the  $r^{-12}$  interaction is “smoothed out” in a  $\lambda$  dependent way. The most common form of the pairwise potential is:

$$H(\lambda, r) = 4\epsilon\lambda^n \left[ \left( \alpha(1 - \lambda)^m + \left( \frac{r}{\sigma} \right)^6 \right)^{-2} + \left( \alpha(1 - \lambda)^m + \left( \frac{r}{\sigma} \right)^6 \right)^{-1} \right] \quad (8)$$

where  $\varepsilon$  and  $\sigma$  are the standard Lennard-Jones parameters,  $\alpha$  is a constant (usually 0.5), with the original choice of  $n = 4$  and  $m = 2$  (32). Further research has shown that using  $n = 1$  and  $m = 1$ , with  $\alpha$  fixed at 0.5, noticeably improves the variance (28, 29, 34); this particular choice is now the default in GROMACS, for example. Additional research has shown that there are even more efficient pairwise pathways, though these have not yet been carefully studied to identify unforeseen problems (35).

To turn off intermolecular interactions between a molecule and its surroundings requires decoupling both the charge and the Lennard-Jones interactions. One highly reliable, relatively high-efficiency pathway for annihilation or decoupling of atoms is to turn off the charges of these atoms linearly and then afterwards turn off the Lennard-Jones terms of the uncharged particles using the soft core approach. The same pathway can be followed in reverse for atomic sites that are introduced (23, 29). This ensures that when the repulsive cores with infinite positive energy at  $r = 0$  are eliminated, there are no negative infinity energies at  $r = 0$  due to Coulombic attraction between unlike charges. Note that soft-core methods should not be used to change from one set of Lennard-Jones potentials to another, as the infinite core is first removed and then reintroduced, which is very noisy. Another similar approach is to turn off both the Coulombic and the dispersion terms first and then in a separate step turn off the repulsive term. There appears to be little difference in efficiency between these two approaches. It is possible to turn off both the Coulombic term and the van der Waals term at the same time using soft-core potentials (32, 36), but it can be difficult to choose parameters for these approaches that are transferable between systems (37). We recommend using the soft core for only the van der Waals interactions, after charges have been turned off separately and linearly.

Constructing alchemical pathways between two molecular end states involves one of two main approaches. These are the *single topology* approach and the *dual topology* approach (see Fig. 2). In the single topology approach (a, upper), a single topology has sites that correspond to atoms in both molecules. In the specific example given in the figure, at one end state, two hydrogens are turned into “dummies” which have no nonbonded interactions with the rest of the system, and the upper heavy atom is an oxygen. At the other end state, all atoms are interacting, and the upper heavy atom is now a carbon. The alternative dual topology approach (b, lower) differs in that no atoms change their type; they merely change back and forth from being dummies to being fully interacting particles. In this case, at the ethanol end state, the methyl group is noninteracting, while in the ethane end state, the hydroxyl group is noninteracting. One advantage to dual topology approach is that the moieties which change are free to sample the configurational space while being decoupled. This can help increase the sampling

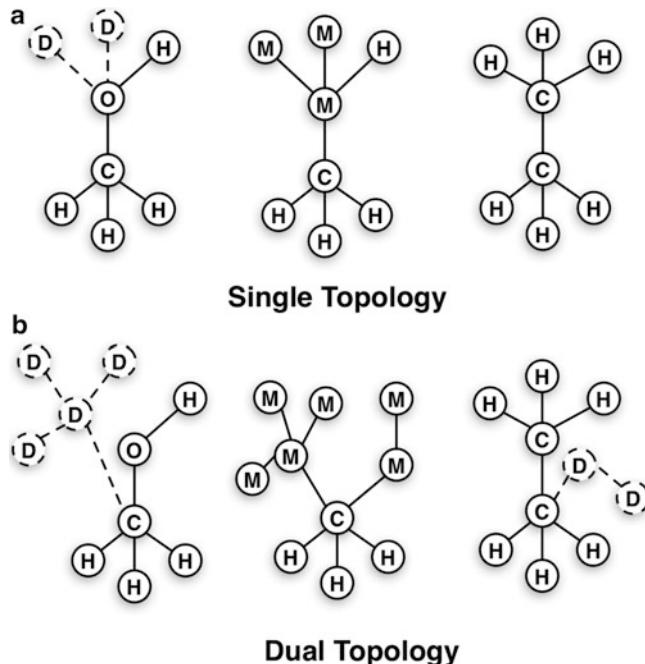


Fig. 2. Single topology (**a**, upper) and dual topology (**b**, lower) approaches to constructing an alchemical path between ethane and ethanol. D represents noninteracting dummies, while M represents nonphysical intermediate atoms. In a dual topology approach, no atoms change type, only have their interactions turned off from the rest of the system; however, more atoms need to be altered to go from the initial to the final state.

if the simulations at different intermediates are coupled in a way that allow exchanges, such as expanded ensemble or Hamiltonian exchange simulations. However, for a dual topology approach, more atoms or molecules must be annihilated or decoupled from the environment, which will therefore require more intermediates. But in many cases, even with such large molecular changes occurring, the convergence time may be the limiting factor, so a dual topology approach can be more efficient.

Dummy atoms can in principle affect free energies, but handled correctly, their effects can often be neglected. Although the end states shown in Fig. 2 have the correct nonbonded interactions for both ethane and ethanol, they are clearly different molecular objects, as they have nonphysical dummies bonded to the carbon or oxygen. Because these dummy atoms affect the system through bonded interactions, we need to account for their free energy contributions. The easiest solution is to perform the transformation in both vacuum and in the molecular surroundings. In the rigid-rotor approximation, where all bonds are fixed in length, the effect on the free energy of these nonphysical dummies cancel out (38). If the bonds are not constrained, then there will be slight differences, but they appear to be small enough (less than 0.01 kcal/mol) to be neglected in any problem of real interest.

In many cases, we also need to modify the bonded interactions of the molecule. For example, in the single topology transformation of ethane to ethanol, the angle and dihedral terms involving the changing heavy atom are clearly different in the two end states. We must change these bonded interactions in addition to the nonbonded interactions. This transformation can be handled in a straightforward way. The variance due to changes in the bonding terms is not generally a problem; although the energy changes for these terms can be quite large, the time scale of the motions means that they converge quite quickly. Pathways that are linear in the bonded parameters (like harmonic spring constants and equilibrium bond lengths or angles) are perfectly adequate. However, care must also be taken for constrained bonds. There is no phase-space overlap between bonds constrained to two different lengths, and so an approach that only constrains covalent bonds to hydrogens is preferred to avoid correction terms that can be difficult to compute (39).

The choice of single versus dual topology will depend on the simulation code used—individual simulation packages may only support one or the other. Both cases will lead to correct final results if performed correctly.

Opening or closing rings is very complicated in performing free energy calculations. Performing such ring changes requires removing bonds in both single and dual topology methods, which is problematic because of changes in phase-space overlap. It is much better to appear or disappear rings entirely even though they are somewhat large. The proper use of soft-core potentials makes decoupling of five or six atom rings relatively straightforward, as evidenced by the high accuracy of numerous small-molecule solvation calculations (34, 40–43).

### 2.1.3. Pulling Methods

A completely different choice of pathway for the free energy of protein-ligand association is to physically pull the molecule away from the protein. If the final state is sufficiently far from the original protein, the free energy of this process can be equated to the free energy of binding. This can be done either by nonequilibrium simulations, using the Jarzynski equation as discussed earlier (44), or by computing a PMF using umbrella sampling with different overlapping harmonic oscillators at specified distances from the binding site (45–47).

There are a number of complications with pulling methods. Pulling a ligand out of a completely buried site can have high statistical error because of the lack of a direct pathway, and it can be difficult to pull the ligand sufficiently far away from the protein as to be decoupled with a simulation box of tractable size. Additionally, a pulling pathway must be chosen. In the case of reasonable box sizes, some analytical or mean-field approximation must be applied for the free energy of pulling the ligand to infinity, and

there has not been extensive research on the reliability of such corrections. Some researchers have argued that pulling may be more efficient for highly charged ligands (46). However, because of the difficulty of choosing such a pathway, pulling approaches are not recommended for beginners.

#### *2.1.4. Rules of Thumb for Constructing Intermediate States*

There are a number of other small points that are worth taking into account when deciding on a series of intermediate states, not all of which can be fully described in limited space, but we list as many as possible here, as well as summaries of the discussions above that require emphasis:

- Bonded terms, such as angle or bond force constants can be changed or turned off linearly. Changes in unconstrained bond distances, can also be performed linearly.
- Constrained bonds should not generally change length, as there are free energy terms associated with these changes that cannot be neglected (48).
- Choose a pathway that maximizes the similarity between two states. Remove or decouple fewer atoms when possible.
- Do not open or close rings. There are some fundamental theoretical problems with changing the number of degrees of freedom in changing thermodynamic states. It is much better to make entire rings disappear and appear, even if it involves more atoms changing.
- Given a fixed number of intermediate states, the states should be chosen such that the statistical uncertainty of the free energy difference between any neighboring pair of states is equal. This is not simply an empirical rule of thumb; it is also proven to lower the overall variance (49).
- *Changes* in parameters can be calculated using a simple linear function. Introduction or deletion of atoms should always be done with a “soft-core” type potential on Lennard-Jones terms.
- Charges on any atoms to be created or annihilated should be completely off before the atomic repulsive terms are turned off. Otherwise, the simulation will rapidly crash as charges of opposite charge will approach to zero distance and go to a negative infinite potential.
- The variance shrinks very quickly as a function of state spacing. Until the free energy differences between intermediates are lowered to approximately  $2\text{--}3 k_B T$ , and if sufficient CPUs are available, it is better to use more states than fewer states. If limited by the number of CPUs available, fewer states can be used. However, it may end up being less statistically efficient, as more uncorrelated samples will be required from each simulation.

- For a given scheme, the shape of the variance curve as a function of  $\lambda$  does not change significantly with the number of atoms that are changing (40), though the magnitude obviously does change significantly. The same relative spacing density for a given code and  $\lambda$  pathway will be approximately the same a molecule with ten atoms will also work well for a molecule with 50 atoms, though of course, more total intermediates will be required for a larger molecule.
- Quickly prototyping possible intermediate states with short simulations is highly recommended. The rough magnitude of variance of free energy differences can be estimated with very short simulations, frequently as quickly as 100 ps, though occasionally this variance may be artificially low if it does not include important configurations that have not yet been sampled.
- The total charge of the simulation should be maintained across all values of  $\lambda$ . Free energy calculations with charged molecules are fine, as long as the total charge of the system remains the same. Most methods for computing long-range electrostatics make approximations, such as a uniform neutralizing charge, which are reasonable if the total charge of the system remains the same. However, when the overall charge of the system changes as a function of  $\lambda$ , these approximations can lead to significant differences in the overall free energy. Simulations with changing charges require complicated corrections to give quantitatively reliable answers (50, 51).

#### 2.1.5. Perform Simulations of the States of Interest

The heart of the free energy calculation is conducting *equilibrium* simulations of *the states of interest* and any required intermediate states to collect *uncorrelated, independent* samples. There are several important topics to cover to ensure reliable, repeatable results:

- The simulations must be at equilibrium. Even for non-equilibrium work simulations, the initial states must be in equilibrium. Sufficient time must be given for the system to reach equilibrium before samples are collected. Because many free energy methods effectively give large weight to favorable rare events, a small amount of unequilibrated data can have an outsized contribution to the overall free energies.
- The system must reach equilibrium *at each value* of  $\lambda$ . One efficient way to start each system is to run a series of short (10–100 ps) simulations at each  $\lambda$  state, restarting the next state from the final configuration of the simulation of the previous state. This gives the system time to partly relax to the new intermediate state's potential and avoid instabilities in simulations. Changes in the volume occupied by the changing molecule or molecules can affect the total energy. As  $\lambda$  changes, the volume should be allowed to adjust as well so that the

solvent density of the system does not change as the effective volume of the molecule changes. Small changes in  $V$  can cause problems, not because the  $PV$  term becomes significant, but because liquids are nearly incompressible, and a small change in average volume leads to a large change in thermodynamic properties. To obtain the most consistent results, if the final simulations at each  $\lambda$  are run at NVT, they should use the average volume of the system from an NPT equilibration, as different fluctuations in the box volume can lead differences of 0.1–0.3 kcal/mol in the final free energy for typical box sizes for small-molecule solvation. However, it can take 100–1,000s of ps, or even longer in some cases (52), to relax to the new intermediate state, so significant time should be allowed for this relaxation to occur. This time varies drastically from system to system, and no hard and fast rule can be given. For solvation of smaller molecules, it may take only 100–500 ps, but for systems that are started out of equilibrium and have long correlation times, it could take hundreds of nanoseconds. The average energy of the simulation,  $\langle \frac{dU}{d\lambda} \rangle$ , as well as structural observables, must be carefully monitored for convergence. The number of hydrogen bonds to a small molecule is one useful observable to watch for convergence of a simulation because it can exhibit a relatively slow equilibration rate (53).

- The samples must be collected at the state of interest. In all simulation codes, different choices of simulation parameters can result in changes in the potential energy surface. If such a change moves the entire potential energy surface up by a constant amount, or affects the relative depths of wells by less than a few tenths of  $k_B T$ , then a simulation at a given intermediate may appear to be unaffected. But if these choices result in energies that change in the potential surface as a function of  $\lambda$ , it can lead to significant free energy differences.

To give just one example, for simulations done with the standard particle mesh Ewald (PME) treatment of long-range electrostatics, PME parameters that are sufficient for “standard” MD can give significant errors in the free energy for modifying partial charges on a molecule, up to 4 kcal/mol for some small molecules. So when doing free energy calculations, it is in general not a good idea to assume that particular settings are not important. If the potential could possibly be affected, the dependence on this parameter should be checked.

- The samples must be *independent*, meaning they are uncorrelated in time. All of the analysis methods presented here assume independent samples. But for all but the simplest of systems, completely independent samples can be very difficult to generate. For protein-ligand binding affinities, the time scale for some motions may be 100s of ns, meaning truly uncorrelated samples may be impossible to generate in a reasonable amount

of time with today’s simulation technology. In this case, free energy calculations *might* provide some useful information, but will only be approximations to the correct free energy for that model, and cannot be considered reliable.

### 2.1.6. Extract Information from the Samples

Once samples and energies are obtained, then we can apply the analysis methods discussed above. The data required from the sample will depend on the method used:

- TI requires the value of  $\frac{dU(\vec{q})}{d\lambda}$ .
- EXP requires either the energy difference  $\Delta U_{k,k+1}(\vec{q})$  or  $\Delta U_{k,k-1}(\vec{q})$ , where  $k$  is the state of that sample, depending on which direction the free energy is calculated.
- BAR requires both the energy difference  $\Delta U_{k,k+1}(\vec{q})$  and  $\Delta U_{k,k-1}(\vec{q})$  between all pairs of states.
- WHAM and MBAR both require the set of energy differences  $\Delta U_{k,j}(\vec{q})$ , where  $j = 1 \dots K$  runs over all states along the pathway, though this information must be binned for WHAM.

For BAR, MBAR, and WHAM, this information can either be computed directly during the simulation, or in post-processing. It is obviously preferable to have this information automatically computed during the simulation, as it eliminates additional work incurred from additional simulation setup and analysis, avoids errors that might result from these additional steps, and reduces the amount of data that must be kept. It is recommended to use information computed during the simulation if at all possible.

However, if configurations from each simulated state  $k$  are stored sufficiently frequently, and with sufficient precision, then single point energy calculations can be run using each of these configurations as input to produce the quantities  $\Delta U_{k,j}(\vec{q})$ . For BAR, only three single point calculations (at  $k + 1$ ,  $k$ , and  $k - 1$ ) need to be performed for each saved configuration, while for MBAR or WHAM,  $K$  single point calculations need to be performed for each configuration. Although technically  $\Delta U_{k,k}(\vec{q})$  does not need to be computed, as it should be zero, it is highly recommended to compute this quantity. First, it allows a check of whether the energy obtained for that configuration during the original simulation at state  $k$  is the same as the energy obtained in the reevaluation. If the difference between the two energies is greater than could be explained by numerical precision issues, then the simulation setup should be rigorously checked for self-consistency; such errors can easily lead to large free energy differences. The precision in the coordinates of the output files must be greater than the precision in standard PDB files. Coordinates stored as binary format are of course greatly preferred, but precision to within  $10^{-5}\text{\AA}$  may be a sufficient compromise depending on the software used. In any case, specific choices must be carefully validated.

In some special cases where  $U(\lambda, \vec{q})$  is a separable function of  $\lambda$  and  $\vec{q}$  like the linear case,  $U(\lambda, \vec{q}) = (1 - \lambda)U_0(\vec{q}) + \lambda U_1(\vec{q})$ , then TI can be computed in post-processing using the single point energies of the endpoints. In other cases, like for soft-core potentials,  $\frac{dU(\vec{q})}{d\lambda}$  cannot be computed in post-processing energy components and must be computed directly in code.

Once the data have been assembled, we must extract independent subsets of the data, which involves an analysis of autocorrelation times. The autocorrelation time measures the time between effectively uncorrelated samples, and there are a number of approaches for computing it. Most start with the calculation of the normalized autocorrelation function of the observable  $A$  gathered over a simulation of time  $T$ . If we write  $\delta A(t) = A(t) - T^{-1} \int_{t=0}^T A(t) dt$ , the instantaneous value minus the average over the interval, then

$$C_A(\Delta t) = \frac{\int_{\tau=0}^T \delta A(\tau) \delta A(\tau + \Delta t) d\tau}{\int_{\tau=0}^T \delta A(\tau)^2 d\tau}.$$

If the  $C_A(\Delta t) = 0$  at and after  $\Delta t$ , then two samples separated by  $\Delta t$  are uncorrelated, and can be treated as independent.

For a series of  $N$  samples, occurring time  $\delta t$  apart, then  $C_A(\delta t)$  will be defined at  $i$  distinct points. Defining  $\delta A(i) = A(i) - \frac{1}{N} \sum_{i=0}^N A(i)$ , then:

$$C_A(i) = \frac{\sum_{j=0}^N \delta A(j) \delta A(j + i)}{\sum_{j=0}^N \delta A(j)^2}.$$

The autocorrelation time  $\tau$  is then defined as the integral under  $C_A$ . In many circumstances, the autocorrelation function can be fit to an exponential, in which case  $\tau$  is simply the relaxation time of the exponential function. Alternatively,  $\tau$  can be computed as the integral under the  $C_A(t)$  curve, though care must be taken as it becomes noisy at long times, especially at more than half of the total simulation time. As a rule of thumb, a total time of  $50\tau$  should be simulated in an estimate of  $\tau$ , as long correlation times may not be detected by shorter simulations. Once  $\tau$  is computed, then under standard statistical assumptions, samples can be considered effectively uncorrelated if they are spaced by  $2\tau$ .

Many mature simulation packages have tools to compute these correlation times, sometimes at a more sophisticated level than presented here. In any case, we emphatically encourage practitioners to use some tool for computing correlation times, or the calculated statistical uncertainty will be lower than the true uncertainty.

It appears that for solvation free energies of small molecules, the time scales involved are often not particularly long. The longest time scales are those for water rearrangement and torsions. In unpublished test in our group, the correlation times of  $\frac{dU}{d\lambda}$  for

small rigid molecules were all around 5–30 ps. However, if there are explicit torsional barriers in the molecule which are particularly high, such as boat/chair transitions or slow rotations of internal torsions (such as the hydroxyl orientation in carboxylic acids), this correlation time can be many nanoseconds (54).

Once the correlation time is calculated, there are two possible ways to use the information to obtain answers from independent data. For methods that compute averages from single states, like TI, the average over all samples can be used for the mean, and the variance then multiplied by  $\sqrt{2\tau/\Delta t}$ , where  $\Delta t$  is the time between samples, to obtain an effective variance. Alternatively, the data set can be *subsampled*, with a set of samples separated by  $2\tau$  being selected to analyze. For example, assume we are using BAR to compute the free energy between the first and second states, and we have collected 5 ns of simulation, with snapshots every 10 ps, for a total of 500 samples. Then we need to take the time series  $\Delta U_{1,2}$  and  $\Delta U_{2,1}$ , and compute the autocorrelation function and correlation time. If we assume the correlation time for  $\Delta U_{1,2}$  is 20 ps, and the correlation time for  $\Delta U_{2,1}$  is 40 ps, then we should take every fourth sample (or  $2\tau$ ) from the  $\Delta U_{1,2}$  data series and every eighth sample from the  $\Delta U_{2,1}$  data series, and do subsequent analysis only with this reduced data set. If the correlation time is estimated accurately, we are not actually throwing away information by discarding data, since the discarded data duplicates information contained in the retained data.

Technically, we are only sampling independent configurations if *all* coordinates are uncorrelated between samples, not just the energies. In most cases, independent sampling of the energies also implies uncorrelated sampling of the configurations. However, there are a number of situations in which energies appear to be sampled approximately independently within the limit of the noise, but the configurational space is only partly sampled. For example, if there is a second binding pose that has similar binding affinity, but which the ligand only travels to occasionally, this might not show up when inspecting the correlation time of the energetic components alone.

This problem can be partially solved by also monitoring structural correlation times. For example, for a small-molecule solvation energy, the correlation times of slow dihedrals can be computed. For a binding affinity problem, the autocorrelation time of the distance between a given point on the protein or the ligand dihedral angle between a bond in the protein and a bond in the ligand can be computed to verify that sufficient sampling is indeed happening on the time scale of the simulation.

### 2.1.7. Analyze the Information from the Simulations to Obtain a Statistical Estimate for the Free Energy

Once we have a set of independent samples of energy data from a series of equilibrium simulations, we can analyze these data to obtain an estimate of the free energy and the error associated with its estimate. The exact form of the analysis will depend on the method being used, so we will look at different methods individually.

*Data Analysis for TI.* Given a set of  $N_k$  samples of  $\frac{dU}{d\lambda}$  from equilibrium at each of  $k$  states,  $\langle \frac{dU}{d\lambda} \rangle_k$  can be computed from the simple averages  $\langle \frac{dU}{d\lambda} \rangle_k = N_k^{-1} \sum_{i=1}^{N_k} \frac{dU}{d\lambda}$  at each state  $k$ . To compute the free energy  $\Delta A$ , we then perform numerical integration

$$\Delta A \approx \sum_{k=1}^K w_k \left\langle \frac{dU}{d\lambda} \right\rangle_k$$

where the  $w_k$  are weighting factors corresponding to different types of numerical integration. If we are using simple trapezoid rule, with equal  $\lambda$  spacing,  $w_1 = w_K = \frac{1}{2(K-1)}$ , while  $w_k = \frac{1}{K-1}$  for  $i \neq 1, K$ . For the trapezoid rule with uneven spacing, then  $w_1 = \frac{\lambda_2 - \lambda_1}{2}$ ,  $w_K = \frac{\lambda_K - \lambda_{K-1}}{2}$ , and  $w_k = \frac{\lambda_{k+1} - \lambda_{k-1}}{2}$  for  $k \neq 1, K$ . As discussed previously, the trapezoid rule is the most robust and most recommended for beginners since it easily allows for unequal spacing in  $\lambda$ , which is required in order to minimize the variance. Although alternative methods can yield lower integration error, these methods require significant problem-specific information, and are not recommended for beginners. In almost all cases, it is simpler to identify regions of high curvature and run more simulations in these areas.

Computing the overall variance of TI is straightforward, though it involves one pitfall. It is important to calculate the overall variance of the total integral, rather than calculating the variance of each individual  $\Delta A_{i,i+1}$  and assuming the variances of the individual transformations add independently. Each transformation involves data from neighboring states, so the uncertainty in these intervals do not add independently. Each of the  $\langle \frac{dU}{d\lambda} \rangle$  results is independent of the others since they are generated from different simulations, and therefore,  $\text{var}(\delta A) = \sum_{i=1}^K w_k^2 \text{var}(\frac{dU}{d\lambda})_k$ . In the case of simple trapezoid rule, we can see that

$$\begin{aligned} \text{var}(\Delta A_{1,K}) &= \sum_{k=1}^K w_k^2 \text{var}\left(\frac{dU}{d\lambda}\right)_i \\ &= \frac{1}{4} \text{var}\left(\frac{dU}{d\lambda}\right)_1 + \text{var}\left(\frac{dU}{d\lambda}\right)_2 + \cdots + \text{var}\left(\frac{dU}{d\lambda}\right)_{K-1} \\ &\quad + \frac{1}{4} \text{var}\left(\frac{dU}{d\lambda}\right)_K \end{aligned}$$

This is very different than if we calculated the variance for each  $\Delta A_{i,i+1}$ , and then added these variances directly! That process would look like:

$$\begin{aligned}\text{var}(\Delta A_{i,i+1}) &= \frac{1}{4} \text{var}\left(\frac{dU}{d\lambda}\right)_i + \frac{1}{4} \text{var}\left(\frac{dU}{d\lambda}\right)_i, \\ \text{var}(\Delta A_{1,N}) &= \sum_{i=1}^{N-1} \text{var}(\Delta A_{i,i+1}) \\ &= \frac{1}{4} \text{var}\left(\frac{dU}{d\lambda}\right)_1 + \frac{1}{2} \text{var}\left(\frac{dU}{d\lambda}\right)_2 + \cdots + \frac{1}{2} \text{var}\left(\frac{dU}{d\lambda}\right)_{K-1} \\ &\quad + \frac{1}{4} \text{var}\left(\frac{dU}{d\lambda}\right)_K\end{aligned}$$

As discussed above, each of the individual averages should then multiplied by  $\sqrt{2\tau}$  to obtain a corrected variance that corresponds to the correlation time for each state. The standard error can then be computed as  $\sqrt{\text{var}(\Delta A_{1,N})}$  from all samples. Alternatively, averaging and integrating can be performed on the subsampled data set. For alchemical changes that result in smooth, low curvature sets of  $\frac{dU}{d\lambda}$ , TI can be accurate using a relatively small number of points. However, if the curvature becomes large, as is frequently the case for alchemical simulations where Lennard-Jones potentials are turned on or off, then the bias introduced by discretization of the integral can become large (3, 32, 40). Even in the case of small curvature (i.e. charging of small polar molecule in water) reasonably large errors can be introduced (i.e., 5–10% of the total free energy with 5 $\lambda$  values). The basic conclusion is that TI is an adequate method for most purposes, but a researcher *must* verify that enough states are included such that the free energy is essentially independent of the number of states. If a molecule is being annihilated, TI might require a large number of states to give accurate results, as the curvature is large. Large variance at a given state indicates large curvatures, so  $\lambda$  should be chosen to minimize variance.

*Data analysis for EXP.* Free energy propagation from EXP can be analyzed in the same way as TI, using the correlation time to either subsample or to calculate an effective sample number. Since EXP produces free energy differences between intermediates that depend only on samples from one state, variance estimates for individual  $\Delta A_{i,i+1}$  values are independent, and the total variances will add.

*Data analysis for BAR and MBAR.* For BAR, the mathematical details are somewhat more complicated, since they involve solving a set of iterative equations. The variance estimate from BAR computes the variance in the free energy between two states. As with TI, the variances of consecutive intervals  $k-1$  to  $k$  and  $k$  to  $k+1$  are correlated since they both involve samples from the state  $k$ . However, the relationship between these values is more complicated than with TI. Alternative methods, such as bootstrap sampling described below, must be used to obtain an accurate error estimate. MBAR also involves solving complex systems of linear equations to compute the variances. However, for MBAR, all correlations

between data are taken into account. Implementations of both BAR and MBAR, with examples for free energy calculations, can be found at <http://www.simtk.org/pymbar>. Subsampling must be done to produce proper error estimates, as the statistical efficiency cannot be added at the last step as it is for EXP and TI, as different states can have different correlation times  $\tau$ .

### 2.1.8. Bootstrap Sampling

A simple, general, and robust method to calculate the statistical uncertainty is the *bootstrap method* (55), which works with all the free energy methods discussed here. Assume we have some function  $F(x)$  that is computed from  $n$  data samples,  $x_1, \dots, x_n$ , with the property that as  $n$  goes to infinity,  $F(x)$  converges to some constant value  $\bar{F}$ .  $F$  could be something as simple as the average of all the  $x$  values =  $\frac{1}{n} \sum_{i=1}^n x_i$ , or the average of some function  $\frac{1}{n} \sum_{i=1}^n A(x_i)$ , but it could be something as complicated as the MBAR or WHAM free energy estimate. To compute a bootstrap variance, we:

1. Pick  $n$  samples from the list of samples  $\vec{x}\{x_1, \dots, x_n\}$ , *with replacement*, giving a new set of samples  $\vec{x}_i$ . Since when picking randomly from  $\vec{x}$ , we can pick any sample for all  $n$ , there will almost certainly be repeated samples. This process is called sampling from the *empirical distribution*—that is, the distribution we measured, rather than true distribution from which we sampled. For example, if  $\vec{q} = \{1, 2, 6, 4, 3\}$ ,  $\vec{q}_i$  might be  $\{6, 3, 3, 1, 1\}$ .
2. Compute  $\hat{F}_i = F(x_i)$ . For each bootstrap sample, compute the function from the bootstrapped data set  $i$ . For example, if our function was a simple average, we would compute the average of our new bootstrap sample. For MBAR, we would generate bootstrap samples from *each* of  $K$  states each time and compute the full set of free energy differences from these  $K$  bootstrap samples of each state.
3. Return to step 1, repeating the process  $M$  times, with  $M$  at least 50–200 to obtain accurate variances (55). If the calculation in step 2 is cheap,  $M$  can be significantly larger to obtain more reliable uncertainties; the relative uncertainty scales as  $M^{-1/2}$ , it may require more than 1,000 steps for the variance to converge to 1%.
4. Compute the variance  $\text{var}(\hat{F})$ , the sample variance over the  $M$  bootstrap values  $\hat{F}_i$ . This variance is the bootstrap variance. The bootstrap standard deviation is the square root of the bootstrap variance.

The  $\hat{F}_i$  values are the bootstrap distribution of the function  $F(x)$ . In the limit of large numbers of samples, under most conditions, it will converge to the true distribution of  $\bar{F}$ , but is a good estimate even for moderate values of  $n$ . If  $n$  is too low, then the bootstrap distribution will be somewhat different than the true distribution.

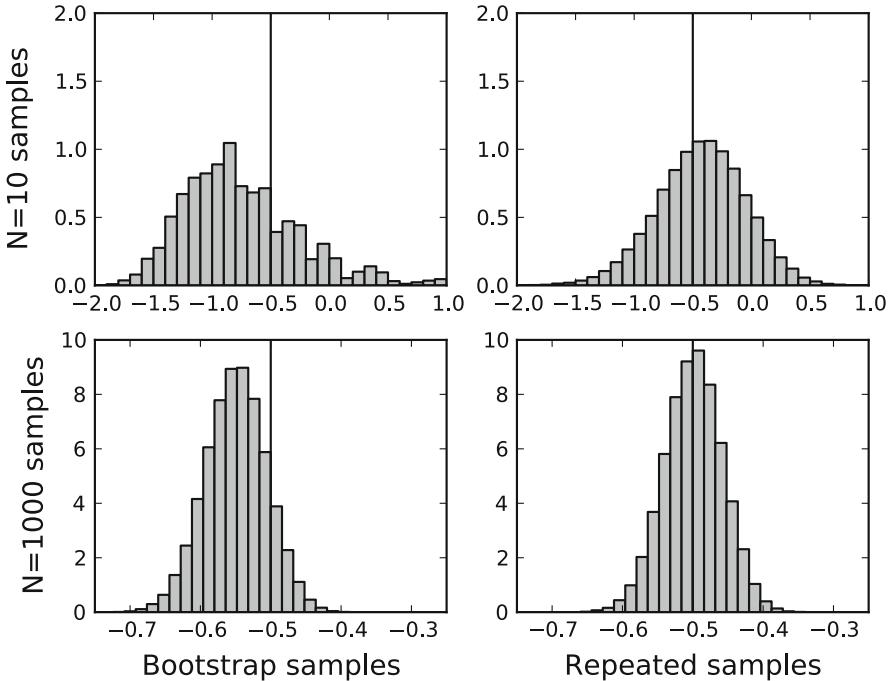


Fig. 3. A comparison of the estimated distribution of exponential average (EXP) results using bootstrap sampling to the true distribution obtained using multiple draws. Each draw consists of  $N$  samples from a normal distribution with mean 0 and variance 1. For  $N = 10$  samples, bootstrap sampling does not give the true distribution; for  $N = 1,000$  samples, bootstrap sampling yields a very close match to the distribution, with standard error 0.0437 vs. 0.0414. Either 100,000 bootstrap samples or 100,000 independent draws were performed in each case.

A simple example will illustrate its use. We sample values of  $x$  from a 1D Gaussian with mean 0 and variance 1, then compute  $F(x_i) = -\ln N^{-1} \sum_{i=1}^N \exp(-x_i)$ . This is, of course, the Zwanzig formula with  $k_B T = 1$ , representing a typical case for free energy calculations. We then compare the distribution of free energies obtained with 100,000 bootstrap samples of a single sample of  $N$  points with the true distribution obtained by calculating the quantity from 100,000 independent samples (see Fig. 3). For small  $N$ , the distributions are not close, but for larger  $N$ , the distributions almost converge. In the case of  $N = 1000$ , the variances of this new distribution can be seen by inspection to be essentially the same. Indeed, the entire distribution is extremely close to the true distribution. We emphasize again: the second distribution uses data from  $1000 \times 100,000$  points collected, while first distribution captures the main features with only *one* sample of 1000 points, resampled 100,000 times. The same amount of analysis is done in both cases, but the second case requires only 0.001% of the data collection. Clearly, rare events that require more than 1000 samples to observe will not be properly treated, but if  $N$  is sufficiently large to include examples of significant rare events with the proper relative distribution, bootstrap sampling will give a good estimate of the statistical uncertainty of the calculation.

The great power of bootstrap sampling is that it can be used with *any* statistical estimator, no matter how complicated. However, it does require the additional overhead of calculating the function of the samples  $F(x)$   $M$  times. In most cases, this will be negligible compared to the time used to generate the data, perhaps 5–10 min for MBAR, seconds for TI. For further information on bootstrapping, see Ref. (55). One particular method to point out is the method of moving block bootstrapping, in which blocks of length  $\tau$  are randomly selected in order to properly treat time correlations without subsampling.

## **2.2. Accelerating Sampling**

At this stage, we have presented a series of methods and their robust implementation for calculating the free energy differences between two systems. However, in many cases of interest, this may require significant investment of computational resources, beyond that which can be obtained by most researchers. We will also therefore examine additional tools for accelerating the sampling. Because of space limitations, we will not go deeply into all these methods. They are not needed to carry out free energy calculations, but may be extremely useful to converge calculations in complex systems with slow dynamics.

### *2.2.1. Using Umbrella Sampling for Convergence*

One standard method for improving sampling in atomistic simulations is umbrella sampling (56), where bias terms are added to constrain the simulation in some way, and the effect of these restraints is then removed. This procedure can be used either to lower potential energy barriers or to restrain simulations to slow-interconverting configurations that are relevant to the binding affinity (e.g., different torsional states), allowing each of the component free energies to be properly computed and then combined (23, 57, 58). Sometimes this can even be necessary for hydration free energy calculations with slow degrees of freedom (54). Another application of umbrella sampling is computing the free energy of constraining the free ligand into the bound conformation directly before computing the free energy of binding and then computing the free energy of releasing these restraints in solution. This usually decreases the correlation times for sampling of the intermediate states and thus can increase the efficiency of the simulation (23, 46).

### *2.2.2. Expanded Ensemble, Hamiltonian Exchange, and $\lambda$ -Dynamics*

It is possible to bring all the intermediates together in a single simulation system, either as series of coupled simulations of the intermediate states, usually called a Hamiltonian exchange simulation, or as a single simulation that simultaneously samples both intermediate states and separate coordinates, called expanded ensemble simulation. A number of studies have shown that Hamiltonian exchange simulations can speed up simulations by allowing the system to avoid kinetic barriers by going through alchemical states where those barriers are not as pronounced, significantly

speeding up free energy simulations (59–64). Alternatively, the alchemical variable  $\lambda$  can be treated as a dynamical variable, which adds complications by introducing a fictitious mass corresponding to the  $\lambda$  degree of freedom, but is essentially equivalent to Monte Carlo techniques (62, 65–67). There are a number of variations of sampling in  $\lambda$  that may show promise in the future, but such methods are still in the preliminary stages of development (68–72).

At the current time, although they are extremely promising, we cannot recommend expanded ensemble and  $\lambda$ -dynamics methods to most practitioners. The methodology and implementations are not always robust, and require tweaking additional parameters to obtain proper convergence. However, we do recommend Hamiltonian exchange methods. Most codes implementing Hamiltonian exchange methods do so on top of well-tested temperature replica exchange routines, and no additional analysis is needed; the outputs of Hamiltonian exchange simulations can be analyzed in the same way as the outputs of  $K$  uncoupled simulations. These simulations are guaranteed to decorrelate as fast or faster than standard simulations, though the exact amount of improvement depends on the system; they are guaranteed to not be any worse! The analysis of correlation times can be somewhat complicated by such simulations; computation of correlation times should be computed along trajectories that sample different states, not along single states that might be switching back and forth along very different trajectories.

### **2.3. Verification, Verification, Verification**

There are a number of ways that free energy simulations can go wrong, and the lists presented here cannot cover all possible problems. The best defense is to consistently evaluate the validity of each step of the process. For example, it is generally a very good idea to start out by calculating free energies that are well-known. For example, the free energy of solvation of OPLS methane in TIP3P water is known to be  $-2.2 \pm 0.1$  kcal/mol, and has been replicated a number of times in different simulation programs. The hydration free energy of toluene in TIP3P water with AMBER/AMBER-GAFF parameters and HF 6-31G\* RESP charges has also been the object of multiple studies and has been reported as  $-0.41 \pm 0.2$  and  $-0.7 \pm 0.1$  kcal/mol (29, 41). The website <http://wwwalchemy.org> maintains a number of these examples to test.

One of the most common problems that can occur is that the input files and/or options used to perform the free energy calculations are different than the input files used to perform standard calculations. In virtually every free energy enabled code, this leads to the possibility that the state set up for free energy calculations no longer corresponds to the same state when free energy options are turned off. To avoid this, you should always verify that the potential energy of the system with free energy options turned off in the initial state is *exactly* the same as the potential energy at  $\lambda = 0$  with

the free energy options turned on. Likewise, you should always verify that the potential energy of the system with free energy options turned off in the final state is *exactly* the same as the potential energy at  $\lambda = 1$  with the free energy options turned on. By “exactly,” we mean that any differences should be no more than those caused by numerical rounding from differences in order of operations. Anything larger indicates some breakdown in the computation that could potentially result in propagated error significantly affecting the results.

Another common problem is human error in setting up simulations. If humans are involved in editing topology and other input files for the initial and final states, it is easy to accidentally set up one atom to have an incorrect final state, or mis-type a key parameter. This typically means human input is a bad idea and calculation setup should be done by script or program, since at least then bugs are reproducible. New tools for calculation setup should be carefully tested on cases with known results to ensure that setup is functioning correctly.

Poor convergence, undetected by uncertainty analysis, can also wreak havoc on results. There are several methods for validating convergence, such as checking that thermodynamic cycles sum to zero when conducting relative free energy calculations, or ensuring computed free energies are consistent when beginning from markedly different starting structures, which is applicable for both relative and absolute binding calculations, as well as for hydration free energy calculations (54, 58, 73).

As a simple but very useful check, simulation trajectories should always be visually inspected. Visual inspection of trajectories can often catch errors that are hard to otherwise notice. For example, if the calculation is of the relative binding affinity of ligands that are tight binders, but the composite ligand is somehow ejected into the solvent, or adopts unnaturally high-energy configurations, then there is likely an error in the simulation setup. If the molecules move visibly between very different configurations on a long time scale, it indicates either the system is not yet equilibrated or that the correlation times for the system may be slow in a way that does not yet show up in the energetic analysis. Visual inspection during and not just after simulations have completed allows errors to be recognized before too much computational time is wasted.

#### **2.4. Example 1: Solvation of OPLS-AA Ethanol in TIP3P Water**

We now give some specific examples for implementing free energy calculations as described in this chapter. We will look specifically at the free energy of solvation of ethanol and the relative free energy of phenol and toluene binding to T4 lysozyme (57, 58, 74):

- *What is the thermodynamic cycle?* There are two possible thermodynamic cycles: (a) is a direct calculation where only intramolecular interactions are eliminated and (b) an indirect calculation

where all atomic charges and Lennard-Jones  $\epsilon$  are turned to zero, and a second calculation where the same parameters are turned off in vacuum.

- *What are the end states?* In case A, one end state is a periodic box of TIP3P water interacting with OPLS-AA ethanol, and the other end state is a OPLS-AA ethanol in the vacuum state and a box of the same number of TIP3P waters. Both end states can be done in the same simulation, since the two subsystems do not interact, and the final volume for both simulations is the same as in the initial state. In case (b), two calculations need to be performed. The initial state is a periodic box of TIP3P water interacting with ethanol, and the end state is ethanol with its partial charges and Lennard-Jones terms turned off in the same box of water. Depending on the capabilities of the code being used, this second pathway might be the only one available. It has the additional complication of requiring two simulations, but the total computational time is essentially the same, as simulation of a single small molecule in vacuum state is extremely cheap.
- *What are the intermediates?* In case (a), a robust pathway would be to turn off the intermolecular charge interactions with the solvent linearly, then turn off the intermolecular van der Waals interactions with a soft-core potential. Depending on the program used, these two steps might require two different sets of input files, or might only require one. For case (b), a robust pathway would be to linearly scale the charges to zero, and then turn off the Lennard-Jones with a soft-core potential, performing this transformation both in solution and in vacuum.

The number of intermediates required would depend on the length of the simulations, the free energy method used, and the exact details of the soft-core potential. As an example, for BAR or MBAR, 3–5 charge terms and 5–8 Lennard-Jones terms for a small molecule like benzene or toluene would be sufficient to give errors in the 0.1–0.05 kcal/mol range using 5 ns simulation at each  $\lambda$  state. This level of precision was obtained in previous large-scale studies (34), with the caveat that  $\lambda$  spacing was adjusted to give approximately equal variance in free energy over each interval.

If equal spacing in  $\lambda$  is used, the number of states might be significantly higher, perhaps 10 for the Coulombic terms and 20 or more for Lennard-Jones terms. Therefore, finding an appropriate spacing equalizing variance is important for efficiency. There is a tradeoff in  $\lambda$  spacing; it obviously requires more processors to sample more intermediate states, but variance decreases make up for this until spacing is relatively close. TI would likely require even more intermediates, the exact number depending on the level of uncertainty required.

In general, the choice of spacing will depend on the availability of processors, the correlation times of the system under study, and the level of precision required. In both cases, bonded interactions stay the same in the beginning and end state, so nothing special needs to be done for these components.

- *What simulations to run?* Equilibrium simulations must be run at each of the intermediates. Typically one could start with the fully interacting state at all intermediates, and run for perhaps 1 ns, to allow the system to equilibrate, or alternatively prepare initial configurations at each intermediate sequentially, and then equilibrate these for 1 ns or so. The simulation box should be large enough for the molecule not to interact with itself, so the width of the box should be at least twice the cutoff plus the longest width of the molecule. The simulation time required will depend on the accuracy of the simulation; for a molecule this size, for an uncertainty less than 0.1 kcal/mol, then 5 ns at each  $\lambda$  value will likely be sufficient if the time scales for torsional rotation are not too large.
- *What analysis do we perform?* In this case, we assume we are using BAR and that the code does not automatically print out the energy differences. In that case, the potential energy differences must be generated by single point simulations. This can be done by saving configurations every  $N$  steps, where  $N$  will depend on the correlation times of the potential energy. Typically, for a small rigid molecule, it would be around 1–10 ps, though if there are slow degrees of internal freedom, it could be longer. We would then take those configurations, run a series of single point energy calculations. These calculations should be identical to the ones performed to generate the runs, but each configuration will be evaluated at the  $\lambda$  value of the neighboring intermediate. For each interval, we will have two energy differences, from state  $i$  to  $i + 1$  sampled from state  $i$  and from  $i + 1$  to  $i$  sampled from state  $i + 1$ . The BAR calculation is performed for each interval, giving an estimate for the free energy difference. We would then apply bootstrap sampling to the data set of evaluated energies to obtain an error estimate. If the energy differences are printed out, then we can skip all but the bootstrap sampling and BAR calculation, greatly simplifying the analysis.
- *Anything else to watch out for?* At the end state, the ethanol will be isolated from the rest of the system and so cannot exchange its kinetic energy with the rest of the system. With some choices of thermostat, kinetic energy can get transferred between the isolated subsystems in a way that violates equipartition. The degrees of freedom of the water and the ethanol should either be thermostated independently or more robustly with a Langevin or other stochastic thermostat, such as the Andersen thermostat (75).

Although one might not think long-range dispersion effects would cause a significant difference, since dispersion interactions fall off as  $r^{-6}$ , they can play a significant role, adding up to 0.5–1 kcal/mol for small-molecule solvation free energies. If not supported in the code, an analytical correction must be made to integrate the attractive part of the van der Waals term out to infinity. If the system is nonisotropic, then even the analytical correction may be incorrect; however, solutions have been found (76).

### 2.5. Example 2: Relative Free Energies of Binding

As a second example, we look at the difference in the relative free energies of binding of toluene to phenol in the apolar cavity of T4 lysozyme:

- *What is the thermodynamic cycle?* We compute the free energy to turn phenol into toluene in the protein cavity, and compute the free energy to turn phenol into toluene in solution, as described in Fig. 1.
- *What are the end states?* The end states for the first is T4 lysozyme, in water, with an intermediate molecule with non-bonded parameters that look like toluene at one end state and that look like phenol in the another end state. There are a number of choices for even this small change. It is perhaps simplest to choose a dual state topology, such as in Fig. 4; an *ortho*- or *meta*- arrangement could be used just as easily.
- *Which series of intermediates?* Arbitrarily selecting phenol as the initial state, the OH moiety must disappear, and the methyl must appear. A good approach would be to first turn the charge

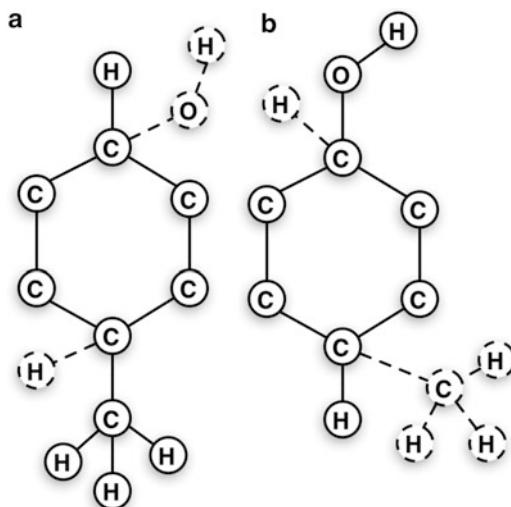


Fig. 4. A sample dual topology design for the transformation of phenol (a) to toluene (b) as described in the text. The choice of *para* arrangement is arbitrary; *ortho*- or *meta*-arrangements would work as well.

on the OH and the *para* H to zero. It should be done keeping the overall system at the same total charge at each intermediate. Once these charges are turned, then the Lennard-Jones  $\epsilon$  of the hydroxyl and that *para* can be turned to zero. At the same time, and angle, bonded, and torsional terms can be turned off linearly. Then the LJ  $\epsilon$  of the methyl group and its *para* H can be turned on, while its bonded terms are turned on, and finally, the charges of the methyl and its *para* H can be turned on.

How many intermediates will this require? In practice, using BAR or MBAR, perhaps 2–4 intermediates for turning off the charges and 4–6 intermediates for turning off the Lennard-Jones terms and bonded terms, slightly less than for ethanol solvation, since fewer molecules are changing, and thus the phase-space overlap is lower. Thermodynamic integration would require perhaps 2–3 times as many intermediates. It is possible to perform both Lennard-Jones transformations simultaneously, but in this case, it would be necessary to remove the improper torsions on these molecules, as they would force the substituents to collide with each other.

- *What simulations to run?* Again, equilibrium simulations must be run at each of the intermediates. The choice of which starting points to use is more complicated, since the environment is a protein binding site, not a homogeneous liquid. The ideal starting structure is a crystal structure of the ligand bound to the site, or at least a homologous ligand to which the ligand of interest can be modeled without distorting the structure of either the ligand or the protein. If the binding site is not known, then obtaining an accurate free energy is not likely; docking is not necessarily reliable for picking the single true experimental binding site. If the binding site is known but a crystal structure is not available, then docking can be used to generate a range of potential starting locations. Initial simulations several nanoseconds in length, tens of nanoseconds if possible, should be used to test if these configurations interconvert. If they do not, it may be necessary to run multiple simulations for each of the binding sites (58).

Once starting configurations are selected, one would again generally start with the fully interacting state at all intermediates, and run short simulations at each  $\lambda$  to allow the system to partially equilibrate at each new  $\lambda$  value, followed by long equilibrations for each  $\lambda$  state with constant pressure simulations to find the equilibrium density. In this case, at least 2–4 ns should be used to equilibrate, since the correlation times are guaranteed to be longer than in the small-molecule solvation case. Even for a crystal structure, several starting configurations (perhaps obtained from multiple crystal structures, if available) should be used, and examined to see if the ligands sample the

same conformational states in all simulations (58, 73). It is possible that even closely related ligands may not bind in the same orientation, presenting some sampling problems (73).

The simulation box should again be large enough for the protein and ligand to not interact with themselves, so the width of the box should be at least twice the cutoff plus the longest width of the protein plus small molecule. The simulation time required will depend on the accuracy of the simulation; for a molecule this size, simulation times of perhaps 50 ns may be necessary to get consistent results. At the present time, even with relatively rigid proteins, getting results that have statistical uncertainty of less than 0.5 kcal/mol is difficult.

- *How do we analyze the data?* Assume we are now performing thermodynamic integration. In this case, we expect that the values of  $\frac{dU}{d\lambda}$  are printed at each step. It would be impossible to generate this TI data in post-processing, assuming we are using the recommended soft-core potentials. We simply average the values from each simulation, and perform numerical integration and error estimation from the formulas above. The free energies of two transformations, in aqueous solution and in the presence of the protein, are then subtracted to obtain the final result.
- *Anything else to watch out for?* Visualizing the simulation is always a good idea just to make sure nothing strange is happening. Note that the details of the protein were hardly mentioned in the discussion; the protein, in most respects, is just different external environment than the water. One difference that occasionally has some relevance is the location of the binding state. Assuming the ligand is a tight binder, then the ligand will always remain tightly localized around the binding site, and the definition of the binding site becomes pretty much irrelevant. A ligand that is a weak binder ( $K_d > 100 \mu\text{m}$ ) spends more time outside the binding site, and the term “binding affinity” becomes more difficult to define. However, this difficulty of definition occurs in both experiment and simulation. For weak binder, careful comparison of exactly the physical phenomena leading to signaling of binding must be explored, and getting quantitative results will be difficult. In most standard cases, determining precise binding affinities of weak binders is not required. Rather, the scientific questions will be to distinguish between tight binding ligands, or to tell whether a ligand is a tight or weak binder. It can be useful to perform multiple simulations starting from different initial configurations, as it can potentially identify when the simulations have not yet converged.

## 2.6. Conclusions

Free energy calculations are a sophisticated, powerful set of tools for finding properties such as solvation free energies in arbitrary solvents and binding free energies. However, they give only a statistical estimate of free energy differences between two thermodynamic states whose accuracy and precision depend on careful choices of parameters, pathways of intermediates, and methods. Additionally, they can only give the free energies of the model, not the true experimental system; the molecular parameters of the system under study must be sufficiently accurate for the correct free energy for the model to match the free energy of the system.

For complicated systems with long correlations times, free energy methods are not always reliable, because of the difficulty of collecting uncorrelated samples. Relative or absolute free energies of binding to proteins must therefore be taken with some degree of caution. As can be seen from the extensive set of notes and qualifications in the methods presented here, free energy “black-box” methods will not “automagically” determine free energies without significant investment in the physics, chemistry, and biology of the system under study.

However, such calculations are certainly closer to providing utility to biophysical researchers than they were in the past. As we have presented in this chapter, the methods used for free energy calculations are changing rapidly. Major biomolecular simulation packages, such as AMBER, CHARMM, NAMD, GROMACS, and GROMOS, are all undergoing major improvements and changes in the features used to compute binding free energies. Although these changes will likely greatly improve the ability to perform free energy calculations in the near future, ongoing changes make it difficult to put together stable work flows for preparing ligands and simulation structures and determining ideal free energy protocols without significant human effort. It is difficult to recommend particular codes for the easiest use at the present time; we instead recommend using the code with which the user is most comfortable, as long as it supports one of the methods discussed here.

Because of the scope of free energy calculations, a single review cannot hope to capture all possible problems or issues; for further information, readers are encouraged to read a number other of reviews on the subject of free energy calculations (77–84), particularly several very recent reviews (27, 85–87), as well as several useful books (77, 88–91).

## Acknowledgements

The authors wish to thank John Chodera (UC-Berkeley) for ongoing discussions of reliability and usability for free energy calculations as well as TriPham (UVa) and Justin Lemkul (Virginia Tech), Joe Allen (SUNY Stonybrook) for careful review of the manuscript.

## References

1. Woods CJ, Manby FR, Mulholland AJ (2008) An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *J Chem Phys* 128:014109
2. Zwanzig RW (1954) High-temperature equation of state by a perturbation method. i. non-polar gases. *J Chem Phys* 22:1420–1426
3. Shirts MR, Pande VS (2005) Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J Chem Phys* 122:144107
4. Lu ND, Singh JK, Kofke DA (2003) Appropriate methods to combine forward and reverse free-energy perturbation averages. *J Chem Phys* 118:2977–2984
5. Lu ND, Kofke DA (1999) Optimal intermediates in staged free energy calculations. *J Chem Phys* 111:4414–4423
6. Wu D, Kofke DA (2005) Phase-space overlap measures. I. Fail-safe bias detection in free energies calculated by molecular simulation. *J Chem Phys* 123:054103
7. Wu D, Kofke DA (2005) Phase-space overlap measures. II. Design and implementation of staging methods for free-energy calculations. *J Chem Phys* 123:084109
8. Jarzynski C (2006) Rare events and the convergence of exponentially averaged work values. *Phys Rev E* 73:046105
9. Resat H, Mezei M (1993) Studies on free energy calculations. I. Thermodynamic integration using a polynomial path. *J Chem Phys* 99:6052–6061
10. Jorge M, Garrido N, Queimada A, Economou I, Macedo E (2010) Effect of the Integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration. *J Chem Theor Comput* 6:1018–1027
11. Shyu C, Ytreberg FM (2009) Reducing the bias and uncertainty of free energy estimates by using regression to fit thermodynamic integration data. *J Comput Chem* 30:2297–2304
12. Crooks GE (2000) Path-ensemble averages in systems driven far from equilibrium. *Phys Rev E* 61:2361–2366
13. Bennett, C. H. (1976) Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* 22:245–268
14. Shirts MR, Bair E, Hooker G, Pande VS (2003) Equilibrium free energies from non-equilibrium measurements using maximum-likelihood methods. *Phys Rev Lett* 91:140601
15. Ytreberg FM, Swendsen RH, Zuckerman DM (2006) Comparison of free energy methods for molecular systems. *J Chem Phys* 125:184114
16. Rick SW (2006) Increasing the efficiency of free energy calculations using parallel tempering and histogram reweighting. *J Chem Theor Comput* 2:939–946
17. Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo Data Analysis. *Phys Rev Lett* 63:1195–1198
18. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations. I. The method. *J Comput Chem* 13:1011–1021
19. Bartels C, Karplus M (1997) Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations. *J Comput Chem* 18:1450–1462
20. Gallicchio E, Andrec M, Felts AK, Levy RM (2005) Temperature weighted histogram analysis method, replica exchange, and transition paths. *J Phys Chem B* 109:6722–6731
21. Hub JS, de Groot BL, van der Spoel D (2010) g\_wham—a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J Chem Theor Comput* 6:3713–3720
22. Souaille M, Roux B (2001) Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput Phys Commun* 135:40–57
23. Wang J, Deng Y, Roux B (2006) Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophys J* 91:2798–2814
24. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:129105
25. Oostenbrink C, van Gunsteren WF (2006) Calculating zeros: non-equilibrium free energy calculations. *Chem Phys* 323:102–108
26. Oberhofer H, Dellago C, Geissler PL (2005) Biased sampling of nonequilibrium trajectories: can fast switching simulations outperform conventional free energy calculation methods? *J Phys Chem B* 109:6902–6915
27. Pohorille A, Jarzynski C, Chipot C (2010) Good practices in free-energy calculations. *J Phys Chem B* 114:10235–10253
28. Pitera JW, van Gunsteren WF (2002) A comparison of non-bonded scaling approaches for free energy calculations. *Mol Simulat* 28:45–65

29. Steinbrecher T, Mobley DL, Case DA (2007) Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *J Chem Phys* 127:214108
30. Pearlman DA, Connelly PR (1995) Determination of the differential effects of hydrogen bonding and water release on the binding of FK506 to native and TYR82 → PHE82 FKBP-12 proteins using free energy simulations. *J Mol Biol* 248:696–717
31. Wang L, Hermans J (1994) Change of bond length in free-energy simulations: algorithmic improvements, but when is it necessary? *J Chem Phys* 100:9129–9139
32. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett* 222:529–539
33. Zacharias M, Straatsma TP, McCammon JA (1994) Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J Phys Chem* 100:9025–9031
34. Shirts MR, Pande VS (2005) Solvation free energies of amino acid side chains for common molecular mechanics water models. *J Chem Phys* 122:134508
35. Pham TT, Shirts MR (2011) Identifying low variance pathways for free energy calculations of molecular transformations in solution phase. *J Chem Phys* 135:034114
36. Blondel A (2004) Ensemble variance in free energy calculations by thermodynamic integration: theory, optimal Alchemical path, and practical solutions. *J Comput Chem* 25:985–993
37. Steinbrecher T, Joung I, Case DA (2011) Soft-core potentials in thermodynamic integration: comparing one- and two-step transformations. *J Comput Chem ASAP* 32:3253–3263
38. Boresch S, Karplus M (1999) The role of bonded terms in free energy simulations. 2. Calculation of their influence on free energy differences of solvation. *J Phys Chem A* 103:119–136
39. Boresch S, Tettinger F, Leitgeb M, Karplus M (2003) Absolute binding free energies: a quantitative approach for their calculation. *J Phys Chem A* 107:9535–9551
40. Shirts MR, Pitera JW, Swope WC, Pande VS (2003) Extremely precise free energy calculations of amino acid side chain analogs: comparison of common molecular mechanics force fields for proteins. *J Chem Phys* 119:5740–5761
41. Mobley DL, Dumont È, Chodera JD, and Dill KA (2007) Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. *J Phys Chem B* 111:2242–2254
42. Nicholls A, Mobley DL, Guthrie PJ, Chodera JD, Pande VS (2008) Predicting small-molecule solvation free energies: an informal blind test for computational chemistry. *J Med Chem* 51:769–779
43. Mobley DL, Bayly CI, Cooper MD, Dill KA (2009) Predictions of hydration free energies from all-atom molecular dynamics simulations. *J Phys Chem B* 113:4533–4537
44. Ytreberg F (2009) Absolute FKBP binding affinities obtained via nonequilibrium unbinding simulations. *J Chem Phys* 130:164906
45. Lee MS, Olson MA (2006) Calculation of absolute protein-ligand binding affinity using path and endpoint approaches. *Biophys J* 90:864–877
46. Woo H-J, Roux B (2005) Calculation of absolute protein-ligand binding free energy from computer simulation. *Proc Natl Acad Sci* 102:6825–6830
47. Gan W, Roux B (2008) Binding specificity of SH2 domains: insight from free energy simulations. *Proteins* 74:996–1007
48. Boresch S, Karplus M (1996) The Jacobian factor in free energy simulations. *J Chem Phys* 105:5145–5154
49. Shenfeld DK, Xu H, Eastwood MP, Dror RO, Shaw DE (2009) Minimizing thermodynamic length to select intermediate states for free-energy calculations and replica-exchange simulations. *Phys Rev E* 80:046705
50. Kastenholz MA, Hünenberger PH (2006) Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids. *J Chem Phys* 124:124106
51. Kastenholz MA, Hünenberger PH (2006) Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. *J Chem Phys* 124:224501
52. Fujitani H, Tanida Y, Ito M, Shirts MR, Jayachandran G, Snow CD, Sorin EJ, Pande VS (2005) Direct calculation of the binding free energies of FKBP ligands. *J Chem Phys* 123:084108
53. Smith LJ, Daura X, van Gunsteren WF (2002) Assessing equilibration and convergence in biomolecular simulations. *Proteins: Struct Funct Bioinf* 48:487–496
54. Klimovich PV, Mobley DL (2010) Predicting hydration free energies using all-atom molecular dynamics simulations and multiple starting conformations. *J Comp Aided Mol Design* 24:307–316

55. Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall/CRC, Boca Raton, FL
56. Torrie GM, Valleau JP (1977) Non-physical sampling distributions in monte-carlo free-energy estimation : umbrella sampling. *J Comput Phys* 23:187–199
57. Mobley DL, Chodera JD, Dill KA (2007) Confine-and-release method: obtaining correct binding free energies in the presence of protein conformational change. *J Chem Theor Comput* 3:1231–1235
58. Mobley DL, Graves AP, Chodera JD, McReynolds AC, Shoichet BK, Dill KA (2007) Predicting absolute ligand binding free energies to a simple model site. *J Mol Biol* 371:1118–1134
59. Okamoto Y (2004) Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J Mol Graph Model* 22:425–439
60. Roux B, Faraldo-Gómez JD (2007) Characterization of conformational equilibria through Hamiltonian and temperature replica-exchange simulations: assessing entropic and environmental effects. *J Comput Chem* 28:1634–1647
61. Woods CJ, Essex JW, King MA (2003) Enhanced configurational sampling in binding free energy calculations. *J Phys Chem B* 107:13711–13718
62. Banba S, Guo Z, Brooks III CL (2000) Efficient sampling of ligand orientations and conformations in free energy calculations using the lambda-dynamics method. *J Phys Chem B* 104:6903–6910
63. Bitetti-Putzer R, Yang W, Karplus M (2003) Generalized ensembles serve to improve the convergence of free energy simulations. *Chem Phys Lett* 377:633–641
64. Hritz J, Oostenbrink C (2008) Hamiltonian replica exchange molecular dynamics using soft-core interactions. *J Chem Phys* 128:144121
65. Guo Z, Brooks III CL, Kong X (1998) Efficient and flexible algorithm for free energy calculations using the  $\lambda$ -dynamics approach. *J Phys Chem B* 102:2032–2036
66. Kong X, Brooks III CL (1996)  $\lambda$ -dynamics: a new approach to free energy calculations. *J Chem Phys* 105:2414–2423
67. Li H, Fajer M, Yang W (2007) Simulated scaling method for localized enhanced sampling and simultaneous “alchemical” free energy simulations: a general method for molecular mechanical, quantum mechanical, and quantum mechanical/molecular mechanical simulations. *J Chem Phys* 126:024106
68. Zheng L, Carbone IO, Lugovskoy A, Berg BA, Yang W (2008) A hybrid recursion method to robustly ensure convergence efficiencies in the simulated scaling based free energy simulations. *J Chem Phys* 129:034105
69. Zheng L, Yang W (2008) Essential energy space random walks to accelerate molecular dynamics simulations: convergence improvements via an adaptive-length self-healing strategy. *J Chem Phys* 129:014105
70. Min D, Yang W (2008) Energy difference space random walk to achieve fast free energy calculations. *J Chem Phys* 128:191102
71. Li H, Yang W (2007) Forging the missing link in free energy estimations: lambda-WHAM in thermodynamic integration, overlap histogramming, and free energy perturbation. *Chem Phys Lett* 440:155–159
72. Min D, Li H, Li G, Bitetti-Putzer R, Yang W (2007) Synergistic approach to improve “alchemical” free energy calculation in rugged energy surface. *J Chem Phys* 126:144109
73. Boyce SE, Mobley DL, Rocklin GJ, Graves AP, Dill KA, Shoichet BK (2009) Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. *J Mol Biol* 394:747–763
74. Mobley DL, Chodera JD, Dill KA (2006) On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J Chem Phys* 125:084902
75. Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72:2384–2393
76. Shirts MR, Mobley DL, Chodera JD, Pande VS (2007) Accurate and efficient corrections for missing dispersion interactions in molecular simulations. *J Phys Chem B* 111:13052–13063
77. Shirts MR, Mobley DL, Chodera JD (2007) Alchemical free energy calculations: ready for prime time? *Ann Rep Comput Chem* 3:41–59
78. Huang N, Jacobson MP (2007) Physics-based methods for studying protein-ligand interactions. *Curr Opin Drug Di De* 10:325–331
79. Gilson MK, Zhou H-X (2007) Calculation of protein-ligand binding affinities. *Ann Rev Bioph Biom* 36:21–42
80. Meirovitch H (2007) Recent developments in methodologies for calculating the entropy and free energy of biological systems by computer simulation. *Curr Opin Struc Bio* 17:181–186
81. Rodinger T, Pomès R (2005) Enhancing the accuracy, the efficiency and the scope of free energy simulations. *Curr Opin Struc Bio* 15:164–170
82. Jorgensen, W. L. (2004) The many roles of computation in drug discovery. *Science* 303:1813–1818

83. Chipot C, Pearlman DA (2002) Free energy calculations. The long and winding gilded road. *Mol Simulat* 28:1–12
84. Brandsdal BO, Österberg F, Almlöf M, Feierberg I, Luzhkov VB, Aqvist J (2003) Free energy calculations and ligand binding. *Adv Prot Chem* 66:123–158
85. Steinbrecher T, Labahn A (2010) Towards accurate free energy calculations in ligand protein-binding studies. *Curr Med Chem* 17:767–785
86. Michel J, Essex JW (2010) Prediction of protein-ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *J Comput Aided Mol Des* 24:639–658
87. Christ CD, Mark AE, van Gunsteren WF (2010) Basic ingredients of free energy calculations: a review. *J Comp Chem* 31:1569–1582
88. Chipot C, Pohorille A (eds) (2007) Free energy calculations: theory and applications in chemistry and biology, vol 86. Springer, New York
89. Frenkel D, Smit B (2002) Understanding molecular simulation: from algorithms to applications. Academic Press, San Diego, CA
90. Reddy MR, Erion MD (eds) (2001) Free energy calculations in rational drug design. Kluwer Academic
91. Leach AR (1996) Molecular modelling: principles and applications. Addison Wesley Longman Limited, Harlow, Essex, England
92. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693

# Chapter 12

## Recipes for Free Energy Calculations in Biomolecular Systems

**Mahmoud Moradi, Volodymyr Babin, Celeste Sagui,  
and Christopher Roland**

### Abstract

During the last decade, several methods for sampling phase space and calculating various free energies in biomolecular systems have been devised or refined for molecular dynamics (MD) simulations. Thus, state-of-the-art methodology and the ever increasing computer power allow calculations that were forbidden a decade ago. These calculations, however, are not trivial as they require knowledge of the methods, insight into the system under study, and, quite often, an artful combination of different methodologies in order to avoid the various traps inherent in an unknown free energy landscape. In this chapter, we illustrate some of these concepts with two relatively simple systems, a sugar ring and proline oligopeptides, whose free energy landscapes still offer considerable challenges. In order to explore the configurational space of these systems, and to surmount the various free energy barriers, we combine three complementary methods: a nonequilibrium umbrella sampling method (adaptively biased MD, or ABMD), replica-exchange molecular dynamics (REMD), and steered molecular dynamics (SMD). In particular, ABMD is used to compute the free energy surface of a set of collective variables; REMD is used to improve the performance of ABMD, to carry out sampling in space complementary to the collective variables, and to sample equilibrium configurations directly; and SMD is used to study different transition mechanisms.

**Key words:** Free energy, Replica exchange, Proline, Carbohydrate

---

### 1. Introduction

During the last decade, new algorithms combined with advances in sampling techniques and increasing computational power have enabled both accurate and robust (within a given model) calculations of molecular free energies. One can expect that future biomolecular simulations will make use of these new free energy methods and related technology with ever increasing frequency. The purpose of this chapter is to briefly review three complementary simulation techniques in current use for calculating biomolecular free energies

and to illustrate their working with two prototypical examples. Specifically, we will cover the adaptively biased molecular dynamics (ABMD) (1), replica-exchange molecular dynamics (REMD) (2), and steered molecular dynamic (SMD) (3) methods, as applied to the puckering of sugars and the conformational space of short polyproline peptides (4–6). We proceed to the examples after a brief overview of the methodology.

Consider a complex multi-particle system. In order to investigate its equilibrium properties, it is often convenient to identify a suitable *collective variable*  $\sigma = \sigma(\mathbf{r}_1, \dots, \mathbf{r}_N)$ , which is understood to be a smooth function of the particle positions  $\sigma : \mathbf{r}_1, \dots, \mathbf{r}_N \mapsto \mathbb{Q}$ , and then to study its equilibrium probability density:

$$p(\xi) = \langle \delta[\xi - \sigma(\mathbf{r}_1, \dots, \mathbf{r}_N)] \rangle, \quad \xi \in \mathbb{Q}. \quad (1)$$

Here, the angular brackets indicate an ensemble average. This probability density  $p(\xi)$  provides useful information as to the relative stability of the states corresponding to different values of  $\xi$  (7). In practice the Landau free energy or *potential of mean force* (PMF)

$$f(\xi) = -k_B T \log p(\xi) \quad (2)$$

(here  $T$  is temperature and  $k_B$  the Boltzmann constant) is typically preferred over  $p(\xi)$ . Either  $p(\xi)$  or  $f(\xi)$  provides for a reduced-dimensionality description of the system in terms of the collective variable  $\xi$  with other degrees of freedom integrated out. Typically, the collective variable is chosen in such a way as to bring out the relevant physics in a natural and transparent manner. It is often associated with the slowest degrees of freedom of the original system, although this is not a formal requirement.

In most circumstances, the direct computation of the free energy from an equilibrium trajectory is impractical, because the sampling is limited to the vicinity of the local minima of  $f(\xi)$ . The problem may be resolved by means of *umbrella sampling* methods (8), in which the trajectory is biased in order to explore the desired range of  $\xi$ . Later, this bias is “undone,” and the free energy is recovered. Umbrella sampling schemes in combination with the so-called weighted histogram analysis method (WHAM) (9–11) offer a practical route to computing free energies. Most recently, free energy methods based on biased nonequilibrium simulations have become popular. These methods all estimate the PMF from an evolving trajectory (12, 13) and use that estimate to modify the system dynamics so as to ultimately flatten the effective free energy surface. Methods that explore this idea include the local elevation method (LEM) (14), the Wang-Landau approach (15), the non-equilibrium metadynamics (16, 17), etc. Collectively, all these methods may be considered as umbrella sampling in which the biasing potential eventually reproduces the initially unknown free energy surface in the long time limit, as illustrated in Fig. 1.

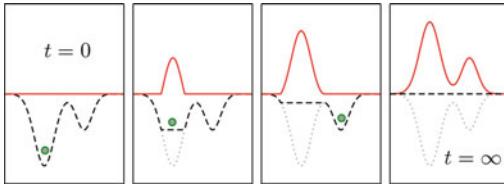


Fig. 1. A schematic illustration for the time-dependent biasing potential used to estimate an unknown free energy surface. Here, the *dotted line* represents the unknown free energy of the system as a function of a suitable collective variable. The solid line (*red*) represents the evolving biasing potential, and the *dashed line* the “effective” free energy of the system. The latter is given by the sum of the unknown free energy and the biasing potential. Initially, at  $t = 0$ , the effective free energy coincides with the free energy of the system. As time evolves towards  $t \rightarrow \infty$ , the effective free energy of the system is flattened and the biasing potential gives the negative of the free energy.

---

## 2. The Adaptively Biased Molecular Dynamics Method

The ABMD method is formulated in terms of the following equations:

$$m_a \frac{d^2 \mathbf{r}_a}{dt^2} = \mathbf{F}_a - \frac{\partial}{\partial \mathbf{r}_a} U[t|\sigma(\mathbf{r}_1, \dots, \mathbf{r}_N)], \quad (3)$$

$$\frac{\partial U(t|\xi)}{\partial t} = \frac{k_B T}{\tau_F} G[\xi - \sigma(\mathbf{r}_1, \dots, \mathbf{r}_N)], \quad (4)$$

where the first one contains the Newton’s equations that govern ordinary molecular dynamics (MD) (temperature and pressure regulation terms are not shown), augmented with an additional force coming from the time-dependent biasing potential  $U(t|\xi)$  (with  $U(t=0|\xi) = 0$ ), whose time evolution is given by the second equation. In the following, we refer to  $\tau_F$  as *flooding time scale* and to  $G(\xi)$  as the *kernel* (in analogy to the *kernel density estimator* widely used in statistics (18)). The kernel should be positive definite ( $G(\xi) > 0$ ) and symmetric ( $G(-\xi) = G(\xi)$ ). It may be viewed as a smoothed Dirac delta function. For large enough  $\tau_F$  and small enough kernel width, the biasing potential  $U(t|\xi)$  converges toward  $-f(\xi)$  as  $t \rightarrow \infty$  (12, 13) (see Fig. 1).

Our numerical implementation of the ABMD method involves the use of cubic B-splines (19) (or products thereof) in conjunction with a biweight kernel for  $G(\xi)$  (18). The method has moderate storage requirements and does not introduce any noticeable performance penalty. In addition, ABMD is characterized by only *two* control parameters: the flooding time scale  $\tau_F$  and the kernel width  $4\Delta\xi$ .

We have also implemented two simple extensions to ABMD. The first is identical in spirit to the *multiple walkers metadynamics* (12, 20).

It amounts to carrying out several different MD simulations biased by the same  $U(t|\xi)$ , which evolves via

$$\frac{\partial U(t|\xi)}{\partial t} = \frac{k_B T}{\tau_F} \sum_{\alpha} G[\xi - \sigma(\mathbf{r}_1^{\alpha}, \dots, \mathbf{r}_N^{\alpha})], \quad (5)$$

where  $\alpha$  labels different MD trajectories. A second extension is to gather several different MD trajectories, each bearing its own biasing potential and, if desired, its own distinct collective variable, into a generalized ensemble for “replica exchange” with modified “exchange” rules (21–23). Both extensions are advantageous and lead to a more uniform flattening of  $U(t|\xi) + f(\xi)$ . This enhanced convergence to  $f(\xi)$  is due to the improved sampling of the “evolving” canonical distribution.

We have implemented the ABMD method in the AMBER package (24), with support for both replica exchange and multiple walkers. In pure “parallel tempering” replica exchange (same collective variable for all replicas),  $N_r$  replicas are simulated at different temperatures  $T_n$ ,  $n = 1, \dots, N_r$ . Each replica has its own biasing potential  $U^n(t|\xi)$ ,  $n = 1, \dots, N_r$ , that evolves according to its own dynamical equation. Exchanges between neighboring replicas are attempted at a prescribed rate, with an exchange probability given by (21)

$$w(m|n) = \begin{cases} 1, & \Delta \leq 0, \\ \exp(-\Delta), & \Delta > 0, \end{cases} \quad (6)$$

$$\begin{aligned} \Delta = & \left( \frac{1}{k_B T_n} - \frac{1}{k_B T_m} \right) (E_p^m - E_p^n) \\ & + \frac{1}{k_B T_m} [U^m(\xi^n) - U^m(\xi^m)] \\ & - \frac{1}{k_B T_n} [U^n(\xi^n) - U^n(\xi^m)], \end{aligned} \quad (7)$$

where  $E_p$  denotes the atomic potential energy. The biasing potentials are temperature-bound and converge in the  $t \rightarrow \infty$  limit to the free energies at their respective temperatures.

In a *general* replica-exchange scheme, we allow different replicas to have different collective variables and/or temperatures. We also allow for either an evolving or a static biasing potential (the latter obviously includes the case of  $U^n(t|\xi) = 0$ ) on a per-replica basis (when all biasing potentials are static, the method reduces to a particular variant of the “Hamiltonian Replica Exchange” described in (21)). Exchanges between random pairs of replicas are then attempted at a prescribed rate.

Finally, we note that in order to improve the accuracy of the free energies as calculated with the ABMD method, follow-up equilibrium umbrella sampling runs may be necessary. Such a procedure is very much in the spirit of adaptive umbrella sampling (25)

and makes use of the biasing potential  $U(\xi)$  as is (26). With such runs, one calculates the biased probability density:

$$p^B(\xi) = \langle \Delta[\xi - \sigma(\mathbf{r}_1, \dots, \mathbf{r}_N)] \rangle_B. \quad (8)$$

The idea here is that if, as a result of an ABMD run,  $f(\xi) + U(\xi)$  is a constant, then the biased probability density  $p^B(\xi)$  will be flat. In a typical simulation, however, this is not the case. Using the results of the follow-up equilibrium runs, one can correct the calculated free energy surface via:

$$f(\xi) = -U(\xi) - k_B T \ln p^B(\xi). \quad (9)$$

This corrective stage can also involve many independent equilibrium runs whose outcomes are “joined” using WHAM (9, 10). Finally, we note that approximate biasing potentials computed by the ABMD may be utilized for enhanced unbiased equilibrium sampling through the replica-exchange scheme (5, 27). All the samples generated by such replica-exchange runs can also be fed to WHAM to “correct” the free energies.

### 3. Replica-Exchange Schemes Utilizing Approximate Free Energies

Although the free energy maps can be used to characterize the properties of a system in terms of one or more collective variables, it is often desirable to sample unbiased equilibrium configurations directly. In this case, the approximate free energies computed via ABMD may be used to set up the Hamiltonian replica-exchange molecular dynamics (HREMD) simulations for enhanced equilibrium sampling (27). Typically, several replicas running at the lowest temperature  $T_0$ , are introduced into the setup. One of the replicas is completely unbiased and therefore samples the unbiased Boltzmann distribution at  $T = T_0$ . The other replicas are subject to a reduced bias at  $T_0$ . These “proxy” replicas ensure adequate exchange rates and hence enhance the mixing.

In the most general case, the different replicas enhance the crossing of barriers by means of a temperature ladder and/or biasing potentials: the former is not specific to any particular mode of the system, while the latter allows to emphasize some modes (e.g., cis/trans transitions of the prolyl bonds for the poly-proline system to be discussed later). In many cases, both devices may be necessary for robust REMD simulations. The ergodic time scale of a REMD simulation is largely determined by two factors: the efficiency of the random walk between the replicas, and the fastest ergodic time scale present among the individual replicas. The first factor can be described in terms of the exchange rates between the pairs of replicas and, as long as the rates are sufficiently high, need not be considered. The replica characterized by the fastest

ergodic time scale is usually referred to as the “hottest” one. Typically, it is the replica running at the highest temperature. Its mixing properties can be assessed by simulating it alone, making sure that it is ergodic, and that its autocorrelation time is sufficiently small so as to generate enough independent samples in a feasible runtime. In practical setting, the number of replicas is typically limited, and hence, the “hottest” replica must be constructed with care. In principle it is always possible to use just the temperature alone (however, the higher the temperature of the “hottest” replica, the more replicas are needed to span the temperatures down to lowest  $T_0$ ) or to use the biasing potentials alone (running all the replicas at the same temperature). The second scenario is not always practical, as it is considerably less generic than the first one. A combination of Hamiltonian and temperature REMD methods (HT-REMD (5, 28)) could provide a practical way to reduce the computational costs of the REMD sampling, as it permits one to facilitate the sampling in the “hottest” replica by using both means, thus allowing for better “tuning” of the entire setup.

## 4. Steered Molecular Dynamics

The SMD (3), is particularly useful for examining select pathways and mechanisms between two equilibrium states. It is based on the following notions. Consider a thermodynamic process that changes the system by means of altering a control parameter  $\xi$ , which is changed from  $\xi_0$  to  $\xi_t$  over a time interval  $t$ . Because of the second law of thermodynamics, the average work  $W$  performed on the system cannot be less than the free energy difference:

$$\Delta F = f(\xi_t) - f(\xi_0) \leq W, \quad (10)$$

with the equality holding if the process is reversible. Thus, a non-equilibrium process can only provide us with an upper bound for  $\Delta F$ . However, the Jarzynski (29) equality

$$\exp\left(-\frac{\Delta F}{k_B T}\right) = \left\langle \exp\left(-\frac{W}{k_B T}\right) \right\rangle \quad (11)$$

holds even when the transformation is irreversible (the angular brackets denote an average over all possible realizations of an external process that takes the system from its initial equilibrium state to a nonequilibrium state under the same external conditions as that of the final equilibrium state). Thus, in principle,  $\Delta F$  can be estimated using *N*finite-timenonequilibrium simulations:

$$\Delta F \approx -k_B T \ln \left[ \frac{1}{N} \sum_{i=1}^N \exp\left(-\frac{W_n}{k_B T}\right) \right]. \quad (12)$$

Here,  $W_n$  represents the work obtained for the  $n$ <sup>th</sup> realization of the process. Unfortunately, because the average of the exponential function is dominated by rare realizations, convergence is often poor unless  $N$  is very large. It is, however, possible to improve the estimates of the calculated free energy differences using modifications based on the maximum likelihood estimator (MLE) and Bennett's acceptance ratio (BAR) method, which are equivalent (30–32). These methods make use of both forward (F) and reverse (R) nonequilibrium pulling processes.

In practical implementations of the SMD method, a restraining harmonic potential is added to the potential energy:

$$U_{\text{SMD}}(\xi) = \frac{K}{2} [\sigma(\mathbf{r}_1, \dots, \mathbf{r}_N) - \xi]^2, \quad (13)$$

so that the system can be “steered” toward the states with the desired value of the collective variable  $\sigma(\mathbf{r}_1, \dots, \mathbf{r}_N)$  by “pulling” on the ends of the spring  $\xi(t)$ . This harmonic potential should be evaluated in the stiff-spring limit, with  $K \gtrsim U_{\text{max}}/(\Delta\xi)^2$ , where  $U_{\text{max}}$  is the highest PMF barrier to be explored and  $\Delta\xi$  sets the desired precision in  $\xi$ . This stiff-spring approximation ensures that the PMF as a function of  $\xi$  of the unbiased original system is well represented by the free energy of the system with the harmonic restraint. The latter can be conveniently thought of as an approximation of the delta function  $\Delta(\xi - \sigma(\mathbf{r}_1, \dots, \mathbf{r}_N))$  in the partition function. Generally, a number of nonequilibrium “pulling” and “pushing” simulations are then carried out by changing the values of  $\xi$  over a prescribed pathway. The numerical values of the nonequilibrium work for the forward ( $W^F$ ) and reverse ( $W^R$ ) pullings are collected, and  $\Delta F$  is estimated from the transcendental equation (30–32):

$$\sum_{i=1}^{n_F} \frac{1}{1 + n_F/n_R \exp(W_i^F - \Delta F)} - \sum_{i=1}^{n_R} \frac{1}{1 + n_R/n_F \exp(W_i^R + \Delta F)} = 0, \quad (14)$$

with  $n_{F, R}$  denoting the numbers of forward and reverse simulations, respectively. As expected, we find that  $\Delta F$  calculated with this two-sided method displays an improved convergence over estimates obtained with the straightforward one-sided application of the Jarzynski equality.

The SMD method is especially easy to implement with one-dimensional collective variables: one simply repeats the same run with different initial configurations and/or random seeds, and collects the work. The Jarzynski equality (or MLE-based formula if both forward and reverse runs are performed) is then used to estimate the free energy difference. Turning to systems described by more than one collective variable, the formalism additionally

requires one to average over all possible paths that join the points  $\xi(0)$  and  $\xi(t)$ . This is rather difficult to achieve for larger systems, and therefore, only selected pathways are typically examined in a practical setting. For these systems, SMD can be also used as a tool to examine and compare pathways associated with different transition mechanisms in a *qualitative* way. One can always “handpick” pathways that appear to be the most interesting. Particular attention is typically given to the ones that require the least amount of work, since these are mostly to dominate the sum for  $\Delta F$  in the Jarzynski equality. Another approach for investigating mechanisms is to steer the system along selected pathways over the two-dimensional free energy landscapes, which have been selected by other means such as the least free energy path (LFEP) method (33).

## 5. Examples

Having described the free energy methods and approaches in the previous sections, we illustrate their utility by considering two prototypical examples, involving the puckering of a select sugar molecule, and the conformational properties of several proline oligopeptides. All the computations were performed using the AMBER 10 (24) simulation package.

### 5.1. Puckering of $\beta$ -D-Fructose

Here, we consider the calculation of the free energy landscape which characterizes the puckering of  $\beta$ -D-fructofuranose (fructose), shown in Fig. 2.

As per instruction figures should be in black and white but color ‘grey’ is mentioned in Fig. 2. Please check.

Computational studies of carbohydrates have evolved along several directions. Considerable work has been devoted to quantum chemical studies of the energetics of relatively small compounds (see, e.g., (34) and references therein). While very accurate, such *ab initio* calculations cannot typically handle explicit solvent over nanosecond time scale. Yet, it is known that the solvation plays a crucial role in the conformational preferences of carbohydrates (35, 36). Until recently, only coarse-grained models were able to breach the long time scales involved in the conformational sampling.

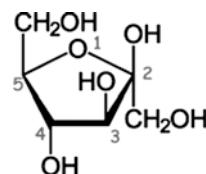


Fig. 2. Structure of the  $\beta$ -D-fructofuranose molecule. The atoms on the ring are numbered in grey, and used to compute the puckering coordinates as discussed.

However, methodological advances coupled to growing computational power have allowed the realization of all-atom, explicit solvent classical molecular dynamics over hundreds of nanoseconds (37). Naturally, there is a growing interest for development and improvement of the force fields (38). An accurate determination of the properties of the compounds in explicit solvent (39) is indispensable for the validation of such model potentials, and also interesting in its own right.

### 5.1.1. Collective Variables

The puckering states of a five-membered ring can be conveniently described in terms of just two numbers  $q_c$  and  $q_s$  (40) defined as follows. Consider the positions of the atoms of the ring  $\mathbf{R}_1, \dots, \mathbf{R}_5$  translated to their geometrical center so that

$$\mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3 + \mathbf{R}_4 + \mathbf{R}_5 = 0.$$

One computes the vector  $\mathbf{n}$  normal to the plane of the ring:

$$\begin{aligned}\mathbf{R}_c &= \sum_{a=1}^5 \mathbf{R}_a \cos \left[ \frac{2\pi(a-1)}{5} \right], \\ \mathbf{R}_s &= \sum_{a=1}^5 \mathbf{R}_a \sin \left[ \frac{2\pi(a-1)}{5} \right], \\ \mathbf{n} &= \frac{\mathbf{R}_s \times \mathbf{R}_c}{|\mathbf{R}_s \times \mathbf{R}_c|},\end{aligned}$$

and uses it to calculate the displacements  $z_a = \mathbf{n} \cdot \mathbf{R}_a$ ,  $a = 1, \dots, 5$  for the vertices of the ring out of the plane. The displacements  $z_a$  are subject to three constraints (40), so that for the five-membered ring, only two independent combinations can be constructed. The puckering coordinates ( $q_c$ ,  $q_s$ ) used in this study are then defined as

$$q_c = \sqrt{\frac{2}{5}} \sum_{a=1}^5 z_a \cos \left[ \frac{4\pi(a-1)}{5} \right] \quad (15)$$

$$q_s = \sqrt{\frac{2}{5}} \sum_{a=1}^5 z_a \sin \left[ \frac{4\pi(a-1)}{5} \right] \quad (16)$$

It is customary (40) to use the puckering amplitude  $Q$  and phase angle  $\varphi$  instead of the coordinates ( $q_c$ ,  $q_s$ ) (i.e.,  $q_c = Q\cos\varphi$  and  $q_s = Q\sin\varphi$ ). We chose to retain the coordinate representation because it is somewhat easier to deal with numerically.

### 5.1.2. Simulation Details

We used the GLYCAM 06 (38) parameters to model the fructose molecule along with the TIP3P (41) water model for explicit solvent simulations and the GB/SA (42–45) approximation for the implicit solvent simulations. The integrator time step was 1 fs. The Langevin thermostat (46) with collision frequency  $\gamma = 1 \text{ ps}^{-1}$  was used to keep the temperature at 300 K. The lengths of all bonds containing

hydrogen were fixed via the SHAKE algorithm with the tolerance set to  $10^{-6}$ . The GB/SA simulations were implemented using no cutoff on the nonbonded (van der Waals and electrostatics) interactions. For the explicit water simulations, we used periodic boundary conditions with fixed box size (30.5 Å in all three directions) chosen to render the density equal to  $\approx 1 \text{ g/cm}^3$ . The particle mesh Ewald (PME) (47, 48) method was used with the direct space cutoff set to 9 Å and a  $32 \times 32 \times 32$  grid for the smooth part of the Ewald sum. In all instances, the systems were thoroughly equilibrated, and new random number generator seeds were set for every run.

### 5.1.3. Calculation of the Free Energy Landscape

Simulations were started in implicit GB/SA solvent with “coarse” flooding given by  $4\Delta q_s = 4\Delta q_c = 0.12 \text{ \AA}$  and  $\tau = 1 \text{ ps}$  for 1 ns. This short run turned out to be enough to cover the relevant region in the  $(q_c, q_s)$  space as shown in Fig. 3. Although the nonequilibrium nature of the run influences the accuracy of the resulting free energy in a rather complicated way, it is expected that the closer to equilibrium the system is, the more accurately the real, equilibrium free energy is approximated by the slowly growing biasing potential. Therefore, we carried out additional 5 ns long run using the same “spatial” resolution ( $\Delta q_{c,s}$ ) but flooding the system at a slower pace with  $\tau = 100 \text{ ps}$ . Again, it turned out that such run was long enough to repeatedly visit all the relevant points of the phase space, thereby correcting for the inaccuracies associated with the initial run.

In a typical ABMD calculation, the biasing potential “grown” in vacuum or with an implicit solvent generally becomes the starting point for considerably more demanding explicit solvent simulations. One generally continues with a “fine” flooding under explicit solvent and then uses these results for the corrective equilibrium sampling. For this particular system, however, we found only very minor differences between the implicit and explicit solvent results. We could therefore safely skip the flooding under the explicit solvent step and proceed directly with 100 ns long biased equilibrium run in

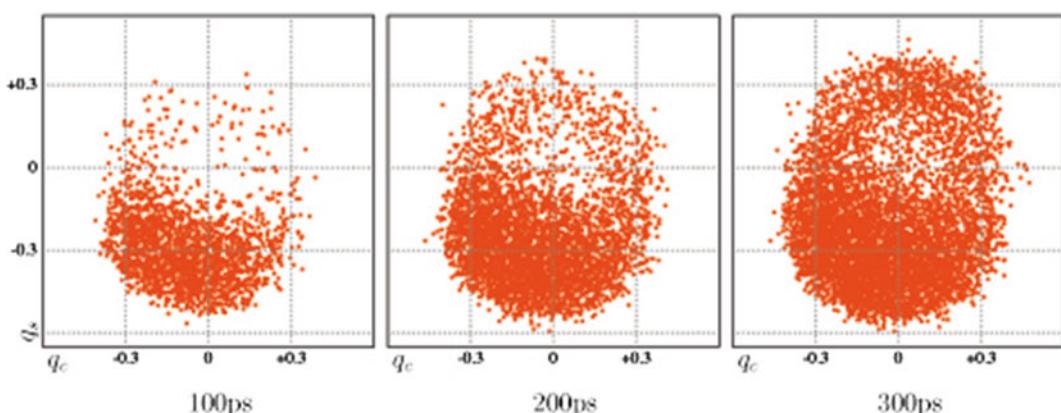


Fig. 3. Snapshots marking the progression of the ABMD simulation in the  $(q_c, q_s)$  plane.

explicit water. We collected the samples of ( $q_c$ ,  $q_s$ ) every 0.1 ps. While this is way too often for the biased density reconstruction, it allowed us to accurately compute the autocorrelation functions  $q_c$ ,  $q_s(0)q_{c,s}(t)$  of the puckering coordinates. The autocorrelations, as expected, decay exponentially, with a characteristic time of 10 ps for water (1 ps for GB/SA, 100 ps for ethanol). With these results, we used samples separated by 10 ps to reconstruct the biased *equilibrium* probability density as described in (18):

$$p^B(q_c, q_s) = \frac{1}{N_S h_c h_s} \sum_{l=1}^{N_S} K\left(\frac{q_c - q_c^l}{h_c}\right) K\left(\frac{q_s - q_s^l}{h_s}\right), \quad (17)$$

where  $N_S = 10,000$  is the number of samples ( $q_c^l$ ,  $q_s^l$ ),  $l = 1, \dots, N_S$ ,  $h_{c,s}$  are the window widths, and  $K(\xi)$  are the kernels. We used Gaussian kernels

$$K(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\xi^2\right] \quad (18)$$

for both  $q_c$  and  $q_s$ . The smoothing widths  $h_{c,s}$  along with the number of samples determine the precision of the resulting estimator. It is important to understand that the density is an infinite-dimensional object and hence cannot be reconstructed from any finite number of samples; the nonzero widths  $h_{c,s}$  effectively “discretize” the density, i.e., they set the scale below which the density remains unknown. To estimate the widths, we followed Silverman’s guidelines (p. 56, (18)): we examined the plots of the Laplacian of the density for several decreasing values of the widths “by eye”; below some “optimal” choice of  $h_{c,s}$  the Laplacian of the reconstructed density becomes noisy; values of  $h_{c,s}$  slightly larger are then deemed to be “optimal.” In our particular case, we settled at  $h_c = h_s = 0.1\text{\AA}$ . Having the  $h_{c,s}$ , we computed the estimator using the first 5,000 samples and the last 5,000 samples and compared these two with the one based on all 10,000 samples: the absolute difference in the region  $q_c^2 + q_s^2 < (3/5)^2$  did not exceed 0.1; this translates into an *at most* 0.1 kcal/mol difference between the corrections computed using the whole sample and either of its halves. We therefore conclude that the final free energy plot that is shown in Fig. 4 is accurate to at least 0.1 kcal/mol. The global minimum in the Fig. 4 is at the pseudo-rotation angle  $\varphi = \tan^{-1} q_s/q_c \approx 250^\circ$  which corresponds to the E<sub>3</sub> (C3'-*exo*) conformation.

## 5.2. Proline Oligopeptides

Proline is the only natural amino acid in which the side chain is cyclized to the backbone, setting the backbone dihedral angle to  $\phi = -75^\circ$ . This restricts proline’s conformational freedom considerably, as compared to other amino acids. Polyproline peptides exhibit two well-characterized helical structures: (a) a left-handed polyproline helix (PPII), with all peptide bonds in the trans-isomer conformation (i.e., backbone dihedral angle  $\omega = 180^\circ$ ), and (b) a more compact right-handed polyproline helix (PPI), with all the

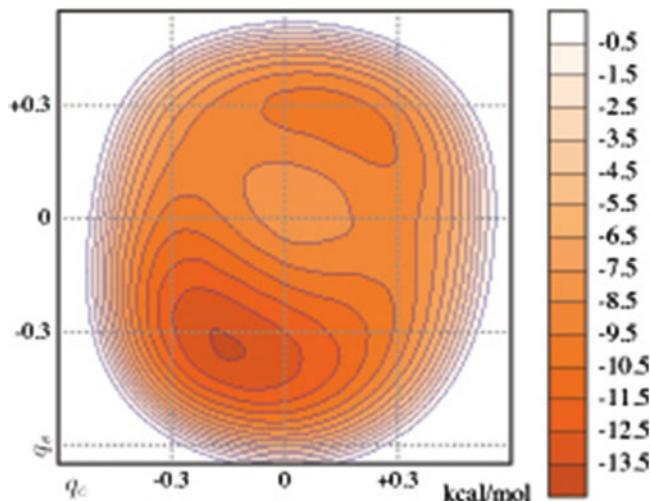


Fig. 4. Free energy landscape of  $\beta$ -D-fructofuranose in terms of the puckering coordinates ( $q_c$ ,  $q_s$ ) solvated by 920 TIP3P water molecules at  $T = 300$  K.

peptide bonds in the cis-isomer conformation (i.e., backbone dihedral angle  $\omega = 0^\circ$ ). Of the 20 natural amino acids, only proline is “comfortable” in the cis-isomer conformation. In addition, the cis-trans isomerization of X-Pro peptide groups is one of the rate-determining steps for folding and unfolding of various proteins (49–53). The probability distribution of cis/trans prolyl bonds is affected by neighboring amino acids (54, 55), pH and ionic strength (56), solvent (57–63), and chain length (64). Experimentally, it is known that the left-handed PPII structure is favored in water, benzyl alcohol, trifluoroethanol, organic acids, and most of the other solvents, while the PPI structure is favored in the presence of aliphatic alcohols like propanol (64–67).

Traditionally, the relative rigid structure of PPII has been used as a “molecular ruler” in structural biology, especially for the validation of spectroscopic rulers in Föster resonance energy transfer (FRET) experiments (68). However, recent studies have called this traditional picture of polyproline peptide as a rigid rod into question. (69–75). Specifically, recent calculations (4–6) based on the methodology outlined in this chapter indicate that pure PPI and PPII are just one of the possible minima of the phase space, with many other minima being associated with other stable or metastable structures which are characterized in terms of the cis or trans nature of the prolyl bonds. These results complement the PET experimental results (74), with regard to the existence of stable heterogeneous subpopulation of polyproline conformers. More importantly, by providing detailed free energy landscapes in terms of different collective variables it has become possible to characterize the probability distribution of the different isomers, along with the associated most probable transition pathways and mechanisms.

### 5.2.1. Simulation Details

From a computational point of view, the characterization of the conformations of proline-based peptides is rather difficult, since the *cis/trans* isomerization is much slower than the formation and rupture of hydrogen bonds in an  $\alpha$ -helix coil transition. Indeed, characteristic time scales for prolyl isomerization range in tens to hundreds of seconds at room temperature (54, 76), leading to an energy barrier of the order of 10–20 kcal/mol (77–83). Due to these huge energy barriers, traditional MD simulations cannot explore the relevant conformational space and the whole range of *cis/trans* transitions. Hence, we made use of ABMD, HT-REMD, and SMD methods to explore the free energy landscapes of proline systems (4–6, 84). In particular, to exemplify the methodology, we discuss simulations of the polyproline pentamer (i.e., peptide *Ace*–(*Pro*)<sub>*n*</sub>–*Nme* with *n* = 5 denoted as *P*<sub>5</sub>) in an implicit water environment based on the Generalized Born approximation (85, 86), as well as simulations a corresponding polyproline nanomer (denoted *P*<sub>9</sub>) in vacuo and in explicit hexane and 1-propanol.

Initial configurations in all cases consisted of the unfolded peptides, which were generated using LEAP. The *P*<sub>5</sub> simulations made use of the ff99SB version of the Cornell et al. force field (87), while *P*<sub>9</sub> simulations used ff99 (88). The integrator time step was set to 1 fs, and use was made of the Langevin thermostat (46) with a collision frequency of  $\gamma = 1 \text{ ps}^{-1}$ . The lengths of all bonds containing hydrogen atoms were fixed using the SHAKE algorithm, with the tolerance set to  $10^{-5}$ . There was no cutoff associated with the nonbonded interactions in vacuo and implicit water simulations. For the *P*<sub>9</sub> explicit solvents simulations, the peptide was bathed separately in 426 molecules of the nonpolar solvent hexane, and in 912 molecules of 1-propanol. The explicit solvent calculations were carried out in a truncated octahedron cell using periodic boundary conditions, with a box length of 79.800 Å in all directions. The PME method (47, 48) was used with the cubic spline approximation, a cutoff of 8.0 Å, a nonbonded “skin” of 1.0 Å, a direct sum tolerance of  $10^{-5}$ , and a grid size of 64 in each of the three-dimension.

### 5.2.2. Collective Variables

It is convenient to cast the collective variables of the proline systems in terms of changes associated with the *cis/trans* isomerization of the  $\omega$  prolyl bonds (see Fig. 5), which also happens to be the most relevant slow mode of the system. Since  $\omega$  varies between  $\omega = 0^\circ$  (*cis*) and  $\omega = 180^\circ$  (*trans*) one can “capture” different conformations of proline peptides with the following collective variable:

$$\Omega = \sum_{i=1}^n \cos \omega_n, \quad (19)$$

where the sum runs over all prolyl bonds. Clearly, for a proline-rich peptide with *n* prolyl bonds,  $\Omega$  takes on the values –*n*, –*n* + 2, ..., *n*, and describes the net balance of the *cis/trans* states.

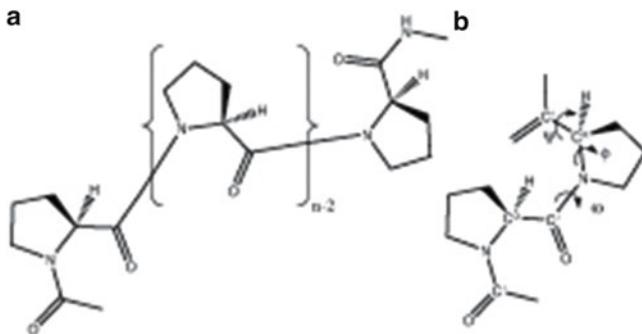


Fig. 5. (a) Schematic two-dimensional representation of Ace-(Pro)<sub>n</sub>-Nme; (b) Definition of backbone dihedral angles  $\phi$ ,  $\psi$ , and  $\omega$ .

Although very useful, there are still a considerable number of degeneracies associated with this collective variable. To remove the most relevant ones (i.e., the ones associated with the prolyl bonds), we define additional collective variables that may be used in conjunction with  $\Omega$ . For instance, one can consider the “interface” between bonds and define a collective variable  $\Lambda$  as

$$\Lambda = \sum_{i=1}^{n-1} \cos(\omega_i + \omega_{i+1}). \quad (20)$$

For a peptide with  $n$  prolyl bonds, there are  $n - 1$  interfaces. If two neighboring bonds have the same dihedral angle  $\omega$ , then their interface is  $\cos(\omega_i + \omega_{i+1}) = +1$ ; otherwise, it is  $-1$ . Correspondingly,  $\Lambda$  can take on any of the values  $-n + 1, -n + 3, \dots, n - 1$ . For  $|\Omega| = n$ , there is only one value of  $\Lambda$  ( $\Lambda = n - 1$ ); for  $\Omega = 0$  (that exists only for even values of  $n$ ), there are  $n - 1$  values of  $\Lambda$ ; for  $|\Omega| = m$  (with  $m \neq n$  and  $m \neq 0$ ), there are  $n - m$  values of  $\Lambda$ . Thus, a typical free energy landscape in the  $(\Omega, \Lambda)$  plane has a triangular shape. Note that larger the numerical value of  $\Lambda$ , the more the structure corresponds to the ideal PPI or PPII helix, since the number of cis/trans interfaces is minimized.

### 5.2.3. Free Energy Maps

In this section, we discuss the free energy results obtained from the ABMD runs for the selected examples. Specifically, one-dimensional free energy profile based on  $\Omega$  for  $P_5$  and two-dimensional free energy maps as a function of the collective variables  $(\Omega, \Lambda)$  for  $P_9$  in different environments will be briefly discussed.

ABMD simulations for  $P_5$  in implicit water ( $P_9$  in vacuo) were carried out using 20 (24) replicas. Initial calculations of the free energy landscapes involved all the replicas at the same temperature  $T = 1, 200$  K without any exchanges between the replicas. These short and inaccurate multiple-walker ABMD runs of 10 ns duration provided us with a rough estimate of the biasing potentials to be encountered during the next phase of the runs. The second

step in the simulation protocol involved long ABMD runs with replica exchange, using the results from the first step. However, this time, the temperature of all the replicas was varied from 300 to 1,200 K. Specifically, these were set to 300, 322, 347, 373, 401, 432, 464, 499, 537, 578, 622, 669, 720, 774, 833, 896, 964, 1, 037, 1, 115, and 1, 200 K for  $P_5$  in implicit water and 300, 318, 338, 359, 381, 405, 430, 457, 485, 516, 548, 582, 618, 656, 697, 740, 786, 835, 887, 942, 1, 001, 1, 063, 1, 129, and 1, 200 K for  $P_9$  in vacuo. Each replica had its own biasing potential. Note that the temperature distribution between the lowest and highest values was chosen to be exponential, which usually leads to an optimal mixing (i.e., equal rates of exchange between the neighboring replicas). Although the number of replicas was somewhat influenced by peculiarities of our computer setup, the highest temperature chosen needs to be sufficiently high as to ensure flexibility in moving between different conformations (i.e., a small autocorrelation time). At the same time, because it determines the entire temperature distribution of the replicas, it needs to be small enough as to give a high rate of exchange between the replicas (i.e., good mixing). Both conditions need to be satisfied in order to ensure that the free energy results converge to the desired accuracy within a feasible runtime. Turning to the explicit solvent simulation, these were performed using the multiple-walker extension only. This was because test runs showed that parallel tempering runs were characterized by a very low measured rate of exchange between different replicas and therefore could not be considered useful.

In terms of the ABMD parameters, for the one-dimensional  $P_5$  simulations, a kernel width of  $4\Delta\xi = 0.2$  was used along with a flooding time scale of  $\tau_F = 5.0, 25.0, 100.0$  ps with a runtime of 10 ns for each of the different  $\tau_F$  simulations. For the two-dimensional free energy maps, a kernel width of  $4\Delta\xi = 0.5$  was used for both collective variables. The flooding time scales were  $\tau_F = 1.0, 2.0, 5.0$  ps, with each simulation lasting 20 ns. We note that the initial flooding time scale was chosen to be small enough as to generate rough estimates of the biasing potential in a short period of time. The larger flooding time scales were used to obtain more accurate results. However, the price one pays for this accuracy is that these simulations take considerably longer to reach the desired accuracy. Finally, we repeated the simulations using the calculated biasing potential for follow-up equilibrium umbrella sampling runs for another 30 ns. The kernel estimator method was then used to recover the biased probability density  $p^B(\xi)$  (18).

Figure 6 shows the resulting free energy profile for  $P_5$  in implicit water before (red) and after (black) the corrective runs. The difference is always smaller than 0.3 kcal/mol. The free energy difference between all-trans ( $\Omega = -5$ ) and all-cis ( $\Omega = 5$ ) structures was estimated to be  $4.9 \pm 0.3$  kcal/mol. The structures associated with  $\Omega = -3, -1, 1, 3$  have calculated free energies

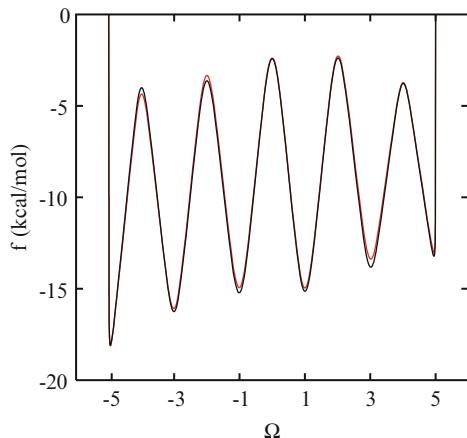


Fig. 6. Free energy profile of the polyproline pentamer peptide in implicit water as a function of  $\Omega$  at  $T = 300$  K.

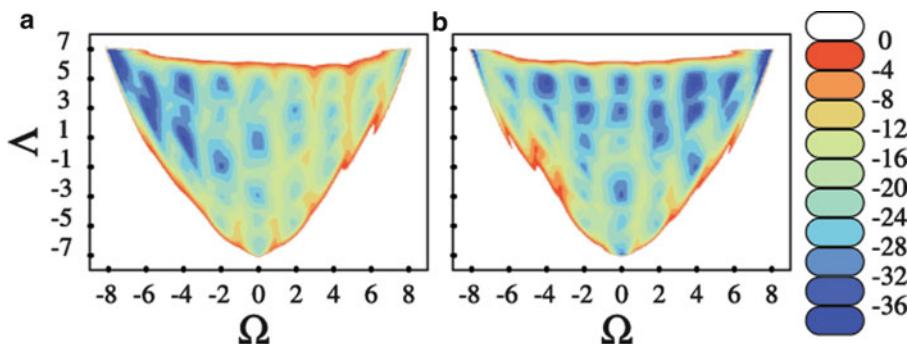


Fig. 7. Two-dimensional ( $\Omega, \Lambda$ ) free energy landscapes of the  $P_9$  peptide in explicit solvents (a) hexane and (b) propanol.

of  $-3.0, -2.0, -1.9, -0.6$  kcal/mol, relative to the all-cis conformation. Overall, the results indicate that the more cis prolyl bond a structure has the greater its free energy.

We now discuss the  $(\Omega, \Lambda)$  maps, shown in Figs. 7 and 8a. For these results, we used the  $n=9$  prolyl bonds only to define the  $(\Omega, \Lambda)$  collective variables, so that  $(-8, 7)$  and  $(8, 7)$  correspond to the PPII and PPI structures, respectively. Associated with  $\Omega = -6$  are two values of  $\Lambda$ : the  $\Lambda = 5$  structure involves a terminal cis bond, while  $\Lambda = 3$  involves an inner cis bond. For  $\Omega = 0$ , there are  $n - 2 = 6$  values of  $\Lambda$ , with the lowest  $\Lambda = -7$  corresponding to a situation where all prolyl bonds alternate with respect to each other. Overall, the free energy minima on these maps are distributed in a triangular or wedge-like pattern.

The left (right) side of the triangle represents the trans (cis)-rich conformations. It is clear that the solvent has a great influence on the distribution of the conformers. Figure 7a shows the free energy map associated with a nanomer in the nonpolar solvent hexane.

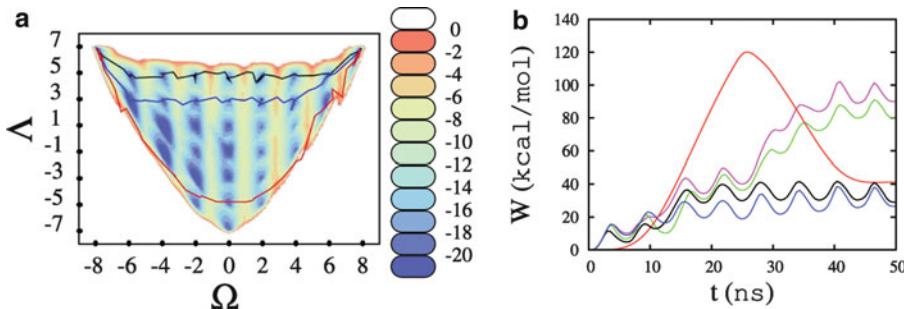


Fig. 8. (a) The  $(\Omega, \Delta)$  free energy landscapes in kcal/mol of a polyproline nanomer *in vacuo*, along with selected SMD pathways over these free energy landscapes. Here, the black, blue and red lines correspond to ZR, ZI0 and “en-masse” pathways, respectively. (b) Here, we show the work ( $W$ ) (required for a typical  $PPII \rightarrow PPI$  transition) versus time for a nanomer *in vacuo*. Different mechanisms are characterized by different colors: *black* (ZR), *blue* (ZI0), *pink* (ZF), *green* (ZOI) and *red* (“en-masse”), respectively.

Here, the trans-rich structures dominate and the ideal PPII structure represents the global minimum with a free energy of about 20 kcal/mol lower than that of the ideal PPI structure. In the solvent 1-propanol, the proline nanomer has a very different behavior, and it is the cis-rich structures that are favored (Fig. 7b). In particular, the free energy difference between PPI and PPII was found to be  $\approx -8$  kcal/mol. We also note, that in general, the larger values of  $\Delta$  are usually associated with a lower free energy. This trend is almost reversed for the proline nanomer *in vacuo*, as shown in Fig. 8a. The deeper minima are associated with lower  $\Delta$  values, and neither PPII nor PPI represents the global minimum. The latter is associated with some of the more compact globular structures.

#### 5.2.4. SMD Results

All the SMD simulations were carried out at  $T = 300$  K, with a harmonic constant set to 100 kcal/mol. SMD simulations were used for two different purposes:

1. For a quantitative estimation of the free energy difference between the PPI and PPII structures for short peptides implicit water. Specifically, we discuss here results for the  $P_5$  peptide.
2. For a qualitative comparison of the different mechanisms associated with the  $PPII \rightarrow PPI$  transitions. Specifically, we discuss here results for  $P_9$  *in vacuo*.

To estimate the free energy difference between PPII and PPI for  $P_5$  using both the Jarzynski equality and BAR method, we carried out 100 forward and reverse nonequilibrium simulations by pulling on the  $\Omega$  variable. Each of these simulations involved an initial equilibration run lasting 0.5 ns, followed by a 2.0 ns SMD run during which the  $\Omega$  variable was pulled at a constant velocity. The simulations were carried out at  $T = 300$  K, with the harmonic constant set to 100 kcal/mol.

In terms of the convergence, the results based on the Jarzynski equality turn out not to be particularly accurate. However, the results based on 100 runs using the MLE method yield a free energy difference between PPI and PPII of 4.6 kcal/mol, which is in good agreement with the 4.9 kcal/mol estimate obtained from the ABMD runs.

For a qualitative comparison of the different transition mechanisms associated with  $\text{PPII} \rightarrow \text{PPI}$  transition of  $\text{P}_9$  in vacuo, we steered the system via the successive flipping of the  $\omega$  dihedral angles. Since it is not possible to exhaustively examine all possible transition pathways and mechanisms, we concentrated on a select group of mechanisms. Specifically, we focused on different zipper-like mechanisms, for which changes take place via the successive switching of neighboring prolyl bonds. These include: a zipper-like mechanism starting from the forward, acetylated end for the  $\text{PPII} \rightarrow \text{PPI}$  transition which we label ZF (zipper-forward), such that the intermediate conformation in either direction would take the form  $\text{Ace} - \text{CCC} \cdots \text{TTT} - \text{Nme}$ ; a zipper-like mechanism starting from the reverse, amidated end for the  $\text{PPII} \rightarrow \text{PPI}$  transition which we label ZR (zipper-reverse), such that the intermediate conformation would take the form  $\text{Ace} - \text{TTT} \cdots \text{CCC} - \text{Nme}$ ; a zipper-like mechanism starting from the center and moving outward for the  $\text{PPII} \rightarrow \text{PPI}$  transition which we label ZIO (zipper-inside-outwards), such that the intermediate conformation would take the form  $\text{Ace} - \text{TT} \cdots \text{CCC} \cdots \text{TT} - \text{Nme}$ ; a zipper-like mechanism starting from both ends, and working its way inward for the  $\text{PPII} \rightarrow \text{PPI}$  transition which we label ZOI (Zipper-Outside-Inwards), such that the intermediate conformation would take the form  $\text{Ace} - \text{CC} \cdots \text{TTT} \cdots \text{CC} - \text{Nme}$ . Additionally, we considered mechanisms that involve changing the prolyl bonds in an “en-masse” mechanism where all the bonds are changed simultaneously. Clearly, there is some ambiguity in how the ZIO and ZOI mechanisms can be implemented. For example, one can flip either two bonds at a time *simultaneously*, or only one at a time. For the latter, the order of bond flipping provides for a further degree of freedom. In our simulations the alternative flipping of bonds was used.

Figure 8a plots the pathways associated with some of these mechanisms in the  $(\Omega, \Lambda)$  space. Specifically, the black, blue, and red lines mark pathways taken by the ZR, ZIO, and “en-masse” mechanisms. From this map, it is clear that ZR mechanism always chooses the largest value of  $\Lambda$ , while for ZIO, the path is always one minimum below the ZR path, except for the end points. Although not shown here, ZF (ZOI) mechanism follows a path similar to ZR (ZIO) in the  $(\Omega, \Lambda)$  space, due to the symmetries associated with both of the collective variables. In an “en-masse” mechanism,  $\Lambda$  does not stay close to its numerically largest value. The work performed by the system for these simulations is shown in

Fig. 8b. In contrast to the many short duration SMD runs performed for estimating the free energy differences, the results here are based on a few long runtime simulations lasting 50 ns.

In terms of the most to least amount of work required for the transition, the ordering is ZF, ZOI, “en-masse,” ZR, and ZIO. It is also of interest to figure out how the transition is likely to be initiated. For the *PPII* → *PPI* transition, the flipping at the amidated end and/or middle bonds appear to be good options. To probe this question from a different perspective, we performed SMD simulations for a *PPII*  $P_9$ , in order to calculate the amount of work required to flip each of the bonds from trans to cis. The results indicate that the prolyl bonds at the acetylated and amidated ends require the least amount of work. However, the next preferred ones are not their neighbors. Rather, it is the middle bonds and their neighbors that require the least amount of work for flipping. Thus, it appears that while the first bond switch might well take place at either end, the subsequent bond switches favor a ZIO mechanism. This behavior suggests ZIO mechanism can be an appropriate mechanism for *PPII* → *PPI* transition, while ZOI is very unlikely. Turning back to Fig. 8b, it is clear that ZR mechanism requires less work than ZF, indicating that the intermediate structures would take the form of *Ace* – *TT* · · · *CC* – *Nme* rather than *Ace* – *CC* · · · *TT* – *Nme*. This is consistent with the results from enzymatic hydrolysis kinetics experiments on polyproline in aqueous solution (89, 90), but opposite to the results from the NMR experiments on polyproline in  $D_2O$  solution (91).

#### 5.2.5. Enhanced Equilibrium Sampling

In order to exemplify a population analysis carried out with REMD, we discuss here the cis versus trans population distribution for  $P_5$  in implicit water. For such an analysis, the one-dimensional  $\Omega$  free energy maps as obtained prior to the “umbrella correction” step (see Fig. 6) formed the basis of the HT-REMD runs. Initially, we considered running only temperature REMD. However, in assessing the mixing properties of the hottest replica at the highest temperature of  $T = 1,200$  K (by simulating it alone), we found that the temperature (without the biasing potential) was not enough to induce frequent cis/trans transitions. In other words, the hottest replica did not show enough flexibility and its autocorrelation time turned out to be too large (tens of nanoseconds). We also attempted to use biased replicas alone, all at the lowest temperature of  $T = 300$  K. But again, the autocorrelation time turned out to be large. Therefore, we were forced to combine both features in our REMD simulations. We used the highest temperature replica with the same biasing potential as obtained from ABMD in order to ensure that (a) all possible cis/trans states get visited not in more than a few hundred picoseconds; and (b) the autocorrelation time is reasonably small (less than 2 ns) so as to produce a reasonable number of independent samples in a feasible runtime.

Four more replicas were added all at the lowest temperature  $T = 300$  K, one with no biasing potential and three with the ABMD-resulted biasing potential of the lowest temperature replica, scaled-down by the factors of 0. 49, 0. 76, 0. 9. The temperatures, the scaling factors, and the ratio of the temperature-varying versus Hamiltonian-varying replicas (20 vs. 4) were chosen in a way as to ensure a similar rate of exchange between all neighboring replicas. We ran 100 ns HT-REMD simulations, and collected the trajectory of the unbiased replica every 1 ps that resulted in  $10^5$  equilibrium samples. We note that the rest of the simulation parameters were identical to those used for the ABMD runs.

Concerning the cis/trans pattern, the all-trans structure proved to be the most probable conformation by a population of 55%, followed by structures with only one cis bond that form 28% of the whole population, combined. Labelling the proline residues  $P_1$ ,  $P_2$  etc from the *Ace* – end of the peptide, we found that structures with only a single cis bond at *Ace* –  $P_1$ ,  $P_5$  – *Nme*,  $P_1$  –  $P_2$ ,  $P_2$  –  $P_3$ , or  $P_4$  –  $P_5$  have the populations of 11%, 8%, 5%, 4%, or 4%, respectively. The rest of the structures with more than one cis bond form only 17% of the entire population. A residue-based analysis of the polyproline structure, as opposed to a sequence-

**Table 1**  
**Probability (as a percentage) of prolines in the trans (as opposed to cis) conformation ( $P(T)$ ), to be in the F (as opposed to ) region of the Ramachandran plot ( $P(F)$ ) for  $P_5$  in implicit water as obtained bu HT-REMD simulations. Conditional probabilities (as a percentage) of being in F conditioned on being either trans or cis are denoted as  $P(F|T)$  and  $P(F|C)$ , respectively, are also listed. Finally, the probability (as a percentage) of being in a PPII or PPI configuration are also listed. These probabilities are all listed for the individual residues along with the average value for the peptide**

bond	$p(T)$	$p(F)$	$p(F T)$	$p(F C)$	$p(PPII)$	$p(PPI)$
Ac- $P_1$	81	98	98	97	82	18
$P_1$ - $P_2$	88	98	98	99	85	12
$P_2$ - $P_3$	90	99	98	99	88	10
$P_3$ - $P_4$	94	56	53	80	49	5
$P_4$ - $P_5$	85	2	2	1	2	0
Average	<b>88</b>	<b>71</b>	<b>70</b>	<b>75</b>	<b>62</b>	<b>9</b>

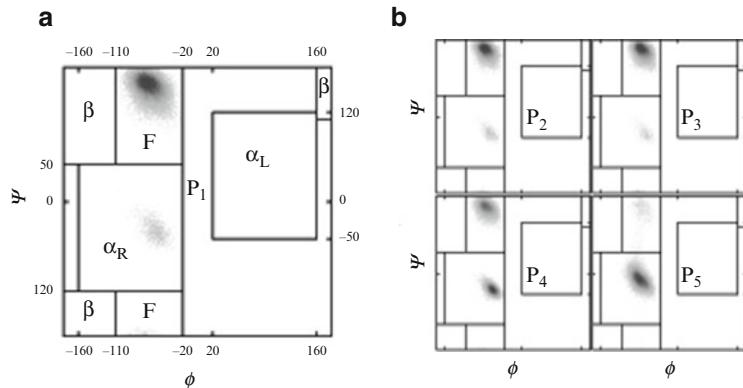


Fig. 9. Ramachandran plots of all proline residues for a polyproline pentamer peptide: (a)  $P_1$  residue results along with the definition of different regions (b) results for  $P_2, 3, 4, 5$  residues.

based analysis, is given in Table 1. In brief, one can determine the cis/trans content of each residue. The *in vacuo*  $P(T)$  results indicate that the trans content is about 88% on average, varying somewhat with the position of the residue. The highest cis content can be found at the acetylated and amidated ends by 19% and 15%, respectively.

Please provide the significance of bold in Table 1.

Having determined coordinates for a large number of equilibrium structures, one can investigate not only the cis/trans patterns but also the Ramachandran plots of each residue (5, 84). Figure 9 shows Ramachandran plots for all  $P_5$  residues. It turns out, that the dihedrals for the PPI and PPII structure fall into the F-region of the Ramachandran plots, as defined by the Zimmerman code (92). We have marked different regions of the Ramachandran plot in Fig. 9a as  $F$ ,  $\beta$ ,  $\alpha_R$ , and  $\alpha_L$ . The F region here is defined to be slightly broader than its original formulation, in order to encompass all the different fluctuations. Table 1 also gives the probability of finding the proline residue in the F region ( $p(F)$ ) and the conditional probabilities of finding the proline residue in the F region and having either a trans or cis form (denoted by  $p(F|T)$  and  $p(F|C)$ , respectively). Since an ideal PPII proline residue is in a trans conformation and also in the F region, one can find the probability of each residue being in the PPII conformation as  $p(\text{PPII}) = p(T) \times p(F|T)$ . Similarly, the same holds true for a residue in the PPI configuration so that  $p(\text{PPI}) = p(C) \times p(F|C)$ . The average PPII (PPI) content of proline pentamer is estimated as 62(9)%. We note that, the last proline residue at the amidated end is mostly found to be in the  $\alpha$ -helix region rather than F region (see Fig. 9b). This effect is also evident in the neighboring “next-to-the last” residue, particularly when it is in trans conformation.

## 6. Summary

In summary, we have briefly outlined three MD-based methods (ABMD, REMD, and SMD) useful for calculating biomolecular properties. Their workings, along with simulation details, was presented in terms of two examples drawn from sugar puckering and polyproline peptide folding. A practical implementation of these methods is currently to be found in the SANDER module of the AMBER versus10 (24) simulation package.

## Acknowledgements

This research was supported by NSF via FRG-0804549 and 1021883, and DOE. We also thank the NC State HPC for extensive computational support.

## References

- Babin V, Roland C, Sagui C (2008) Adaptively biased molecular dynamics for free energy calculations. *J Chem Phys* 128:134101
- Geyer CJ (1991) Markov chain Monte Carlo maximum likelihood. Computing Science and statistics: the 23rd symposium on the interface. Interface Foundation, Fairfax, pp 156–163
- Izrailev S, Stepaniants B, Kosztin D, Lu H, Molnar F, Wriggers W, Schulten K (1998) Steered molecular dynamics. Computational molecular dynamics: challenges, methods, ideas. Springer, Berlin, Germany
- Moradi M, Babin V, Roland C, Darden T, Sagui C (2009) Conformations and free energy landscapes of polyproline peptides. *Proc Natl Aca Sci USA* 106:20746
- Moradi M, Babin V, Roland C, Sagui C (2010) A classical molecular dynamics investigation of the free energy and structure of short polyproline conformers. *J Chem Phys* 133:125104
- Moradi M, Lee JG, Babin V, Roland C, Sagui C (2010) Free energy and structure of polyproline peptides: an ab initio and classical molecular dynamics investigation. *Int J Quant Chem* 110:2865–2879
- Frenkel D, Smit B (2002) Understanding molecular simulation. Academic Press, Orlando, USA
- Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J Comput Phys* 23:187–199
- Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo data analysis. *Phys Rev Lett* 63:1195–1198
- Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg J (1992) The weighted histogram analysis method for free energy calculation of biomolecules: I The Method. *J Comput Chem* 13:1011
- van Duijneveldt JS, Frenkel D (1992) Computer simulation study of free energy barriers in crystal nucleation. The method. *J Chem Phys* 96:4655–4668
- Lelièvre T, Rousset M, Stoltz G (2007) Computation of free energy profiles with parallel adaptive dynamics. *J Chem Phys* 126:134111
- Bussi G, Laio A, Parrinello M (2006) Equilibrium free energies from nonequilibrium metadynamics. *Phys Rev Lett* 96:090601
- Huber T, Torda AE, van Gunsteren WF (1994) Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J Comput Aided Mol Des* 8:695–708
- Wang F, Landau DP (2001) Phys Rev Lett 115:2050–2053
- Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci* 99:12562–12566

17. Iannuzzi M, Laio A, Parrinello M (2003) Efficient exploration of reactive potential energy surfaces using car-parrinello molecular dynamics. *Phys Rev Lett* 90:238302–238311
18. Silverman BW (1986) Density estimation for statistics and data analysis. Chapman and Hall, London
19. De Boor C (1978) A practical guide to splines. Springer, New York
20. Raiteri P, Laio A, Gervasio FL, Micheletti C, Parrinello M (2006) Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J Phys Chem* 110:3533–3539
21. Sugita Y, Kitao A, Okamoto Y (2000) *J Chem Phys* 113:6042
22. Bussi G, Gervasio FL, Laio A, Parrinello M (2006) Free-energy landscape for  $\beta$  hairpin folding from combined parallel tempering and metadynamics. *J Am Chem Soc* 128:13435–13441
23. Piana S, Laio A (2007) A bias-exchange approach to protein folding. *J Phys Chem B* 111:4553–4559
24. Case D et al (2008) AMBER 10. University of California, San Francisco
25. Hooft RWW, van Eijck BP, Kroon J (1992) An adaptive umbrella sampling procedure in conformational analysis using molecular dynamics and its application to glycol. *J Chem Phys* 97:6690–6694
26. Babin V, Roland C, Darden TA, Sagui C (2006) The free energy landscape of small peptides as obtained from metadynamics with umbrella sampling corrections. *J Chem Phys* 125:204909
27. Babin V, Sagui C (2010) Conformational free energies of methyl- $\alpha$ -L-iduronic and methyl- $\beta$ -D-glucronic acids in water. *J Chem Phys* 132:104108
28. Laghaei R, Mousseau N, Wei G (2010) Effect of the disulfide bond on the monomeric structure of human amylin studied by combined Hamiltonian and temperature replica exchange molecular dynamics simulations. *J Phys Chem B* 114:7071–7077
29. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693
30. Crooks GE (2000) Path-ensemble averages in systems driven far from equilibrium. *Phys Rev E* 61:2361–2366
31. Shirts MR, Bair E, Hooker G, Pande VS (2003) Equilibrium free energies from non-equilibrium measurements using maximum-likelihood methods. *Phys Rev Lett* 91:140601
32. Kosztin I, Barz B, Janosi L (2006) Calculating potentials of mean force and diffusion coefficients from nonequilibrium processes without Jarzynski's equality. *J Chem Phys* 124:064106
33. Ensing B, Laio A, Parrinello M, Klein M (2005) A recipe in the computation of the free energy barrier and the lowest free energy path of concerted reactions. *J Phys Chem B* 109:6676–6687
34. Rao VSR, Qasba PK, Balaji PV, Chandrasekaran R (1998) Conformation of carbohydrates. Harwood Academic Publishers, Amsterdam
35. Kirschner KN, Woods RJ (2001) Solvent interactions determine carbohydrate conformation. *Proc Natl Acad Sci* 98:10541–10545
36. Almond A, Sheehan JK (2003) Predicting the molecular shape of polysaccharides from dynamic interactions with water. *Glycobiology* 13:255–264
37. Kräutler V, Müller M, Hünenberger PH (2007) Conformation, dynamics, solvation and relative stabilities of selected [beta]-hexopyranoses in water: a molecular dynamics study with the gromos 45A4 force field. *Carbohydr Res* 342:2097–2124
38. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, Woods RJ (2008) GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J Comp Chem* 29:622–655
39. Christen M, van Gunsteren WF (2008) On searching in, sampling of, and dynamically moving through conformational space of biomolecular systems: a review. *J Comp Chem* 29:157–166
40. Cremer D, Pople JA (1975) General definition of ring puckering coordinates. *J Am Chem Soc* 97:1354–1358
41. Jorgensen WL, Chandrasekhar J, Madura J, Klein ML (1983) Comparison of Simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
42. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc* 112:6127–6129
43. Hawkins GD, Cramer CJ, Truhlar DG (1996) Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J Phys Chem* 100:19824–19839
44. Hawkins GD, Cramer CJ, Truhlar DG (1995) Pairwise solute descreening of solute charges from a dielectric medium. *Chem Phys Lett* 246:122–129
45. Tsui V, Case D (2001) Theory and applications of the generalized Born solvation model in

- macromolecular simulations. *Biopolymers* (Nucl Acid Sci) 56:275–291
46. Brünger A, Brooks CL, Karplus M (1984) Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem Phys Lett* 105:495–500
  47. Darden TA, York DM, Pedersen LG (1993) Particle mesh Ewald: an  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
  48. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593
  49. Brandts JF, Halvorson HR, Brennan M (1975) Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* 14:4953–4963
  50. Tanaka S, Scheraga HA (1977) Hypothesis about the mechanism of protein folding. *Macromolecules* 10:291–304
  51. Schmid FX, Mayr LM, Mücke M, Schönbrunner ER (1993) Prolyl isomerases: role in protein folding. *Adv Protein Chem* 44:25–66
  52. Houry WA, Scheraga HA (1996) Nature of the unfolded state of Ribonuclease A: effect of cis-trans X-Pro peptide bond isomerization. *Biochemistry* 35:11719–11733
  53. Wedemeyer WJ, Welker E, Scheraga HA (2002) Proline cis-trans isomerization and protein folding. *Biochemistry* 41:14637–14644
  54. Reimer U, Scherer G, Drewello M, Kruber S, Schutkowski M, Fischer G (1998) Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J Mol Biol* 279:449–460
  55. Grathwohl C, Wuthrich K (1976) The X-Pro peptide bond as an NMR probe for conformational studies of flexible linear peptides. *Biopolymers* 15:2025–2041
  56. Grathwohl C, Wuthrich K (1976) Nmr studies of the molecular conformations in the linear oligopeptides H-(L-Ala)<sub>n</sub>-L-Pro-OH. *Biopolymers* 15:2043–2057
  57. Steinberg IZ, Harrington WF, Berger A, Sela M, Katchalski E (1960) The configurational changes of poly-L-proline in solution. *J Am Chem Soc* 82:5263–5279
  58. Gornick F, Mandelkern L, Diorio AF, Roberts DE (1964) Evidence for a cooperative intramolecular transition in poly-L-proline. *J Am Chem Soc* 86:2549–2555
  59. Mandelkern L (1967) Poly-L-proline. In: Fasman GD (ed.) *Poly- $\alpha$ -amino acids*. Marcel Dekker, New York
  60. Strassmair H, Engel J, Zundel G (1969) Binding of alcohols to the peptide CO-group of poly-L-proline in the I and II conformation. I. Demonstration of the binding by infrared spectroscopy and optical rotatory dispersion. *Biopolymers* 8:237–246
  61. Tanaka S, Scheraga HA (1975) Theory of the cooperative transition between two ordered conformations of poly(L-proline). II. Molecular theory in the absence of solvent. *Macromolecules* 8:504–516
  62. Tanaka S, Scheraga HA (1975) Theory of the cooperative transition between two ordered conformations of poly(L-proline). III. Molecular theory in the presence of solvent. *Macromolecules* 8:516–521
  63. Kofron JL, Kuzmic P, Kishore V, Colon-Bonilla E, Rich DH (1991) Determination of kinetic constants for peptidyl prolyl cis-trans isomerasers by an improved spectrophotometric assay. *Biochemistry* 30:6127–6134
  64. Kakinoki S, Hirano Y, Oka M (2005) On the stability of polyproline-I and II structures of proline oligopeptides. *Poly Bull* 53:109–115
  65. Traub W, Shmueli U (1963) Structure of poly-L-proline I. *Nature* 198:1165–1166
  66. Cowan PM, McGavin S (1955) Structure of poly-L-proline. *Nature* 176:501–503
  67. Mutter M, Wohr T, Gioria S, Keller M (1999) Pseudo-prolines: induction of cis/trans-conformational interconversion by decreased transition state barriers. *Biopolymers* 51:121–128
  68. Stryer L, Haugland RP (1967) Probing polyproline structure and dynamics by photoinduced electron transfer provides evidence for deviations from a regular polyproline type II helix. *Proc Natl Acad Sci USA* 58:719–726
  69. Jacob J, Baker B, Bryant GR, Cafiso DS (1999) Distance estimates from paramagnetic enhancements of nuclear relaxatio n in linear and flexible model peptides. *Biophys J* 77:1086–1092
  70. Watkins LP, Chang H, Yang H (2006) Quantitative single-molecule conformational distributions: a case stud y with poly-(L-proline). *J Phys Chem A* 110:5191–5203
  71. Sahoo H, Roccatano D, Hennig A, Nau WM (2007) A spectroscopic ruler for short polyprolines. *J Am Chem Soc* 129:9762–9772
  72. Schuler B, Lipman EA, Steinbach PJ, Kumke M, Eaton WA (2005) Polyproline and the spectroscopic ruler revisited with single-molecule fluorescence. *Proc Natl Acad Sci USA* 102:9754–9759

73. Best RB, Merchant KA, Gopich IV, Schuler B, Bax A, Eaton WA (2007) Effect of flexibility and *cis* residues in single-molecule FRET studies of polyproline. *Proc Natl Acad Sci USA* 104:18964–18969
74. Doose S, Neuweiler H, Barsch H, Sauer M (2007) Probing polyproline structure and dynamics by photoinduced electron transfer provides evidence for deviations from a regular polyproline type II helix. *Proc Natl Acad Sci USA* 104:17400–17405
75. Dolghih E, Ortiz W, Kim S, Krueger BP, Krause JL, Roitberg AE (2009) Theoretical studies of short polyproline systems: recalibration of molecular ruler. *J Phys Chem A* 113:4639–4646
76. Grathwohl C, Wuthrich K (1981) Nmr studies of the rates of proline cis-trans isomerization in oligopeptides. *Biopolymers* 20:2623–2633
77. Phillips WD (1955) Restricted rotation in amides as evidenced by nuclear magnetic resonance. *J Phys Chem* 23:1363–1264
78. Pauling L (1960) The nature of the chemical bond, 3rd edn. Cornell Univ. Press, Ithaca, NY
79. Tonelli AE (1973) Estimate of the barriers hindering rotation about the C. $\alpha$ -C' bond between the *cis*' and *trans*' conformations in an isolated L-proline residue. *J Am Chem Soc* 95:5946–5948
80. McDonald DC, Still WC (1996) Molecular Mechanics parameters and conformational free energies of proline-containing peptides. *J Org Chem* 61:1385–1391
81. Venkatachalam CM, Price BJ, Krimm S (2004) A theoretical estimate of the energy barriers between stable conformations of the proline dimer. *Biopolymers* 14:1121–1132
82. Kang YK, Choi HY (2004) Cis-trans isomerization and puckering of proline residue. *Bioophys Chem* 111:135–142
83. Kang YK, Jhon JS, Park HS (2006) Conformational preferences of proline oligopeptides. *J Phys Chem B* 110:17645–17655
84. Moradi M, Babin V, Sagui C, Roland C (2010) A statistical analysis of the PPII propensity of amino acid guests in proline-rich peptides. *Biophysical J* 100:1083
85. Onufriev A, Bashford D, Case DA (2000) Modification of the generalized Born model suitable for macromolecules. *J Phys Chem B* 104:3712–3720
86. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins* 55:383–394
87. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65:712–725
88. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J Am Chem Soc* 117:5179–5197
89. Lin LN, Brandts JF (1979) Role of cis-trans isomerism of the peptide bond in protease specificity. Kinetic studies on small proline-containing peptides and on polyproline. *Biochemistry* 18:5037–5042
90. Lin LN, Brandts JF (1980) Kinetic mechanism for conformational transitions between poly-L-prolines I and II: a study utilizing the cis-trans specificity of a proline-specific protease. *Biochemistry* 19:3055–3059
91. Torchia DA, Bovey FA (1971) A nuclear magnetic resonance study of poly(L-proline) in aqueous and aqueous salt solutions. *Macromolecules* 4:246–251
92. Zimmerman SS, Pottle MS, Némethy G, Scheraga HA (1977) Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules* 10:1–9

# Chapter 13

## Molecular Docking Methodologies

**Andrea Bortolato, Marco Fanton, Jonathan S. Mason, and Stefano Moro**

### Abstract

Molecular docking represents an important technology for structure-based drug design. Docking is a computational technique aimed at the prediction of the most favorable ligand–target spatial configuration and an estimate of the corresponding complex free energy, although as stated at the beginning accurate scoring methods remain still elusive. In this chapter, the state of art of molecular docking methodologies and their applications in drug discovery is summarized.

**Key words:** Molecular docking, Scoring functions, Structure-based drug design

---

### 1. Molecular Docking

In the past decades we have seen an explosion of the available protein structures and in the number of real and virtual drug-like molecules. In particular the Protein Data Bank (PDB) (1) stimulated the development of computational methods in structural biology. The PDB is a repository of protein structures determined using X-ray crystallography and nuclear magnetic resonance spectroscopy. In this context structure-guided design has contributed to the discovery of a number of very important drugs, such as inhibitors of protein kinases, HIV-protease, carbonic anhydrase, neuraminidase, and thrombin. Target-based methods provide important tools to help the identification of “bioactive regions” of chemical space relevant in drug discovery projects. The challenge is that these are tiny spots not evenly distributed in an almost infinite universe limited only by the chemist’s imagination (2, 3).

Structure-based molecular modelling approaches had a difficult start, with important limitations linked to the computational power available, and the relatively few high-resolution three-dimensional (3D) structures available. In 1985 the program GRID developed by Goodford represented an important step forward in

understanding and predicting (designing) ligand interactions. The method, still widely used today, is based on using chemical probes to energetically survey a protein binding site to generate Molecular Interaction Fields (MIFs), thus identifying regions where a particular chemical functional group is favorable (4). In 1988 Kuntz et al. provided the first rigid body docking software considering only geometric complementarity, without searching for potential interactions (5). Since that time docking algorithms have remarkably improved and at the moment more than 40 docking programs and even more scoring functions are available. However despite these advances, accurate prediction of binding affinities for a diverse set of molecules turns out to be genuinely difficult, and protein flexibility still cannot be readily taken into account.

Molecular docking represents an important technology for structure-based drug design. Docking is a computational technique aimed at the prediction of the most favorable ligand–target spatial configuration and an estimate of the corresponding complex free energy, although as stated at the beginning accurate scoring methods remain still elusive. In the first step a conformational search algorithm explores the possible ligand conformations (poses) inside the protein binding pocket. In the second step a scoring function is applied to evaluate and select the most favorable pose. In many programs the two parts are coupled and the scoring function drives the ligand poses generation. Docking is often used to mine a database of compounds (or idea structures) for those most likely to be active, with a ranking of the ligand molecules by the docking score, a process usually referred to as (structure-based) virtual screening. Due to various possible errors in the docking or scoring process, a visual inspection of the “best” scoring hits and final selection is needed.

---

## 2. Methodology

### 2.1. Compound Preparation

For correct docking it is important to have both the protein binding site and the putative ligands correctly protonated, and for the ligands to sample relevant tautomers. If this is not done, it may not be possible to even dock compounds, due to steric or electrostatic clashes. Many approaches used in software for protonation and tautomeric produce irrelevant structures (e.g., double protonated pyrazines, weird tautomers etc.), but the MoKa suite from Molecular Discovery (6, 7) has been found to produce generally sensible results, with the option to generate multiple structures that would have significant abundances at pH 7 etc. (e.g., both neutral and protonated pyridines) The program BLABBER generates the ionization states at a given pH range, and TAUTHOR the tautomers with defined significant abundance.

## 2.2. Conformational Search Algorithm

Docking a ligand into a binding site needs to compute several degrees of freedom. These are the six degrees of translational and rotational freedom of one body relative to another and then the conformational degrees of freedom of the ligand and of the protein.

In the rigid docking algorithms are considered only six degrees of freedom of the small organic molecules, corresponding to translation and rotation, with both ligand and protein treated as rigid bodies. Today the standard is the semi-flexible docking, where the conformational flexibility of ligands is also taken in account while the protein is kept rigid. A systematic search of all the rotatable bonds of a drug-like molecule is not efficient from a computational viewpoint because the number of the possible combinations of the rotamers increases exponentially with the number of rotatable bonds. The search algorithms address this problem and aim to explore the conformational space of the ligands inside the protein active site in an efficient and fast fashion. In the approaches based on systematic methods the result is exactly reproducible and the conformational space is somehow reduced and simplified. An example is the incremental construction method of the docking software FlexX (8), where the conformation space of the ligand is sampled on the basis of a discrete model using a tree-search procedure for placing the ligand incrementally into the active site. More common are random or stochastic approaches like the genetic algorithm based on an evolutionary strategy (e.g., GOLD (9)) or the tabu search method able to avoid to visit repeatedly already determined ligand poses (10) (e.g., MOE-Dock (11)). Simulation methods based on molecular mechanics force fields, like molecular dynamics or energy minimization, are often combined with search algorithms (e.g., Glide (12)).

Other approaches are based on the drug design key concept of pharmacophore. The IUPAC definition of a pharmacophore is “an ensemble of steric and electronic features that is necessary to ensure the optimal supra-molecular interactions with a specific biological target and to trigger (or block) its biological response”. An important refinement of this definition is that a pharmacophore defines only a necessary, but not necessarily sufficient ensemble of features; thus many molecules that match a simple pharmacophore will not be biologically active at the desired target as they contain additional electronic or steric features that are incompatible with the target, or lack an additional required feature. Ph4Dock is based on the mapping of pharmacophoric features of the ligand on corresponding pharmacophores of the active site of the protein (13). The more recent algorithm FLAP (Fingerprints for Ligands And Proteins) exploits four-point pharmacophore fingerprints for both ligand conformations and the protein active site (14–17). Molecular interaction fields (MIFs) for the target binding pocket are calculated using the software GRID (18) and “hotspots” converted to pharmacophores. Ligands are described either by

atom-based pharmacophores (using the GRID atom types) or MIFs (e.g., for ligand–ligand similarity, pharmacophore perception, etc.). The four-point pharmacophores are able to select promising ligand binding modes that are filtered using a molecular-cavity shape derived from the protein target (three-point pharmacophores performed less well). Protein flexibility can be included in the protocol using “on-the-fly” generation of side-chain conformations while the protein site points are being generated or by using multiple protein conformations. FLAP in protein cavities is able to work in an automatic fashion from protein importation, GRID-parameterization, cavity detection, and virtual similarity with a cavity database.

Such algorithms where the protein conformational space is also in part explored are called flexible docking methods. In the attempt to minimize the high computational cost generally only conformation that are close to the experimentally determined target structure are evaluated. The less computationally demanding possibility is to include amino acids side-chain flexibility exploiting rotamer libraries. Flexibility at the target level can also be included coupling the docking with a molecular dynamics program such AMBER (19) or GROMACS (20) to post-process docking results. The flexibility of the backbone in such protocols is included, but with the risk of possible inaccuracies due to the molecular mechanics force fields.

### 2.3. Scoring Function

Energy scoring functions are mathematical functions used to estimate the binding free energy of the small organic molecule to the protein target active pocket. Unfortunately, scoring is the weakest step in docking methodologies. In fact, in the majority of the cases it is unable to even rank-order a hit list.

Common scoring functions used in the molecular docking software simplify dramatically the thermodynamics of the binding event. The principal parts of ligand–protein interactions are taken in account to estimate in a fast way the most important energy contributions. Electrostatic and steric energy terms are generally included together with an explicit evaluation of the hydrogen bonding interaction. An internal energy part could also be included, while entropy and desolvation effects are neglected. The scoring process can also be a multistep procedure composed by a first fast analysis followed by a more accurate and computational demanding rescoring phase.

Scoring functions can be grouped in three families: molecular mechanics force field, empirical and knowledge-based scoring functions. In molecular mechanics the energy includes intra-molecular and inter-molecular contributions. Molecules are represented using force field specific atom and bond types with atom-centered partial charges. Bond energy derives from a bond stretching, bond angle, torsion angle, and improper torsion angle energy terms. The electrostatic energy is estimated using the Coulomb equation, while for

the van der Waals contribution the Lennard-Jones energy term is used. The AMBER (19) and OPLS (21) force field are well parameterized for protein and small organic molecules, but the disadvantage is that they are more computationally demanding than the knowledge-based and empirical scoring functions.

Empirical scoring functions approximate the binding energy as a sum of uncorrelated energy terms. Coefficients are obtained from a regression analysis of a set of ligands with known experimental binding energy to the target and with available X-ray structures of the complex. They have the role to compensate for possible error of the energy terms used. Examples are ChemScore (22), the Piecewise Linear Potential (PLP) (23) and X-Score (24). Their accuracy depends on how well the ligand and receptor were represented in the training data used to fit the coefficients. They can be optimized for particular tasks, like binding mode prediction, ranking of a particular set of inhibitors or to study a particular target.

Knowledge-based scoring functions are composed of multiple weighted molecular features related to ligand–receptor binding modes. The features are often atom–atom distances between protein and ligand in the complex, but also the number of intermolecular hydrogen bonds or atom–atom contact energies. A large number of X-ray diffraction crystals of protein–ligand complexes are used as a knowledge base. A putative protein–ligand complex can be assessed on the basis of how similar its features are to those in the knowledge base. These contributions are summed over all pairs of atoms in the complex and the resulting score is converted into a pseudo-energy function estimating the binding affinity. The coefficients of the features can be fitted using a linear regression analysis, but also other nonlinear statistical approaches can be used, like neural network, Bayesian modelling or machine learning technique like Random Forest analysis. Examples are PMF (25), DrugScore (26), LUDI (27), and RF-Score (28). Disadvantages with this class of scoring functions are difficulties in the evaluation of the chemical–physical meaning of the score and the risk of errors when trying to predict ligands not included in the training set.

#### **2.4. Physic-Based Endpoint Scoring Methods**

In the last step of the computational protocol, when the most promising ligands have been selected is possible to further evaluate their interaction with the target with more demanding computational approaches. For example the top ranked compounds from a virtual screening study can be rescored before the final selection is done. It is also possible to apply these techniques in a project in optimization phase to the most promising derivatives of the lead compound. There are different high-quality methods based on a rigorous physical framework; however, they still have to be further evaluated to better understand the potential and limits. Additional improvements are still needed to correctly model the high complexity of the ligand binding event (29).

Endpoint methods consider only the bound and unbound states of the ligand–protein complex without taking in account the intermediate states. This simplification reduces sensibly the computational cost compared to other physic-based methods. The free energy of binding is therefore estimated as follow:

$$\Delta G_{bind} = G_{complex} - (G_{ligand} + G_{protein})$$

The most popular endpoint method is probably the molecular mechanics Poisson-Boltzmann surface area method (MM-PBSA) (30). In this approach the individual energy terms are decomposed in a gas phase component calculated using the force field and a solvation energy term. For ligands also an entropic contribution is included:

$$G_{ligand} = G_{gas} + G_{solvation} - TS_{ligand}$$

The electrostatic contribution to the free energy of solvation is evaluated using an implicit solvent model: the Poisson-Boltzmann equation in MM-PBSA or the generalized Born equation in MM-GBSA. The hydrophobic contribution to the free energy of solvation is taken in account evaluating the solvent-accessible surface area (SASA) of the molecule.

$$G_{solvation} = G_{PB \text{ or } GB} + G_{SASA}$$

Only in some studies the vibrational entropy is taken in account for the ligand using normal mode analysis (31) because of the high computational cost and the risk of producing large errors. The final free energy of binding is estimated comparing the energy terms of the ligand and protein alone with the complex. The approach has been used to calculate absolute and relative binding affinities with error frequently of 1 or 2 kcal/mol. Protein flexibility is taken in account using molecular dynamics (MD) simulation or a faster energy minimization protocol. MM-GBSA using simply energy minimization can evaluate one ligand per minute. Still too slow to be applied in virtual screening studies, but order of magnitude faster than MM-PBSA using molecular dynamics with accuracy sometimes comparable or even higher (32).

The Linear Interaction Energy (LIE) method developed by Åqvist represents a plausible compromise between accuracy and computational speed in determining free energy of binding (33). The LIE approach is based on the assumption that the inhibitor free energy of binding to a macromolecule is linearly correlated to several energy terms that can be calculated using a molecular mechanic force field. In the original version the LIE binding free energy is approximated using the following equation (34):

$$\Delta G_{bind} = \alpha \Delta \langle G^{vdw} \rangle + \beta \Delta \langle G^{el} \rangle + \gamma$$

where the van der Waals ( $G^{vdw}$ ) and the electrostatic ( $G^{el}$ ) interaction energy of the ligand with its surrounding environment are evaluated for the bound and unbound state. The  $\Delta$  in the equation

denotes that the difference between these two states is calculated. For both states the averages of these energy contributions are computed on a population of conformations sampled by a molecular dynamics or a Monte Carlo procedure. Using a training set of molecules with known activity, a semi-empirical energy model is built by fitting the energy terms to the experimental free energy of binding. The LIE method assumes that the intra-molecular strain, entropy and desolvation effects are embedded in this linear response, and can be cancelled out by the empirically determined scaling parameters. The constant term  $\gamma$  can be substituted with a third energy term containing the difference in solvent-accessible surface area of the ligand, scaled by an empirical coefficient (35). The molecular dynamics sampling method can be substituted by simple energy minimization with a sensible decrease in the calculation times (36). The LIE method demonstrated to result in accurate predictions of relative and absolute free energy of binding with error around 1 kcal/mol (36, 37).

The endpoint methods seem so far the best approaches to bridge the gap between simple docking scoring methods and more rigorous free-energy calculations to improve accuracy at a practicable computational cost.

## 2.5. Limitations

Modern sophisticated docking methods allow a fast evaluation of a large number of ligands poses corresponding to different conformations and orientations of the small organic molecule in the protein target. Recently also receptor flexibility has started to be considered (38). Although the sensible improvement in the speed of calculation and efficacy of the conformational search algorithm, several limits are still challenging the predicting capability of these approaches especially affecting the scoring functions.

The calibration of the scoring functions is generally based on the data available in the PDB of small organic molecules in complex with proteins. These X-ray crystal structures could include important uncertainties as result of the subjective interpretation of the experimental electron-density map concerning in particular: (I) the identity of the isoelectric nitrogen and oxygen of the side chains of asparagine and glutamine, (II) the position of whole flexible residues, like lysine and glutamate, especially at the protein surface, or of mobile loops (III) also ligand atoms can be ambiguous, for example the position of pyridine nitrogen of asymmetrical substituted pyridine, (IV) the identification and location of water molecules, that are often isoelectronic to common buffer constituents in crystallization media, (V) the influence of crystallization media can affect crystal morphology, but also the ligand and the active site conformation (39), (VI) since hydrogen are not experimentally observed the ionization and tautomeric states cannot be determined and could be difficult to evaluate. The degree of confidence in the position of a particular atom or residue can be assessed using

the temperature factors and examining the structure together with the electron-density map.

The data set used to calibrate the scoring function tends to be unbalanced as consequence of the smaller number in the PDB of low-affinity ( $K_i > 1$  mM) ligands compared to the high-affinity molecules. As consequence the effects of unfavorable geometries of ligands in the protein pocket are not considered. The dipole moment of the ligand and the molecular electrostatic potential of the protein are often not included in the scoring functions. Residual flexibility of the protein or ligands is also not considered and entropic effects are often neglected. Some docking algorithms try to approximate such important contribution analyzing the number of rotatable bonds affected by the binding event. The desolvation event is roughly evaluated by the area of the interacting hydrophobic surfaces. Generally an inadequate evaluation of the desolvation effect can result in an overestimation of the affinity of polar compounds. When protein flexibility is considered in the search algorithm especially with a minimization step the risk is that locally introduced strain is dissipated by other part of the protein to such an extent to become unrecognizable by the scoring function. Non classical types of interactions are often neglected or not accurately evaluated (40): cation- $\pi$  interactions, charge transfer interactions, hydrogen bonding to  $\pi$ -systems, halogen bonding, orthogonal dipolar alignment, dipolar antiperiplanar interactions,  $\pi$ -stacking,  $\pi$  edge-to-face contacts, hydrogen bonding involving CH groups.

The limits of the scoring functions are the direct consequence of our not complete understanding of the energetic contributions of individual interactions. Formulating rules is possible only within certain boundaries especially if we consider that molecular interactions behave in a highly not additive fashion (41, 42). The link between thermodynamics and geometry of ligand-protein complexes still remains elusive.

Additionally, as clearly stated by Tirado-Rives and Jorgensen (43), the “window of activity” is very tiny. Thus, the free energy difference between the best ligand that one might reasonably expect to identify using virtual screening (potency,  $\approx 50$  nM) and the experimental detection limit (potency,  $\approx 100$   $\mu$ M) is only about 4.5 kcal/mol. The free energy contributions due to conformational factors alone for typical drug-like ligands (which are usually neglected in most scoring functions) can be as large as this.

---

### 3. The Role of Molecular Docking

Molecular docking is now a widely used approach in pharmaceutical and agrochemical companies. A number of drugs have been developed using docking methods at some stage of the development

(44–46). It is generally applied for three different purposes: (1) binding mode prediction, (2) virtual screening and (3) free energy of binding prediction. The performance of molecular docking in especially the first two tasks has been now widely tested. Even if these studies are not perfect, in the literature is possible to find empirical evaluations and comparative data that can help to have a general assessment of the docking technology.

### **3.1. Binding Mode Prediction**

Many of the published evaluation studies state that docking studies can reproduce the crystallographic ligand binding mode given sufficient computational time (47). Small organic molecules with low complexity are predicted in the right binding orientation in more than 50% of the cases, reaching in some studies values around 90% (48). When the number of rotatable bonds is high ( $>8$ ) or the binding site is large the error increases sensibly, with the risk that the docked pose is predicted within a rmsd of 2 Å from the crystallographic pose in only 30% of the cases. The success of the prediction is depending from the target type and the scoring function used. Systems with few specific hydrophobic interactions between ligand and protein are particularly challenging. On other hand scoring functions are not reliable in the identification of the pose more similar to the crystallographic one even if this is sampled by the search algorithm. We can therefore assume that the search problem is solved in docking programs, while the scoring remains the molecular docking limit.

### **3.2. Virtual Screening**

Molecular docking is often applied to large database of virtual compounds ( $10^5$ – $10^6$ ) to identify new promising ligands. The parameters of the algorithm are in this case optimized to increase the speed of the calculation. One pose is in the end selected per compound and evaluated by the scoring function. The top ranking molecules are chosen and tested experimentally. Virtual screening examples reported in the literature show that enrichment above random are very often obtained. The results show however that the performance of docking algorithms is inconsistent among protein families and it is very difficult to predict which molecular docking protocol will result in the best enrichment. Generally a 5% hit rate is considered substantial, where a “hit” shows a mid-micromolar affinity.

In the perfect case the docking algorithm is able to predict the correct binding mode, and the scoring function will result in enrichment thanks to a correct ranking of the library compounds. However in the virtual screening mode can be not crucial the ability of the algorithm to predict the correct ligand pose. Good binding mode predictions do not always correspond to active compounds enrichment due to the limit of the scoring functions. The source of a better-than-random result in virtual screening is often the consequence of the precision of the scoring function in discriminating

possible binders from totally inactive compounds that cannot bind at all. Such scoring function can still result in enrichment even if the predicted binding modes are wrong. In this case however the ranking will be probably also wrong, since the scoring function will act more as a classifier able to identify inactive molecules.

Recently molecular docking has been used also for virtual screening to prioritize fragments for testing (49, 50). Fragment-based drug discovery is focused on low molecular weight compounds (<250 Da) able to target subpockets of the active site (20). Screening of small library of 1,000–10,000 fragments resulted often in high hit rate (5%) and in an efficient sampling of the chemical space. Indeed the number of fragments necessary to cover all the chemotypes is much smaller: the number of potential fragments with up to 12 heavy atoms has been estimated at  $10^7$  (51), whereas the number of potential drug-like molecules with up to 30 heavy atoms is estimated at more than  $10^{60}$  (52). High-throughput methods for experimental testing are not feasible as consequence of the low affinity of the fragments and the high concentration needed for experimental testing. Virtual screening is therefore an attractive approach to select the most promising fragments. Such application of molecular docking could be challenging as consequence of possible promiscuous binding modes of the fragments and scoring functions parameterized on larger “drug-like” molecules. Nevertheless the first studies are encouraging, for example Chen et al. docked a library of ~67,000 fragments to the active site of  $\beta$ -lactamase, 69 have been selected for experimental testing and 10 resulted to have an IC<sub>50</sub> in the millimolar range (hit rate of 14.5%). 5 of these 10 have been crystallized in complex with the target and the predicted binding modes resulted in a rmsd range from 1.1 to 2.8 Å conserving the same polar and nonpolar contacts (53).

### **3.3. Free Energy of Binding Prediction**

Free energy of binding prediction is the real challenge of molecular docking methods. The assumption is that if the right binding pose is predicted, it will be possible to evaluate the interactions between the ligand and the target to try to estimate the binding energy. The limits are that the scoring functions are unable to take in account important chemical–physical aspects of the binding event like solvation and entropic contributions. Therefore even the ranking of compounds by affinity remains a difficult task. A solution to this limit is crucial to change the application of docking from a qualitative, descriptive method to a more quantitative tool able to drive lead optimization projects. Scientists from Glaxo-Smith-Kline evaluated the affinity prediction of 37 scoring functions and found no statistically significant correlation between docking score and measured affinity for most of the eight proteins examined (48). Only few publications showed good correlations between scoring function and compounds affinity and often there was low chemical diversity.

An evaluation of the LIE method showed that this approach can be superior to empirical scoring functions in a lead optimization context (37). On the other hand there is still a potential to improve current empirical methods if more reliable and statistically robust data will be available. It is believed that with sufficient computing power and the application of rigorous methods like free energy perturbation approach or quantum mechanical scoring could be possible to improve the docking ranking capabilities (54–57). However at the contrary of common docking methods, these approaches have not been extensively tested yet and this assumption has still to be proven. For example different studies show very different performance of the MM-GBSA method going from very good predictions of binding affinities (58) to quite poor results ( $r^2 = 0.2$ ) (59). Often a true comparison among published data is not possible because of differences in the protocols used.

---

#### 4. Application of Molecular Docking to GPCRs

Human G Protein-Coupled Receptors (GPCRs) with at least 800 unique full-length members comprise the largest family of cell-surface receptors (60). They are ubiquitous biological control points of the cell. External signals are translated by this membrane protein family into readable stimuli resulting in precise cell behaviors (61). Examples of physiological responses controlled by GPCRs are cell growth and differentiation, cardiovascular function, metabolism, immune responses, and neurotransmission. They also represent the largest family of drug targets with about 50% of the existing drugs currently targeting GPCRs for their beneficial action (62), and their therapeutic potential might be even larger (63). Structurally, GPCRs are cell receptors characterized by seven transmembrane (TMs) helices clustered in the form of a bundle and linked by three intracellular (IL) and three extracellular (EL) loops.

The mechanisms that control ligand–GPCR binding, receptor activation, and signal transduction are multifactorial and highly complex. Understanding the structural determinants required for receptor function can play an important role in the design of new ligands of potential pharmaceutical interest. The first crystal structure of prototypic class A GPCR bovine rhodopsin covalently linked to *cis*-retinal as an inverse agonist was solved in year 2000. The high abundance of rhodopsin in rod outer segments, the conformational homogeneity and the stability in the inactive state made this protein the first choice for GPCR crystallography. At the contrary the other members of the family represent a real challenge for X-ray diffraction crystallography as consequence of their intrinsic conformational flexibility, their inherent instability resulting in short half-life in nonmembranous environment and the

low abundance from both natural and recombinant sources. Important new developments in methodology like protein engineering and antibody/fusion protein cocrystallization finally resulted starting from 2007 in new crystal structures of GPCRs of pharmaceutical interest. The structural features of important drug targets, i.e., the  $\beta 2$  adrenergic receptor,  $\beta 1$  adrenergic receptor, and adenosine receptor A2A occupied by antagonist/inverse agonist, provided important information for the rational design of new drugs.

#### **4.1. Molecular Docking to GPCRs X-Ray Crystal Structures**

The breakthroughs in GPCR crystallography improve dramatically the potential of GPCRs structure-based ligand design approaches. In virtual screenings against the  $\beta 2$  adrenergic crystal structure unusual high and promising hit rate have been observed in several different studies. In one of them about a million “lead-like” molecules were evaluated in silico and 25 compounds were chosen for experimental testing (64). Six resulted having an affinity between 4  $\mu\text{M}$  and 9 nM (hit rate of 24%). In another study a remarkable hit rate of 36% was obtained from a virtual screening of 400,000, with the best compound having a subnanomolar activity (65). Further computational studies suggested the possibility to exploit the  $\beta 2$  adrenergic crystal structure in the inactive state to apply molecular docking also for the identification of agonist and partial-agonist after applying precise modifications of the protein structure (66, 67).

Virtual screening studies against the adenosine receptor A2A resulted also in very high hit rates. A four million “drug-like” and “lead-like” compounds virtual screening resulted indeed in a hit rate of 41% (68). Out of 56 compounds experimentally tested, 23 showed antagonist activity under 10  $\mu\text{M}$ , among them 11 had submicromolar activity and two compounds had affinity under 60 nM. Nine novel chemotypes were identified supporting high diversity of the hits generated through structure-based virtual screening. In another study (69) a total of 1.4 million compounds have been screened in silico against the same A2A X-ray structure, 20 high-ranking novel compounds have been selected and tested experimentally resulting in a hit rate of 35%. The activity range was between 200 nM and 9  $\mu\text{M}$ .

These studies suggest practical applicability of receptor-based virtual screening in GPCR drug discovery. Furthermore extraordinary high hit rates and high activity have been identified suggesting the high potential of the X-ray diffraction crystals compared to the earlier homology models. This is in agreement with the community-wide, blind assessment of the prediction of the structure of the human adenosine A2A in complex with the ligand ZM241385 (70). Twenty-nine groups have been involved and a total of 206 structural models submitted, the average rmsd of the ligand predictions resulted of 9.5 Å, the best model had a ligand rmsd of 2.8 Å. On the other hand the high hit rate can also be

explained by the commercially available database bias toward GPCR-like ligands as result of more than 50 years intense research in this field.

---

## 5. GPCRs Flexibility in Molecular Docking

The recent detailed structural information of class A GPCR has considerably improved our understanding of this receptor family and represents a structural context to efficiently integrate experimental data derived from molecular biology, biophysics, bioinformatics, pharmacology and organic chemistry. In the last years these joined efforts have demonstrated that both active and inactive states can be described by several different receptor conformations depending on the nature of both ligand and receptor (67). In particular the structural flexibility of GPCRs and the ability of ligands to induce or to stabilize specific conformations can result in functional plasticity. The elaboration of successful methods to explore the multiconformational space of GPCRs is crucial to fully understand the intimate mechanism of ligand–receptor recognition and also to speed up the discovery of new potent and selective GPCR-ligands.

### 5.1. Ligand-Based Homology Modelling

Ligand-based homology modelling is a computational strategy developed as an attempt to model the high conformational complexity of GPCRs (71). This modelling technique is an evolution of a conventional homology modelling algorithm. It is based on a Boltzmann weighted randomized modelling procedure adapted from Levitt and combined with specialized logic for the proper handling of insertions and deletions, where any selected atom is included in the energy tests and in the minimization stages of the modelling procedure. The homology model is built in the presence of a small organic molecule docked to the primary template. In this specific case both model building and refinement take into account the presence of the ligand in terms of specific steric and chemical features. The final aim is the simulation of possible receptor structural reorganization induced by the antagonist binding.

The approach has been validated on more than 200 known human Adenosine Receptor A3 antagonists in the corresponding putative ligand binding site (72–76). A starting homology model based on rhodopsin X-ray crystal structure has been exploited to perform docking studies of a class of pyrazolo-triazolo-pyrimidines derivatives. The resulting binding mode of this class of antagonists occurred in the upper region of the TM helical bundle. In particular TMs 3, 5, 6 and 7 seemed to be crucial for the recognition of the antagonists. The experimental effect on ligand binding affinity of the mutation of several residues predicted to be involved in the

antagonist recognition confirmed the theoretical binding mode (77). However the model could not explain the observed activity when bulky substituents were present as consequence of a strongly destabilizing van der Waals energy component avoiding the sampling of reasonable antagonist-receptor complexes. Therefore this starting homology model was not appropriated to guarantee a good complementarity among the topology of the receptor's cleft and the shape of the bulkier derivatives. Starting from this conventional rhodopsin-based homology model and applying the ligand-based homology modelling implementation has been possible to generate other two antagonist-like conformational states of human A3 receptor in which the ligand recognition cavity has been expanded from 670 Å<sup>3</sup> to 770 and 1,120 Å<sup>3</sup> respectively. These two new antagonist-like conformational states were able to rationalize the observed activities for all reported compounds.

Other similar techniques including ligand information to improve the accuracy of the active site binding pocket conformation are the ligand-supported homology model approach (78) and the ligand-steered homology model method (79). In the first the ligand binding is considered in the homology model building by user-defined ligand–protein restraints derived from manual placement of the ligands in the active pocket. The included set of ligands and a set of predefined interactions are used to generate receptor models in agreement with mutagenesis data and structure–activity relationship of the ligands. Further refinement through energy minimization of the active site is possible and the final models are ranked using the DrugScore scoring function to evaluate the interaction between the pocket and the ligand. This approach combined with interaction fingerprinting-based similarity has been applied in a retrospective virtual screening experiment to detect antagonist of the metabotropic glutamate receptor subtype 5. The enrichment obtained was significantly higher than other conventional scoring functions. In the similar ligand-steered homology modelling method ligands are used to shape and optimize the binding site through a docking-based stochastic global energy minimization procedure. The method has been applied to the melanin-concentrating hormone receptor 1 and the application of the virtual screening procedure of a library of more than 180,000 compounds. Six out of 129 compounds tested experimentally resulted active with Ki values ranging within 7–20 μM.

## **5.2. Metadynamics-Based Flexible Ligand Docking**

The balloon potential method developed at Bristol-Myers Squibb is another approach to try to solve the problem of the GPCR models with too small active site unable to accommodate known ligands (80). This method applies an increasing pressure by expanding balloon in the cavity walls of the active site during a MD simulation. This is accomplished using a set of Lennard-Jones particles initially placed in the binding site with radii that are slowly increased to

mimic an increasing pressure. The method has been applied to the chemokine receptor CCR2 and the docking of three known ligands resulted in orientation consistent with published mutagenesis data.

Even if this method is an unbiased procedure that does not depend on the use of experimental information, it results in a priori pocket modification not caused by ligand binding. Metadynamics on the contrary can include induce fit events caused by the binding event to the receptor, like side chains reorganization, resulting in an refinement of the pocket based on specific chemical–physical interactions with the small organic molecule in study. The final ligand bound receptor conformation can be validated with a comparison *a posteriori* to the experimental data available. Metadynamics is an enhanced sampling algorithm within the framework of classical MD allowing the efficient exploration of the phase space of biological systems. A non-Markovian (history-dependent) bias is added to the interaction potential allowing the estimation of the relative free-energy of the biological event in study (81). Metadynamics method has been proven to be a useful method to the flexible docking of small organic molecules to protein active pockets (82–84). It is based on a careful choice of collective variables (CVs) able to describe the slow events that are relevant to ligand binding.

The method has been applied in the study of the docking event of the antagonist Naloxone (Fig. 1a) to the  $\delta$  opioid receptor (85). Opioid receptors bind a variety of ligands like endogenous peptides, opiates and thousand of synthetic compounds. In particular the clinical interest for this family of GPCR is linked to the development of strong analgesics providing outstanding benefits for relief of severe pain. Among them Morphine, isolated from opium in 1803, is still the most often used analgesic drug even if it is accompanied by important side effects like tolerance and both physical and psychological dependence. Naloxone is a nonselective

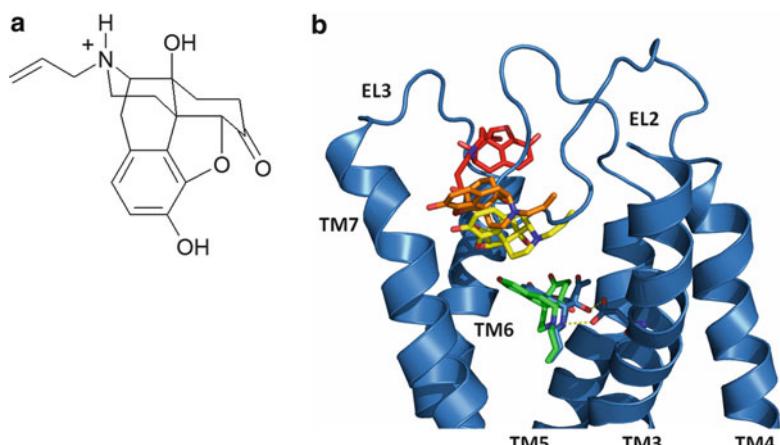


Fig. 1. (a) Chemical structure of the antagonist Naloxone. (b) Entrance path followed by Naloxone to the  $\delta$  opioid receptor active pocket. TM1 and TM2 are hidden for clarity.

opiate antagonist used to counter the effects of opioid overdose, like life-threatening depression of the central nervous system and respiratory system caused by heroin or morphine. The design of new therapies with more desirable properties without potential for abuse or adverse side effects cannot be accomplished without understanding the modes of interactions of drugs/ligands with these receptors.

The  $\delta$  opioid receptor molecular model has been built using the homology modelling approach for the transmembrane region exploiting as template the crystal structure of  $\beta 2$  adrenergic receptor. Loops have been added applying an ab initio method and the system has been equilibrated in a fully explicit water-membrane environment. The metadynamics approach has been exploited to simulate a total of 0.5  $\mu$ s to reconstruct the free energy surface of the ligand binding event including docked, undocked and metastable bound intermediates (Fig. 1b) and to predict a free energy of binding close to experimental values.

An intermediate binding mode has been detected by the simulation where the ligand creates stable interactions with the extracellular surface of the receptor, in particular in the region comprised between the EL2 in its closed conformation and EL3 (Fig. 2a). The two loops can generate a binding cleft characterized principally by apolar residues like Phe202 and Trp284<sup>6.58</sup> and covered by a highly dynamic electrostatic region resulting from the EL3 residues Asp290, Arg291, Arg292, and Asp293. Naloxone interacts with

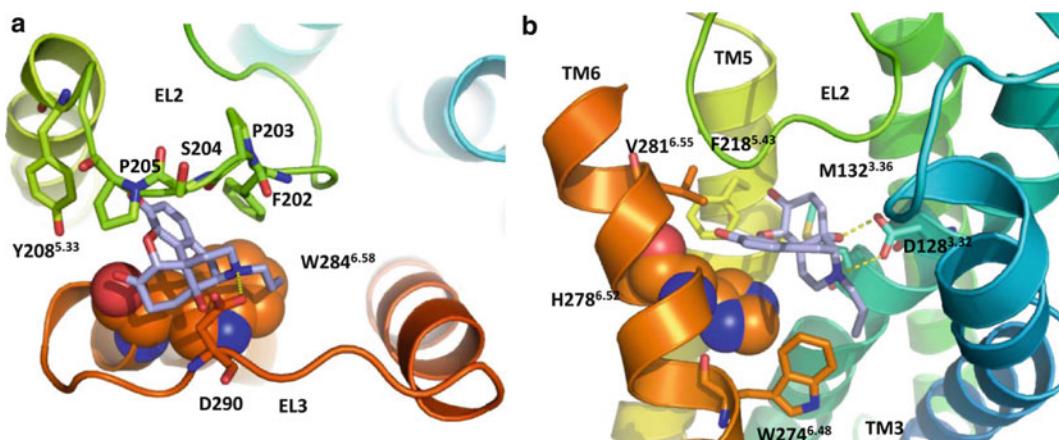


Fig. 2. (a) Intermediate binding mode of Naloxone to the EL2–EL3 region. The interaction is stabilized by a salt bridge with Asp290 and van der Waals contacts with Trp284<sup>6.58</sup>. (b) Most stable Naloxone- $\delta$  opioid receptor complex. The ligand creates strong polar and ionic interaction with Asp128<sup>3.32</sup>, stacking interaction with His278<sup>6.52</sup> and hydrophobic contacts principally with Trp274<sup>6.48</sup>, Phe218<sup>5.43</sup>, Met132<sup>3.36</sup> and Val281<sup>6.55</sup>. Ballesteros-Weinstein numbering (28) is used as superscript to the protein numbering for TMs residues. Specifically, the first number (from 1 to 7) of this two-number identifier refers to the TM helix number, whereas the second number indicates the residue position in the helix relative to the most conserved residue (assigned index of 50) in that helix, with numbers decreasing toward the N-terminus of the helix and increasing toward its C-terminus.

this region of DOR principally through a salt bridge with Asp290 and van der Waals contacts with Trp284<sup>6,58</sup>. It creates a transient hydrogen bonding with Tyr208<sup>5,33</sup> mediated by a water molecule and weaker apolar interactions with Pro205 and Phe202.

The most stable binding mode of Naloxone to the  $\delta$  opioid receptor is placed among TM3, TM5 and TM6 (Fig. 2b). The antagonist strongly interacts with Asp128<sup>3,32</sup> creating a salt bridge through its ammonium group and a polar interaction with its tertiary alcohol moiety. Naloxone orientation in the pocket allows a stacking interaction with His278<sup>6,52</sup> and apolar contact between Trp274<sup>6,48</sup> and Tyr308<sup>7,43</sup> to the antagonist propenyl tail. Further van der Waals interactions are created with Phe218<sup>5,43</sup> and Met132<sup>3,36</sup>.

Notably several residues, resulted important for the described docking mechanism of Naloxone to the  $\delta$  opioid receptor, are experimentally known to influence ligand binding. For example mutagenesis experiment showed the importance of Trp284<sup>6,58</sup> at the beginning of the EL3, and its effect on binding affinity of several ligands and also on receptor activation. The mutation of the highly conserved Asp128<sup>3,32</sup> to Lys decreases the affinity of Naloxone, and additional data from two-dimensional mutagenesis in the  $\mu$  opioid receptor supported the notion that this residue acts as counterion for opioid ligands. Other residues proven experimentally to be important for ligand binding and in agreement with the prediction are Trp274<sup>6,48</sup>, Tyr308<sup>7,43</sup>, Phe218<sup>5,43</sup>, and His278<sup>6,52</sup>.

Flexibility is now commonly believed to be a key element in GPCR function and can be essential to be taken in account to study ligand binding. In this contest metadynamics represents a promising approach to refine the GPCR active pocket directly as consequence of the chemical–physical interactions with the molecule in study. This approach is still very computational demanding; however, the refined final structure can then be used to perform standard virtual screening protocols, even if the expected improved performance compared to standard homology models has still to be demonstrated.

---

## 6. Conclusions

Undoubtedly, molecular docking is now used by almost every major pharmaceutical company and academic institution. But it is also true that docking seems to have reached a performance plateau and is waiting for a real and important breakthrough. Like most scientific breakthroughs, such an advance is hard to predict. Some criticisms are summarized as follows.

As stated at the beginning, despite all the advances in molecular docking, scoring functions are still very primitive from a thermodynamic point of view. Thus even if a correct pose is generated (requiring that the correct “bioactive” conformation of the molecule has been generated and a protein conformation that does not hinder the correct docking has been used) it may be scored as less favorable than other “incorrect” poses and be discarded. Even if the most accurate approaches such as thermodynamic integration/free energy perturbation methods, which can sometimes calculate the differences in enthalpic affinities between related molecules to within “chemical accuracy” about  $1 \div 2$  kcal/mol, suffer of the remarkable inaccuracy of the entropic factor calculation. Moreover, these methods only compare close analogues; they do not predict absolute binding affinities nor can they compare affinities among the diverse, unrelated molecules found in a typical screening library. They still demand so much computation time as to be infeasible for a large library.

Receptor flexibility remains one of the major challenges for the field. In fact, the interactions among associating molecules are dynamic, which adds to the complexity of molecular recognition. Hopefully, the regular exponential growth in number of protein structures in PDB, the improved understanding of the induced fit mechanisms, and improved conformational sampling methods might improve the explicitly accounting for the flexibility of both the receptor and the docked ligands. Indeed, these types of the full flexibility generators can be used in the absence of a suitable diversity of the crystal structures or even starting from homology models.

In terms of recent more “transformational” advances, a better understanding of the energetics of the whole ensemble of binding site waters is providing new insights into structure–activity relationships and docking pose evaluation, where effects on the water network and the displacement of waters that are either “happy” or “unhappy” relative to bulk solvent can drive binding energy differences and binding orientations. A new software program from Schrodinger called WaterMap uses molecular dynamics simulations and clustering of results to readily provide a stable network and the relative energy of each water in the site, with the enthalpic and entropic contributions, calculated with or without bound ligands (86–88).

From an informatics infrastructure point of view, both grid and cloud computing might sensibly cut down the computational time requested for the application of the most accurate scoring functions in any screening of a large library of chemical or for virtually in parallel several biological targets as requested for the design of an *in silico* chemical proteomics experiment.

We are positive that the wedding between thermodynamic-driven chemical accuracy and large docking-based screening campaign it is only a matter of time, and not too much of it.

## References

1. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
2. Lipinski C, Hopkins A (2004) Navigating chemical space for biology and medicine. *Nature* 432:855–861
3. Muller G (2003) Medicinal chemistry of target family-directed masterkeys. *Drug Discov Today* 8:681–691
4. Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* 28:849–857
5. DesJarlais RL, Sheridan RP, Seibel GL, Dixon JS, Kuntz ID, Venkataraghavan R (1988) Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J Med Chem* 31:722–729
6. Milletti F, Storchi L, Sforna G, Cross S, Cruciani G (2009) Tautomer enumeration and stability prediction for virtual screening on large chemical databases. *J Chem Inf Model* 49:68–75
7. Milletti F, Storchi L, Sforna G, Cruciani G (2007) New and original pKa prediction method using grid molecular interaction fields. *J Chem Inf Model* 47:2172–2181
8. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 261:470–489
9. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol* 245:43–53
10. Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD (1998) Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins* 33:367–382
11. Hall SB, Venkitaraman AR, Whitsett JA, Holm BA, Notter RH (1992) Importance of hydrophobic apoproteins as constituents of clinical exogenous surfactants. *Am Rev Respir Dis* 145:24–30
12. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. I. Method and assessment of docking accuracy. *J Med Chem* 47:1739–1749
13. Goto J, Kataoka R, Hirayama N (2004) Ph4Dock: pharmacophore-based protein-ligand docking. *J Med Chem* 47:6804–6811
14. Baroni M, Cruciani G, Scialbo S, Perruccio F, Mason JS (2007) A common reference framework for analyzing/comparing proteins and ligands. Fingerprints for Ligands and Proteins (FLAP): theory and application. *J Chem Inf Model* 47:279–294
15. Scialbo S, Stanton RV, Mills JE, Flocco MM, Baroni M, Cruciani G, Perruccio F, Mason JS (2010) High-Throughput Virtual Screening of Proteins Using GRID Molecular Interaction Fields. *J Chem Inf Model* 50:155–169
16. Carosati E, Budriesi R, Ioan P, Urgenti MP, Frosini M, Fusi F, Corda G, Cosimelli B, Spinelli D, Chiarini A, Cruciani G (2008) Discovery of novel and cardioselective diltiazem-like calcium channel blockers via virtual screening. *J Med Chem* 51:5552–5565
17. Carosati E, Mannhold R, Wahl P, Hansen JB, Fremming T, Zamora I, Cianchetta G, Baroni M (2007) Virtual screening for novel openers of pancreatic K(ATP) channels. *J Med Chem* 50:2117–2126
18. Park K, Kim D (2006) A method to detect important residues using protein binding site comparison. *Genome Inform* 17:216–225
19. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
20. Congreve M, Chessari G, Tisi D, Woodhead AJ (2008) Recent developments in fragment-based drug discovery. *J Med Chem* 51:3661–3680
21. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
22. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP (1997) Empirical scoring functions: I The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J Comput Aided Mol Des* 11:425–445
23. Verkhivker GM (2004) Computational analysis of ligand binding dynamics at the intermolecular hot spots with the aid of simulated tempering and binding free energy calculations. *J Mol Graph Model* 22:335–348
24. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comput Aided Mol Des* 16:11–26

25. Muegge I (2006) PMF scoring revisited. *J Med Chem* 49:5895–5902
26. Velec HF, Gohlke H, Klebe G (2005) Drug-Score(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. *J Med Chem* 48:6296–6303
27. Bohm HJ (1994) The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J Comput Aided Mol Des* 8:243–256
28. Ballesteros JA, Weinstein H (1995) Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci* London 25:366–366
29. Huang N, Jacobson MP (2007) Physics-based methods for studying protein-ligand interactions. *Curr Opin Drug Discov Devel* 10:325–331
30. Gouda H, Kuntz ID, Case DA, Kollman PA (2003) Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and thermodynamic integration methods. *Biopolymers* 68:16–34
31. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE 3rd (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 33:889–897
32. Kuhn B, Gerber P, Schulz-Gasch T, Stahl M (2005) Validation and use of the MM-PBSA approach for drug discovery. *J Med Chem* 48:4040–4048
33. Aqvist J, Medina C, Samuelsson JE (1994) A new method for predicting binding affinity in computer-aided drug design. *Protein Eng* 7:385–391
34. Aqvist J, Luzhkov VB, Brandsdal BO (2002) Ligand binding affinities from MD simulations. *Acc Chem Res* 35:358–365
35. Jones-Hertzog DK, Jorgensen WL (1997) Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method. *J Med Chem* 40:1539–1549
36. Bortolato A, Moro S (2007) In silico binding free energy predictability by using the linear interaction energy (LIE) method: bromobenzimidazole CK2 inhibitors as a case study. *J Chem Inf Model* 47:572–582
37. Stjernschantz E, Marelius J, Medina C, Jacobson M, Vermeulen NP, Oostenbrink C (2006) Are automated molecular dynamics simulations and binding free energy calculations realistic tools in lead optimization? An evaluation of the linear interaction energy (LIE) method. *J Chem Inf Model* 46:1972–1983
38. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3:935–949
39. Stubbs MT, Reyda S, Dullweber F, Moller M, Klebe G, Dorsch D, Mederski WW, Wurziger H (2002) pH-dependent binding modes observed in trypsin crystals: lessons for structure-based drug design. *Chembiochem* 3:246–249
40. Bissantz C, Kuhn B, Stahl M (2010) A medicinal chemist's guide to molecular interactions. *J Med Chem* 53:5061–5084
41. Mark AE, van Gunsteren WF (1994) Decomposition of the free energy of a system in terms of specific interactions Implications for theoretical and experimental studies. *J Mol Biol* 240:167–176
42. Dill KA (1997) Additivity principles in biochemistry. *J Biol Chem* 272:701–704
43. Tirado-Rives J, Jorgensen WL (2006) Contribution of conformer focusing to the uncertainty in predicting free energies for protein-ligand binding. *J Med Chem* 49:5880–5884
44. Garman E, Laver G (2004) Controlling influenza by inhibiting the virus's neuraminidase. *Curr Drug Targets* 5:119–136
45. von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, Van Phan T, Smythe ML, White HF, Oliver SW et al (1993) Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* 363:418–423
46. Kaldor SW, Kalish VJ, Davies JF 2nd, Shetty BV, Fritz JE, Appelt K, Burgess JA, Campanale KM, Chirgadze NY, Clawson DK, Dressman BA, Hatch SD, Khalil DA, Kosa MB, Lubbenhusen PP, Muesing MA, Patick AK, Reich SH, Su KS, Tatlock JH (1997) Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *J Med Chem* 40:3979–3985
47. Chen H, Lyne PD, Giordanetto F, Lovell T, Li J (2006) On evaluating molecular-docking methods for pose prediction and enrichment factors. *J Chem Inf Model* 46:401–415
48. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS (2006) A critical assessment of docking programs and scoring functions. *J Med Chem* 49:5912–5931

49. Murray CW, Callaghan O, Chessari G, Cleasby A, Congreve M, Frederickson M, Hartshorn MJ, McMenamin R, Patel S, Wallis N (2007) Application of fragment screening by X-ray crystallography to beta-secretase. *J Med Chem* 50:1116–1123
50. Huang JW, Zhang Z, Wu B, Cellitti JF, Zhang X, Dahl R, Shiao CW, Welsh K, Emdadi A, Stebbins JL, Reed JC, Pellecchia M (2008) Fragment-based design of small molecule X-linked inhibitor of apoptosis protein inhibitors. *J Med Chem* 51:7111–7118
51. Fink T, Bruggesser H, Reymond JL (2005) Virtual exploration of the small-molecule chemical universe below 160 Daltons. *Angew Chem Int Ed Engl* 44:1504–1508
52. Bohacek RS, McMurtin C, Guida WC (1996) The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* 16:3–50
53. Chen Y, Shoichet BK (2009) Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat Chem Biol* 5:358–364
54. Jorgensen WL (2009) Efficient drug lead discovery and optimization. *Acc Chem Res* 42:724–733
55. Brown SP, Muchmore SW (2006) High-throughput calculation of protein-ligand binding affinities: modification and adaptation of the MM-PBSA protocol to enterprise grid computing. *J Chem Inf Model* 46:999–1005
56. Fujitani H, Tanida Y, Ito M, Jayachandran G, Snow CD, Shirts MR, Sorin EJ, Pande VS (2005) Direct calculation of the binding free energies of FKBP ligands. *J Chem Phys* 123:084108
57. Fedorov DG, Kitaura K (2007) Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *J Phys Chem A* 111:6904–6914
58. Lyne PD, Lamb ML, Saeh JC (2006) Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J Med Chem* 49:4805–4808
59. Yang CY, Sun H, Chen J, Nikolovska-Coleska Z, Wang S (2009) Importance of ligand reorganization free energy in protein-ligand binding-affinity prediction. *J Am Chem Soc* 131:13709–13721
60. Gloriam DE, Fredriksson R, Schioth HB (2007) The G protein-coupled receptor subset of the rat genome. *BMC Genomics* 8:338
61. Marinissen MJ, Gutkind JS (2001) G-protein-coupled receptors and signaling networks: emerging paradigms. *Trends Pharmacol Sci* 22:368–376
62. Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5:993–996
63. Strange PG (2008) Signaling mechanisms of GPCR ligands. *Curr Opin Drug Discov Devel* 11:196–202
64. Kolb P, Rosenbaum DM, Irwin JJ, Fung JJ, Kobilka BK, Shoichet BK (2009) Structure-based discovery of beta2-adrenergic receptor ligands. *Proc Natl Acad Sci USA* 106:6843–6848
65. Sabio M, Jones K, Topiol S (2008) Use of the X-ray structure of the beta2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. *Bioorg Med Chem Lett* 18:5391–5395
66. de Graaf C, Rognan D (2008) Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor. *J Med Chem* 51:4978–4985
67. Katritch V, Reynolds KA, Cherezov V, Hanson MA, Roth CB, Yeager M, Abagyan R (2009) Analysis of full and partial agonists binding to beta2-adrenergic receptor suggests a role of transmembrane helix V in agonist-specific conformational changes. *J Mol Recognit* 22:307–318
68. Katritch V, Jaakola VP, Lane JR, Lin J, Ijzerman AP, Yeager M, Kufareva I, Stevens RC, Abagyan R (2010) Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *J Med Chem* 53:1799–1809
69. Carlsson J, Yoo L, Gao ZG, Irwin JJ, Shoichet BK, Jacobson KA (2010) Structure-based discovery of A2A adenosine receptor ligands. *J Med Chem* 53:3748–3755
70. Michino M, Abola E, Brooks CL 3rd, Dixon JS, Moult J, Stevens RC (2009) Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat Rev Drug Discov* 8:455–463
71. Moro S, Deflorian F, Bacilieri M, Spalluto G (2006) Ligand-based homology modeling as attractive tool to inspect GPCR structural plasticity. *Curr Pharm Des* 12:2175–2185
72. Moro S, Deflorian F, Spalluto G, Pastorin G, Cacciari B, Kim SK, Jacobson KA (2003) Demystifying the three dimensional structure of G protein-coupled receptors (GPCRs) with the aid of molecular modeling. *Chem Commun (Camb)* 24:2949–2956
73. Maconi A, Pastorin G, Da Ros T, Spalluto G, Gao ZG, Jacobson KA, Baraldi PG, Cacciari B, Varani K, Moro S, Borea PA (2002) Synthesis, biological properties, and molecular modeling investigation of the first potent, selective, and

- water-soluble human A(3) adenosine receptor antagonist. *J Med Chem* 45:3579–3582
74. Pastorin G, Da Ros T, Spalluto G, Deflorian F, Moro S, Cacciari B, Baraldi PG, Gessi S, Varani K, Borea PA (2003) Pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as adenosine receptor antagonists. Influence of the N5 substituent on the affinity at the human A3 and A2B adenosine receptor subtypes: a molecular modeling investigation. *J Med Chem* 46:4287–4296
75. Colotta V, Catarzi D, Varano F, Calabri FR, Lenzi O, Filacchioni G, Martini C, Trincavelli L, Deflorian F, Moro S (2004) 1,2,4-triazolo[4,3-a]quinoxalin-1-one moiety as an attractive scaffold to develop new potent and selective human A3 adenosine receptor antagonists: synthesis, pharmacological, and ligand-receptor modeling studies. *J Med Chem* 47:3580–3590
76. Moro S, Braiuca P, Deflorian F, Ferrari C, Pastorin G, Cacciari B, Baraldi PG, Varani K, Borea PA, Spalluto G (2005) Combined target-based and ligand-based drug design approach as a tool to define a novel 3D-pharmacophore model of human A3 adenosine receptor antagonists: pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as a key study. *J Med Chem* 48:152–162
77. Gao ZG, Chen A, Barak D, Kim SK, Muller CE, Jacobson KA (2002) Identification by site-directed mutagenesis of residues involved in ligand recognition and activation of the human A3 adenosine receptor. *J Biol Chem* 277:19056–19063
78. Evers A, Gohlke H, Klebe G (2003) Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. *J Mol Biol* 334:327–345
79. Cavasotto CN, Orry AJ, Murgolo NJ, Czarnecki MF, Kocsi SA, Hawes BE, O'Neill KA, Hine H, Burton MS, Voigt JH, Abagyan RA, Bayne ML, Monsma FJ Jr (2008) Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J Med Chem* 51:581–588
80. Kimura SR, Tebben AJ, Langley DR (2008) Expanding GPCR homology model binding sites via a balloon potential: A molecular dynamics refinement approach. *Proteins* 71:1919–1929
81. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99:12562–12566
82. Masetti M, Cavalli A, Recanatini M, Gervasio FL (2009) Exploring complex protein-ligand recognition mechanisms with coarse metadynamics. *J Phys Chem B* 113:4807–4816
83. Gervasio FL, Laio A, Parrinello M (2005) Flexible docking in solution using metadynamics. *J Am Chem Soc* 127:2600–2607
84. Vargiu AV, Ruggerone P, Magistrato A, Carloni P (2008) Dissociation of minor groove binders from DNA: insights from metadynamics simulations. *Nucleic Acids Res* 36:5910–5921
85. Provasi D, Bortolato A, Filizola M (2009) Exploring molecular mechanisms of ligand recognition by opioid receptors with metadynamics. *Biochemistry* 48:10020–10029
86. Pearlstein RA, Hu QY, Zhou J, Yowe D, Levell J, Dale B, Kaushik VK, Daniels D, Hanrahan S, Sherman W, Abel R (2010) New hypotheses about the structure-function of proprotein convertase subtilisin/kexin type 9: analysis of the epidermal growth factor-like repeat A docking site using WaterMap. *Proteins* 78:2571–2586
87. Guimaraes CR, Mathiowitz AM (2010) Addressing limitations with the MM-GB/SA scoring procedure using the WaterMap method and free energy perturbation calculations. *J Chem Inf Model* 50:547–559
88. Chrencik JE, Patny A, Leung IK, Korniski B, Emmons TL, Hall T, Weinberg RA, Gormley JA, Williams JM, Day JE, Hirsch JL, Kiefer JR, Leone JW, Fischer HD, Sommers CD, Huang HC, Jacobsen EJ, Tenbrink RE, Tomasselli AG, Benson TE (2010) Structural and thermodynamic characterization of the TYK2 and JAK3 kinase domains in complex with CP-690550 and CMP-6. *J Mol Biol* 400:413–433

# **Chapter 14**

## **Simulation Studies of the Mechanism of Membrane Transporters**

**Giray Enkavi, Jing Li, Paween Mahinthichaichan, Po-Chao Wen,  
Zhijian Huang, Saher A. Shaikh, and Emad Tajkhorshid**

### **Abstract**

Membrane transporters facilitate active transport of their specific substrates, often against their electrochemical gradients across the membrane, through coupling the process to various sources of cellular energy, for example, ATP binding and hydrolysis in primary transporters, and pre-established electrochemical gradient of molecular species other than the substrate in the case of secondary transporters. In order to provide efficient energy-coupling mechanisms, membrane transporters have evolved into molecular machines in which stepwise binding, translocation, and transformation of various molecular species are closely coupled to protein conformational changes that take the transporter from one functional state to another during the transport cycle. Furthermore, in order to prevent the formation of leaky states and to be able to pump the substrate against its electrochemical gradient, all membrane transporters use the widely-accepted “alternating access mechanism,” which ensures that the substrate is only accessible from one side of the membrane at a given time, but relies on complex and usually global protein conformational changes that differ for each family of membrane transporters. Describing the protein conformational changes of different natures and magnitudes is therefore at the heart of mechanistic studies of membrane transporters. Here, using a number of membrane transporters from diverse families, we present common protocols used in setting up and performing molecular dynamics simulations of membrane transporters and in analyzing the results, in order to characterize relevant motions of the system. The emphasis will be on highlighting how optimal design of molecular dynamics simulations combined with mechanistically oriented analysis can shed light onto key functionally relevant protein conformational changes in this family of membrane proteins.

**Key words:** Alternating access mechanism, Molecular dynamics, Conformational change, Conformational coupling, Outward-facing (OF) state, Inward-facing (IF) state, Occluded state, State transition, ABC transporters, Maltose transporter, Nucleotide binding domains (NBDs), ATP hydrolysis, Biased simulation,  $\text{Na}^+$ -coupled galactose transporter, Ion release, Substrate release, Betaine, Glycerol-3-phosphate (G3P), Inorganic phosphate ( $\text{P}_i$ ), Major facilitator superfamily (MFS), Transmembrane helices, Protonation state, Titration state, *Apo* state, Rocker-switch model, Salt bridge, Normal mode analysis (NMA), Anisotropic network model (ANM), Glycerol-3-phosphate transporter (GlpT), Glutamate transporter, Primary transporter, Secondary transporter, Extracellular gate, Intracellular gate, Coupling, Dipole moment,  $\text{Na}^+/\text{betaine}$  symporter (BetP), Binding pocket, Binding site, Solvent-accessible, Putative binding site, Unbinding pathway

---

## 1. Introduction

Dynamical description of macromolecular systems is fundamental to our understanding of molecular biology. Complementing experimental biophysical methodologies such as single-molecule measurements and other spectroscopic methods, molecular dynamics (MD) offers a powerful method for capturing functionally relevant conformational dynamics of biomolecular systems such as proteins and nucleic acids at various spatial and temporal resolutions, sometimes unparalleled by experiments. In this chapter, we describe several examples of the application of the method to studying membrane transporters. These molecular machines mediate one of the most fundamental processes in all living cells, that is, transport of materials across the barrier of the cellular membrane in a selective and efficient manner. These proteins often facilitate the translocation of the substrates against their electrochemical gradients, a process which relies on delicate coupling mechanisms between substrate transport and various sources of cellular energy exploited by the transport mechanism.

The transport cycle in membrane transporters is usually composed of an unknown number of steps, often resulting in very slow turnover of the process, and therefore, long time scales that are beyond the reach of traditional MD simulations. These steps usually include protein conformational changes of different natures and magnitudes that are triggered by various molecular events, for example, binding, unbinding, and stepwise translocation of the substrate and its cotransported ions, hydrolysis of ATP, and hydration and dehydration of various compartments within the transporter protein. Specific protein conformational changes coupled to these events are at the heart of the mechanisms employed for energy-coupling and efficient transport by membrane transporters.

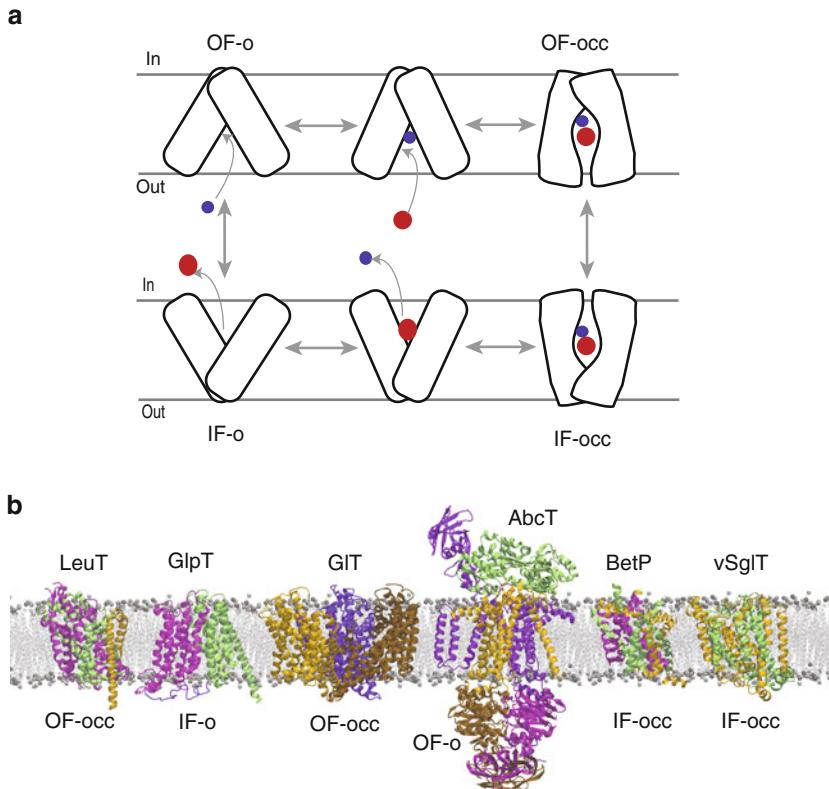
Membrane transporters couple various sources of cellular energy to vectorial translocation of their substrates, often against the chemical gradient (*pumping*). The required energy is provided by ATP, redox reactions, or photons in *primary* transporters, while secondary transporters couple substrate transport to cotransport (*symport* or *antiport*) of ionic species (most prominently H<sup>+</sup> and Na<sup>+</sup> ions). The energy-coupling mechanism relies on a complex set of molecular processes within the transporter protein engaging many of its structural elements. The dynamics of membrane transporters is therefore highly relevant to their function. Relevant structural motions of membrane transporters span a wide spectrum of scales, ranging from localized side-chain conformational changes, to loop flipping motions, and up to extensive subdomain/domain structural transitions. MD simulation offers a method with sufficient temporal and

spatial resolutions to characterize functionally relevant molecular events in proteins. Employing MD simulations, our lab has described several functionally relevant motions in membrane channels and transporters (1–9).

A widely accepted model for the function of membrane transporters is the “alternating access mechanism” (10). This model ensures that the substrate is only accessible from one side of the membrane at any given time, thus, preventing the formation of a *leak*, that is, a channel-like structure that would allow free diffusion of the substrate or other molecular species, during the transport cycle. The alternating access model requires at least two major conformational states of the protein, namely, the inward-facing (IF) and outward-facing (OF) states, whose interconversion switches the substrate accessibility from one side of the membrane to the other (Fig. 1a). The complete transport cycle naturally involves many other intermediate states; for instance, it has been demonstrated for a number of transporters that both the IF and OF states can exist in either open (IF-o and OF-o) or occluded (IF-occ and OF-occ) states. Due to various technical difficulties, however, for most transporters, only one of these major conformational state has been experimentally characterized. The other state(s), and the conformational changes involved in their transitions, therefore, have to be studied using other methodologies that would yield a dynamical description of these complex proteins.

The complexity and time scale of the function of membrane transporters pose a great challenge also to computational studies. Simulation of membrane transporters requires the inclusion of the embedding lipid bilayer, water, and ions explicitly in the system, as often these elements directly participate in the mechanism. This usually results in large system sizes (100,000–500,000 atoms), requiring large computational resources to simulate. More importantly, characterizing the complete transport cycle in transporters would require simulations on the order of at least  $\mu$ s-ms time scales, which cannot be achieved routinely based on the current computational resources available. Despite these technical limitations, we have demonstrated that MD simulations of membrane transporters when designed and analyzed properly can be very effective in describing key molecular events and processes involved in their function (1–9, 11–29). These studies provide ample evidence for MD simulations to be a powerful method in significantly advancing our understanding of the molecular mechanisms of energy coupling and transport phenomena in membrane transporters.

In this chapter, we describe the application of MD to investigating the mechanism of function of several membrane transporters. These systems include the glutamate transporter (GltT), ATP-binding cassette transporters (ABCT), glycerol-3-phosphate transporter (GlpT), leucine transporter (LeuT), the  $\text{Na}^+$ -coupled galactose



**Fig. 1.** (a) The alternating access model proposed for the transporter function, including two major open states, outward facing-open (OF-o) and inward facing-open (IF-o), and two intermediate substrate-occluded states (OF-occ and IF-occ). “Out” and “In” represent the extracellular/periplasmic and cytoplasmic sides of the membrane, respectively. The access to the substrate (*blue circle*) from the two sides of the membrane is controlled by protein conformational changes of different nature and magnitudes. (b) The six membrane transporters discussed in this chapter along with the conformational state captured in the shown structure: leucine transporter (LeuT), glycerol-phosphate transporter (GlpT), glutamate transporter (GIT), ATP binding cassette transporter (AbcT), betaine transporter (BetP), and the  $\text{Na}^+$ /galactose transporter (vSGLT). The proteins are shown embedded in a lipid bilayer (grey), representing the typical simulation setup used for their simulation.

transporter (vSGLT), and the betaine transporter (BetP). These systems are shown in Fig. 1b. While each section will naturally spend some time describing some of the major discoveries made by these simulations in order to convince the reader of the power of the method, we will dedicate a substantial fraction of our discussion to describing the design rationale and strategies adopted to address certain functional questions and to the protocols adopted for the simulations and their analysis. Thus, in each section, we will emphasize and describe in some detail a number of specific simulation and analysis methodologies that have been employed to the particular problem at hand. We hope that the reader will find such expanded descriptions useful for their own specific problems.

---

## 2. Methods

In all the systems, the simulation setup uses experimentally solved, atomic-resolution protein structures, while water, lipid (membrane), and ions are added by modeling. The program NAMD (30), an efficiently parallelized code for MD simulations of biomolecules, was used for all the simulations. The system temperature is maintained at 310 K via Langevin dynamics with a damping coefficient,  $\gamma$ , of  $0.5 \text{ ps}^{-1}$ . The pressure is maintained at 1 atm using the Langevin Nosé-Hoover method (31, 32). The simulations employ particle mesh Ewald (PME) (33) for the calculation of long-range electrostatic forces without truncation. Missing parameters and topology files for ligands are included either by adopting similar parameters from the available force field or by quantum mechanical calculations in a manner consistent with the employed force field.

The protocols used for the simulations involve (a) placing the transporter in the membrane; (b) performing short MD simulations (typically 0.5 ns) on the acyl chains of the lipid molecules while constraining other atoms, in order to “melt” (introducing disorder into) the lipid tails; (c) restrain the protein atoms while equilibrating the membrane and the solvent, also typically 0.5 ns; (d) remove the protein restraints, allowing the adjustment of protein and membrane lipids during further equilibration steps, typically on the order of  $\sim 5$  ns; and (e) the production simulations. Some of the simulations reported in this chapter used an earlier version of the CHARMM force field (34) for lipids, which had a tendency to overshrink lipid bilayers (smaller area per lipid values when compared to experimental values). To prevent the membrane from excessive shrinkage during the simulations which could potentially affect the protein conformation, a constant-area constraint was imposed on the system during the production phase. In the latest revision of the CHARMM force field (C36) (35), this issue has been largely addressed, allowing one to perform simulations using routine constant pressure algorithms. We also note that when simulating anisotropic systems such as lipid bilayers, it is necessary to decouple the x and y components from the z component (the membrane normal) during pressure adjustments, in order to allow them to change independently. The initial equilibration period typically ranges between 1–5 nanoseconds (ns), while the production simulations are carried out for 50–200 ns. As will be demonstrated, on these time scales, we have been able to capture not only fast motions such as water and ion diffusion, side-chain reorientations, and loop movements, but also functionally relevant larger-scale domain motions of the studied transporters.

### 3. ATP-Driven Transport in ABC Transporters

ATP is a universal energy source in all cells, and a large number of “primary active transporters” are fueled by ATP. Among several families within this category, the ABC transporter family is the most abundant and exhibits the greatest versatility of substrate specificity. Also, many ABC transporters are highly biomedically relevant, for example, several are involved in the multidrug resistance in cancer cells, and some are directly responsible for diseases such as cystic fibrosis.

Members of the ABC transporter family include both importers and exporters, that is, some ABC transporters move their substrates into the cytoplasm, while others pump the substrates out of the cell. Despite different directions of substrate translocation, all ABC transporters share the same transport mechanism—an alternating access model—in which the substrate is carried across the membrane during the conformational change of the transporter between an extracellular accessible (OF) state and a cytoplasmic open (IF) state (Fig. 2).

All ABC transporters share the same structural architecture: two MSDs forming the pathway for substrate transport and two cytoplasmic NBDs that bind and hydrolyze ATP to provide energy for the active transport. ABC importers, additionally, are equipped with one or more SBPs at the extracellular side to aid the substrate acquisition. The basic two-plus-two architecture has been seen in many crystal structures in recent years (36–47), which were captured in several different conformations and are generally in agreement with the alternating access model (Fig. 2).

Based on the conformations revealed by these crystal structures, as well as numerous biochemical and biophysical studies of different ABC transporters, it has been established that the transport is controlled by the opening and closing motions of NBDs: the open NBDs are coupled to the inward-facing MSDs, whereas the closure of NBDs results in the MSDs to open toward the extracellular side (48, 49). ATP plays the central role in these processes, in that the binding of ATP causes the NBDs to dimerize, while its hydrolysis opens the dimer. Each of the NBDs contains two essential, highly conserved ATP-binding motifs, each contributing to half of the nucleotide binding sites. Therefore, upon NBD dimerization, the two ATP molecules are both “sandwiched” in between two ATP-binding motifs, each provided by one NBD monomer. However, many ABC transporters were shown to have only one ATP hydrolysis per transport (50–52); hence, some interesting questions were posed regarding the doubly ATP-bound NBD dimers: how ATP hydrolysis triggers the dimer opening, and whether both ATP molecules are required to be hydrolyzed for the opening.

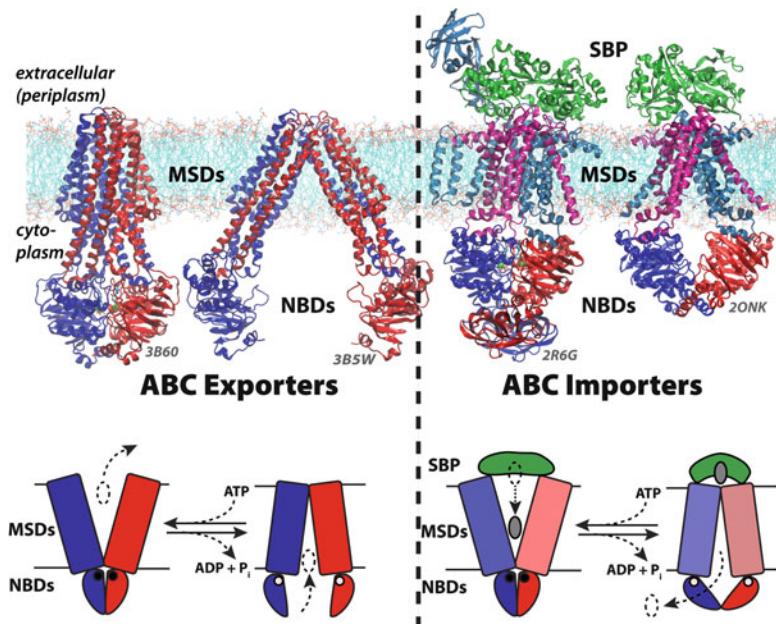


Fig. 2. Structure and mechanism of ABC transporters. ABC transporters can be divided into exporters (*left*) and importers (*right*) based on their directionality and their transport mechanisms. The architectures of both ABC exporters and importers are similar: two nucleotide binding domains (NBDs, half-ovals in the schematic representations) and two membrane-spanning domains (MSDs, rectangles in the schematic representations), with exception of a substrate binding protein (SBP, irregular shaped) only attributes to the importers. Many ABC transporters have been crystallized, of which two homologous structures in each subtype (with the PDB IDs attached below their structures) were selected to demonstrate their two major conformations (extracellular-open and the cytoplasmic-open, i.e., OF and IF states). The interconversion of the two conformational states is controlled by the dimerization and opening of the NBDs, a process dependent on the presence or absence of ATP and Mg<sup>2+</sup> (the *filled* or *open circles* in the schematic representations) in the NBDs. Meanwhile, the substrates (*dashed* or *filled ovals*) are transported across the membrane during the conformational changes between the two major states.

To study the ATP dependency of the NBDs, the crystal structure of the ATP-bound, dimeric NBDs of the maltose transporter (53) was adopted as the initial structure of a set of simulations. It should be noted that, in order to capture the NBDs in their ATP-bound, dimeric form, the Mg<sup>2+</sup> cofactors of ATP were removed from the crystallizing buffer to prevent ATP hydrolysis. Thus, the coordinates of the missing Mg<sup>2+</sup> ions, as well as those for the water molecules bound to the Mg<sup>2+</sup> ions, must be modeled in the simulation systems. As a reference for Mg<sup>2+</sup> coordinates, the crystal structure of an archaeal homolog of the NBD dimer (54) bound with ATP and Na<sup>+</sup> was used, where the position of the Na<sup>+</sup> is consistent with the Mg<sup>2+</sup> coordinates in another NBD structure of ABC exporter (55), as well as the water/ATP coordination of Mg<sup>2+</sup> generally observed in most nucleoside

triphosphate binding proteins. The structure of the NBD dimer was then solvated in a periodic water box with ~200 mM NaCl. Equilibrium simulation was carried out for 8.25 ns after 3,000 steps of energy minimization, during which the coordinating environment around the modeled Mg<sup>2+</sup> was closely monitored to ensure stable binding.

To simulate the effect of ATP hydrolysis on the NBD conformations, the bound ATP was converted to an ADP and an inorganic phosphate. The phosphate species chosen here was HPO<sub>4</sub><sup>2-</sup> (P<sub>i</sub><sup>2-</sup>) instead of H<sub>2</sub>PO<sub>4</sub><sup>-</sup> (P<sub>i</sub><sup>-</sup>). Although the pK<sub>a</sub> between the two species of P<sub>i</sub> is around the neutral pH, here it was assumed as a nascent phosphate from ATP hydrolysis and was considered still in close proximity of the Mg<sup>2+</sup> ion, which is known to increase the pK<sub>a</sub> of its bound species significantly. Also, due to the presence of the bound Mg<sup>2+</sup>, both ATP and ADP in our simulations were modeled in completely deprotonated states. In a modeled post-hydrolytic nucleotide binding site, both ADP and P<sub>i</sub><sup>2-</sup> initially preserve the original coordinates as ATP, thus overlapping at the position where a bridging oxygen atom connects two phosphate groups of ATP (56).

Four simulation systems were created by hydrolyzing either one of the two, both, or none of the ATP molecules, all of which were further simulated for 70 ns following a 3,000-step energy minimization. The simulation results show that the NBD dimer can only exist when both nucleotide binding sites are occupied with ATP and Mg<sup>2+</sup>, and hydrolysis in any of the two sites (or both) triggers the opening of the NBD dimer (56), as clearly demonstrated by monitoring the distances between the two NBDs (Fig. 3a). However, the effect is not instantaneous, and a lag exists between the hydrolysis and the major dimer separation event. Moreover, the time of the dimer opening varies in each system, suggesting that the dimer opening is a stochastic process, in which a certain criterion needs to be fulfilled by a combination of several random events before the major phenomenon can take place. Further examination of the simulation trajectories pointed out that several transient hydrogen bonds between one NBD monomer and the P<sub>i</sub><sup>2-</sup> prevent the dimer from prompt opening after ATP hydrolysis, and the simultaneous rupture of all of them accounts for the sudden opening to a greater degree later on. The simulations here exemplify a typical scenario in which only long enough trajectories are able to capture the desired molecular events. For example, had those simulations been less than 30 ns, no dimer opening could be observed and different conclusions might have been made. Similar simulations performed by several other groups all generally agree that the removal/replacement of ATP in the nucleotide binding sites results in the separation of the NBDs (57–60). Our simulations, in addition, show that the NBD dimer opening can be induced by a single hydrolysis in the two nucleotide binding sites, and the opening occurs before the dissociation of either hydrolysis products (ADP or P<sub>i</sub><sup>2-</sup>) (56).

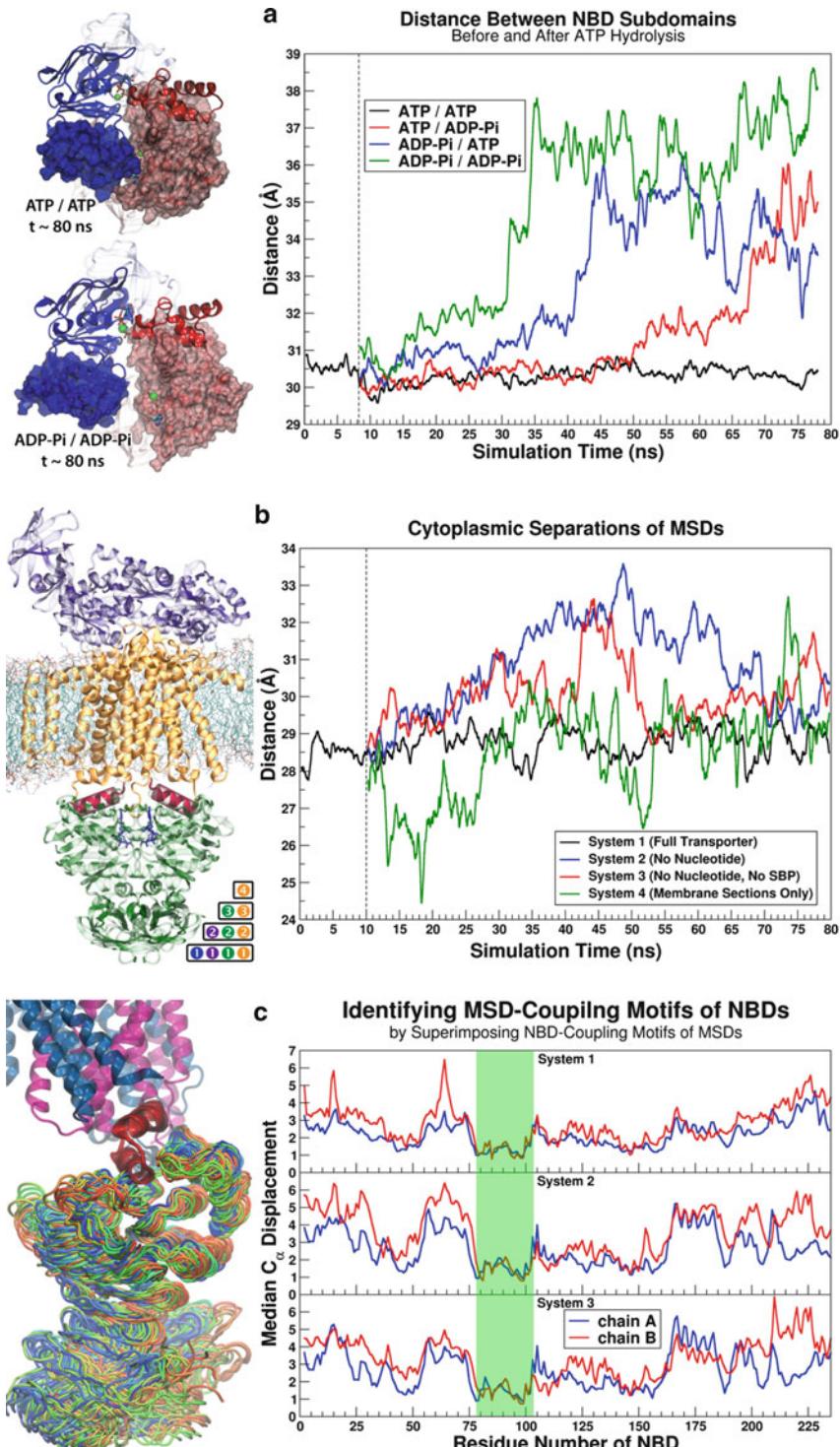


Fig. 3. Conformational dynamics of ABC transporters captured by MD simulations. (a) The separation of NBD monomers after ATP hydrolysis. Shown on the *left* are the NBD conformations (viewed from the periplasmic side) at the end of two representative simulations, where both ATP molecules were either both retained or both replaced by ADP and  $\text{Pi}^{2-}$ . Shown on the *right* are the distances during all four simulations (two heterogeneous systems with one ATP and one set of ADP- $\text{Pi}^{2-}$  in different nucleotide binding sites, besides the two homogeneous systems described above) between two specific NBD subdomains (both shown in surface representations on *left*), highlighting the degree of NBD openings at one side of the dimer. The doubly ATP-bound

The next important question regarding the transport mechanism is how the conformational changes occurring in the NBDs are transmitted to the MSDs. This is especially important for ABC importers: unlike ABC exporters, the NBDs and MSDs of ABC importers are usually coded in different genes rather than expressed as a single polypeptide chain. The conformational coupling between different domains are hard to capture with crystal structures due to their highly dynamic nature. Analyzing the trajectories obtained from MD simulations, therefore, provides powerful means to elucidate the coupling mechanism.

The crystal structure of an intact maltose transporter (42) was used to study the coupling mechanism. Mg<sup>2+</sup> placement was performed as described above, since the transporter was crystallized without Mg<sup>2+</sup>. To simulate the entire transporter together in the membrane environment, the general protocol of membrane system construction described in previous section is employed.

Four sets of simulations were carried out to investigate the NBD-MSD coupling (61): (1) an intact maltose transporter bound with ATP, Mg<sup>2+</sup>, and maltose; (2) a nucleotide-free transporter, that is, removing ATP and Mg<sup>2+</sup> from system 1; (3) a nucleotide-free transporter without the SBP—further removing the entire periplasmic section of the structure; and (4) a reduced model of the transporter containing only the membrane-spanning sections. Each system was simulated for 70 ns. It is found that the ATP-bound transporter (system 1) maintains its conformation, whereas in the absence of the nucleotide (systems 2 and 3), the NBDs open up and trigger the separation of the MSDs at their cytoplasmic ends, as measured by the change in distance between two conserved motifs of MSDs at their cytoplasmic ends. When the NBDs are absent (system 4), on the other hand, the cytoplasmic ends of the MSDs fluctuate greatly but exhibit little net movement (Fig. 3b). This implies that the conformational changes in the MSDs require the presence of the NBDs.

---

Fig. 3. (continued) simulation shows lowest degree of inter-subdomain distances, while other simulations (all with one or both ATP hydrolyzed) exhibit a major opening event at a certain time point. (b) The separation of cytoplasmic structural motifs of MSDs during simulations. Shown on the *left* is the compositions of the four simulation systems as the combination of differently colored domains listed at *lower right*: System 1, all parts; System 2, everything except the nucleotide (*blue ball-and-sticks*) in the NBDs; System 3, NBDs and MSDs excluding their extracellular parts (*green, red and orange* structures); System 4, only the membrane spanning sections (*red and orange* structures). On the *right* are the distances between the *red* helices during the four simulations. The two *red* helices maintain a relatively stable distance in System 1, whereas exhibit a slight separation in Systems 2 and 3, and fluctuate largely in System 4. (c) Structural analysis in searching for the MSD-coupled motifs of the NBDs. The simulation trajectories are superimposed using a structural motif of the MSDs (colored *red* shown on the *left*), and the relative displacements of the C<sub>α</sub> atoms in the NBDs are measured on the *right*. The region shown least relative displacements (hence highest degrees of MSD-coupling) is highlighted as a box in *green*. The NBD structures on the *left* are shown as the sum in the entire trajectory of System 1, superimposed using the method described above, and colored according to the simulation time when the snapshot is taken (from *blue* to *green* then *red*).

To characterize the coupling mechanism between different domains, one needs information on relative movement within the protein complex. The easiest way is to calculate the correlation between the motions of different residues which can be calculated from the MD trajectories. Lange and Grubmüller provided an easy-to-use computer program for this purpose (62). However, with prior knowledge of the key structural elements involved in the process, one might use faster methods to pinpoint specific interdependent parts, for example, at the domain interfaces, without the burden to deal with the massive data of the cross-correlations among  $\sim 1,900 C_{\alpha}$  atoms in the maltose transporter. Since structurally we know where the NBD-contacting motifs in the MSDs are, detailed cross-domain coupling information can be obtained by superposition of the conserved NBD-contacting motifs at the MSDs in the simulation trajectories, and then comparing the relative displacements of  $C_{\alpha}$  atoms in the NBDs to those in the crystal structure. In such a superimposed trajectory, a highly MSD-coupled residue in the NBDs should display both a lower average deviation and less fluctuations (61).

The data point out that a region spanning  $\sim 20$  residues in the NBDs, which corresponds to two conserved motifs near the interface of the MSDs, exhibits maximal MSD coupling in all NBD-containing trajectories (Fig. 3c) (61). Examining the crystal structure, interestingly, it was found that only one of the two conserved motifs is in contact with the MSD, and the secondary motif couples to the MSD only through a tight association to the primary MSD-binding motif. With the aid of structure-based sequence alignments, this secondary MSD-coupling motif was identified and its structural role was determined for the first time (61). Moreover, it was found that other homologous crystal structures can be superimposed in the same way, and the same region in the NBDs also shows lower relative deviations, indicating the MSD-coupling motifs are indeed conserved within this ABC importer sub-family homologous to the maltose transporter, and is likely a unique feature not exhibited by other sub-families of ABC transporters (61).

---

#### **4. Ion-Coupled Neurotransmitter Uptake by the Glutamate Transporter**

Communication between neurons in the brain is accomplished primarily by neurotransmitters. These chemicals are released into the synaptic cleft by presynaptic neurons in response to electrical activities and then detected and converted back into electrical signals by postsynaptic neurons. For this process to be efficient, the neurotransmitters must be rapidly removed from the synaptic cleft

after release (63). Glutamate is the predominant excitatory neurotransmitter in the brain that plays critical roles in learning and memory. Glutamate transporters (GLTs; also termed excitatory amino acid transporters, EAATs (64)) are membrane proteins in neurons and glial cells that are responsible for clearing the neurotransmitter glutamate/aspartate from the synapse (65, 66).

GLTs belong to the family of secondary membrane transporters, which couple “uphill” translocation of the substrate across the membrane to the energetically favorable flow of ions down their concentration gradient. Recent crystallographic studies (67–75) have greatly promoted our understanding of the structural principles that underlie the function of secondary membrane transporters. However, the crystal structures (67–75) are static pictures and do not provide the dynamic information involved in the transport process. Various questions regarding the fundamental nature of the transport cycle remain unanswered, for example, what structural changes occur in response to the binding of ions, how the binding of ions is coupled to the translocation of the substrate, and what is the sequence of events that lead to the binding of substrate? MD simulations permit the study of complex, dynamic processes that occur in biological systems, such as conformational changes and ion transport in biological systems. MD simulations might provide detailed information on the fluctuations and conformational changes of proteins based on calculations of the time-dependent behavior of a molecular system.

By the cotransport of three  $\text{Na}^+$ , a bacterial GLT homolog ( $\text{Glt}_{\text{ph}}$ ) transports one negatively charged substrate (aspartate) across the membrane during each transport cycle. The crystal structures of  $\text{Glt}_{\text{ph}}$  (73–75), solved in both the OF and the IF states, reveal a trimeric architecture with each monomer composed of eight transmembrane helices (TM1–TM8) and two highly conserved helical hairpins (termed HP1 and HP2, respectively) forming the binding sites for the substrate and  $\text{Na}^+$  ions. The architecture of a single monomer in the OF state (74) along with the two structurally resolved  $\text{Na}^+$  binding sites (73–75) (Na1 and Na2 sites) is shown in Fig. 4 a–c. The crystal structures, however, do not provide any information on the third  $\text{Na}^+$  binding site (Na3). A mutagenesis study (76) on a mammalian GLT has shown that a conserved acidic residue corresponding to Asp312 in  $\text{Glt}_{\text{ph}}$  is involved in the coordination of one of the  $\text{Na}^+$  ions during the transport cycle, suggesting that this residue is likely involved in the Na3 binding site.

Transporter-mediated transport of solutes across the membrane involves the alternating access mechanism by which the substrate binding site is alternatively exposed to the extracellular and intracellular sides via large-scale conformational changes of the transporter (10). The substrate and ion binding sites are flanked by two gates: one (extracellular gate) controlling access to the outside

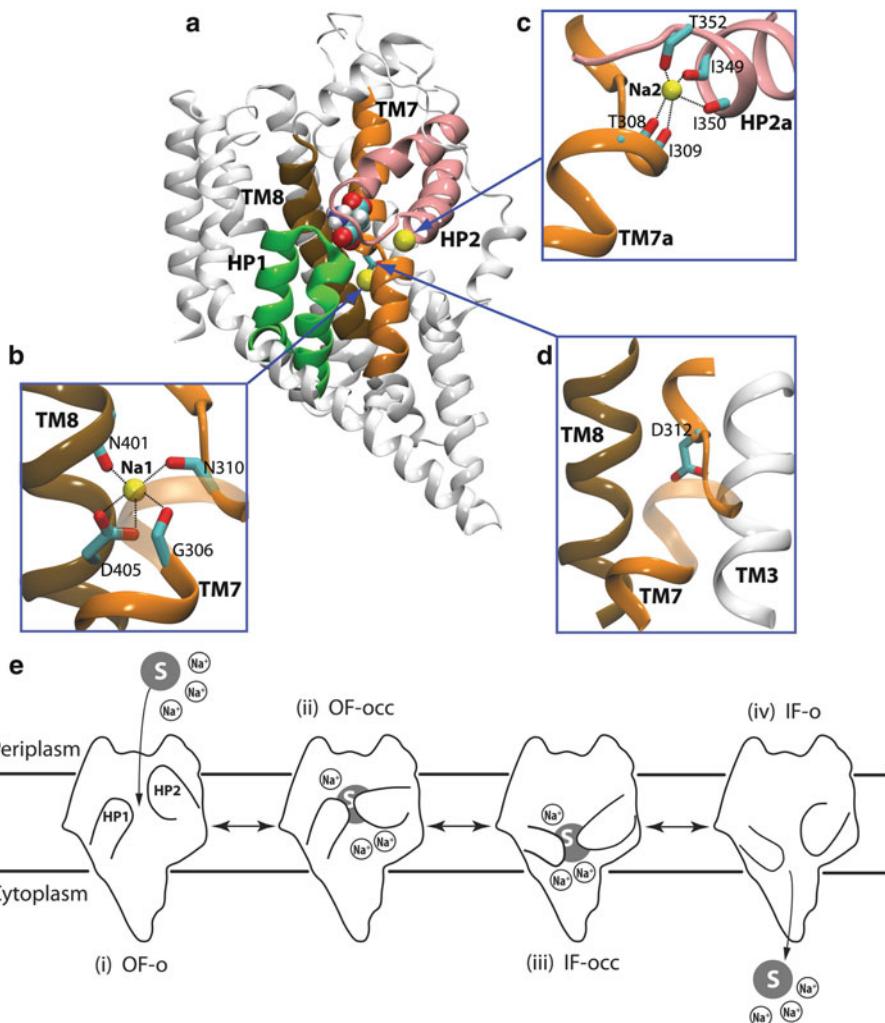


Fig. 4. Structural features and hypothetical transport cycle of Glt<sub>ph</sub>. (a) The structure of Glt<sub>ph</sub> monomer with bound substrate (shown in VDW) and the two structurally resolved Na<sup>+</sup> ions (at sites Na1 and Na2) (spheres). Helical hairpins HP1 and HP2, and transmembrane helices TM7 and TM8 together form the substrate and Na<sup>+</sup> binding sites. (b) The Na1 site. (c) The Na2 site. Note that the residues involved in the Na<sup>+</sup> sites are labeled in the figure. (d) The Asp312 neighborhood (potential Na3 site). (e) Schematic transport cycle: (i) the apo state with HP2 in an open conformation; (ii) binding of three Na<sup>+</sup> ions and the substrate induces the closure of HP2, yielding the OF-occ state; (iii) the IF-occ state is formed by large-scale conformational changes in the protein that switches the accessibility of the binding site to the cytoplasmic side and (iv) opening of the cytoplasmic gate(s) allows the release of the Na<sup>+</sup> ions and the substrate into the cytoplasm.

of the cell and the other (intracellular gate) controlling access to the inside. Only one of these gates can open at a time, allowing substrates and ions to reach the binding sites. Therefore, how the gates work (gating mechanism) is a fundamental question to understand how secondary transporters work. What principles govern the alternating opening and closing of the extracellular and intracellular gates when substrates bind from the outside and unbind to the inside? To address these questions, the conformational changes that occur during transport must be determined.

Taking GIT as an example, here we elucidate how to investigate some of these questions by MD simulations. The binding of substrate/ $\text{Na}^+$  ions to GIT induces local conformational changes (76–78), thus facilitating other species to bind to the protein. That is, substrate binding and  $\text{Na}^+$  ions binding to GIT are coupled. Although the binding processes of substrate and  $\text{Na}^+$  ions cannot be captured by MD simulations, we can perform various simulations, in which the substrate and/or  $\text{Na}^+$  ions is/are placed into their binding sites in GIT, to study how the binding of  $\text{Na}^+$  ions is coupled to each other and the binding of the substrate by investigating the substrate/ $\text{Na}^+$ -induced conformational changes. In order to investigate the mechanism of the extracellular gate and the coupling between binding of the substrate and three  $\text{Na}^+$  ions from the extracellular side to  $\text{Glt}_{\text{ph}}$ , we have performed a series of MD simulations on membrane-embedded trimeric models of the OF  $\text{Glt}_{\text{ph}}$  (74) with different combinations of the substrate and the bound  $\text{Na}^+$  ions (5, 8) to characterize the specific conformational changes and to investigate the dynamics and energetics of the binding sites of the substrate and ions. The simulation system was constructed by embedding the crystal structure of the outward-facing, occluded (OF-occ) state of the trimeric  $\text{Glt}_{\text{ph}}$  into a POPE membrane surrounded by water molecules and neutralized with NaCl.

Comparison of the dynamics of the substrate-bound and the substrate-free (*apo*) simulations clearly indicates that the helical hairpin HP2 plays the role of the extracellular gate (5). Invariably in all the simulations performed in the presence of the substrate, HP2 stays closed (displays a very stable conformation) while removing the substrate results in its large opening motion and complete exposure of the substrate binding site to the extracellular solution (Fig. 5a) (5). These results suggest that HP2 plays the role of the extracellular gate, and that, more importantly, its opening and closure are controlled by substrate binding (5). A gating role for HP2 is supported by the structure of  $\text{Glt}_{\text{ph}}$  in the presence of an inhibitor (74). An independent MD study has also provided support for this role of HP2 (13). Interestingly, despite its apparent structural symmetry to HP2, helical hairpin HP1 was found to exhibit a high level of conformational stability regardless of the presence of the substrate (Fig. 5a) (5). This result, which might be attributed to the shorter length of the loop of HP1, suggests that, at least during the extracellular half of the transport cycle, HP1 does not play a functional role.

The second major consequence of substrate binding revealed by the simulations is the formation of a  $\text{Na}^+$  binding site (the Na2 site) (5) at a position between two half-helical structures (Fig. 5a). In the *apo* state, the dipole moments of these half-helices were found to be completely misaligned (Fig. 5a). Upon substrate binding, these two opposing half-helices align such that their dipole

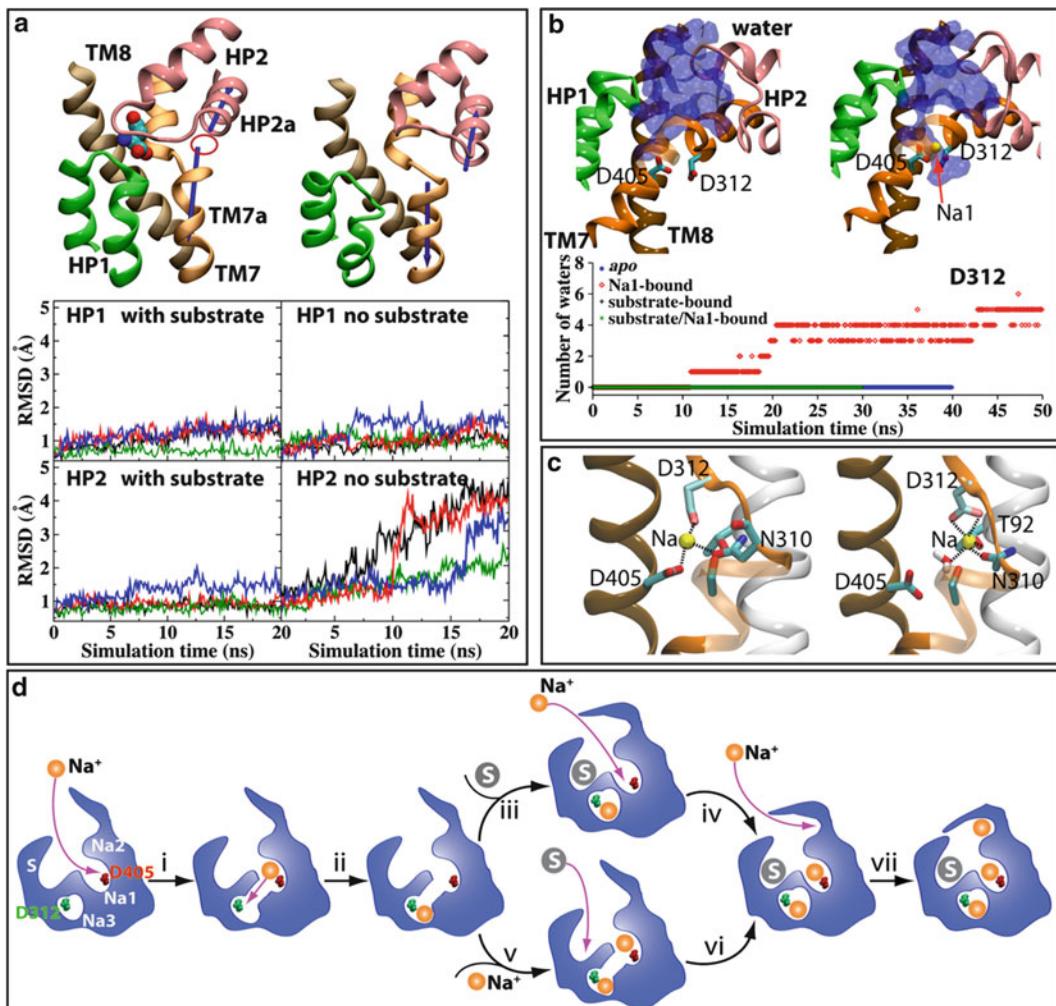


Fig. 5. Dynamics of the extracellular gate and the sequence and coupling between substrate and three  $\text{Na}^+$  ions in Glt<sub>ph</sub>. (a) Dynamics of the extracellular gate in Glt<sub>ph</sub>. Left and right panels show the results of the simulations performed in the presence and in the absence of the substrate, respectively. (left, top panel) Substrate-bound state. Closing of HP2 (the extracellular gate) and formation of the Na2 binding site (marked with an open circle). (right, top panel) Substrate-free state. Opening of HP2 and exposure of the binding site. Note significant misalignment of the dipole moments of helices TM7a and HP2a (arrows). The bottom panel of (a) shows the time evolution of the RMSDs of the helical hairpins HP1 and HP2 in the presence and absence of the substrate. (b) Water accessibility of Asp312 in the apo state (left, top panel) and Na1-bound state (right, top panel), and time series of the number of water molecules within 3.5 Å of the carboxylate groups of Asp312 (bottom panel) in various bound states. The only system in which Asp312 can be accessed by extracellular water is the Na1-bound state. (c) The two sites for Na3. (left) The intermediate site and (right) the putative Na3 site. The residues involved in the sites are labeled in the figure. (d) Schematic mechanism and sequence of binding of the substrate and three  $\text{Na}^+$  ions to Glt<sub>ph</sub>. Na1, Na2, Na3, and S in the first figure mark the position of the Na1, Na2, Na3, and substrate binding sites, respectively. (i) One  $\text{Na}^+$  binds to the Na1 site in the apo transporter resulting in the exposure of Asp312. The same ion likely moves to the putative Na3 site (ii) Next, either the second  $\text{Na}^+$  binds to the Na1 site followed by substrate binding (iii-iv), or the substrate binds to the  $\text{Na}^+$ -bound transporter followed by binding of the second ion to the Na1 site (v-vi), resulting in the half closure of the extracellular gate and the formation of the Na2 site. (vii) One  $\text{Na}^+$  binds to the Na2 site, completing the binding process and resulting in the formation of the OF-occ state.

moments converge on the same region resulting in the formation of the Na<sub>2</sub> site (Fig. 5a). These results strongly suggest that Na<sub>2</sub> binding can only take place after binding of the substrate.

Based on the idea that a Na<sup>+</sup> site has to be accessible by water from outside before it can be reached by a Na<sup>+</sup> ion and the fact that MD simulations can efficiently capture the rapid dynamics of water, we employed MD simulations to investigate the position of the third Na<sup>+</sup> binding site (Na3) in Glt. By monitoring the solvent accessibility of the key acidic residues in the protein and how they are affected by different binding events (8), it was shown that extracellular water can easily access the Na<sub>1</sub> site in all the states (Fig. 5b), indicating the presence of a diffusion pathway for the extracellular Na<sup>+</sup> ions to access the Na<sub>1</sub> site in the *apo* state. However, only in the Na<sub>1</sub>-bound form is the putative Na3 site accessible by extracellular water (Fig. 5b). Binding of the substrate results in occlusion of this site from the water. Therefore, it might be concluded that Na3 binds to the transporter after Na<sub>1</sub> but before the substrate (8). Combined with the results of other simulations, we propose a model for the sequence of the binding events in Glt<sub>ph</sub> (Fig. 5d).

In order to probe the exact binding mode of a Na<sup>+</sup> ion in the putative Na3 site, we took advantage of the position of the water molecules that hydrate this region during the simulations. A series of MD simulations were performed, in which individual water molecules were randomly selected from the pool of water molecules hydrating this region and were replaced by a Na<sup>+</sup> in order to characterize the dynamics of this placed Na<sup>+</sup> ion and the exact site to which it binds (8). Exercising this experiment, we identified two sites to which the ions bind: one characterized to be the putative Na3 site, whereas the second site was suggested to represent an intermediate site (Fig. 5c) (8).

The simulations presented in this section shed light on the mechanism of the opening and closure of the extracellular gating elements in Glt<sub>ph</sub> and the sequence and coupling between the binding of the substrate and the co-transported Na<sup>+</sup> ions from the extracellular side. The mechanisms of the transition between the OF and the IF states and the details of their release from the IF-occ state into the cytoplasm (Fig. 4e) are completely unknown (74, 75) and call for additional simulation studies.

## **5. Rocker-Switch Mechanism in a Major Facilitator Transporter, GltT**

Major facilitator superfamily (MFS) is the largest and the most diverse group of secondary active membrane transporters found in all kingdoms of life. MFS members include a large number of pharmaceutically important transporters (79, 80). Although MFS

transporters show little to no sequence homology, they are structurally closely related. Most of the MFS members consist of 12 transmembrane helices organized into two six-helix halves related by pseudo-twofold symmetry. Only the structures of a small number of MFS members have been experimentally determined (81–84).

GlpT, which is one of the first solved MFS structures (81), facilitates active import of glycerol-3-phosphate (G3P) using counter-transport (antiport) of inorganic phosphate ( $P_i$ ) from the cytoplasm to the periplasm (81). The structure of GlpT (81) has been used to model several homologous proteins (85–88). Similar to other MFS members, it is composed of two six-helix halves, referred to as the N- and the C-terminal domains (Fig. 6a) (81, 86, 89). The two halves show pseudo-twofold symmetry with weak sequence homology. The structure of GlpT manifests a lumen open to the cytoplasm, therefore, representing the IF state of the transporter. Two symmetrically positioned basic residues (Arg45 and Arg269) confer positive electrostatic potential to the apex of the lumen. Along with a conserved histidine (His165), these arginines have been suggested to constitute a putative substrate binding site (Fig. 6b). The protonation state of the binding site histidine is likely to change upon binding of the negatively charged substrate(s), a process that might be, at least partially, responsible for some of the conformational changes induced by substrate in the transporter (81, 86, 89). Mutagenesis studies on GlpT (90) and its close homologs supported the suggested roles for these residues (81, 91).

The alternating access mechanism, which is the generally accepted model for transporter function, has also been suggested for GlpT (Fig. 6c) (92). This mechanism asserts that the accessibility of the binding site from either side of the membrane is alternated during the transport cycle. Based on the crystal structure of GlpT (81), it was suggested that alternating access during the exchange of inorganic phosphate ( $P_i$ ) with organophosphate (G3P) is realized by a “rocker-switch”-like motion of the two halves (81, 86, 89, 93).  $P_i$  binding to the IF state of GlpT triggers transition to the OF state by internal helix motions (94) and/or relative rigid-body rotation of the two halves (95).  $P_i$  is then replaced by G3P on the OF state, switching the transporter back to the IF state (81, 86, 89, 93, 96). Substrate binding was suggested to lower the activation energy of interconversion between the IF and OF states by drawing binding site arginines (Arg45 and Arg269) together (81, 96). A modeling study (95) confirmed that  $\sim 10^\circ$  of rigid-body rotation of each of the two halves can form a functional state, which can conduct substrate to the periplasm (80, 81, 86, 93).

The available structures, especially that of GlpT (81), revealed significant information about the mechanism of MFS transporters. However, the substrate binding site and how the substrate initiates

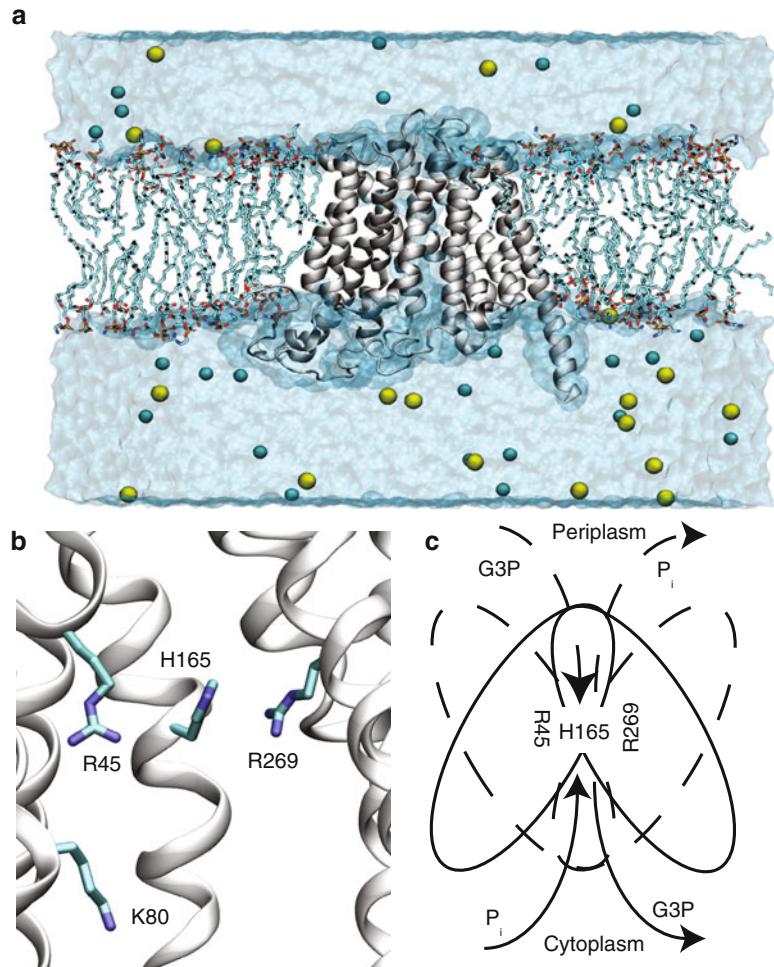


Fig. 6. (a) The simulation system of GlpT showing the protein (cartoon) embedded in a lipid bilayer (licorice) with water (blue surface) and ions (spheres). (b) The putative binding site residues. Two symmetrically positioned arginines (R45 and R269), the histidine (H165) and the lysine (K80) are shown. (c) Schematic representation of the alternating access model of GlpT.  $P_i$  replaces G3P in the IF structure from the cytoplasmic side (solid lines), resulting in formation of the OF structure (dashed lines). G3P binds to the OF structure replacing  $P_i$  and resets the transport cycle. The binding site residues are only accessible from either the cytoplasmic or the periplasmic sides as asserted by the alternating access mechanism.

the transitioning to the other state were not known. To address these questions, MD simulations of a membrane-embedded model of GlpT in the presence of its natural substrates (monovalent or divalent  $P_i$  and G3P) were employed (2, 97). The simulations revealed the unknown substrate binding site (2) and initial substrate-induced conformational changes supporting the rocker-switch model. Not only was the substrate binding site characterized in detail, but also closing at the cytoplasmic side of the lumen

and reorganization of the periplasmic salt bridges upon substrate binding were observed in these simulations (97).

The simulation systems consisted of membrane-embedded GlpT, as well as explicit water and ions (Fig. 6a). Various methods can be employed to predict unknown substrate binding sites and modes, in order to construct a starting configuration for MD simulations. Since the process of substrate binding involves free diffusion of the substrate toward its binding site, usually requiring very long time scales that are inaccessible by MD simulations, docking or non-equilibrium simulations have been used for this purpose. Docking is based on calculating an empirical score for the substrate when placed at different positions in a predefined grid around the protein. Although docking can be fast, a static or semi-flexible representation of the protein is generally used, which can result in inaccurate prediction of the bound state, especially in cases in which substrate binding is expected to induce significant protein conformational changes, for example, in membrane transporters. Pulling the substrate towards its putative binding site, for example, by means of non-equilibrium MD is an alternative approach, but it requires prior knowledge of the binding site, as well as the binding pose of the substrate. Although these methods are the most common approaches and have been used extensively, in a recent study (12), we had shown that there is a good chance for equilibrium MD simulations to capture spontaneous binding of the substrate when the substrate is charged. Once placed in the opening of the lumen, the substrate was indeed recruited into its binding site in a relatively short time (though still long for MD simulations) (12). This approach would provide the least biased and most dynamical description of substrate binding to its site. Given that the substrates are charged in the case of GlpT, we decided to take a similar strategy for identifying its unknown binding site. The substrates were placed at the cytoplasmic mouth of the lumen, and extended equilibrium simulations were performed.

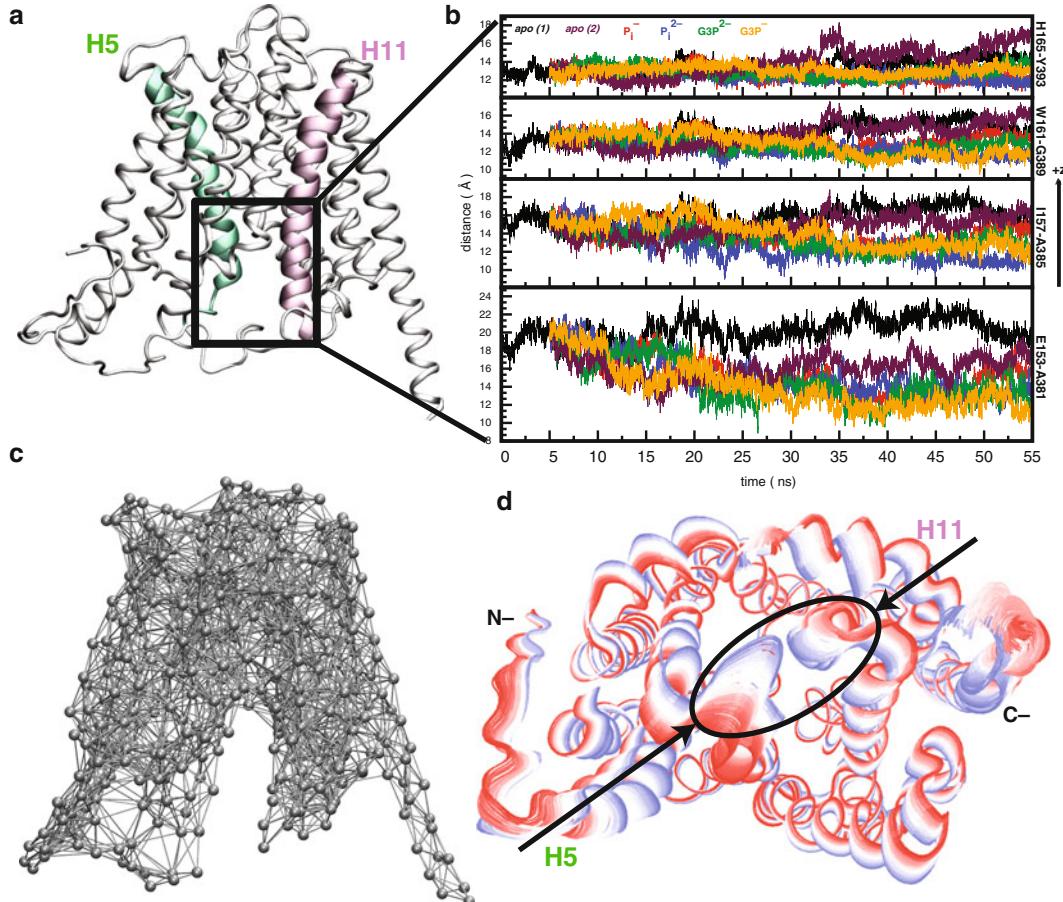
We note that both  $P_i$  and G3P are titratable species, and that the titration state used to represent them in the simulations could have a marked effect on their dynamics. Both  $P_i$  and G3P have  $pK_a$  values around neutral pH, which make both divalent and monovalent states prominent at neutral pH. Moreover, the titration state of the substrate can change during its approach and association to its binding site. Although a dynamic description of the titration state is not present in conventional MD simulations, which impose a fixed topology on the molecular system, one can take advantage of running independent simulations in which different various titration states of the substrates are individually examined. We note that there is an ongoing effort on developing the so-called “constant-pH” simulations which would provide an algorithm to dynamically adjust the titration states of titratable species on the fly (98–106).

The equilibrium “binding” simulations described above were repeated several times using different initial positions and orientations of the substrates in order to test the reproducibility and convergence of the results. The initial positions of the substrates were all chosen to be inside the lumen, since the substrate is expected to approach the lumen of the protein by simple diffusion in the bulk solution, which is not of interest. As a control, the *apo*-GlpT (substrate-free) system was also simulated in order to distinguish between protein conformational changes that are independent of the presence of the substrate from those that are induced by its binding.

These equilibrium simulations successfully captured spontaneous binding of the substrate and identified key residues and events along the translocation pathway of the substrate (97). The substrates form stable interactions with one of the arginines (Arg45), the histidine (His165) in the putative binding site, a lysine (Lys80), and tyrosine residues, forming a cage-like binding pocket (97). Several other interactions that confer higher affinity to G3P through interactions with its glycerol moiety were also identified, and their roles in binding were confirmed by mutagenesis experiments (2). Although it was previously suggested that both apex arginines might be involved in substrate binding (81, 90), our simulations showed that only one of the arginines is involved in the initial binding and that the other arginine might only play a role in the later steps of the transport. The differential roles of the two arginines in initial binding are fully explained based on the luminal electrostatic potential calculated from the MD simulations, displaying a peak around one of the arginines (Arg45) (97). This study exemplifies a successful case in which a more complete description of the substrate binding event and the site was only possible through the dynamic, atomic-scale picture provided by equilibrium MD simulations (2, 97).

The MD simulations capturing spontaneous substrate binding also revealed initial conformational changes induced by substrate binding that are consistent with the alternating access mechanism and result in partial occlusion of GlpT. These structural changes represent early events involved in the formation of the occluded state after substrate binding. Substrate binding results in straightening of two helices (Fig. 7a, b), which are both originally bent in the absence of the substrate, thereby resulting in partial occlusion of the GlpT lumen (97). GlpT exhibited this partial occlusion in all substrate binding simulations, but not in the *apo* simulations, confirming the substrate dependence of the phenomenon (97). Interestingly, the two helices involved in the observed partial occlusion have been also shown to have the highest flexibility among all helices in GlpT (94).

While the observed dynamics in the MD simulations described above clearly described the initial events induced by the binding of the substrate, much larger conformational changes of the protein that are necessary for the complete transition from the IF to the OF



**Fig. 7.** (a) Helices 5 and 11 labeled on the complete structure of GlpT. (b) Distance between the C<sub>α</sub> atoms of representative residues that are approximately on the same plane on helices 5 and 11. The residues on the cytoplasmic side come together in all substrate binding simulations (red, blue, green, and orange), on the other hand, in apo-simulations either closing does not happen at all (black), or the closing is reversed towards the end of the simulation (maroon). (c) Network representation of GlpT used in anisotropic network model (ANM). C<sub>α</sub> atoms (spheres) within a certain distance are connected by harmonic springs (sticks). (d) One of the normal modes calculated by ANM. The picture shows several superimposed structures deformed along the normal modes from positive to negative direction (blue to red). This mode captures helices 5 and 11 coming together.

state are difficult to be obtained from all-atom MD simulations. As an alternative approach, normal mode analysis (NMA) of reduced representations of proteins has been used successfully to predict functionally relevant modes of vibrations in these macromolecular systems (107, 108). The potential energy surface governing the protein dynamics is coarse grained by assumption of harmonicity, namely, stripping of higher order contributions to the potential energy (107, 109), which has been successful in approximating the energy surface around a local minimum. To reduce the number of degrees of freedom and to make the calculation of the normal modes for larger proteins more tractable, reduced representations

of proteins were used in which each residue is represented by a single interaction site (usually the  $C_\alpha$  carbon) and connected by harmonic springs (Fig. 7c) (109–112). Anisotropic network model (ANM) (109–112), which takes advantage of the reduced representation of proteins, has been successfully used in capturing functionally relevant structural changes. The effectiveness of NMA in capturing functional structural changes has been attributed to the assumption that the function of a protein has evolved most likely around its inherent property of undergoing vibrations, and that large-scale protein conformational changes can be best governed by low-frequency modes of vibrations (108). Vibrations around these “soft modes” can make up large deformations at an affordable energy cost. Therefore, conformational changes of a protein at much longer time scales than what can be achieved in MD can be predicted by these modes (108). Although this type of analysis has been fruitful as an inexpensive way of investigating functional dynamics of proteins, identifying the biologically relevant mode(s) is always challenging. Besides, normal modes fail to capture anharmonic nature of the fluctuations in proteins (113).

In order to study long-time scale dynamics of GlpT, for which only the IF state is available, normal modes of GlpT were calculated using ANM. In searching for large global conformational changes of a protein structure, only those modes that engage a major portion of the protein significantly are usually interesting from a functional perspective. Fraction of the protein affected significantly by a certain mode is defined as the “collectivity” of the mode (114). To select functionally relevant modes among the lowest frequency ones, collectivity of the modes can be adopted as a measure. Several collective normal modes that describe the closing motion of the cytoplasmic side, were identified in agreement with the results of the MD simulations (Fig. 7d). The normal modes generally capture swinging of the two halves of GlpT on the cytoplasmic side. This type of motion may result in the formation of the occluded state in which two halves of the protein come together at the cytoplasmic side. On the other hand, the calculated normal modes fail to capture the opening on the opposite (periplasmic) side. This is, indeed, not counterintuitive since the closed periplasmic side is almost held rigid by several harmonic springs gluing the two halves. This may also indicate that the transition from cytoplasmic-open state to periplasmic-open state cannot be represented by a single vibrational mode.

In addition to the partial occlusion of the cytoplasmic side, substrate binding also affects a periplasmic salt-bridge network which we view as an important switch mechanism in the rocker-switch model (90, 97). This salt-bridge network keeps the N- and C-terminal halves together on the periplasmic side (80, 81, 90). It is reasonable to assume, therefore, that breaking or rearranging this network is needed for the opening of the protein on the periplasmic side. The simulations indicated that the substrate significantly

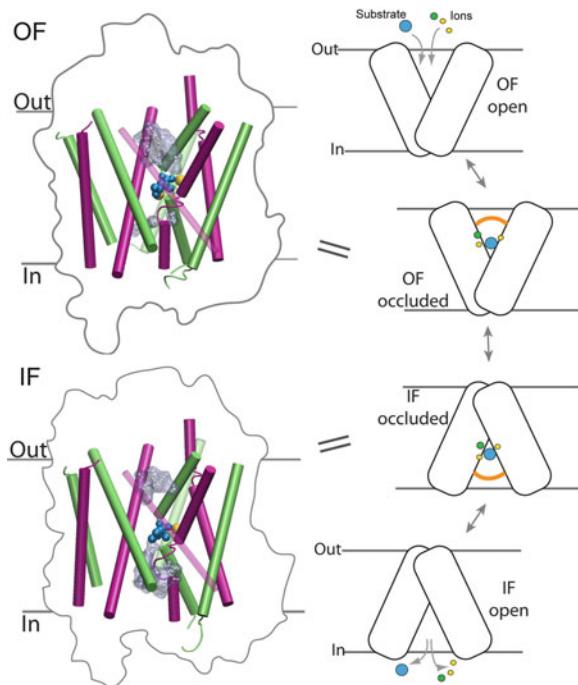
perturbs the interaction of the residues within the salt-bridge network (97), although the periplasmic interface did not peel apart during the simulations. A similar phenomenon was also observed in our simulations of a mitochondrial carrier protein in which binding of ADP to its binding site resulted in significant rearrangement of a preexisting salt-bridge network in the protein (12). Therefore, it appears that at least for charged substrates, perturbation of internal salt-bridge networks in the protein is a common initial step in inducing protein conformational changes, specifically in decreasing the activation energy of interconversion of the IF and OF states in membrane transporters.

In summary, we captured spontaneous substrate binding and, more importantly, structural changes induced by substrate binding in GlpT by equilibrium MD simulations. Some major steps along the “rocker-switch” mechanism resulting in the formation of the occluded state were also identified, namely, partial occlusion of the cytoplasmic mouth of the lumen and substrate-induced reorganization of the periplasmic salt-bridge network. Large conformational changes that involve bending and hinging of domains can, in general, be captured by NMA. Although NMA is a fast and an effective way to obtain insight on protein dynamics, choosing the functionally relevant mode(s) continues to be among the main challenges of applying the method.

---

## 6. Alternating Access Model in Leucine Transporter

Neurotransmitter:sodium symporters (NSS) are a family of specialized transport proteins that move neurotransmitters across the neuronal membrane, deriving energy for the process by coupling it with ion ( $\text{Na}^+/\text{Cl}^-$ ) transport (115). This process is key to clearance of the synaptic cleft of neurotransmitters such as serotonin, dopamine, norepinephrine, and GABA. The proper function of NSSs is therefore central to various physiological processes, for example, sleep, emotions, and social behavior. Abnormality in the function of these transporters may occur due to several reasons including drug abuse or genetic factors, and lead to neurological disorders such as depression, anxiety, and hyperactivity. The NSS family thus constitutes one of the main targets for psychostimulants and antidepressants as well as drugs for other neurological disorders. Elucidating the mechanism of these transporters based on a detailed structural description is important not only to understand the molecular basis of the diseases and disorders associated with their impaired function, but also to design more effective and selective drugs to treat such conditions. While extensive functional data exists for the NSS family, crystal structures are known for only



**Fig. 8.** *Left:* OF and IF states of LeuT. Outlines of the LeuT OF crystal structure (*above*) and IF model (*below*) are shown with its pseudosymmetric 5 + 5 inverted repeat highlighted (TM1 to TM5—green, TM6 to TM10—pink). Substrate (blue), bound Na<sup>+</sup> ions (yellow), and water penetration (grey surfaces) are also shown. *Right:* The alternating access mechanism proposed for transporter function. Substrate (blue) and co-transported ions (Na<sup>+</sup> and Cl<sup>-</sup> in NSS) bind to the transporter in the OF-open state, followed by conformational changes that result in their occlusion (orange curve) from the extracellular solution. Larger conformational changes result in conversion to the IF state, where substrate and ions are initially occluded from the intracellular solution, and eventually exposed, and released. The LeuT structures on the left represent OF and IF occluded states.

one member, that of the LeuT, a bacterial homologue (116). LeuT structures continue to serve as models to interpret as well as obtain, structural and mechanistic information for the NSS family.

Similar to other membrane transporters, LeuT and other NSS members are proposed to function by the “alternating access” mechanism, where large-scale protein conformational changes alternately expose the binding sites for the substrate and the co-transported ions to the exterior or interior of the cell. During these structural changes, the protein adopts various conformational states, including two major open states, namely, the inward-facing open (IF-o) and outward-facing (OF-o) states, as well as two intermediate “occluded” states, IF-occ and OF-occ (Fig. 8). The transport mechanism is believed to include the following steps. The substrate and Na<sup>+</sup> ions bind to the protein in the OF-o state, following which the substrate is occluded from the extracellular

solution, resulting in an intermediate OF-occ state. Subsequently large-scale conformational changes occur that result in closure of the protein on the extracellular side, and opening on the intracellular side, forming an IF state. It is believed that initially, the IF state is occluded, and is followed by an IF-o state where the substrate and ions are exposed to the intracellular solution, and eventually released (Fig. 8). Reported crystal structures for LeuT are in the OF state, thus providing only a partial view of the mechanism of NSS proteins. In order to obtain a complete picture of the mechanism of transport, a description of the IF and intermediate conformational states in the mechanism is necessary. Moreover, crystal structures provide only a static view, and studying the dynamics of these structures is important to pinpoint structural elements involved in the transport function. We have hence employed computational modeling and simulation techniques to model unknown states of the transport cycle and study the dynamics of these states.

The dynamics of the OF state of LeuT have been recently studied by MD simulations (1, 27, 28, 117), providing a description of some conformational changes associated with substrate binding and transport. Some of these studies (1, 27, 117) have also employed biased simulation techniques, namely, steered MD (SMD) (118) and targeted MD (TMD) (119), which allow additional forces and constraints to be included in the simulation system, in order to probe processes that cannot be described by equilibrium MD due to their long time scales. In the first reported SMD study on LeuT (1), forces were applied to pull the substrate through its putative translocation pathway, in order to describe substrate binding/unbinding. The starting point for this study was the OF-occ crystal structure of LeuT, with bound substrate and Na<sup>+</sup> ions (116). As the substrate translocated through the extracellular vestibule, major barriers to unbinding were offered by interaction and/or physical blockage by a bound Na<sup>+</sup> ion, aromatic residues forming the occlusion over the substrate, and a salt bridge. It was seen that a mere rotation of the side chains of the aromatic residues allows opening of the binding site when the substrate exits. The substrate showed a common unbinding pathway in a majority of the simulations, suggesting functional relevance of this pathway. In an independent study, SMD simulations were again adopted to pull the substrate towards the extracellular as well as intracellular side (27). This study revealed two major findings—first, the presence of a second substrate binding site in the extracellular region of the protein, that was supported by experimental studies and suggested to modulate the transport process (27, 120, 121); and second, a putative mechanism for intracellular substrate release. These simulations provided one of the earliest views of the putative substrate translocation pathway in the extracellular pore, and transition from the open to occluded state in an NSS member.

Describing the IF state of LeuT, and the conformational changes associated with transition between OF and IF states would form the next step in studying the NSS transport cycle. In order to investigate the IF state, independent computational studies have been performed to model the IF-o/IF-occ states (117, 122) as well as to simulate the OF-to-IF transition (117). LeuT is composed of 12 transmembrane (TM) helices, of which helices 1–5 and 6–10 form two pseudosymmetric domains, arranged as inverted repeats. Based on the internal symmetry of LeuT, an IF state model was generated where the structures of the two pseudosymmetric domains of LeuT were exchanged (122). Based on these states, the transport mechanism was described in terms of rigid-body motions of a bundle of transmembrane helices in LeuT (122).

The continuing absence of a crystal structure of the IF state for LeuT, as well as that of a dynamic description of this state, spurred a recent project in our group, wherein we developed a computational protocol to structurally model and study the dynamics of the IF state of this protein (schematic in Fig. 9). Interestingly, the typical inverted repeat fold of LeuT has also been observed in several secondary transporters from other families (123), and seeing this as an opportunity to model alternate states of LeuT, we utilized this information to generate IF models (117). The crystal structure of vSGLT, a member of the sugar -sodium symporter family, and exhibiting the LeuT-like inverted repeat fold, but in the IF state (68), was used as the template for building the IF state model. The first challenge presented to modeling undertaken this way, was the lack of high sequence similarity between LeuT and vSGLT. This was addressed by performing structure-based alignment using DALI-Lite (124), that utilizes the similarity between the overall fold of LeuT and vSGLT, to perform a structural as well as sequence alignment. The sequence alignment thus obtained could be used by a threading program such as MODELLER (125), wherein the sequence of LeuT is directed to adopt the IF structure of vSGLT. After performing minor adjustments in the sequence alignment to ensure that the LeuT substrate and  $\text{Na}^+$  binding sites were maintained during the transition from the OF state to the IF state, the threaded structure of LeuT was generated. The next challenge was to address incorrect modeling of regions of LeuT that do not have a direct equivalent (template) in the vSGLT structure, and hence could not be modeled during the threading process. At this stage, we decided to combine the above approach of structure-based alignment and threading, with the simulation technique of TMD. During TMD, the original OF-occ LeuT crystal structure was gradually directed toward the threaded structure of the IF state, by applying biased potentials (forces) only to the inverted repeats, thus preserving the structure of all other regions where there is no structural equivalence between vSGLT and LeuT. In other words, the structure of the region for which a high-quality

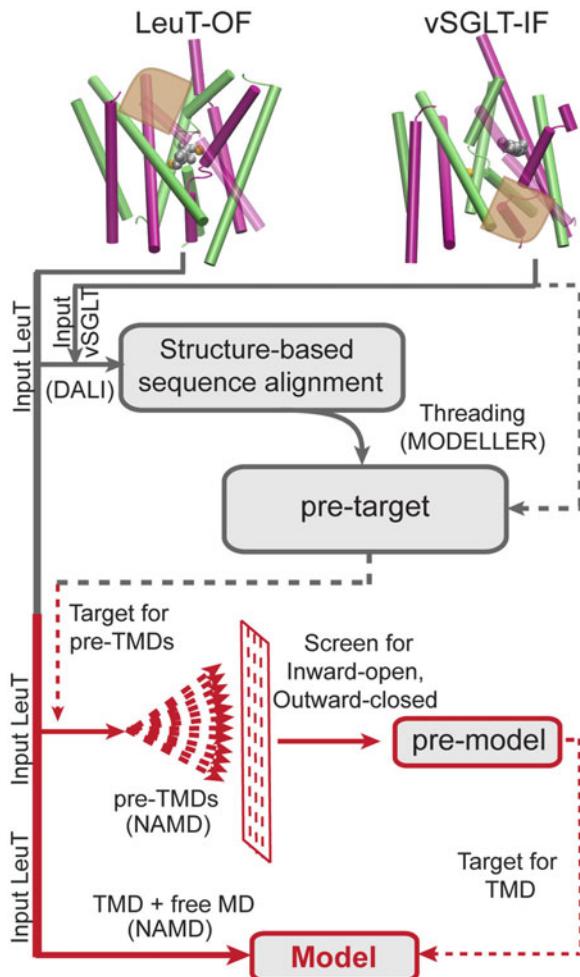


Fig. 9. The computational protocol developed for modeling the inward-facing state of LeuT is presented. The location of the EC and IC halves of the lumen in the LeuT and vSGLT structures, respectively, are marked in brown. The initial phase of structural alignment and threading (grey), and the pre-TMD and final TMD phase (red) resulting in the model, are illustrated. Dotted lines are used for structural input in addition to the LeuT crystal structure.

template existed was deformed toward the template, while the other regions were subjected to free MD and allowed to adapt to the new conformation of the protein. In these TMD simulations, the C<sub>α</sub> atoms of a subset of the helices involved in the inverted repeats were constrained towards the targeted structure, while the remaining parts of the structure, including all side chains, underwent equilibrium MD and adapted to the new conformation, in the presence of a lipid bilayer, solvent, and ions. Long TMD simulations (50 ns each) were used to induce the OF-to-IF transition, followed by 20 ns of free MD for relaxation, finally resulting in models of the IF state of LeuT (117).

The models retained the secondary structural features of LeuT and preserved the substrate/ion binding sites, while adopting an IF state. As a measure of the accessibility of the substrate and ion binding sites, water permeation was monitored with the progress in transition and indicates inward opening. Release of one of the bound  $\text{Na}^+$  ions, an important step in the transport process, was observed during the transition. An interesting revelation was the possible role for pseudosymmetry in LeuT structures; modification of only one of the two pseudosymmetric domains induced an overall transition to the IF state suggesting that each of the two domains may represent a functional unit capable of inducing transition in the full protein. This study thus presented the first detailed description of the IF state dynamics of an NSS homologue and a possibly general transport mechanism in transporters adopting the LeuT fold.

Computational studies of LeuT continue to play a major role in providing missing dynamic and structural information to develop the connection between the OF and the IF states, and possible intermediate states representing alternating access. These studies, coupled with information from new crystal structures of transporters adopting this common fold, and experimental characterization of conformational states of these transporters, have enabled the emergence of a clearer picture of the general transport mechanism of secondary transporters across several families.

---

## 7. Interconversion of the Open and Occluded States in $\text{Na}^+$ -Coupled Galactose Transporter

The simulations described in this section performed on a bacterial  $\text{Na}^+$ -coupled galactose transporter (vSGLT) represent a convincing example of how the results of MD studies can complement experimental structural biological studies by filling the gaps that cannot be always filled with currently available experimental approaches. As a major class of membrane proteins, secondary transporters use energy stored in electrochemical ionic gradients to power the uphill translocation of their substrates across the membrane (126–129). The crystal structure of vSGLT (68) takes a critical step toward a better characterization of the mechanism of secondary transporters. In the crystal structure (68), the  $\text{Na}^+$  binding site could not be unequivocally resolved; however, a  $\text{Na}^+$  ion was modeled in a binding site corresponding to that of LeuT (67) based on structural similarity. Although mutagenesis studies confirmed the involvement of a number of residues from this putative site in  $\text{Na}^+$  binding (68), whether  $\text{Na}^+$  ion is present in the binding site or not in the conformation captured in the crystal remained unresolved. This information is the key to answer an important mechanistic question

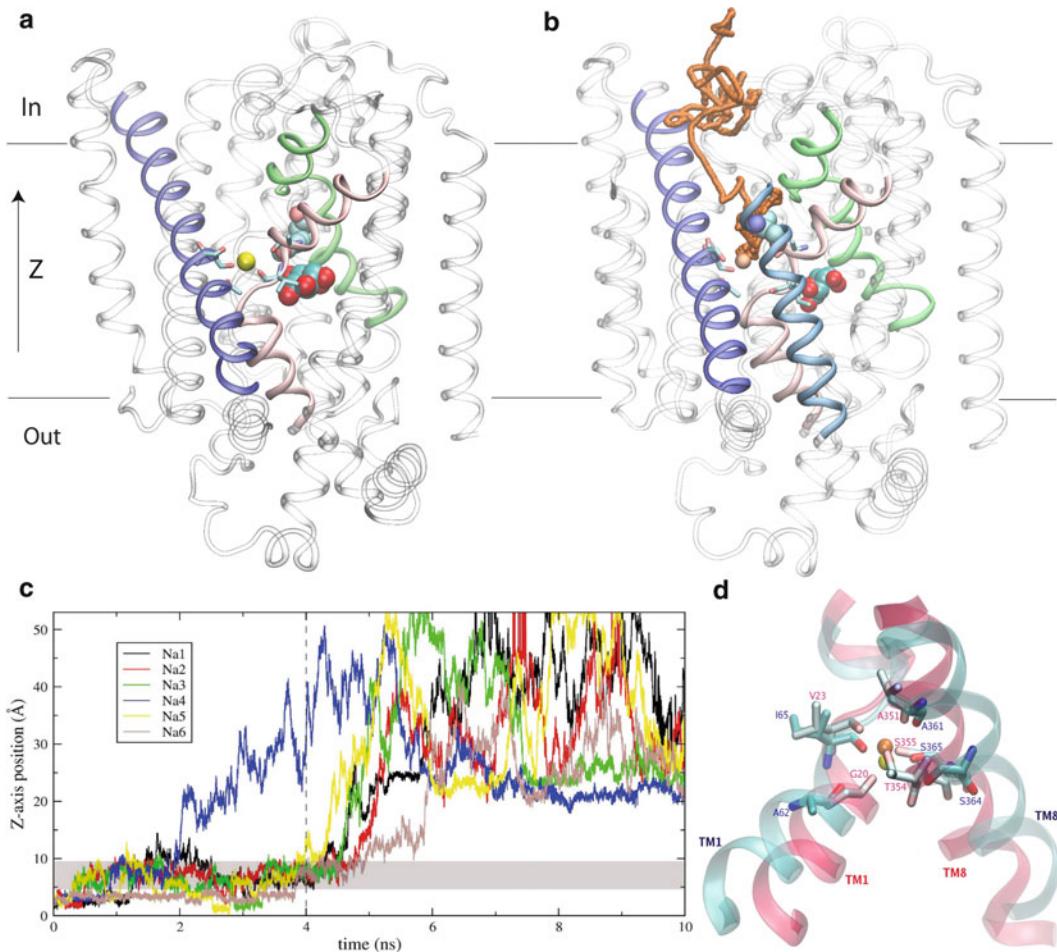


Fig. 10. MD simulation of vSGLT. (a) Overview of the crystal structure of vSGLT. The protein is shown in ribbon representation. The substrate and ion are in van der Waals (vdW). (b) Overview of the release trajectory of the Na<sup>+</sup> ion in MD simulation. Residues in the Na<sup>+</sup>-binding site are displayed as sticks, Asp189 and the substrate are in van der Waals (vdW) representation. (c) Displacement of the Na<sup>+</sup> ion in six repeated MD simulations from its position in the crystal structure. The grey bar highlights the region in the vicinity of Asp189. A rapid Na<sup>+</sup> release can be observed in all of the performed MD simulations. (d) Comparison of the Na<sup>+</sup>-binding sites of vSGLT and LeuT. Alignment was done using TM1 and TM8 helices from vSGLT and LeuT.

regarding the sequence of unbinding events on the cytoplasmic side an aspect that is poorly understood not only for vSGLT but also for a majority of secondary transporters.

To investigate the occupancy of the Na<sup>+</sup> binding site, and to characterize the sequence of unbinding events, we have performed equilibrium MD simulations of a membrane-embedded model of vSGLT (3). In the crystal structure of vSGLT (68), Na<sup>+</sup> is modeled in a site located at the intersection of two transmembrane helices. In this site, several residues corresponding to the residues in the Na<sup>+</sup> binding site of LeuT (67) appear to be optimally positioned to coordinate a Na<sup>+</sup> ion (Fig. 10a). Surprisingly, in all of the

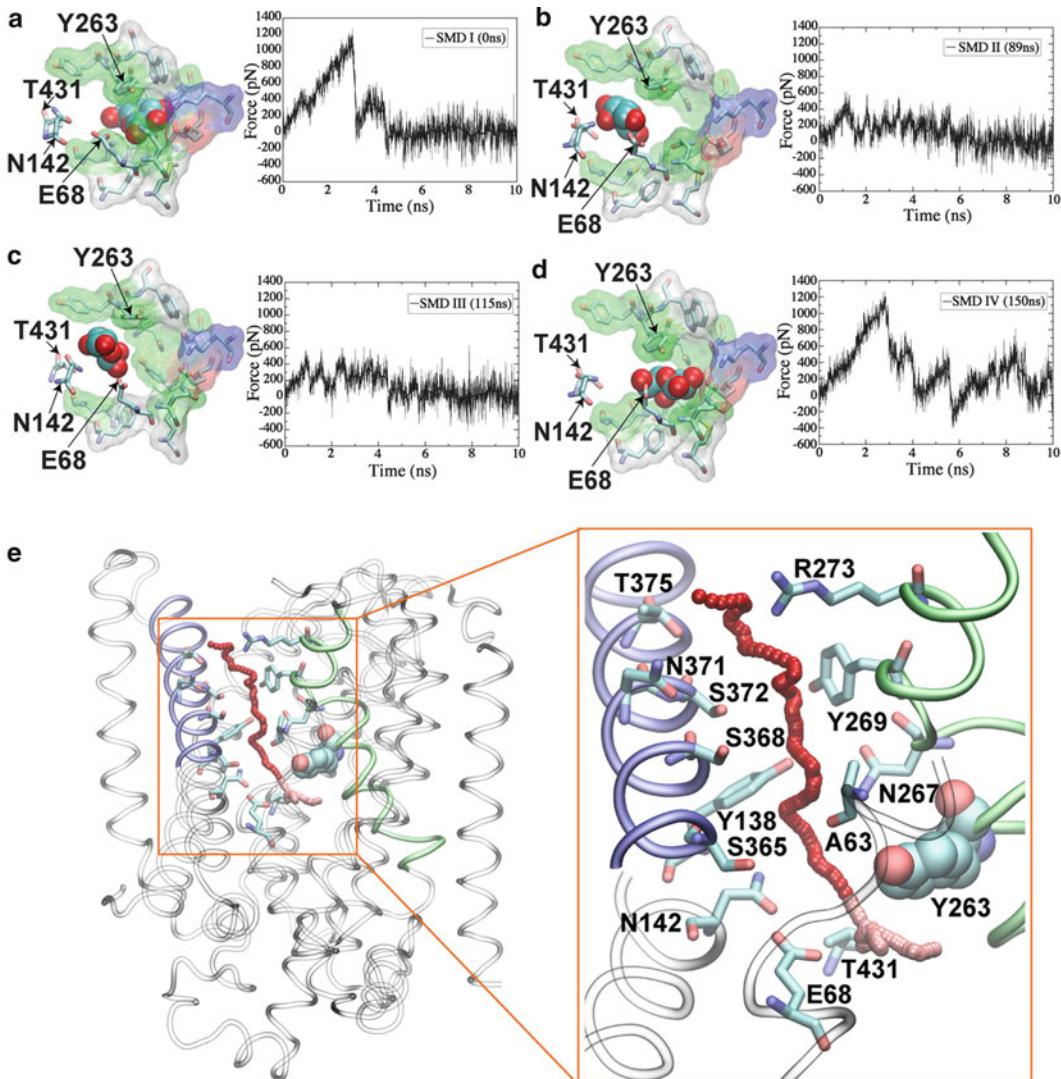
performed MD simulations, we observed a rapid  $\text{Na}^+$  unbinding from the site within a few nanoseconds (Fig. 10c). In addition, in all simulations, the  $\text{Na}^+$  ion was observed to hover around the region close to an aspartic acid after its detachment from the binding site and before completely exiting the transporter (Fig. 10b). Given that in these simulations the protein did not show a major structural change ( $\text{RMSD} \leq 2.5\text{\AA}$ ), the observed spontaneous  $\text{Na}^+$  unbinding suggests that the crystal structure of vSGLT is in an ion-releasing state (3). This result is in sharp contrast to the confirmed presence of an ion in the  $\text{Na}^+$  binding site of LeuT (67), which was further confirmed by MD simulation (1). Close examination and structural alignment of the  $\text{Na}^+$  binding sites in vSGLT and LeuT provide supporting evidence for the observed behavior of the ions in the MD simulations. While at the sequence level, the residues forming the ion binding site of vSGLT are very similar to those in LeuT, there are significant geometrical differences between the two sites. In the crystal structure of LeuT (67), residues coordinating the  $\text{Na}^+$  ion form a typical square pyramidal arrangement with bond distances between the  $\text{Na}^+$  ion and the coordinating oxygen atoms ranging from 2.1 to  $2.4\text{\AA}$ . In vSGLT, on the other hand, all of the corresponding distances are above  $3.1\text{\AA}$ , resulting in an “open” site, which, consistent with the results of the simulations, is not able to hold the ion (3). Thus, MD simulations and characterized differences between the ion binding sites of vSGLT and LeuT successfully capture the type and extent of conformational changes involved in transition between the ion-releasing (open) and the ion-binding (occluded) states in secondary transporters (the terms “open” and “occluded” are usually used with reference to the substrate binding site, but here we use them to describe the  $\text{Na}^+$  binding site).

While MD simulations can capture a wide range of molecular events and processes and even identify certain intermediates, a large number of relevant phenomena are still out of reach by equilibrium MD simulations due to their long time scales. In order to study such processes, one needs to rely on different methodologies. SMD is a computational method that employs external forces to accelerate molecular processes that are too slow to capture in equilibrium MD simulations (118, 130–132). The method is nonequilibrium in nature and therefore should be employed cautiously. However, when applied carefully, it provides an efficient way to describe transitions that are otherwise too slow. In particular, the method is an excellent one for describing processes that rely on random diffusion through flat regions of the potential energy surface, for example, biasing the almost barrier-less diffusion of a substrate in solution toward a certain region. Our simulation study of the complete release of the substrate in vSGLT represents a good example of combining SMD with equilibrium MD simulations to describe the molecular phenomenon at hand (9). In the crystal

structure (68), the substrate is bound at a location about halfway across the membrane and flanked by hydrophobic residues on both the cytoplasmic and extracellular sides. On the cytoplasmic side, a tyrosine (Tyr263) stacks with the pyranose ring of the substrate (galactose), a feature commonly found in sugar-binding proteins (82, 133). This interaction establishes a plug that appears to prevent the exit of the substrate into the cytoplasmic opening of the lumen. Despite detailed characterization of the binding site in the crystal structure, the mechanism of substrate unbinding and its pathway, critical aspects in the cytoplasmic half of the transport cycle, have not been well characterized. Most importantly, it is not known whether a local gating motion (specifically for the side chain of Tyr263) would be necessary for substrate unbinding.

Through the extended equilibrium simulations (~200 ns), we examined whether spontaneous unbinding from the binding site could be captured without biasing the substrate with external forces. Indeed, we were able to capture complete exit of the substrate from its binding pocket on several occasions; however, the substrate did not completely leave the transporter during these equilibrium simulations. To characterize the complete unbinding process of the substrate, we resorted to SMD simulations, however, starting from a configuration in which the substrate was completely out of its binding site.

Starting from this fully unbound state, the substrate could be steered out of the protein without facing any major obstacles by SMD simulations (Fig. 11b, c). These simulations therefore provided a complete and reproducible description of the unbinding pathway for the substrate in vSGLT (Fig. 11e). For comparison, we also performed SMD simulations starting from the fully bound state of the substrate (the original location of the substrate in the crystal structure), as well as starting from several partially unbound states. All these simulations required very large forces to induce substrate unbinding (Fig. 11a, d). The large forces arising in these simulations are primarily due to the steric clash between the substrate and the side chain of Tyr263 which is right in front of the substrate when pulled directly from the binding pocket toward the cytoplasmic side. We note during the initial unbinding captured in our equilibrium simulations, the substrate takes a curved pathway out of the binding site, essentially going around the side chain of Tyr263. This example further highlights the importance of defining an optimal pathway when applying SMD. Otherwise, the simulations can result in artificial mechanism/pathways and inaccurate energetics. In our simulations, however, the initial step of substrate unbinding, which involves a complicated pathway, was fully described in the equilibrium simulations, and we only relied on SMD for the final separation of the substrate from the protein lumen, a process that turned out to be almost barrier-less and merely relying on free diffusion of the substrate within the lumen.



**Fig. 11.** The release of the substrate in MD and SMD simulations. **(a–d)** Initial conformations for SMD simulations captured in different snapshots in MD simulation, and respective force-time profiles in SMDs. Starting from the fully unbound state **(b and c)**, the substrate could be steered out of the protein without facing any major obstacles with SMD simulations, whereas pulling simulations starting from a bound state (original location of the substrate) **(a)** or from several partially unbound states **(d)**, required very large forces for substrate unbinding. **(e)** Overview of the release trajectory of the substrate. The release trajectory is combined by a part of equilibrium trajectory (from 80 ns to 83 ns) and the reproducible trajectory of SMD simulation starting from the full unbound state. The substrate is represented by carbon atom in the pyranose ring of the substrate. The protein is shown in *ribbon* representation, and the residues lined in the substrate-release pathway are displayed as sticks.

It is noteworthy to also mention that during the process of unbinding, the entire protein including the residues lining the unbinding pathway, was found to maintain its original structure, indicating that no further conformational changes are necessary for substrate release in the state captured in the crystal structure. In other words,

the captured structure in the crystal structure appears to represent a IF-open structure. In summary, based on the dynamical descriptions provided by the MD and SMD simulations, we propose a mechanism for substrate unbinding in vSGLT that does not require a gating motion in the cytoplasmic half of the protein (9). Further support of this model is provided by the latest crystal structure of Mhp1, the benzyl-hydantoin transporter from the family of LeuT-fold secondary transporters. This structure is highly similar to the structure of vSGLT (70). While the structure of vSGLT is in its substrate-bound form, Mhp1 is reported to be substrate-free in its crystal form, which suggests that the transition from the substrate-bound state to the substrate-free one in the LeuT-fold transporters might not require major conformational changes of the protein, as also observed in our simulations.

---

## 8. Exploring the $\text{Na}^+$ -Coupled Substrate Transport Mechanism in the $\text{Na}^+/\text{Betaine}$ Symporter

Secondary active transporters utilize ionic gradients as the energy source to drive uphill translocation of molecules across the membrane. In recent years, the structures of several  $\text{Na}^+$ -coupled symporters have been solved in various conformational states, including outward-facing (OF), occluded, and inward-facing (IF) states (123, 134–136). Along with biophysical studies (27, 121, 137), these crystal structures have provided mechanistic insights about the transport cycle and have supported the presence of at least two major conformational states, OF and IF states, in agreement with the universal alternating access model (10).

The  $\text{Na}^+/\text{betaine}$  symporter (BetP) belongs to the family of the betaine/carnitine/choline transporters (BCCTs), which transport choline and/or its derivatives across the membrane. Based on its only crystal structure, obtained at 3.35 Å resolution (71), BetP shares a similar topology with LeuT-fold transporters, of which contain a 5+5 inverted repeat of helices (123, 134, 135). BetP is activated in response to the hyperosmotic stress (136, 138–140) and translocates two  $\text{Na}^+$  ions along with an osmoprotectant, betaine, as the substrate (141). The uptake of betaine helps maintain cellular osmotic balance by preventing dehydration in organisms, such as halophilic bacteria, that inhabit hyperosmotic environments (142).

The crystal structure of BetP is in a substrate-bound state. It has also been suggested to be in an occluded state (71, 134), a state in which the substrate binding pocket is inaccessible from both extracellular and intracellular sides of the membrane. Visual examination of the crystal structure, however, indicates the presence of an elongated cavity extending from the cytoplasmic lumen into the binding pocket, which suggests that the crystal structure might not

necessarily be in an occluded state. BetP, therefore, presents a case in which the determination of the open/occluded state is not trivial and requires more rigorous examination of the structure. Generally, it is difficult to make definitive conclusions about the open/occluded state of a transporter only based on a single snapshot captured in the crystal structure. The dynamics of the protein, which might result in significant breathing of the protein (opening), and the degree of hydration of pathways connecting the binding site to outside solution, which is determined not only by the size but also by the nature of lining residues, cannot be easily assessed without a dynamical description of the system in the presence of explicit solvent. At the obtained resolution ( $3.35\text{\AA}$ ), the crystal structure does not provide information on the water molecules within the elongated cavity. Moreover, the cavity is composed of both hydrophilic and hydrophobic/aromatic residues making it difficult to determine whether the betaine substrate can be accessed by water from the cytoplasmic solution.

Another uncertainty in the structure is with regard to the presence and the position of the cotransported  $\text{Na}^+$  ions. At  $3.35\text{\AA}$  resolution, the position of small ions such as  $\text{Na}^+$  cannot be determined, so it is unclear as to whether there is any ion present in the state of BetP captured in the crystal structure. The only speculation for the presence of an ion is that, in the crystal structure, the carboxylate group of the substrate forms neither salt bridges nor hydrogens bonds with protein residues. Two  $\text{Na}^+$  binding sites, referred to as Na1 and Na2 (Fig. 12a), have been suggested for BetP based on structural comparison to LeuT (71), a transporter for which the position of the bound  $\text{Na}^+$  ions is unequivocally determined (116). Accordingly, in BetP, Na1 is suggested to be bound directly to the carboxylate group of the substrate (Fig. 12a, c), whereas Na2 makes no contact to the substrate and is more than  $3.0\text{\AA}$  away from nearby residues, suggesting a weak binding. We note that unstable ion binding has been recently reported by MD studies performed on the cytoplasmic-open state of vSGLT and Mhp1 (3, 70, 143).

To investigate the dynamics of the conformational state of BetP captured in the crystal structure and to characterize its functional state and examine the stability of ion binding sites, MD simulations were employed. All MD simulations were performed on a membrane-embedded model of BetP constructed from the crystal structure. Both ion-free and several ion-bound states were simulated in order to examine the protein response to the presence of ions and to study the stability of the ion binding sites. The simulations of the membrane-bound BetP follow the general procedures described previously.

Although no water molecules were initially placed inside the BetP protein, in less than 0.5 ns, water molecules entered the elongated cavity and reached the substrate binding pocket from the

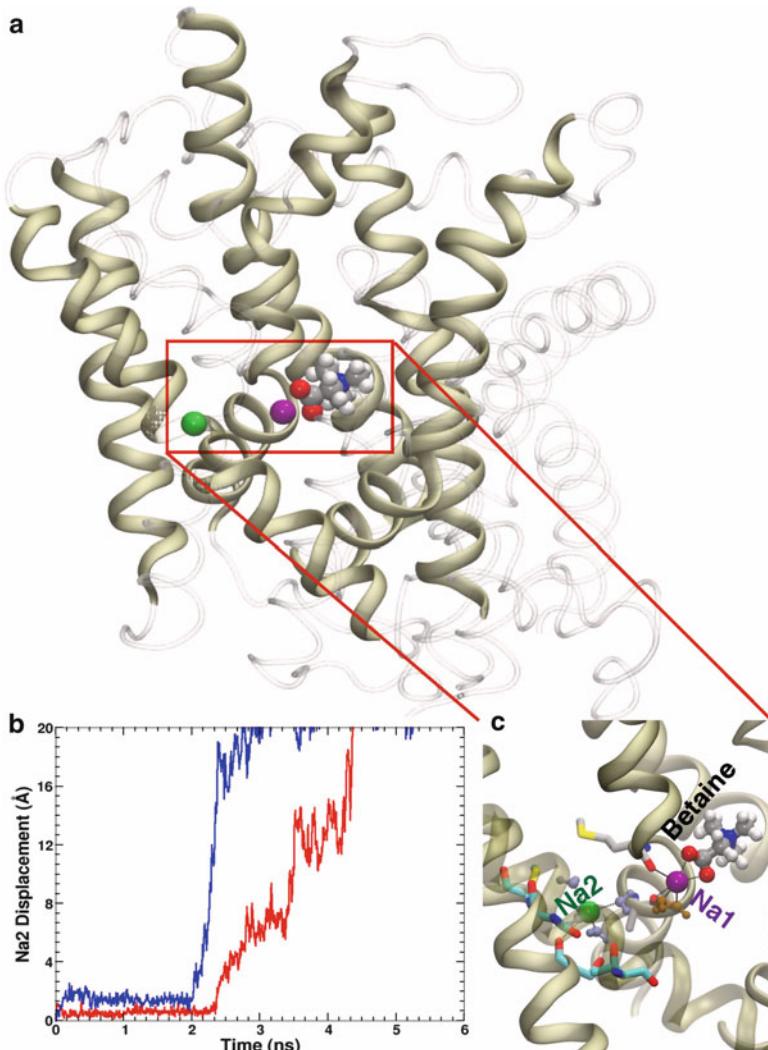


Fig. 12. Ion-substrate bound BetP system. (a) Large image indicates the binding sites for the substrate, Na1 (purple) and Na2 (green) (71). Helices involved in their binding are represented by tan spiral helices. (b) The graph illustrates the unbinding event of Na2 which was restrained by a harmonic potential for 2 ns and was released afterward indicated by its displacement from its initial position at  $t = 0$  ns. Red and blue curves indicate two different trials. (c) Small image indicates the interactions between residues and the substrate at  $t = 2$  ns of constrained simulation. Water molecules that are within  $2.6\text{\AA}$  of Na1 are shown in brown while those interacting with Na2 are shown in blue.

cytoplasmic solution. The cavity becomes fully solvent-accessible at  $t = 1$  ns, characterizing it as a clearly open-hydrated pathway connecting the binding sites of the substrate, and those of the Na1 and Na2 to the cytoplasmic solution (Fig. 12c). It is important to note that the formation of this hydrated pathway did not require the protein to undergo any appreciable conformational changes, thus representing closely the state captured in the crystal structure. The pathway remained open throughout the simulations both in the

ion-bound and in the ion-free systems (144). These observations strongly indicate that the current crystallized state of BetP is solvent-accessible and is in a cytoplasmic-open state.

As pointed out earlier, the position of the modeled  $\text{Na}^+$  ions indicated a tight binding for Na1 but a weak Na2 binding. With both sites becoming fully solvent-accessible during the initial phase of the simulations, both ions have open pathways for potential unbinding from the protein. Indeed, one of the ions (Na2) rapidly escaped from its site and dissociated from the protein during the simulations (Fig. 12b). In order to further stabilize these ions in their putative binding sites during the initial modeling phase, and to prevent their premature release from the protein, we took advantage of various types of constraints in MD simulations. Different constraint schemes including positional and intermolecular-distance constraints were employed to stabilize the binding of Na1 and Na2 in BetP. Several simulation protocols were employed and tested in which the ions were constrained for various initial periods (between 2 to 5 ns), with the rest of the system undergoing free dynamics. The aim was to see whether local relaxation of the protein binding sites could recover a stable ion-bound state. Positional constraints were to keep the ions within their sites while allowing residues to adjust for proper ligating sites. The ions were restrained by harmonic potentials. The purpose of intramolecular-distance constraints is to stabilize ion-ligand interactions by bringing an ion and potential coordinating ligands close to each other. Crystal structures of proteins with  $\text{Na}^+$  ions bound have shown that the average ligand-coordinating distance to a  $\text{Na}^+$  ion is  $\sim 2.6\text{\AA}$  (69, 116, 145, 146), so the ions were kept within 2.5 to 2.8 $\text{\AA}$  from selected oxygen atoms (three or more per ion) of surrounding protein residues (Fig. 12c). The force constant used for these constraints ranged from 35 to 700 pN/ $\text{\AA}$ ; weaker constraint schemes would allow the ion to explore a wider region around its initial position, while stronger constraints can be used to force the binding site to deform for an optimal ion binding. Even after applying such constraint schemes, as soon as the constraints were removed, Na2 left its site (144). This observation is in line with the fact that this ion is significantly hydrated and that it cannot establish strong interaction with the protein. Na1, on the other hand, remained bound to the substrate (Fig. 12b). The simulations therefore suggest that Na2 is likely unbound as BetP becomes accessible to water molecules from the cytoplasmic lumen (Fig. 12c).

To directly probe for possible roles of  $\text{Na}^+$  ions and to confirm Na1 binding in BetP, both ion-bound and ion-free states were equilibrated for additional 50 ns. In the ion-free system, the substrate is displaced from its binding site to a new site located 3.5 $\text{\AA}$  away, where it forms interactions with the protein (Fig. 13a). In the ion-bound simulation, Na1 maintains its contacts with the substrate during the first 30 ns when the substrate remains bound to

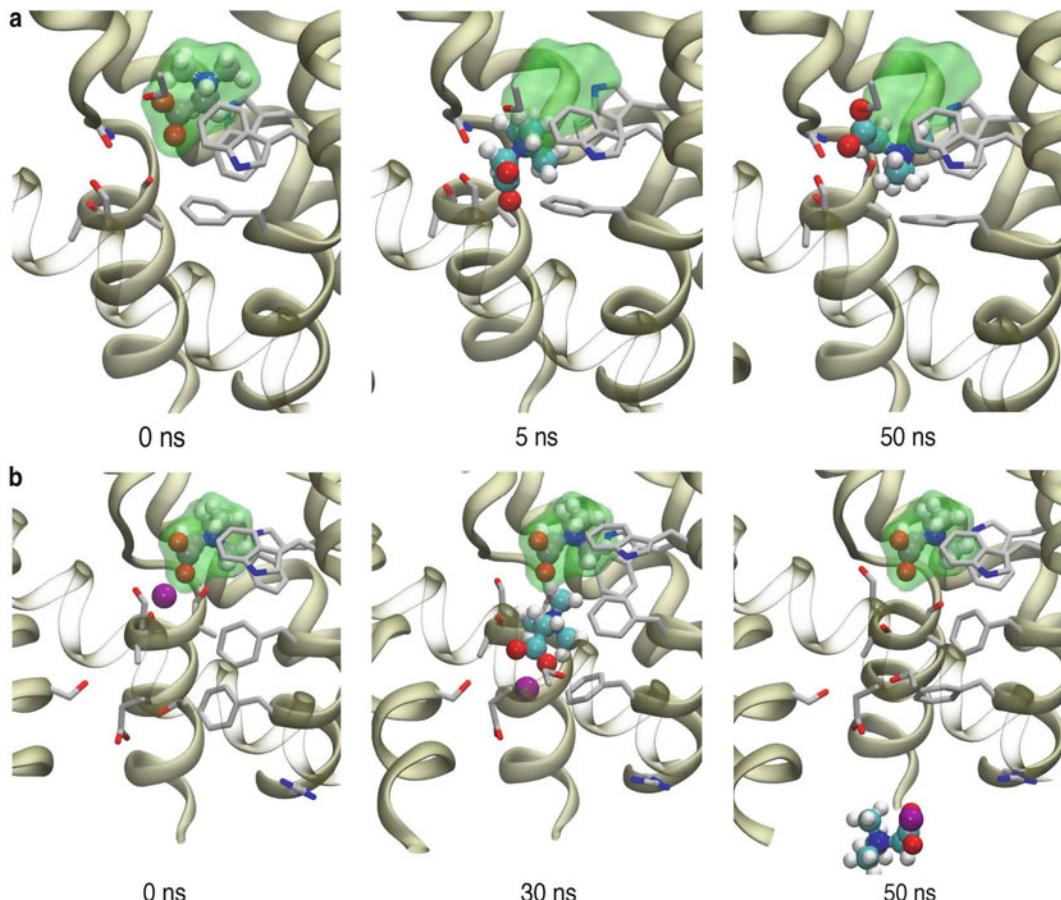


Fig. 13. Dynamics of the ion-coupled substrate dissociation in BetP. (a) Ion-free system and (b) ion-bound system. Carbon atoms of the substrate at  $t = 0$  ns are shown in silver, while those of the substrate taken at moving trajectories are in cyan. Green-transparent surface represents the binding site of the substrate at  $t = 0$  ns. Protein residues that are close in contact with the substrate are represented in sticks.

its original binding site. Interestingly, the cooperative ion-substrate dissociation was observed in the simulation. At  $t = \sim 31$  ns, both Na1 and the substrate spontaneously displaced from their pockets and exited the protein by the end of simulations. They both dissociated along the elongated hydrated cavity described above, which is also the pathway used by Na2 for unbinding (Fig. 13b). Again, BetP did not undergo any appreciable conformational changes during the unbinding of the substrate and Na1 (144). In summary, these simulations suggest that from a conformational perspective, the structure captured in the crystal structure represents the IF-open state of the transporter, in which the substrate has been transiently stabilized under the crystallization conditions. They also clearly describe the unbinding pathway used by the substrate and the ions and suggest a strong coupling between the substrate and one of the co-transported ions.

## 9. Concluding Remarks

Molecular dynamics simulations provide a detailed description of the molecular motions of biomolecular systems at atomic resolution. Despite the limited time scale currently accessible to the method, it has been successfully employed in capturing various dynamical components of the mechanism of function of a wide range of biological molecules. The method is particularly powerful when combined with experiment by which the developed hypotheses and models can be tested and verified. Conformational changes are at the heart of the function of membrane transporters. They rely on various degrees of protein structural transitions not only for translocation of the substrate, but also, and more importantly, for implementing energy-coupling mechanisms which effectively pump the substrate against its concentration gradient across the membrane. Various modes of protein conformational changes are at work in membrane transporters. While localized gating motions are responsible for transition between the open and occluded states of the binding sites, global conformational changes are usually in charge of the change of the overall accessibility of the substrate between the two sides of the membrane, a mechanism broadly referred to as the alternating access model. Unfortunately, describing such structural changes by experimental methods at desirably high resolutions is extremely challenging, if not impossible. These structural changes take the protein through various conformational states involved in the transport mechanism. Given the advances in experimental structural biology, we now have access to an increasing number of membrane transporter structures at atomic resolutions. However, there are only very few transporters for which more than one conformational state has been experimentally resolved. In fact, for most transporters, we do not have resolved structures even for a single state. Applying computational methodologies such as MD to the known structures, we have been able to study functionally relevant conformational changes in membrane transporters.

In this chapter, we have presented the results of MD simulations of several membrane transporter proteins in our laboratory. These studies clearly demonstrate that MD simulations, when designed and applied properly, are able to capture various forms and degrees of conformational changes involved in the function of these mechanistically complex proteins. Combining atomic representations of the proteins and their surrounding (lipid, water, ions, etc.) with extended simulations, we have been able to characterize motions ranging from side-chain rotations and hydrogen-bond forming/breaking events, all the way to subdomain flipping and even domain separations in different membrane transporters. Most importantly, we show that the captured motions are functionally significant, that

is, such motions are triggered in response to specific molecular events. Indeed in most cases, we have also demonstrated that the observed events are reproducible and are seen in independent simulations. The results of these studies have produced novel hypotheses regarding the function of membrane transporters. Many of these hypotheses can be and have been tested by designing new experiments that can verify the proposed mechanisms. Along with the growing computational power, we will be able to extend further the time scale of MD simulations and improve our sampling and statistics. We will look forward to more examples of biomolecular simulations in which key functional dynamical events have been captured. We should also expect to have a larger number of high-resolution structures for various intermediates and functional states of membrane transporters, which, combined with extended simulations and hybrid methods, will allow us to provide a more complete description of the dynamics of the entire transport cycle in membrane transporters. We would like to conclude by noting that many of the protocols and strategies adopted in designing and analyzing the simulations presented in this chapter can be employed in other classes of proteins or any other macromolecular systems for that matter. We hope that the expanded discussion on these protocols in each section will inspire the reader to employ similar approaches to their own problems.

---

## Acknowledgment

The studies reported in this review were supported by grants from NIH (R01-GM086749, R01-GM067887, and P41-RR05969). The authors acknowledge computer time at TeraGrid resources (grant number MCA06N060), as well as computer time from the DoD High Performance Computing Modernization Program at the Arctic Region Supercomputing Center, University of Alaska at Fairbanks.

## References

1. Celik L, Schiott B, Tajkhorshid E (2008) Substrate binding and formation of an occluded state in the leucine transporter. *Biophys J* 94:1600–1612
2. Law CJ, Enkavi G, Wang D-N, Tajkhorshid E (2009) Structural basis of substrate selectivity in the glycerol-3-phosphate:phosphate antiporter GlpT. *Biophys J* 97:1346–1353
3. Li J, Tajkhorshid E (2009) Ion-releasing state of a secondary membrane transporter. *Biophys J* 97:L29–L31
4. Shaikh SA, Tajkhorshid E (2008) Potential cation and H<sup>+</sup> binding sites in acid sensing ion channel-1. *Biophys J* 95:5153–5164
5. Huang Z, Tajkhorshid E (2008) Dynamics of the extracellular gate and ion-substrate

- coupling in the glutamate transporter. *Bioophys J* 95:2292–2300
6. Wen P-C, Tajkhorshid E (2008) Dimer opening of the nucleotide binding domains of ABC transporters after ATP hydrolysis. *Bioophys J* 95:5100–5110
  7. Gumbart J, Wiener MC, Tajkhorshid E (2009) Coupling of calcium and substrate binding through loop alignment in the outer membrane transporter BtuB. *J Mol Biol* 393:1129–1142
  8. Huang Z, Tajkhorshid E (2010) Identification of the third Na<sup>+</sup> site and the sequence of extracellular binding events in the glutamate transporter. *Bioophys J* 99:1416–1425
  9. Li J, Tajkhorshid E (2012) A gate-free pathway for substrate release from the inward-facing state of the Na<sup>+</sup>-galactose transporter. *Biochim Biophys Acta Biomembr* 1818:263–271
  10. Jardetzky O (1966) Simple allosteric model for membrane pumps. *Nature* 211:2406–2414
  11. Gumbart J, Wiener MC, Tajkhorshid E (2007) Mechanics of force propagation in TonB-dependent outer membrane transport. *Bioophys J* 93:496–504
  12. Wang Y, Tajkhorshid E (2008) Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc Natl Acad Sci USA* 105:9598–9603
  13. Shrivastava IH, Jiang J, Amara SG, Bahar I (2008) Time-resolved mechanism of extracellular gate opening and substrate binding in a glutamate transporter. *J Biol Chem* 283:28680–28690
  14. Khalili-Araghi F, Gumbart J, Wen P-C, Sotomayor M, Tajkhorshid E, and Schulten K (2009) Molecular dynamics simulations of membrane channels and transporters. *Curr Opin Struct Biol* 19:128–137
  15. Arkin IT, Xu H, Jensen M, Arbely E, Bennett ER, Bowers KJ, Chow E, Dror RO, Eastwood MP, Flitman-Tene R, Gregersen BA, Klepeis JL, Kolossváry I, Shan Y, Shaw DE (2007) Mechanism of Na<sup>+</sup>/H<sup>+</sup> antiporting. *Science* 317:799–803
  16. Olkhova E, Padan E, Michel H (2007) The influence of protonation states on the dynamics of the NhaA antiporter from *Escherichia coli*. *Bioophys J* 92:3784–3791
  17. Cheng XL, Ivanov I, Wang HL, Sine SM, McCammon JA (2007) Nanosecond-timescale conformational dynamics of the human α7 nicotinic acetylcholine receptor. *Bioophys J* 93:2622–2634
  18. Wang Y, Tajkhorshid E (2007) Molecular mechanisms of conduction and selectivity in aquaporin water channels. *J Nutr* 137:1509S–1515S
  19. Hashido M, Kidera A, Ikeguchi M (2007) Water transport in aquaporins: osmotic permeability matrix analysis of molecular dynamics simulations. *Biophys J* 93:373–385
  20. Huang X, Zhan C-G (2007) How dopamine transporter interacts with dopamine: insights from molecular modeling and simulation. *Biophys J* 93:3627–3639
  21. Holyoake J, Sansom MSP (2007) Conformational change in an MFS protein: MD simulations of LacY. *Structure* 15:873–884
  22. Klauda JB, Brooks BR (2007) Sugar binding in lactose permease: anomeric state of a disaccharide influences binding structure. *J Mol Biol* 367:1523–1534
  23. Cordero-Morales JF, Jogini V, Lewis A, Vasquez V, Cortes DM, Roux B, Perozo E (2007) Molecular driving forces determining potassium channel slow inactivation. *Nat Struct Mol Biol* 14:1062–1069
  24. Sonne J, Kandt C, Peters GH, Hansen FY, Jansen MO, Tieleman DP (2007) Simulation of the coupling between nucleotide binding and transmembrane domains in the ATP binding cassette transporter BtuCD. *Bioophys J* 92:2727–2734
  25. Hub J, de Groot B (2008) Mechanism of selectivity in aquaporins and aquaglyceroporins. *Proc Natl Acad Sci USA* 105:1198–203
  26. Henin J, Tajkhorshid E, Schulten K, Chipot C (2008) Diffusion of glycerol through *Escherichia coli* aquaglyceroporin GlpF. *Bioophys J* 94:832–839
  27. Shi L, Quick M, Zhao Y, Weinstein H, Javitch JA (2008) The mechanism of a neurotransmitter:sodium symporter–inward release of Na<sup>+</sup> and substrate is triggered by substrate in a second binding site. *Mol Cell* 30:667–677
  28. Noskov SY, Roux B (2008) Control of ion selectivity in LeuT: two Na<sup>+</sup> binding sites with two different mechanisms. *J Mol Biol* 377:804–818
  29. Vasquez V, Sotomayor M, Cordero-Morales J, Schulten K, Perozo E (2008) A structural mechanism for MscS gating in lipid bilayers. *Science* 321:1210–1214
  30. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comp Chem* 26:1781–1802
  31. Martyna GJ, Tobias DJ, Klein ML (1994) Constant pressure molecular dynamics algorithms. *J Chem Phys* 101:4177–4189

32. Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) Constant pressure molecular dynamics simulation: the Langevin piston method. *J Chem Phys* 103:4613–4621
33. Darden T, York D, Pedersen LG (1993) Particle mesh Ewald: an  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
34. Feller SE, Yin D, Pastor RW, MacKerell AD Jr (1997) Molecular dynamics simulation of unsaturated lipids at low hydration: parametrization and comparison with diffraction studies. *Biophys J* 73:2269–2279
35. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD Jr, Pastor RW (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114:7830–7843
36. Locher KP, Lee AT, Rees DC (2002) The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* 296:1091–1098
37. Dawson RJ, Locher KP (2006) Structure of a bacterial multidrug ABC transporter. *Nature* 443:180–185
38. Pinkett HW, Lee AT, Lum P, Locher KP, Rees DC (2007) An inward-facing conformation of a putative metal-chelate-type ABC transporter. *Science* 315:373–377
39. Dawson RJ, Locher KP (2007) Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett* 581:935–938
40. Hollenstein K, Frei DC, Locher KP (2007) Structure of an ABC transporter in complex with its binding protein. *Nature* 446:213–216
41. Hvorup RN, Goetz BA, Niederer M, Hollenstein K, Perozo E, Locher KP (2007) Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* 317:1387–1390
42. Oldham ML, Khare D, Quiocco FA, Davidson AL, Chen J (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450:515–521
43. Ward A, Reyes CL, Yu J, Roth CB, Chang G (2007) Flexibility in the ABC transporter MsbA: alternating access with a twist. *Proc Natl Acad Sci USA* 104:19005–19010
44. Gerber S, Comellas-Bigler M, Goetz BA, Locher KP (2008) Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science* 321:246–250
45. Kadaba NS, Kaiser JT, Johnson E, Lee A, Rees DC (2008) The high-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation. *Science* 321:250–253
46. Khare D, Oldham ML, Orelle C, Davidson AL, Chen J (2009) Alternating access in maltose transporter mediated by rigid-body rotations. *Mol Cell* 33:528–536
47. Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL, Chang G (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323:1718–1722
48. Hollenstein K, Dawson RJ, Locher KP (2007) Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* 17:412–418
49. Oldham ML, Davidson AL, Chen J (2008) Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol* 18:726–733
50. Nikaido K, Ames GF (1999) One intact ATP-binding subunit is sufficient to support ATP hydrolysis and translocation in an ABC transporter, the histidine permease. *J Biol Chem* 274:26727–26735
51. Aleksandrov L, Aleksandrov AA, bao Chang X, Riordan JR (2002) First nucleotide binding domain of cystic fibrosis transmembrane conductance regulator is a site of stable nucleotide interaction, whereas the second is a site of rapid turnover. *J Biol Chem* 277:15419–15425
52. Ernst R, Kueppers P, Klein CM, Schwarzmüller T, Kuchler K, Schmitt L (2008) A mutation of the H-loop selectively affects rhodamine transport by the yeast multidrug ABC transporter Pdr5. *Proc Natl Acad Sci USA* 105:5069–5074
53. Chen C, Peng E (2003) Nanopore sequencing of polynucleotides assisted by a rotating electric field. *Appl Phys Lett* 82:1308–1310
54. Smith PC, Karpowich N, Millen L, Moody JE, Rosen J, Thomas PJ, Hunt JF (2002) ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol Cell* 10:139–149
55. Zaitseva J, Jenewein S, Jumperz T, Holland IB, Schmitt L (2005) H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO J* 24:1901–1910
56. Wen P-C, Tajkhorshid E (2008) Dimer opening of the nucleotide binding domains of ABC transporters after ATP hydrolysis. *Biophys J* 95:5100–5110
57. Oloo EO, Fung EY, Tielemans DP (2006) The dynamics of the MgATP-driven closure of MalK, the energy-transducing subunit of the maltose ABC transporter. *J Biol Chem* 281:28397–28407

58. Jones PM, George AM (2007) Nucleotide-dependent allostery within the ABC transporter ATP-binding cassette. *J Biol Chem* 282:22793–22803
59. Jones PM, George AM (2009) Opening of the ADP-bound active site in the ABC transporter ATPase dimer: evidence for a constant contact, alternating sites model for the catalytic cycle. *Proteins: Struct Func Bioinf* 75:387–396
60. Oliveira ASF, Baptista AM, Soares CM (2010) Insights into the molecular mechanism of an ABC transporter: conformational changes in the NBD dimer of MJ0796. *J Phys Chem B* 114:5486–5496
61. Wen P-C, Tajkhорشید E (2011) Conformational coupling of the nucleotide-binding and the transmembrane domains in the maltose ABC transporter. *Biophys J* 101:680–690
62. Lange OF, Grubmüller H (2006) Generalized correlation for biomolecular dynamics. *Proteins: Struct Func Bioinf* 62:1053–1061
63. Bergles DE, Diamond JS, Jahr CE (1999) Clearance of glutamate inside the synapse and beyond. *Curr Opin Neur* 9:293–298
64. Slotboom DJ, Konings WN, Lolkema JS (1999) Structural features of the glutamate transporter family. *Microbiol Mol Biol Rev* 63:293–307
65. Chen NH, Reith ME, Quick MW (2004) Synaptic uptake and beyond: the sodium- and chloride-dependent neurotransmitter transporter family SLC6. *Pflug Arch Eur J Physiol* 447:519–531
66. Grewer C, Rauen T (2005) Electrogenic glutamate transporters in the CNS: molecular mechanism, pre-steady-state kinetics, and their impact on synaptic signaling. *J Membr Biol* 203:1–20
67. Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ (2005) Deconstructing the cadherin-catenin-actin complex. *Cell* 123:889–901
68. Faham S, Watanabe A, Besserer GM, Cascio D, Specht A, Hirayama BA, Wright EM, Abramson J (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into  $\text{Na}^+$ /sugar symport. *Science* 321:810–814
69. Weyand S, Shimamura T, Yajima S, Suzuki S, Mirza O, Krusong K, Carpenter EP, Rutherford NG, Hadden JM, O'Reilly J, Ma P, Saidijam M, Patching SG, Hope RJ, Norbertczak HT, Roach PCJ, Iwata S, Henderson PJF, Cameron AD (2008) Structure and molecular mechanism of a Nucleobase-Cation-Symport-1 family transporter. *Science* 322:709–713
70. Shimamura T, Weyand S, Beckstein O, Rutherford NG, Hadden JM, Sharples D, Sansom MSP, Iwata S, Henderson PJF, Cameron AD (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* 328:470–473
71. Ressl S, van Scheltinga ACT, Vonrhein C, Ott V, Ziegler C (2009) Molecular basis of transport and regulation in the  $\text{Na}^+$ /betaine symporter BetP. *Nature* 458:47–52
72. Fang Y, Jayaram H, Shane T, Kolmakova-Partensky L, Wu F, Williams C, Xiong Y, Miller C (2009) Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature* 460:1040–1043
73. Yernool D, Boudker O, Jin Y, Gouaux E (2004) Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* 431:811–818
74. Boudker O, Ryan RM, Yernool D, Shimamoto K, Gouaux E (2007) Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature* 445:387–393
75. Reyes N, Ginter C, Boudker O (2009) Transport mechanism of a bacterial homologue of glutamate transporters. *Nature* 462:880–885
76. Tao Z, Zhang Z, Grewer C (2006) Neutralization of the aspartic acid residue Asp-367, but not Asp-454, inhibits binding of  $\text{Na}^+$  to the glutamate-free form and cycling of the glutamate transporter EAAC1. *J Biol Chem* 281:10263–10272
77. Larsson PH, Tzingounis AV, Koch HP, Kavanaugh MP (2004) Fluorometric measurements of conformational changes in glutamate transporters. *Proc Natl Acad Sci USA* 101:3951–3956
78. Koch HP, Hubbard JM, Larsson HP (2007) Voltage-independent sodium-binding events reported by the 4B-4C loop in the human glutamate transporter excitatory amino acid transporter 3. *J Biol Chem* 282:24547–24553
79. Pao SS, Paulsen IT, Saier MH Jr (1998) Major facilitator superfamily. *Microbiol Mol Biol Rev* 62:1–34
80. Law CJ, Maloney PC, Wang D-N (2008) Ins and outs of major facilitator superfamily antiports. *Annu Rev Microbiol* 62:289–305
81. Huang Y, Lemieux MJ, Song J, Auer M, Wang D-N (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 301:616–620
82. Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S (2003) Structure and

- mechanism of the lactose permease of *Escherichia coli*. *Science* 301:610–615
83. Hirai T, Subramaniam S (2004) Structure and transport mechanism of the bacterial oxalate transporter oxlt. *Biophys J* 87:3600–3607
84. Yin Y, He X, Szewczyk P, Nguyen T, Chang G (2006) Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science* 312:741–744
85. Salasburgos A, Iserovich P, Zuniga F, Vera J, Fischbarg J (2004) Predicting the three-dimensional structure of the human facilitative glucose transporter glut1 by a novel evolutionary homology strategy: insights on the molecular mechanism of substrate migration, and binding sites for glucose and inhibitory molecules. *Biophys J* 87:2990–2999
86. Lemieux MJ, Huang Y, Wang D-N (2004) Glycerol-3-phosphate transporter of *Escherichia coli*: Structure, function and regulation. *Res Microbiol* 155:623–629
87. Holyoake J, Caulfeild V, Baldwin S, Sansom M (2006) Modeling, docking, and simulation of the major facilitator superfamily. *Biophys J* 91:L84–L86
88. Lemieux MJ (2007) Eukaryotic major facilitator superfamily transporter modeling based on the prokaryotic GlpT crystal structure (review). *Mol Membr Biol* 24:333–341
89. Lemieux M, Huang Y, Wang D (2004) The structural basis of substrate translocation by the glycerol-3-phosphate transporter: a member of the major facilitator superfamily. *Curr Opin Struct Biol* 14:405–412
90. Law CJ, Almqvist J, Bernstein A, Goetz RM, Huang Y, Soudant C, Laaksonen A, Hovmöller S, Wang D-N (2008) Salt-bridge dynamics control substrate-induced conformational change in the membrane transporter glpt. *J Mol Biol* 378:828–839
91. Fann MC, Davies AH, Varadhachary A, Kuroda T, Sevier C, Tsuchiya T, Maloney PC (1998) Identification of two essential arginine residues in uhpt, the sugar phosphate antipporter of *Escherichia coli*. *J Membr Biol* 164:187–195
92. Stroud RM (2007) Transmembrane transporters: an open and closed case. *Proc Natl Acad Sci USA* 104:14445–14446
93. Lemieux MJ, Huang Y, Wang D-N (2005) Crystal structure and mechanism of GlpT, the glycerol-3-phosphate transporter from *E. coli*. *J Electron Microsc* 54:i43–i46
94. D'rozario RSG, Sansom MSP (2008) Helix dynamics in a membrane transport protein: comparative simulations of the glycerol-3-phosphate transporter and its constituent helices. *Mol Membr Biol* 25:571–573
95. Tsigelny IF, Greenberg J, Kouznetsova V, Nigam SK (2008) Modelling of glycerol-3-phosphate transporter suggests a potential ‘tilt’ mechanism involved in its function. *J Bioinformatics Comput Biol* 6:885–904
96. Law CJ, Yang Q, Soudant C, Maloney PC, Wang D-N (2007) Kinetic evidence is consistent with the rocker-switch mechanism of membrane transport by GlpT. *Biochemistry* 46:12190–12197
97. Enkavi G, Tajkhorshid E (2010) Simulation of spontaneous substrate binding revealing the binding pathway and mechanism and initial conformational response of GlpT. *Biochemistry* 49:1105–1114
98. Mertz JE, Pettitt BM (1994) Molecular dynamics at a constant pH. *Supercomputer Appl High Perform Comput* 8:47–53
99. Baptista AM, Teixeira VH, Soares CM (2002) Constant-pH molecular dynamics using stochastic titration. *J Chem Phys* 117:4184–4200
100. Dlugosz M, Antosiewicz JM (2004) Constant-pH molecular dynamics simulations: a test case of succinic acid. *Chem Phys* 302:161–170
101. Dlugosz M, Antosiewicz JM, Robertson AD (2004) Constant-pH molecular dynamics study of protonation-structure relationship in a hexapeptide derived from ovomucoid third domain. *Phys Rev E* 69:021915
102. Bürgi R, Kollman PA, van Gunsteren WF (2002) Simulating proteins at constant pH: an approach combining molecular dynamics and Monte Carlo simulation. *Proteins: Struct Func Gen* 47:469–480
103. Mongan J, Case DA, McCammon JA (2004) Constant pH molecular dynamics in generalized Born implicit solvent. *J Comp Chem* 25:2038–2048
104. Lee MS, Salsbury FR Jr, Brooks III CL (2004) Constant-pH molecular dynamics using continuous titration coordinates. *J Comp Chem* 25:2038–2048
105. Mongan J, Case DA (2005) Biomolecular simulations at constant pH. *Curr Opin Struct Biol* 15:157–163
106. Machuqueiro M, Baptista AM (2007) The pH-dependent conformational states of kytorphin: a constant-ph molecular dynamics study. *Biophys J* 92:1836–1845
107. Chennubhotla C, Rader A, Yang L-W, Bahar I (2005) Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. *Phys Biol* 2:S173–S180
108. Bahar I, Lezon TR, Yang LW, Eyal E (2010) Global Dynamics of Proteins: bridging

- Between Structure and Function. *Ann Rev Biophys* 39:23–42
109. Tirion M (1996) Large amplitude elastic motions in proteins from a single-parameter atomic analysis. *Phys Rev Lett* 77:1905–1908
  110. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 80:505–515
  111. Erman B (2006) The Gaussian network model: precise predictions of residue fluctuations and application to binding problems. *Biophys J* 91:3589–3599
  112. Eyal E, Yang LW, Bahar I (2006) Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics* 22:2619–2627
  113. Ma J (2005) Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* (London, England: 1993) 13:373–380
  114. Brüschweiler R (1995) Collective protein dynamics and nuclear spin relaxation. *J Chem Phys* 102:3396–3403
  115. Kanner BI, Zomot E (2008) Sodium-coupled neurotransmitter transporters. *Chem Rev* 108:1654–1668
  116. Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E (2005) Crystal structure of a bacterial homologue of  $\text{Na}^+/\text{Cl}^-$ -dependent neurotransmitter transporters. *Nature* 437:215–233
  117. Shaikh SA, Tajkhorshid E (2010) Modeling and dynamics of the inward-facing state of a  $\text{Na}^+/\text{Cl}^-$  dependent neurotransmitter transporter homologue. *PLoS Comput Biol* 6(8): e1000905
  118. Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol* 11:224–230
  119. Schlitte J, Engels M, Krüger P, Jacoby E, Wollmer A (1993) Targeted molecular dynamics simulation of conformational change — application to the T $\leftrightarrow$ R transition in insulin. *Mol Sim* 10:291–308
  120. Quick M, Winther A-ML, Shi L, Nissen P, Weinstein H, Javitch JA (2009) Binding of an octylglucoside detergent molecule in the second substrate (S2) site of LeuT establishes an inhibitor-bound conformation. *Proc Natl Acad Sci USA* 106:5563–5568
  121. Zhao Y, Terry D, Shi L, Weinstein H, Blanchard SC, Javitch JA (2010) Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature* 465:188–193
  122. Forrest LR, Zhang Y-W, Jacobs MT, Gesmonde J, Xie L, Honig BH, Rudnick G (2008) Mechanism for alternating access in neurotransmitter transporters. *Proc Natl Acad Sci USA* 105:10338–10343
  123. Krishnamurthy H, Piscitelli CL, Gouaux E (2009) Unlocking the molecular secrets of sodium-coupled transporters. *Nature* 459:347–355
  124. Holm L, Park J (2000) DALI Lite workbench for protein structure comparison. *Bioinformatics* 16:566–567
  125. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779
  126. Kanner BI (1983) Bioenergetics of neurotransmitter transport. *Biochim Biophys Acta* 726:293–316
  127. Kanner BI, Schuldiner S (1987) Mechanism of transport and storage of neurotransmitters. *CRC Crit Rev Biochem* 22:1–39
  128. Kanner BI (1989) Ion-coupled neurotransmitter transport. *Curr Opin Cell Biol* 1:735–738
  129. Poolman B, Konings W (1993) Secondary solute transport in bacteria. *Biochim Biophys Acta* 1183:5–39
  130. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K (1997) Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys J* 72:1568–1581
  131. Bayas MV, Schulten K, Leckband D (2004) Forced dissociation of the strand dimer interface between C-cadherin ectodomains. *Mech Chem Biosystems* 1:101–111
  132. Jensen MØ, Yin Y, Tajkhorshid E, Schulten K (2007) Sugar transport across lactose permease probed by steered molecular dynamics. *Biophys J* 93:92–102
  133. Sujatha M, Balaji PV (2009) Identification of common structural features of binding sites in galactose-specific proteins. *Proteins* 55:44–65
  134. Abramson J, Wright E (2009) Structure and function of  $\text{Na}^+$ -symporters with inverted repeats. *Curr Opin Struct Biol* 19:425–432
  135. Boudker O, Verdon G (2010) Structural perspectives on secondary active transporters. *Trends Pharmacol Sci* 31:418–426
  136. Ziegler C, Bremer E, Kramer R (2010) The BCCT family of carriers: from physiology to crystal structure. *Mol Microbiol* 78:13–34
  137. Claxton DP, Quick M, Shi L, de Carvalho FD, Weinstein H, Javitch JA, Mchaourab HS (2010) Ion/substrate-dependent conformational dynamics of a bacterial homolog of

- neurotransmitter:sodium symporters. *Nat Struct Mol Biol* 17:822–828
138. Wood JM (1999) Osmosensing by bacteria: signals and membrane-based sensors. *Microbiol Mol Biol Rev* 63:230–262
139. Kramer R, Morbach S (2004) Betp of *corynebacterium glutamicum*, a transporter with three different functions: betaine transport, osmosensing, and osmoregulation. *Biochim Biophys Acta* 1658:31–36
140. Kramer R (2009) Osmosensing and osmo-signaling in *corynebacterium glutamicum*. *Amino Acid* 37:487–497
141. Farwick M, Siewe R, Kramer R (1995) Glycine betaine uptake after hyperosmotic shift in *corynebacterium glutamicum*. *J Bacteriol* 177:4690–4695
142. Peter H, Burkovski A, Kramer R (1996) Isolation, characterization, and expression of the *corynebacterium glutamicum* betp gene, encoding the transport system for the compatible solute glycine betaine. *J Bacteriol* 178:5229–5234
143. Zomot E, Bahar I (2010) The sodium/galactose symporter crystal structure is a dynamic, not so occluded state. *Mol Biosyst* 6: 1040–1046
144. Mahinthichaichan P, Tajkhorshid E (2011) Mechanism of ion-couple substrate translocation in an inward-facing secondary membrane transporter. Submitted
145. Varma S, Rempe SB (2008) Structural transitions in ion coordination driven by changes in competition for ligand binding. *J Am Chem Soc* 130:15405–15419
146. Toney M, Hohenester E, Cowan S, Jansonius J (1993) Dialkylglycine decarboxylase structure: bifunctional active site and alkali metal sites. *Science* 261:756–759

# Chapter 15

## Molecular Dynamics Simulations of Lipid Bilayers: Simple Recipe of How to Do It

Hector Martinez-Seara and Tomasz Róz

### Abstract

Simulations based on molecular dynamics (MD) are an important tool in structural biology. In the field of lipid membrane research, MD represents a major tool as it grants access to atomic level features of the membrane difficult to access by experimental means. In this chapter we describe step by step how to simulate a membrane *in silico*. The explanation comprehends from the initial set up of a membrane system to its analysis passing through the simulation itself. Most of the tasks described here are performed using the Gromacs simulation suite, a widely used software for molecular modeling and simulations. In the selected example the attention is focused on a single component lipid bilayer composed of DPPC, one of the most studied lipid species.

**Key words:** Lipid membrane, Lipid bilayer, Membrane simulation, DPPC, Molecular dynamics, Gromacs

---

### 1. Introduction

Biological membranes are a large set of closely interacting macromolecules enclosing cells and organelles. Membranes are involved in many essential cellular processes and functions; most importantly they separate the cell and organelles from the surroundings (1). Biological membranes are mainly composed of lipids, proteins, and sugars. Of all these components, lipids constitute the membrane basic structural matrix. This justifies the extensive set of experimental and theoretical studies focused on pure lipid bilayers.

Molecular dynamics (MD) simulations provide a unique insight into the lipid bilayer structure at atomistic level. Therefore, its growing role in the membrane research field is not surprising (2, 3). Although the matrix of biological membranes is a complex mixture of hundreds lipid species, most of the simulation studies consider only single or two components bilayers. The main reason for that is the

time scale of the lipid diffusional motions and therefore the time needed for mixing properly the different membrane components. As a result, long simulation times to ensure a proper equilibration and sampling of the phase space are needed. In combination with the limited computer resources available, the amount of species simultaneously present in simulations is restricted. Single component lipid bilayers are also easier to validate since most of comparable experiments are typically performed on single component systems.

In this chapter instructions on how to prepare, simulate, and analyze a lipid bilayer are provided. The selected system to serve as example is a single component lipid bilayer composed of dipalmitoyl-phosphatidylcholine (DPPC), one of the most studied lipid species. This example has proven to be a good reference when studying more complex systems.

---

## 2. Software

Several programs have been used in the following example. Most of these programs are multiplatform; however, this example rests entirely on the usage of a machine with a Linux distribution installed. It is important to highlight the convenience of using any modern Linux distribution over other operational systems when dealing with computational science. Here the list of the programs used along with a brief explanation of their specific purpose.

*Gromacs* is a multifunctional package designed for molecular modeling and molecular simulations in the classical context ([www.gromacs.org](http://www.gromacs.org)) (4). It is provided under a GPL license. In practice this means that it is free and more importantly it can be modified or extended if it does not meet your requirements. It consists of a set of the independent programs which allow one to prepare, simulate and analyze a molecular system. At the time of writing the last stable version was V. 4.0.7 which we have used in the examples that follow.

*Vmd* is a molecular graphic software (<http://www.ks.uiuc.edu/Research/vmd/>) (5). It can also be obtained and used free of charge. Its main purpose is the visualization of structures and trajectories obtained in molecular simulations. This will be its main usage in this example. It can be used as well as an analysis tool specially when dealing with proteins.

*Avogadro* is an open-source molecular builder and visualization tool (<http://avogadro.openmolecules.net/>). Here it is used to build the DPPC molecule from scratch.

*Vim* is a highly configurable text editor built to enable efficient text editing (<http://www.vim.org/>). It is used here for writing and

editing the text files. This program can be substituted by any plain text editor program, i.e., emacs, nedit, notepad.

*Xmgrace* is a 2D graph plotting tool which produces publishing quality output (<http://plasma-gate.weizmann.ac.il/Grace/>).

*Style notes.* This example provides many commands that must be typed at the terminal. These commands are presented as follows:

\$ command

where the \$ sign is the terminal prompt and should not be typed. The terminal prompt might differ between systems.

When a command in an example is too long to fit in a single line, it is wrapped and then indented on subsequent lines, like this:

```
$ echo "This sentence was too long to fit
      in a single line"
```

When entered at the terminal, the entire command should be typed on a single line.

*Note.* Several notes are introduced in italics within this text. These notes refer to concepts not directly applicable to the given example but important when dealing with other membranes.

### 3. Simulation of the System

To perform an MD simulation in Gromacs the *mdrun* program is used. This program requires only one file to perform a simulation. This file is the so-called *tpr* file. This is a binary file that contains the initial coordinates and the force field parameters of the system in addition to the specifications of the algorithms and their parameters to be used in the simulation. The *tpr* file is generated by program *grompp* by the following command:

```
$ grompp -c Initial_coordinates.gro -p forcefield_parameters.top -f
algorithm_parametersmdp
-o simulation_file.tpr
```

Before starting the simulation of the DPPC membrane with Gromacs three independent files have to be prepared. Each of them contains an essential part of the simulation details. The first file, “\*.gro,” contains the initial coordinates of each atom present in the simulated system. The second file, “\*.top” contains the descriptions of the forces between the atoms (molecular topology and force field). The third file, “\*.mdp” contains specification of the algorithms used in the calculation and their parameters.

The DPPC simulation used as example in this chapter is based on the recent work of Andreas Kukol (6). The files required to follow the example containing the membrane coordinates (“dppc128\_40ns.pdb”) and force field (“dppc\_53a6.itp”) constitute the

supplementary material of his published work: *J. Chem. Theory Comput.* 5:615. They are freely available from the journal web page.

### 3.1. Setting Up the Simulation System

Although DPPC membrane coordinates are already stored in the “dppc128\_40ns.pdb” file, it is worth showing how they can be obtained from the scratch. This information might be useful when setting up new systems and is rarely found elsewhere. *When working with an already simulated system, it is often better to extend the previous simulation rather than make a new one from scratch. This provides a system that is closer to equilibrium.* A commonly used protocol to build single component lipid bilayers from scratch is presented below. Unless specifically mentioned, the tools used here are found in the Gromacs simulation suite.

1. First, the structure file of a single DPPC lipid molecule has to be prepared (see Fig. 1a). This can be done by drawing the molecular structure in a visual molecular tools program like *Avogadro*. As the force field used in this example does not take into account nonpolar hydrogen atoms, they should be removed from the structure before proceeding to the next step. *It is very important to check the chirality of the chiral atoms and possible isomerizations (e.g., double bonds) as they do not change during the simulation.* This step provides the PDB format file (<http://www.wwpdb.org/docs.html>) of a single DPPC molecule (DPPC.pdb). *The order of the atoms in this file should be the same as the one used in the parameters file (itp file), see below.*

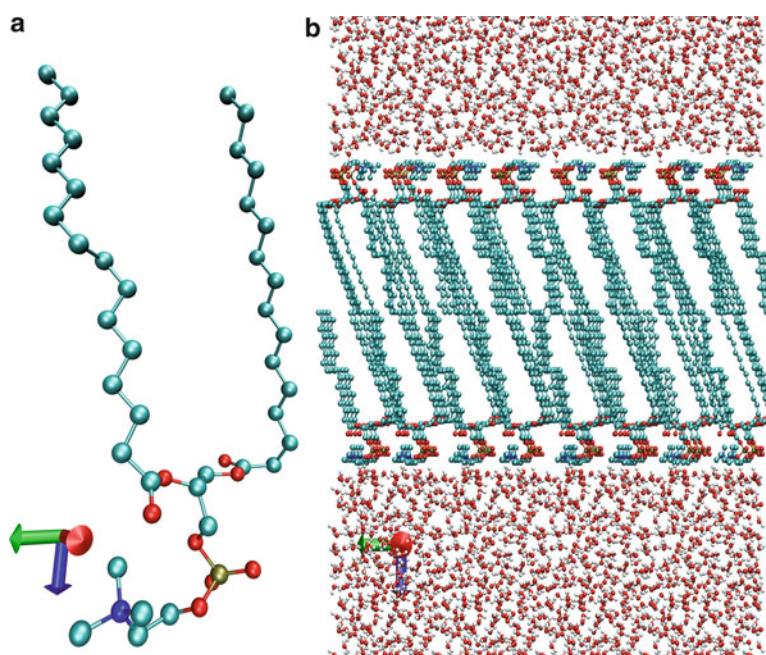


Fig. 1. Initial extended DPPC molecule (a) used for the construction of DPPC membrane (b). The blue arrow indicated the z-direction in the system.

2. Once structure of first molecule is prepared, it must be aligned so that its main axis is parallel to the  $z$ -direction, e.g., acyl chains along  $z$ -axis, so the head group is placed down in the absolute reference system (this direction is the one assumed in this example for the initial DPPC molecule). In order to simplify this and the following steps it is a good practice to draw the acyl chains in their extended positions and parallel one to the other, so that the lipid takes the minimum space in the  $xy$ -plane.

```
$ editconf -f DPPC.pdb -o DPPC_rot.gro -rotate X Y Z -box 0.96 0.96 2.20
```

- (a) “X,” “Y,” and “Z” are degrees of rotation around each axis and they depend on the original positioning of the DPPC molecule when drawn. Often the correct values for “X,” “Y,” and “Z” are found through a process of trial and error.
  - (b) The values in the box option define how close the molecules are to each other in the next step. They can be customized to the particular needs. In this case the initial lipid box is set to an area per DPPC lipid of  $0.922 \text{ nm}^2$  ( $0.96 \text{ nm} \times 0.96 \text{ nm}$ ).
  - (c) The output of this command is the coordinates file of the rotated DPPC molecule in the Gromacs format.
3. In the next step, a monolayer (the future down layer of the membrane) is constructed. From a pure DPPC membrane this is done by replicating the prepared DPPC lipid  $8 \times 8$  times in the  $xy$ -plane (a total of 64 DPPC molecules per layer).

```
$ genconf -f DPPC_rot.gro -o DPPC_down-layer.gro -nbox 8 8 1 -renumber
```

4. The upper layer of the final membrane can be obtained by rotating the previous monolayer in the X direction by  $180^\circ$ .

```
$ editconf -f DPPC_down-layer.gro -o DPPC_up-layer-up.gro -rotate 180 0 0
```

5. The membrane is now easily constructed by placing the up-layer on top of the down-layer. To do so, first, the up-layer has to be translated along the Z coordinate so that its position is on top of the down layer.

```
$ editconf -f DPPC_up-layer.gro -o DPPC_up-layer_trans.gro -translate 0 0 Z
```

- (a) “Z” is the distance in nanometers for which the up-layer needs to be translated, approximately  $2.20 \text{ nm}$  in this example. *It has to be taken into account that layers in their final position should neither be overlapped nor be too separated. If they overlap, the simulations or even the minimization will crash soon after starting. If too much space is left sometimes both layers turn apart. In both cases, this will mean the repetition of everything from this point.*

- (b) By loading DPPC\_down-layer.gro and DPPC\_up-layer\_trans.gro simultaneously using the *vmd* visualization software the effect of the translation can be checked. Start a new VMD session. In the VMD main window, choose *File->New Molecule...* Select DPPC\_down-layer.gro. Repeat the process for DPPC\_up-layer\_trans.gro.
6. Once the up-layer is properly placed, both files can be merged to generate the membrane.

```
$ cat DPPC_down-layer.gro DPPC_up-layer_trans.gro >>
DPPC_membrane.gro
```

- (a) Before executing the command delete the last line of DPPC\_down-layer.gro.
- (b) Before executing the command delete the first two lines of “*DPPC\_up-layer\_trans.gro*.”
- (c) After generating “*DPPC\_membrane.gro*,” open the file and change the number placed in the second line, by the number obtained by multiplying the original number by 2. This number indicates the number of atoms in the system. As both merged layers have identical number of atoms we just need to multiply this number by 2.

7. Finally, the membrane needs to be hydrated. The easiest way of doing that is to generate two slabs of water and place them one on top and one under the membrane in the same way as proceeded above. In fact, it is only needed to generate one slab. By translation of this slab to the desired position, above or below the membrane, both needed slabs are obtained.

```
$ genbox -cs -o water_slab.gro -box X Y Z
```

- (a) “X” and “Y” have to be the same as *x*- and *y*-box sizes in the membrane. They can be found in the last line of “*DPPC\_membrane.gro*” (first two numbers).
- (b) The “Z” value determines the level of hydration of the final system. *In the final system for the phosphatidylcholine membrane 25–30 water molecules per lipid should be enough if no ions or charge lipids are present.*

8. Proceed then with the translation in the same way as above. Use *vmd* to check.

```
$ editconf -f water_slab.gro -o water_slab-up.gro -translate 0 0 Z1
$ editconf -f water_slab.gro -o water_slab-down.gro -translate
0 0 Z2
```

- (a) Do not put the water slabs too close to the membrane as it will generate too many bad contacts. This will result in a useless coordinate set for the next steps due to the large forces artificially created between water and lipids. However, leaving too much space can lead to the above mentioned artifact of layer separation.

9. After building both slabs and placing them properly, the final membrane system can be assembled.

```
$ cat DPPC_membrane.gro water_slab-up.gro water_slab-down.gro > membrane.gro
```

- (a) Before assembling the files delete the first two rows and the last one of each file. First copy the number placed in the second row of each file.
- (b) After assembling insert two lines at the beginning. Add a title in the first line. In the second line add the number of atoms in the system (the sum of the numbers placed in the second row of each file).

10. Finally, the box of the final system must be inserted. This can be performed as follows:

```
$ editconf -f membrane.gro -o membrane_box.gro
```

11. At this stage the coordinate file of the membrane is ready for minimization (see Fig. 1b) to avoid bad contacts, which often are created during this process (not shown). After minimization the system is ready for simulation.

### **3.2. Force Fields**

#### **Parameterization and Force Field Files**

Force field parameterization is the most difficult step in molecular simulations. In the MD context, parameterization means the description of the forces, functional form and associated parameters between each of the atoms in the system. The high intrinsic level of difficulty of the parameterization process often results in long developing and testing times. Therefore, when planning the simulation, only existing parameters or parameterized molecules should be considered. Otherwise, a significantly larger amount of time and resources needs to be allocated to complete the parameterization process. This means that the composition of systems that can be simulated is restricted as often there are no parameters for all desired molecules. *It should be kept in mind that mixing parameters from different force fields might lead to unpredictable results and therefore it is strongly discouraged.*

*There are many force fields available for membranes simulation and new ones are currently being developed, therefore before starting a new project it is appropriate to check the performance of each of the currently available force fields in relation to the goals. One should also consider that not all force fields are available in Gromacs and therefore it might be interesting to consider other simulation software compatible with the desired force field.*

Continuing with the example of the DPPC membrane, the preparation of the topology file, “\*.top,” of the molecules present in the system is needed. This file, which is a plain text file, includes information about type of potential function used, all parameters needed to calculate potential and information about molecular topology (chemical structure of molecule). It is important to note that the top file works as a container. The information inside is

either written explicitly or included as a reference to other external files. A mix of both scenarios is normally usual. For a complete description of such file see the Gromacs manual. The content of the “*dppc128.top*” file used in this example is the following:

```
#include "ffG53a6.itp"
#include "dppc_53a6.itp"
#include "spc.itp"

[ system ]
Bilayer DPPC,
[ molecules ]
DPPC 128
SOL 3655
```

The first line indicates which force field is used. In this case the chosen force field is Gromos96 53a6. The parameters for this particular force field are already incorporated in the Gromacs package in a file called “*ffG53a6.itp*.” This file can always be accessed by Gromacs as it is part of its installation.

The second file includes the “*dppc\_53a6.itp*” file provided by Andreas Kukol. *This file has to be placed in the same directory where the grompp program will be executed.* This contains a list of the atoms in the molecule and lists of interactions present in the DPPC molecule (e.g., list of chemical bonds, valence angles, and torsion angles), in the context of the Gromos96 53a6 force field. *Taking a closer look at this file it can be seen that in most of the cases only atoms which interact are mentioned but no parameters are given. The missing parameters are taken from the previously mentioned “ffG53a6.itp.”*

Similarly, the third line corresponds to the topology file of the water. As the file name suggests, in this case the SPC model has been used, which is compatible with the Gromos96 53a6 force field. *It is important to understand that no other water models should be used unless their compatibility is known. Otherwise a lot of work will have to be performed in order to validate such change.*

After the “include” lines there are two different sections. Each of them can be easily identified as they are preceded by a keyword between brackets. The first one, [ system ], is only followed by a line containing a title. The second one, [molecules], is followed by the kind of molecules and their amount in the exact order as they appear in the coordinate file (the *pdb/gro* file). *Take into account that the names of the molecules are the ones found in the [molecule-type] section in each of the itp files included.*

### 3.3. Simulation Parameters

The last file required by Gromacs to run a simulation is the so-called *mdp* file, which is a plain text file. This file provides information about the kind of simulation being performed (molecular dynamics, minimization, stochastic dynamics, . . .), the algorithms that are used for a given task (temperature and pressure baths, treatment of electrostatic interactions, etc.), and also several other parameters

required by the used algorithms (time step, number of steps, how often to save a snapshot, etc.). *It is important that while some options can be freely adjusted to the desired requirements, others are hard-linked to the used force field or other constrains.* A simplified *mdp* file, sufficient to simulate the DPPC membrane, *dppc128.mdp*, looks like this (excluding the letters at the end of each line):

```

integrator = md (a)
dt = 0.002 (b)
nsteps = 100000000 (c)
nstxout = 50000 (d)
nstvout = 50000 (e)
nstlog = 5000 (f)
nstenergy = 5000 (g)
nstxtcout = 5000 (h)
pbc = xyz (i)
coulombtype = PME (j)
rcoulomb = 1.0 (k)
rvdw = 1.4 (l)
tcoupl = Berendsen (m)
tc-grps = System (n)
tau_t = 0.1 (o)
ref_t = 325 (p)
Pcoupl = Berendsen (q)
Pcoupltype = semiisotropic (r)
tau_p = 2.0 2.0 (s)
compressibility = 4.5e-5 4.5e-5 (t)
ref_p = 1.0 1.0 (u)
constraints = all-bonds (v)

```

The above file indicates that the simulation performed here is a molecular dynamics simulation (a). The time step chosen is 2 fs (b), and the number of steps is 1.000.000 (c). This means that the total length of the simulation is set to 200 ns.

There are several files generated by Gromacs that contain the output of the simulation. Two of them contain the trajectory in binary format. The first one is the *trr* file, which contains the coordinates and velocities in high precision. This type of file is usually large. Since such precision is not required for most of the analyses, there is no need to save frames in this format frequently. The given file sets the saving period to 100 ps for both coordinates (d) and velocities (e) of the atoms. The second trajectory file, *xtc*, is the one typically used in analysis. It contains only atomic coordinates with a default precision of  $10^{-3}$  nm. This file requires a significantly smaller disk space and therefore is recommended. In our example the saving period is set to 10 ps (h), which is the typical saving time period to analyze most of the processes of interest in membranes. The next file type is the *log* file. The content of this text file provides an easy way to check the evolution of the simulation. For each saving point, 10 ps in this

example (f), it shows a summary of the values of selected energy terms, temperature, pressure, etc. Another important file obtained in the simulation is energy file, *edr*. This binary file contains an extended set of energy terms, and other simulation parameters for each saved frame, 10 ps in our example (g).

There are several other options that need to be specified when simulating membranes. The first one is the presence of periodic boundary conditions (i). Another important set of parameters is the treatment of the nonbonded interactions (j) (k) (l), which are hard-linked to the used force field.

Also a very important set of parameters defines the simulation ensemble. The NPT ensemble is the natural ensemble for simulating biological membranes. This ensemble requires two baths to keep the temperature (m) (n) (o) (p) and pressure (r) (s) (t) (u) constant in the system.

Finally in this example all the bonds in the system are constrained so that they do not oscillate during the simulation (v). *This is very important as it allows the increase of the time step from 1 to 2 fs.*

A final note about the *mdp* file. Many options required for the simulations were not explicitly included in this example *mdp* file. This happens because the default values provided by Gromacs are in this case appropriate. Note that this might not be true for other cases. To see a complete list of the parameters used for a given simulation check the *mdp* file output by the *grompp* program, see below. Also it is recommendable to carefully read the Gromacs manual in respect to all possible options available for the simulation.

### **3.4. Running a Simulation**

Running a simulation in Gromacs is a two steps process when all files are ready. The first step generates a single file with all the previously gathered information in addition to the execution of several checking up in the system.

```
$ grompp -f dppc128.mdp -c dppc128_40ns.pdb -p dppc128.top -o dppc128_1.tpr
```

The *grompp* program may give you some warnings or even errors. It is worth checking them all always before proceeding. Once the *tpr* file is correctly generated the simulation can be performed. This is done by running the following command:

```
$ mdrun -v -deffnm dppc128_1
```

This is a very simplified but convenient way of using the *mdrun* program. This command shows an estimated time for the simulation to end. Running *mdrun* also generates the already mentioned files (“*dppc128\_1.trr*,” “*dppc128\_1.xtc*,” “*dppc128\_1.edr*,” “*dppc128\_1.log*”). Additionally, at the end of the simulation it also generates a *gro* file, “*dppc128\_1.gro*,” which contains the coordinates of the last configuration explored during the simulation. For a full explanation over *mdrun* capabilities check the Gromacs manual or execute:

```
$ mdrun -h
```

All Gromacs programs provide an explanation of their capabilities by executing the command followed by `-h`.

An additional step to be done before running the MD simulation is a minimization. This step removes close contacts between atoms, and therefore provides a set of coordinates with more adequate atomic positions, which in turn constitutes a better choice of initial structure for a given MD simulation. The procedure in Gromacs is very similar to the one described in this section. Only some changes are required in the `mdp` file. Please refer to the Gromacs manual to see the needed changes.

### **3.5. Validation of the Simulation**

After the simulation is finished, several aspects of the obtained trajectory have to be examined to ensure its physical reliability. The analysis can be conceptually divided into two parts. The first one examines the trajectory to find technical issues. The second one, which is probably harder, compares the results of the simulation with experimental data. This section deals mostly with the technically associated problems, leaving for later sections the comparison between simulations and experiment.

Before performing any sophisticated test on the trajectory, a visual inspection of the membrane trajectory might be of help in detecting problems. This can be easily done with `vmd` using the following commands:

```
$ echo 0 | trjconv -f dppc128_1.xtc -s dppc128_1.tpr -o dppc128_1_whole.xtc -pbc whole
$ vmd dppc128_40ns.pdb dppc128_1_whole.xtc
```

Large trajectories require a lot of computer memory for being displayed. If the visualization of the trajectory cannot be performed because the computer used does not have enough memory, try to add the “`-b`” option to the first command followed by the amount of time to be skipped. This makes the trajectory shorter and therefore easier to visualize.

By means of visual inspection abnormal interactions between atoms or broken bonds can be easily observed. Another critical property easily observed is the membrane phase behavior. It is fundamental to ensure that the simulated membrane shows the correct phase behavior corresponding to the given simulation conditions. The presence of any of these problems will require additional work in the reparameterization of the system. Another possible problem is the presence of lipids that have escaped from the membrane to the water phase during initial steps of simulations. Due to the limited system size and time scale of the simulations, the effect of the free lipid in the membrane–water interface might significantly change many physical properties of the system. In order to correct this problem few steps of dynamics can be alternated with few steps of minimization, although many other strategies can be used such as reconstructing the whole bilayer again

giving slightly different initial area per lipid. Visual inspection should be performed specially at the beginning of the simulation (first 100 ps) when most of above mentioned problems occur. This can eventually save lots of computational resources.

Once the visual inspection does not indicate any error it should be checked if the simulation has achieved a thermal equilibration. The first property to be checked is the behavior of the temperature and potential energy during the simulation time. This can be done with the *g\_energy* tool that allows the extraction of several simulation parameters as a function of simulation time, which are stored in *edr* file.

```
$ echo "Temperature" |g_energy -f dppc128_1.edr -o Temp.xvg  
$ echo "Potential" |g_energy -f dppc128_1.edr -o Potential.xvg
```

The output files “\*.xvg” can be visualized with the *xmgrace* (i.e., \$ *xmgrace Temp.xvg*). The next fundamental property to determine if the system reached equilibration is the area per lipid. This is the area of the box in the direction of the membrane plane divided by the number of lipid molecules in a layer (64 DPPC in the example). It can be easily calculated as follows (see Fig. 2):

```
$ printf "%s\n%s\n" "Box-X" "Box-Y" | g_energy -f dppc128_1.edr  
-o Box.xvg
```

```
$ xmgrace -nosafe -nxy Box.xvg -pexec "s0.y=(s0.y*s1.y)/64"
    -pexec "kill s1"
    -pexec "autoscale" -pexec "yaxis label \"Area/lipid\""
    -pexec "saveall \"Area.xvg\""
```

Note that the last command outputs a file “Area.xvg” file, which contains the area per lipid necessary for analysis conducted in the following section. The value of the area per lipid is not as stable with

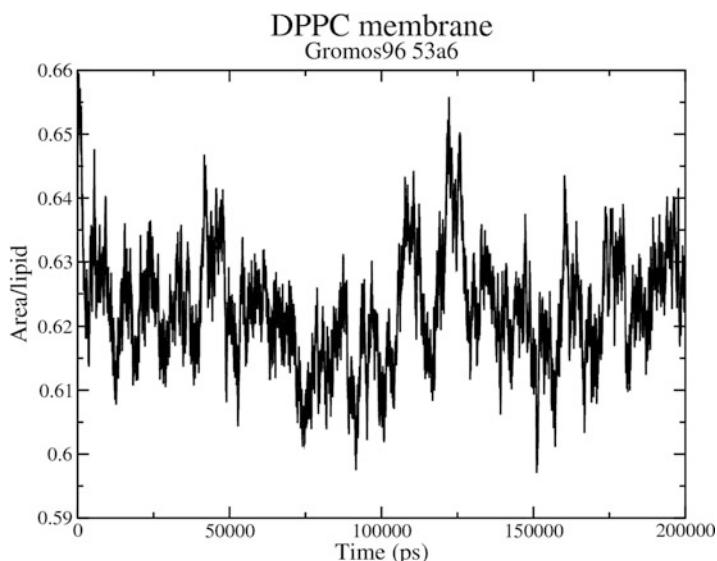


Fig. 2. Time development of the surface area over the simulation run.

the time as the previous two properties. It is characterized by the presence of high oscillations around a mean value. This is why the determination of the equilibration time is not an easy task.

### **3.6. Production Trajectory**

To calculate any membrane property it is not only needed to ensure that the system has reached the equilibrium, but also that it has been simulated long enough (production time) to ensure a proper sampling of the phase space of all the internal degrees of freedom. Proper sampling provides the correct values of the calculated properties that otherwise might be bias. The length of production time depends on the simulated system, therefore it is difficult to predict in advance. The simulation length also depends on the properties to be extracted from the MD trajectory. Some suggestions to estimate the correct simulation length are given below:

1. Due to the increasing computers power, there is no reason to simulate any membrane for less than 150 ns of production time. This is the time when the system loses memory with respect to the initial configuration in case of the autorotation of the glycerol backbone which is the slowest rotation mode observed in the lipids. However, shorter simulations still provide accurate values for many structural properties.
2. For nonionic/neutral membranes 20 ns equilibration times should be added to the desired production time (7). In case of considering ions/charged lipids the equilibration times might be in the 100's ns time scale (8).
3. For calculating diffusion coefficients simulations about 1  $\mu$ s have to be considered. In this case several replicas of the simulations may also be helpful to estimate the associated error of the measures. Additionally, a larger membrane patch should be also considered, since diffusion is a collective process with a correlation range of  $>2.5$  nm (9).
4. Above estimations apply only to single component membrane in the liquid disordered phase. In case of mixtures or gel phase, much longer simulation times are needed to allow proper mixing of the lipids, phase separation or aggregation processes.

---

## **4. Analysis of the Membrane Properties**

In this section we describe how to calculate several important structural and dynamical properties of a pure membrane by means of the Gromacs analysis tools. These tools provide a good starting point to characterize the membrane. *Consult Gromacs manual for a full explanation of the commands in use.* In order to show how these tools work, the analysis of the example DPPC membrane (200 ns) is provided. The error bars can be estimated using the block analysis method (10) provided by the *g\_analyze* program.

#### 4.1. Area per Lipid Molecule

The area per lipid molecule in a pure bilayer corresponds to the total average membrane area divided by the number of lipid molecules in a single leaflet (see Fig. 1). It is a basic variable that characterizes the bilayer structure, and its value depends significantly on the membrane composition, hydration level of the membrane, ionic force of the medium, pH, temperature and pressure. The value of area per lipid summarizes large amounts of structural and dynamic properties of the membrane, specially its ordering (11). The area per lipid is a clear indicator of the membrane's phase and is often used in validation of the particular lipid force field. The calculation of area with Gromacs tools is a three-step process. The first two steps were already seen and commented above:

```
$ printf "%s\n%s\n" "Box-X" "Box-Y" | g_energy -f dppc128_1.edr
-o Box.xvg
$ xmgrace -nosafe -nxy Box.xvg -pexec "s0.y=(s0.y*s1.y)/64"
-pexec "kill s1" -pexec "autoscale"
-pexec "yaxis label \"Area/lipid\" " -pexec "saveall \"Area.xvg\" "
```

The final command computes the average value and the associated error of the area per lipid taking into consideration that the first 20 ns are equilibration time.

```
$ g_analyze -b 20000 -ee -f Area.xvg
```

The output of this command indicates an area per lipid of  $0.6224 \pm (0.0059) \text{ nm}^2/\text{DPPC}$ . The error bar is twice the standard error 0.00298.

#### 4.2. Membrane Thickness

The membrane thickness characterizes the transverse structure of the membrane. It constitutes an essential property of the bilayer that in turn modulates membrane functionality, like the hydrophobic mismatch affecting integral protein functionality. The membrane thickness strongly depends on the lipid composition. Changes in the headgroup, length, or unsaturation level of the acyl chains are reflected in its value.

Membrane thickness is not a well-defined parameter; thus, several alternative definitions exist (11). One possibility is the so-called PP distance, which is relatively easy to calculate for pure membranes. This is the distance between the average positions of phosphorus atoms in opposite membrane leaflets. It is possible to calculate PP distance using *g\_dist* program from Gromacs package as follows:

1. Prepare an index file with two groups each containing the phosphorous atoms separately for each layer. Gromacs provides a tool to help making any required index file, *make\_ndx*. In this example first the P atoms of the down-layer, lipids 1–64, are selected which is followed by P atoms of the up-player, lipids 65–128.

```
$ printf "%s\n" "r 1-64 & a P" "r 65-128 & a P" "q" |  
make_ndx  
-f dppc128_1.gro -o Dpp.ndx
```

2. After that the centering of the membrane is needed to prevent the possible drifting in the  $z$ -direction during the simulation.

```
$ echo "0" |trjconv -f dppc128_1.xtc -s dppc128_1.tpr  
-o dppc128_1-ini.gro -dump 10  
$ echo "0" |trjconv -f dppc128_1.xtc -s dppc128_1-ini.gro  
-o dppc128_1-nj.xtc -pbc nojump  
$ echo "1 0" |trjconv -f dppc128_1-nj.xtc -s dppc128_1.tpr  
-o dppc128_1-nj-ct.xtc -center -boxcenter tric -pbc mol
```

3. *g\_dist* provides the shortest distance between two given points. As only the trans-membrane distance is of interest here, an enlargement of the box is recommended to avoid the possibility that distance is calculated through the water phase instead. This might happen due to the periodic boundary conditions. In case of large hydration this step can be skipped.

```
$ echo "0" |trjconv -f dppc128_1-nj-ct.xtc -s dppc128_1.tpr  
-o dppc128_1-nj -ct -nobox.xtc -box 100 100 100
```

4. Once all preparations are done the calculation of the thickness of the membrane can be performed:

```
$ echo "3 4" | g_dist -f dppc128_1-nj-ct-nobox.xtc -  
s dppc128_1.tpr  
-o dppc128-Dpp.xvg -n Dpp.ndx -b 20000  
$ xmgrace -nosafe -nxy dppc128-Dpp.xvg -pexec "kill s0"  
-pexec "kill s1" -pexec "kill s2" -pexec "s3.y=abs(s3.y)"  
-pexec "autoscale" -pexec "saveall \"Dpp.xvg\" "
```

5. Finally, the thickness value and its associated error can be calculated by *g\_analyze*

```
$ g_analyze -ee -f Dpp.xvg
```

#### 4.3. Density Profiles

A membrane density profile is the time average of some extensive property of the system along the  $z$ -direction, the one perpendicular to the bilayer plane. Typically measured properties are mass, electronic density, charge, scattering length, occurrence, etc. Membrane density profiles are interesting as they can be obtained experimentally as well as from MD simulations. Although different experimental techniques provide access to different profiles, here the attention is focused on mass density profiles and electron density profiles.

To calculate the mass density profile the *g\_density* program can be used. Before calculating density profiles it should be ensured that the membrane has not drifted in the  $z$ -direction over the simulation time. This can be achieved by centering trajectory as described in Subheading 4.2. Once the centered trajectory is

**Table 1**  
**Example of “*electrons.dat*”**

53	O21 = 8	C213 = 8	C18 = 8
C33 = 9	C21 = 6	C214 = 8	C19 = 8
C34 = 9	O22 = 8	C215 = 8	C110 = 8
C35 = 9	C22 = 8	C216 = 9	C111 = 8
N = 6	C23 = 8	C1 = 8	C112 = 8
C32 = 8	C24 = 8	O11 = 8	C113 = 8
C31 = 8	C25 = 8	C11 = 6	C114 = 8
O32 = 8	C26 = 8	O12 = 8	C115 = 8
P = 16	C27 = 8	C12 = 8	C116 = 9
O33 = 8	C28 = 8	C13 = 8	OW = 8
O34 = 8	C29 = 8	C14 = 8	HW1 = 1
O31 = 8	C210 = 8	C15 = 8	HW2 = 1
C3 = 8	C211 = 8	C16 = 8	
C2 = 7	C212 = 8	C17 = 8	

obtained it can be used to calculate the mass density profile with the following command:

```
$ echo "0 1 2" | g_density -ng 3 -f dppc128_1.xtc -s dppc128_1.tpr  
-o density.xvg -sl 100
```

This command calculates three profiles: the first one corresponds to the whole system, the second one just to the membrane and the third one to the water. *The profile corresponding to any group of atoms can be also calculated just by providing an index file (-n option) where this group has been defined.*

Use the following command to visualize all three profiles simultaneously:

```
$ xmgrace -nxy density.xvg
```

The electron density profiles can also be calculated with the *g\_density*. This requires the “*electrons.dat*” file (an example files for the DPPC system is shown in Table 1). This file provides information on the number of electrons on each atom. The first line contains the number of atoms defined subsequently. The following lines contain the name of an atom, as seen in the *itp* files, and number of electrons associated. In this example the number of electrons corresponding to the neutral atom type has been used. Additionally, one electron has been added to the atom for each theoretically attached hydrogen. Remember that a united atom force field is used in this example. Finally, the valence of the atom (+1 Phosphorous and –1 Nitrogen) has been added.

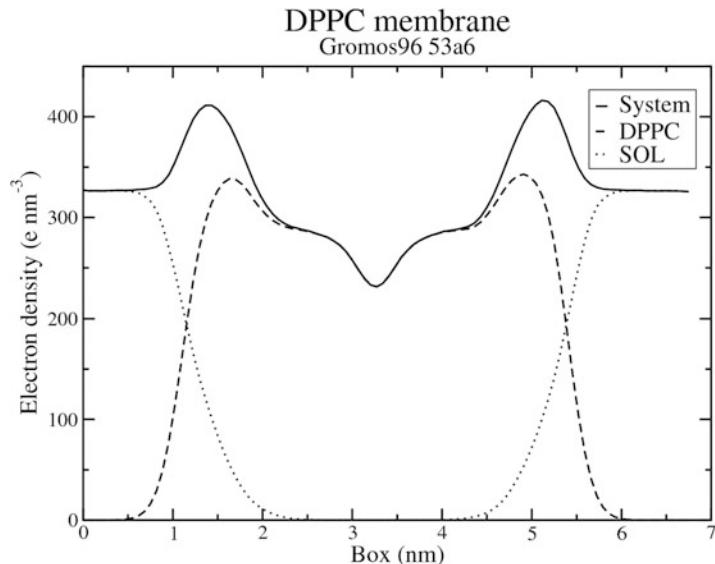


Fig. 3. Electron density profile of the DPPC, water, and all atoms in the system.

Once “*electrons.dat*” file is ready, the corresponding electron density profiles can be calculated with the following command (see Fig. 3):

```
$ echo '0 1 2' | g_density -ng 3 -f dppc128_1.xtc -s dppc128_1.tpr
-o edp.xvg -sl 100 -dens electron -ei electrons.dat
```

#### 4.4. Deuterium Order Parameter

The deuterium order parameter  $S_{CD}$  for pure saturated lipids can also be calculated in Gromacs by using *g\_order*. This parameter can be directly compared with the order parameter measured in NMR experiments, which provides an idea on how ordered is an acyl chain at the chosen segment.  $S_{CD}$  is defined as:

$$S_{CD} = \left(\frac{1}{2}\right) \langle 3\cos^2\theta - 1 \rangle,$$

where  $\theta$  is the angle between the Carbon–Deuterium bond and the bilayer normal, and the angular brackets denote averaging over time and over all phospholipids. The algorithm used by *g\_order* allows the calculation of  $S_{CD}$  without the need of explicit hydrogen atoms in the system. This is very useful when simulating membranes with a united atom force field, like the one used here. In order to calculate  $S_{CD}$  one must provide *g\_order* with the connectivity of the carbon atoms in the tail. This information is provided by means of an index file. In the case of DPPC the index file contains 16 groups, which include the indexes of all carbon atoms located in the acyl chain. The order of groups must be the same as the order of the atoms in the acyl chain. This happens because the  $S_{CD}$  of the  $n$  carbon is calculated using also the positions of the atoms  $n + 1$  and  $n - 1$ . This also means that neither

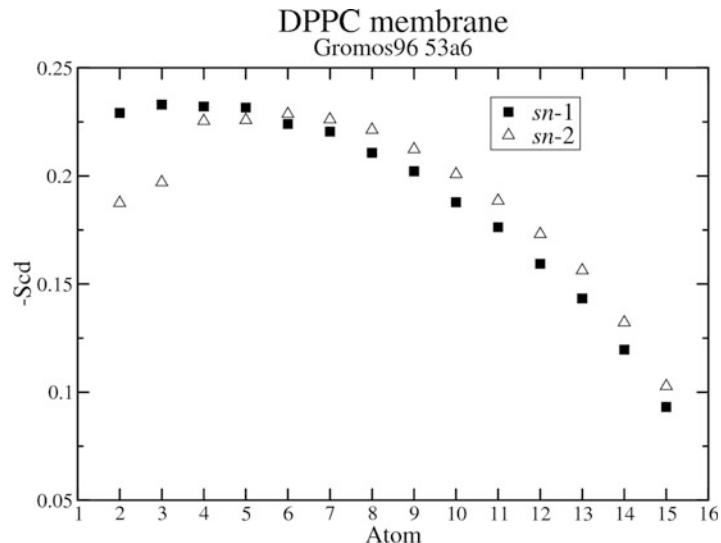


Fig. 4. Profile of the deuterium order parameter  $-S_{CD}$  for *sn*-1 and *sn*-2 chains of DPPC molecule.

the first atom of the chain (the one attached to the oxygen) nor the last (the terminal methyl group) will have an  $S_{CD}$  value calculated while using this method. The index files can be generated as follows:

*For sn-1 chain:*

```
$ printf "%s\n" "aC11" "aC12" "aC13" "aC14" "aC15" "aC16"
      "aC17" "aC18" "aC19" "aC110" "aC111" "aC112"
      "aC113" "aC114" "aC115" "aC116" "del0" "del0" "del0"
      "q" | make_ndx -f dppc128_1.tpr -o sn1.ndx
```

For  $sn\text{-}2$  chain:

```
$ printf "%s\n" "aC21" "aC22" "aC23" "aC24" "aC25" "aC26"
      "aC27" "aC28" "aC29" "aC210" "aC211" "aC212"
      "aC213" "aC214" "aC215" "aC216" "del0" "del0" "del0"
      "q" | make_ndx -f dppc128.l.tpr -o sn2.ndx
```

Once “*sn1.ndx*” and “*sn2.ndx*” are generated, *g\_order* is invoked in the following way:

*For sn-1 chain:*

```
$ g_order -f dppc128_1.xtc -s dppc128_1.tpr -n sn1.ndx  
-o order_sn1.xvg -od deuter_sn1.xvg -b 2000
```

For  $sn\text{-}2$  chain:

```
$ g_order -f dppc128_1.xtc -s dppc128_1.tpr -n sn2.ndx  
-o order sn2.xvg -od deuter sn2.xvg -b 2000
```

To visualize the results just execute (see Fig. 4):

```
$ xmgrace deuter sn1.xvg deuter sn2.xvg
```

#### 4.5. Radial Distribution Functions

The fluid nature of the membrane, due to the balance between the attractive intermolecular interaction and the disordering thermal motion, is shown by the nonexistence of a well-defined (solid-like) positional ordering. Despite this, membranes in their fluid phase still display certain levels of spatial ordering. The analysis of this ordering allows for the identification and description of the fundamental interactions between membrane components that are responsible for the final membrane structure. This analysis is usually performed using a radial distribution functions (RDF).

An RDF describes the probability of finding a particle A at a distance between  $r$  and  $r+dr$  away from a particle B. Limiting the example to 2D RDF, which are very useful to study the ordering in the membrane plane ( $xy$ -plane),  $\text{RDF}_{xy}$ 's can be obtained as follows:

$$\text{RDF}_{xy}(r) = \langle (N_{\text{seg}}(r)/A_{\text{seg}}(r)) \rangle / (N_{\text{total}}/A_{\text{total}}),$$

where the angular brackets denote time average,  $N_{\text{seg}}(r)$  is the number of A particles in an annulus of radius  $r$  and width  $dr$  around the particle B,  $A_{\text{seg}}(r)$  is the area of this annulus,  $A_{\text{total}}$  is area of the simulation box, and  $N_{\text{total}}$  is the total number of A particles. In Gromacs, RDFs can be calculated by using `g_rdf`.

Next there is an example of how to calculate the RDF of the center of mass of the *sn*-1 acyl chains with respect to themselves. First it is needed to construct an index file containing all atoms located in the *sn*-1 acyl chain of the DPPC lipids located in the down layer.

```
$ echo "aC11 C12 C13 C14 C15 C16 C17 C18 C19 C110 C111
    C112 C113 C114 C115 C116 &r1-64" |make_ndx
    -f dppc128_1.tpr -o sn1_r1-64.ndx
```

Finally the RDF in the membrane plane (see Fig. 5) can be calculated as follows:

```
$ g_rdf -f dppc128_1.xtc -s dppc128_1.tpr -o rdf_CM_sn1.xvg
    -xy -n sn1_r1-64.ndx -rdf mol_com -b 2000
```

#### 4.6. Rotational Autocorrelation Functions

Until now it has been seen how Gromacs can be used to calculate the structural properties of the membrane. In addition, Gromacs can calculate several dynamical parameters of the membrane, which help completing the picture.

One interesting dynamical property corresponds to the rotational modes of the lipids. These modes can be obtained by analyzing the so-called autocorrelation functions (ACF) of any intramolecular vector. These functions measure the temporal memory of the system to retain a particular configuration (vector direction). Therefore, the longer time the configuration is retained, the slower is the analyzed mode. In Gromacs, the available ACF are based on the associated Legendre polynomials,

$$\text{ACF}_l(\tau) = \langle P_l[n(t_0)n(t_0 + \tau)] \rangle,$$

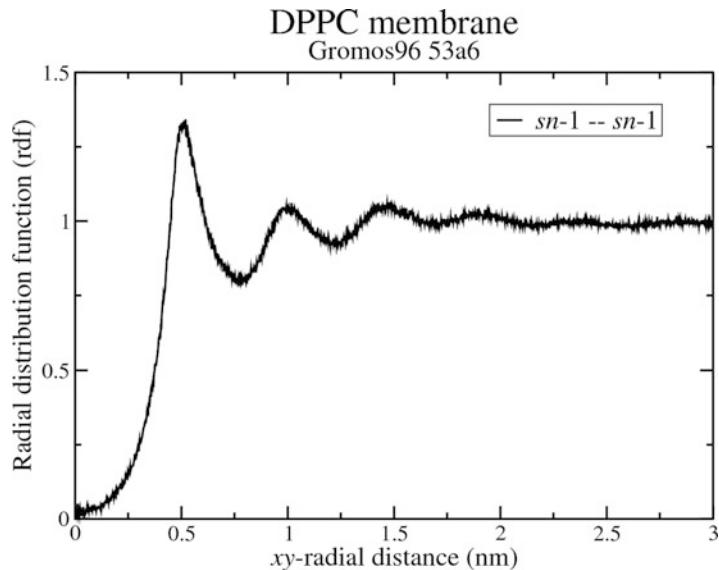


Fig. 5. Radial distribution function of the center of mass of *sn-1* tail relative to the other *sn-1* tails.

where  $P_l$  is the  $l$ -th Legendre polynomial,  $\mathbf{n}$  is a fixed unit intermolecular vector,  $t_0$  and  $\tau$  are the initial and elapsed times, respectively, and  $\langle \dots \rangle$  denotes an averaging over different initial times within a simulation run and over all molecules in the system. This correlation function usually decays as a stretched exponential with time. As an example of how to calculate ACF with Gromacs, the ACF of the  $m-1$  acyl chain (C11–C16 vector) and glycerol (C1–C3 vector) will be calculated. First an index file is generated. In this case it contains two groups. The first one containing the C11 and C16 atoms and the second one containing the C1 and C3 atoms.

```
$ printf "%s\n" "a C11 C116" "a C1 C3" "del 0" "del 0" "del 0"
      "q" |make_ndx -f dppc128_1.tpr -o acf.ndx
```

Finally, the ACFs are calculated as follows:

```
$ echo 0 | g_rerotacf -f dppc128_1.xtc -s dppc128_1.tpr -n acf.ndx  
-b 20000 -P 1 -d -o acf-chain.xvg
```

```
$ echo 1|g_rotacf -f dppc128_1.xtc -s dppc128_1.tpr -n acf.ndx  
-b 20000 -P 1 -d -o acf-Gly.xvg
```

The results are displayed with (see Fig. 6):

```
$ xmgrace acf-chain.xvg acf-Gly.xvg
```

## **4.7. Lateral Diffusion**

Lateral diffusion is an important dynamic property. It measures the capacity of the lipids to move along the leaflet. The lipid lateral diffusion is characterized by measuring the diffusion coefficient.

$$D(\tau) = \lim_{\tau \rightarrow \infty} \left( \frac{\text{MSD}(\tau)}{4\tau} \right),$$

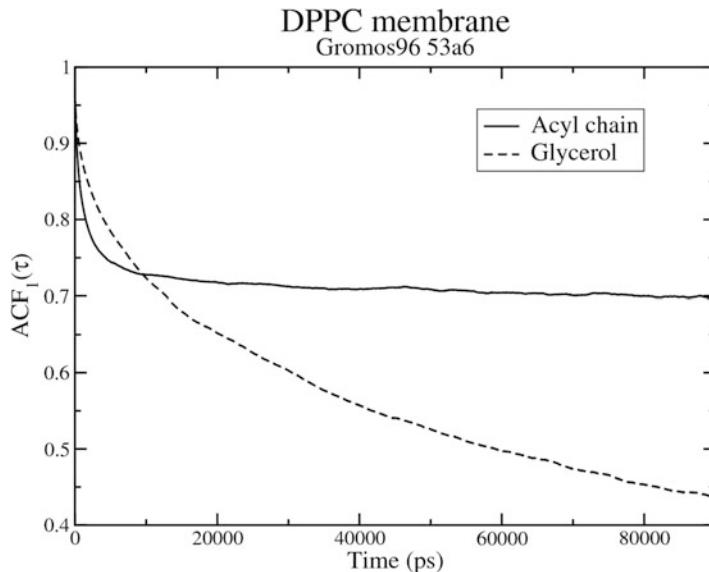


Fig. 6. Rotational autocorrelation function of two selected intramolecular vectors in DPPC molecule.

where  $\tau$  is the elapsed time and MSD is the mean-square displacement of the center of mass of the selected lipids in the  $xy$ -plane averaged over different initial times and the number of lipids. It can be calculated as follows:

$$\text{MSD}(\tau) = \langle |\mathbf{r}(t_0) - \mathbf{r}(t_0 + \tau)|^2 \rangle_{t_0, n}.$$

A linear (diffusive) regime of the MSD is identified in the simulations after some equilibration time, allowing the calculation of the diffusion coefficient. The MSD can be calculated in Gromacs with the `g_msd` program (see Fig. 7):

```
$ g_msd -f dppc128_1.xtc -s dppc128_1.tpr -lateral z
-o dppc_msd.xvg -mol diff_mol.xvg
```

It is clearly seen that the linear regime is not well captured by this simulation. Therefore, in order to properly calculate the diffusion coefficient of the DPPC lipids longer simulations have to be performed. Just as an example of how to proceed once the MSD curve is obtained, the extraction of the diffusion coefficient is shown. To begin with, a linear fit has to be performed in the linear zone, see the horizontal line in the plot. The obtained fitted straight line is shown with dots. Then the slope of the fitted line is calculated,  $24.627 \mu\text{m}^2/\text{s}$  in the example. To obtain the two-dimensional translational diffusion coefficient this number has to be divided by 4. The result of this example is then  $6.157 \mu\text{m}^2/\text{s}$  for the DPPC lipids in the simulated system. Note also that each layer in the membrane can drift independently during the simulation. Therefore, the motion of the center of mass of individual leaflets has to be

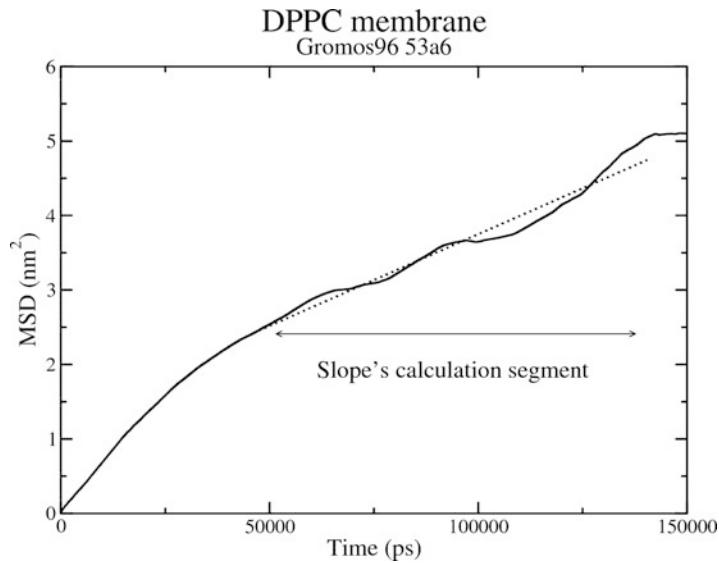


Fig. 7. Mean-square displacement of center of mass of DPPC lipid molecules in the membrane plane. Linear fit to calculate the diffusion coefficient is shown with a *dotted straight*.

*removed before computation of the lipid MSD. This is hard to do in Gromacs and therefore some programming will be needed to extend the capabilities of g\_msd.*

## Acknowledgments

We wish to acknowledge the Academy of Finland for financial support and CSC—IT Centre for Science (Finland) for computational resources. We also want to thank the Danish Centre for Scientific Computing for providing us access to the Horseshoe supercluster at the University of Southern Denmark (DCSC/SDU).

## References

- Mouritsen OG (2004) Life—as a matter of fat: the emerging science of lipidomics. Springer, Berlin
- Rög T, Pasenkiewicz-Gierula M, Vattulainen I, Karttunen M (2009) Ordering effects of cholesterol and its analogues. *Biochim Biophys Acta* 1788:97–121
- Marrink SJ, de Vries AH, Tieleman DP (2009) Lipids on the move: simulations of membrane pores, domains, stalks and curves. *Biochim Biophys Acta* 1788:149–168
- Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447
- Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. *J Mol Graph* 14:33–38
- Kukol A (2009) Lipid models for united-atom molecular dynamics simulations of proteins. *J Chem Theory Comput* 5:615–626

7. Falck E, Patra M, Karttunen M, Hyvönen MT, Vattulainen I (2004) Lessons of slicing membranes: interplay of packing, free area, and lateral diffusion in phospholipid/cholesterol bilayers. *Biophys J* 87:1076–1091
8. Zhao W, Rög T, Gurtovenko AA, Vattulainen I, Karttunen M (2007) Structure and electrostatics of POPG lipid bilayers with  $\text{Na}^+$  counterions: molecular dynamics simulation. *Biophys J* 92:1114–1124
9. Falck E, Rög T, Karttunen M, Vattulainen I (2008) Lateral diffusion in lipid membranes through collective flows. *J Am Chem Soc* 130:44–45
10. Hess B (2002) Determining the shear viscosity of model liquids from molecular dynamics simulations. *J Chem Phys* 116:209–217
11. Nagle JF, Tristram-Nagle S (2000) Structure of lipid bilayers. *Biochim Biophys Acta* 1469:159–195

# Chapter 16

## Simulations of Lipid Monolayers

Svetlana Baoukina and D. Peter Tieleman

### Abstract

A lipid monolayer lining a boundary between two immiscible phases forms a complex interface with inhomogeneous distribution of forces. Unlike lipid bilayers, monolayers are formed in asymmetric environment and their properties depend strongly on lipid surface density. The monolayer properties are also affected significantly by the representation of the pure interface. Here we give a brief theoretical introduction and describe methods to simulate lipid monolayers starting from force-fields and system setup to reproducing state points on the surface tension (pressure)–area isotherms and transformations between them.

**Key words:** Lipid monolayer, Langmuir monolayer, Monolayer collapse, Lung surfactant, Interface, Molecular dynamics

---

### 1. Introduction

A boundary between two immiscible phases is called an interface (or a surface if one of the phases is a gas). The molecules at interfaces have a higher potential energy than in the interior due to loss of favorable contacts with their neighbors and unfavorable or less favorable contacts with the molecules in the dissimilar phase. The interface is characterized by an excess free energy with respect to the bulk. This excess energy divided per unit area gives a surface tension at the interface. Surfactants (surface active agents) are typically amphiphilic molecules: they contain a hydrophobic and a hydrophilic part. Surfactants adsorb at a polar–apolar interface and orient the hydrophilic group towards the polar phase and hydrophobic towards the apolar forming a monomolecular layer (monolayer, see scheme in Fig. 1 and snapshots from simulations in Fig. 2). The monolayer reduces the surface tension by shielding unfavorable polar–apolar contacts (1). Surfactants can also self-assemble in the bulk phases into, e.g., micelles, vesicles, tubes, bilayers, hexagonal and cubic phases, depending on type of surfactant, hydration level,

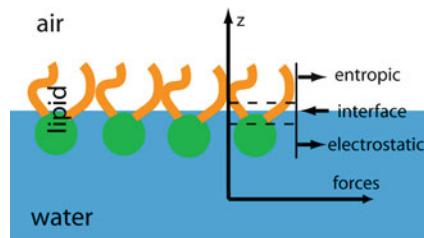


Fig. 1. Schematic representation of a lipid monolayer at an air–water interface.

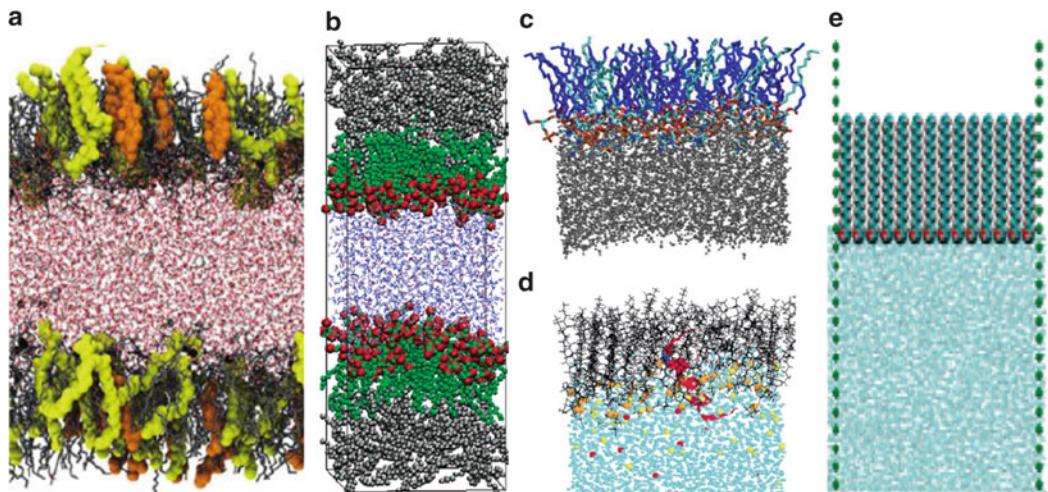


Fig. 2. Typical system setups used for simulations of lipid monolayers: two monolayers at the two air–water (a) and oil–water interfaces (b); a monolayer with a water slab underneath (c), a wall potential applied to the second interface (d), and a monolayer surrounded by walls (e). Reprinted with permission from (47, 50, 51, 58, 70). Copyright 2012.

temperature, etc. The distribution of surfactant between the interface and bulk phases depends on the difference in chemical potential and also kinetically on energy barriers for transfer.

Important properties of monolayers are the asymmetric conditions at the boundary between two different media, and a strong dependence of monolayer properties on the surface density of surfactant molecules. The surface density depends on the available interfacial area and surfactant concentration, and determines the resulting surface tension at the interface. The basic characteristic of surfactant monolayers is their surface tension (or surface pressure)–area dependence (normally measured at constant temperature and called an isotherm) which shows monotonic decrease of surface tension with increasing surfactant density and has so-called plateau (or constant surface tension) regions corresponding to phase transitions (see Fig. 3).

By reducing the surface tension at the interface surfactants can significantly improve material properties. Chemical surfactants have numerous technological applications as detergents, coatings, emulsifiers, wetting and foaming agents (2). Biological

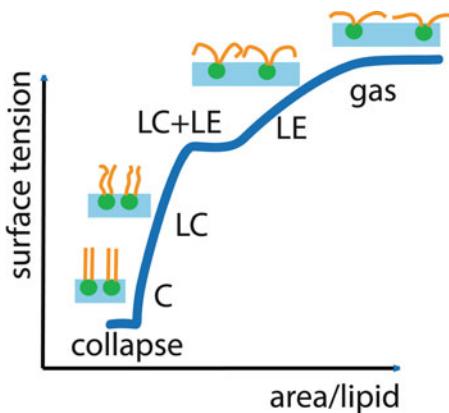


Fig. 3. Schematic representation of a surface tension–area isotherm of a monolayer. As the area per lipid is reduced, the monolayer adopts gas, liquid-expanded (LE), liquid-condensed (LC), condensed/solid (C) phases before collapsing from the interface.

surfactants—lipids—cover the outer layer of the eyes and the inner surface of the lungs (3).

Surfactant monolayers are convenient to form and manipulate and it is relatively straightforward to characterize their structure and viscoelastic properties. For over a century, they have attracted scientific interest leading to discoveries of a variety of 2D and surface phenomena in physics and chemistry (4–6). Lipid monolayers at air–water interfaces are often used to study the properties of lipid bilayers constituting the basis for cell membranes and lipid–protein or lipid–small solute interactions in cell membranes. For example, the Langmuir trough setup can be combined with light scattering, fluorescence and Brewster angle microscopy to visualize monolayer morphology; monolayers can be transferred to a substrate to analyze with a higher resolution scanning probe microscopy (7–11). In computer simulations, lipid bilayers have been studied extensively while monolayers have received less attention (12–14).

Monolayer simulations encounter difficulties with treating gas–liquid interfaces as many molecular building blocks have historically not been explicitly parameterized for the gas phase. As a result, such properties of pure interfaces as the surface tension or electrostatic potential are not accurately reproduced. The interfacial properties affect adsorption and phase behavior when surfactants are added, and careful comparison with experimental data is required (15–18).

In this chapter we focus on monolayer simulations starting from system setup and choice of a force-field to methods targeted to reproduce various properties: from selected points on the surface tension–area isotherm to phase transformations in 2D and 3D, potential difficulties arising and possible ways to address them. The discussed approaches are applicable in general to surfactants at gas–liquid and liquid–liquid interfaces.

---

## 2. Theory

The properties of surfactant monolayers vary strongly with surface density, which can be quantified by area per surfactant molecule. The area can be changed by changing surfactant concentration at the interface or by lateral compression/expansion of the interface. As the area is reduced at constant temperature, the monolayer undergoes transformations from gas to liquid-expanded (LE), liquid-condensed (LC), and solid/condensed (C) phases, with increasing translational and orientational order (19, 20). The phases adopted depend on the monolayer composition, temperature, the media forming the interface, and the range of surface tensions explored.

The surface tension at an interface covered by a monolayer,  $\gamma_m$  is determined as:

$$\gamma_m(A_L) = \gamma_0 - \Pi(A_L), \quad (1)$$

here  $\gamma_0$  is the surface tension at the pure interface (in the absence of surfactant),  $A_L$  is the area per lipid molecule, and  $\Pi(A_L)$  is the surface pressure arising from the interactions between the molecules. Idealized surface tension-area dependence is plotted in Fig. 3. The parameter measured experimentally is the surface pressure, as a function of the area per lipid, at constant temperature, and the dependence is called the surface pressure-area isotherm. As the monolayer thickness (~molecular length ~2–3 nm) is typically much smaller than its size in the lateral direction ( $\mu\text{m}$ –mm), its behavior to a large extent has a 2D character. The surface tension of the pure interface can be lowered by surfactants dramatically, and reach near-zero values. Below a threshold surface tension, which depends on the type of the surfactant, the monolayer becomes unstable at the interface and collapses to explore the third dimension. Despite its small thickness, monolayer-covered boundary between two phases represents a complex interface with the forces varying in nature and magnitude along the molecular axis (21, 22), see Fig. 1. This variation results in an inhomogeneous distribution of lateral pressure (23) along the monolayer normal defined as the difference between the lateral pressure component  $P_L$  and the normal component  $P_N$  of the pressure tensor:

$$\pi(z) = P_L(z) - P_N. \quad (2)$$

The pressure distribution summed over the interface gives the surface tension with opposite sign:

$$\int_0^b \pi(z) dz = -\gamma_m. \quad (3)$$

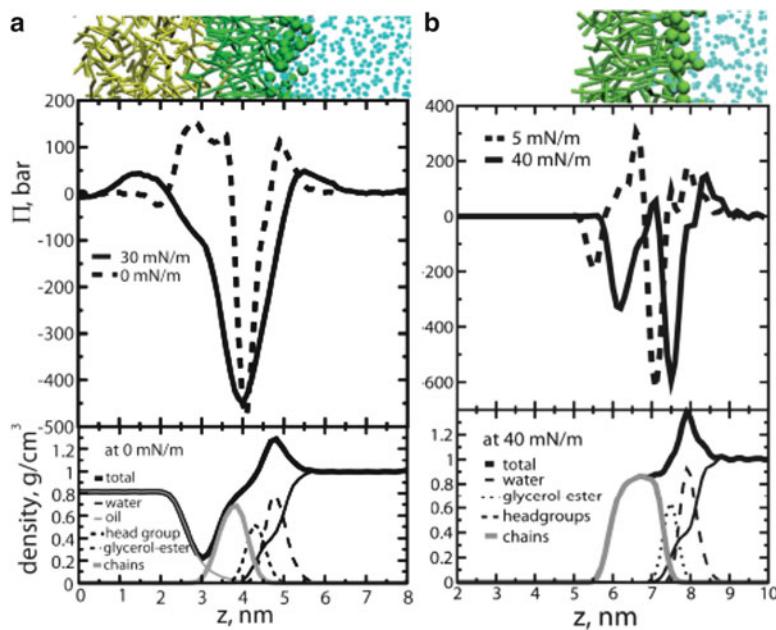


Fig. 4. Lateral pressure profiles across monolayer-covered oil–water (a) and air–water (b) interfaces calculated from simulations using a coarse-grained force-fields, Reprinted with permission from (24). Copyright 2012.

An example of the lateral pressure profile in a lipid monolayer at the air–water and oil–water interfaces calculated from simulations (24) is shown in Fig. 4. Positive pressure in the headgroup and surrounding water region results from net repulsive electrostatic interactions between charges and dipoles, while interactions between the chains are of entropic origin. These forces are counteracted by the surface tension at the chain–water and chain–air boundaries.

### 3. Materials

Here we discuss commonly used all-atom (AA), united atom (UA), and coarse-grained (CG) force-fields retaining significant amount of detail and chemical specificity.

#### 3.1. Pure Interfaces

As the balance of forces in monolayer includes the contributions from pure interfaces, the choice of a model for the water–oil or water–air interface will affect the monolayer properties in simulations. While the experimental values for the surface tension at the chain–water interface (~52 mN/m) and chain–air interface

(~22 mN/m) (25) are reasonably well reproduced, the air–water surface tension (~72 mN/m at 300 K) is often underestimated: ~55 mN/m in SPC water (26), ~53 mN/m in TIP3P AA water (27), ~32–36 mN/m for standard, and ~30 mN/m for polarizable MARTINI CG water (28), although there are recent improvements providing: ~68 mN/m for TIP4P/2005 AA water (29), ~76 mN/m for CG big multipole water (BMW) water (30), and ~72 mN/m in the CG model by Shinoda (31). Surface tension values are sensitive to the cutoffs used for truncation of nonbonded interactions, in particular for liquid–vacuum interfaces (15, 32). The solvent model is usually matched with a lipid force-field: (modified) TIP3P water with the atomistic CHARMM (33) and AMBER (34) force-fields, SPC water with GROMOS (35), and BMW is compatible with MARTINI. These factors are important to consider when choosing a force-field for lipids, in particular for simulations of expanded monolayers (in the LE/gas phases), in which the fraction of pure interface is large, as it could lead to a number of artifacts, see Notes 1,2. If the surface tensions at pure interfaces are not known for a selected force field and conditions, it is worthwhile calculating them from simulations. If electrostatic interactions are expected to be significant, it is also important to consider the interfacial potentials at pure interfaces (28, 30, 36).

### 3.2. Lipid Monolayers

In monolayer simulations, lipid composition is usually simplified with respect to real systems and mimics experimental model systems. Lung surfactant is a multicomponent mixture containing zwitterionic (phosphatidylcholine) and anionic (phosphatidylglycerol), saturated and unsaturated (oleoyl chains) lipids (37). Previous simulations considered pure dipalmitoyl phosphatidylcholine (DPPC) monolayers (24, 32, 38–46), as it is the main component of lung surfactant and one of the most studied lipids, or binary mixtures of DPPC with dipalmitoyl phosphatidylglycerol (DPPG) or palmitoyloleoyl phosphatidylglycerol (POPG) (17, 47, 48), as anionic lipids constitute a significant fraction of lung surfactant and are important for interactions with positively charged surfactant-associated proteins. More complex mixtures with phospholipids, cholesterol and fatty acids were used for simulations of lung surfactant and tear film (49–53). The most commonly used force-fields for lipid monolayers include AA CHARMM (33) force-field, UA Berger model (54) combined with GROMOS/OPLS (55) UA parameters, CG MARTINI force-field (56), and CG model by Shinoda (43). While the atomistic force-fields provide a higher level of detail, their ability to reproduce the phase behavior can be a concern. This is the case in particular if the phase transition temperatures in bilayers are shifted significantly with respect to experimental, or if an application of external surface tension is

required to reproduce experimental areas per lipid for (tensionless) bilayers, as it would shift the surface tension–area dependence in monolayers as well. Limited sampling, however, remains a major issue making many phenomena inaccessible to atomistic simulations due to inherent long times and/or large scales of transformations between (metastable) self-assembled aggregates.

### **3.3. Simulation Conditions**

Simulation parameters normally match physiological conditions or experimental setup. Room temperature 298–300 K, physiological temperature of 310 K and the higher value of 323 K (above the experimental main phase transition of DPPC lipid) are often used. The surface tensions can vary between that of a pure interface (in simulations, see Note 1) and zero, or negative values if destabilization of the interface (monolayer collapse) is of interest. The surface tension (total in the simulation box containing usually more than one interface, see Fig. 2) is determined from the difference between the normal  $P_N$  and lateral,  $P_L$  pressures:

$$\gamma = \langle (P_N - P_L) \cdot L_z \rangle, \quad (4)$$

where  $L_z$  is the box size in the normal to the interface(s) direction.

## **4. Methods**

### **4.1. System Setup**

Unless the process of monolayer self-assembly by adsorption is in focus (57), the monolayers should be prepared by placing all lipids at interface(s) in the starting configuration. This is because self-assembly is likely to lead to asymmetric distribution of lipids between the interfaces and to formation of lipid aggregates in the bulk water/oil phases. Starting configuration can be obtained by selecting a leaflet from a bilayer or generating lipid confirmations in a monomolecular layer. Typical setups (Fig. 2) include a water slab in vacuum or oil (58, 59) with two monolayers at the two interfaces (16, 60), or a monolayer at one of the interfaces (61), whereas a wall potential (17) or position restraints (42) can be applied to the second (pure) interface to prevent water evaporation. The total surface tension of the system,  $\gamma$  (see Note 1), calculated using a formula (4) then represents either twice the monolayer surface tension:

$$\gamma = 2\gamma_m \quad (5)$$

or a sum of the monolayer and the wall/water,  $\gamma_w$ , surface tensions:

$$\gamma = \gamma_m + \gamma_w. \quad (6)$$

Depending on the problem of interest other possible setups include monolayers on solid supports (62, 63) and monolayers

surrounded by walls (57), represented by particles with fixed positions or effective potentials. As the monolayer setups are anisotropic with significantly different properties in the normal and lateral directions, 2D instead of 3D long-range electrostatic (64) and Lennard-Jones (65) potentials/corrections can be used.

#### 4.2. Surface Tension/ Area Control

The target surface tension can be achieved by using a surface tension or semi-isotropic pressure coupling schemes (see Note 1), or by varying the monolayer surface density by changing the number of lipids at the interface or interfacial area (simulation box size) (48). In the former case, setting the system compressibility in the normal to the interface direction to zero will prevent compression of the vacuum region, while in the latter case after achieving the required density/area per molecule, constant volume simulations are often used which fix the normal and the lateral box size. Using this approach, entire surface tension (or pressure, see Note 2)—area isotherm can be calculated from simulations, see Fig. 5a, b. Due to short times accessed in simulations, compression/expansion rates upon variation of the box size are much faster compared to experimental, and the tension-coupling setups provide almost instantaneous changes. Slow monolayer compression can be modeled by equilibrations at constant area/surface tension following compression runs. If dynamic properties are of interest, it is worthwhile comparing monolayer and water viscosities (66).

#### 4.3. Phase Transformations

Transformations between the gas, LE and LC phases (see Fig. 6a) can be reproduced (32, 38, 39, 61) by setting the monolayer area to the plateau regions of the surface tension–area isotherms (see scheme in Fig. 2 and Note 2). This is straightforward if the phase coexistence region is known from previous simulations with

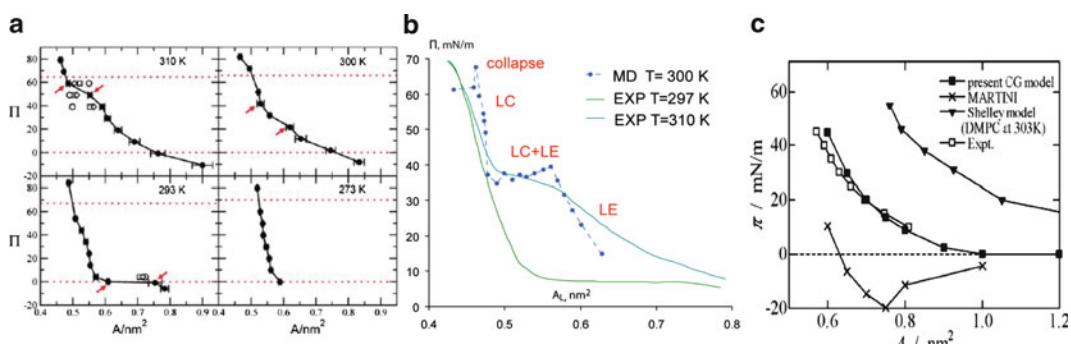


Fig. 5. Surface pressure–area isotherms of lipid monolayers calculated from simulations using atomistic (a) and coarse-grained (b, c) force-fields. Isotherms calculated at different temperatures (a), fitted to experimental data (b, c), and compared for different force-fields (c), Reprinted with permission from (32, 38, 43). Copyright 2012.

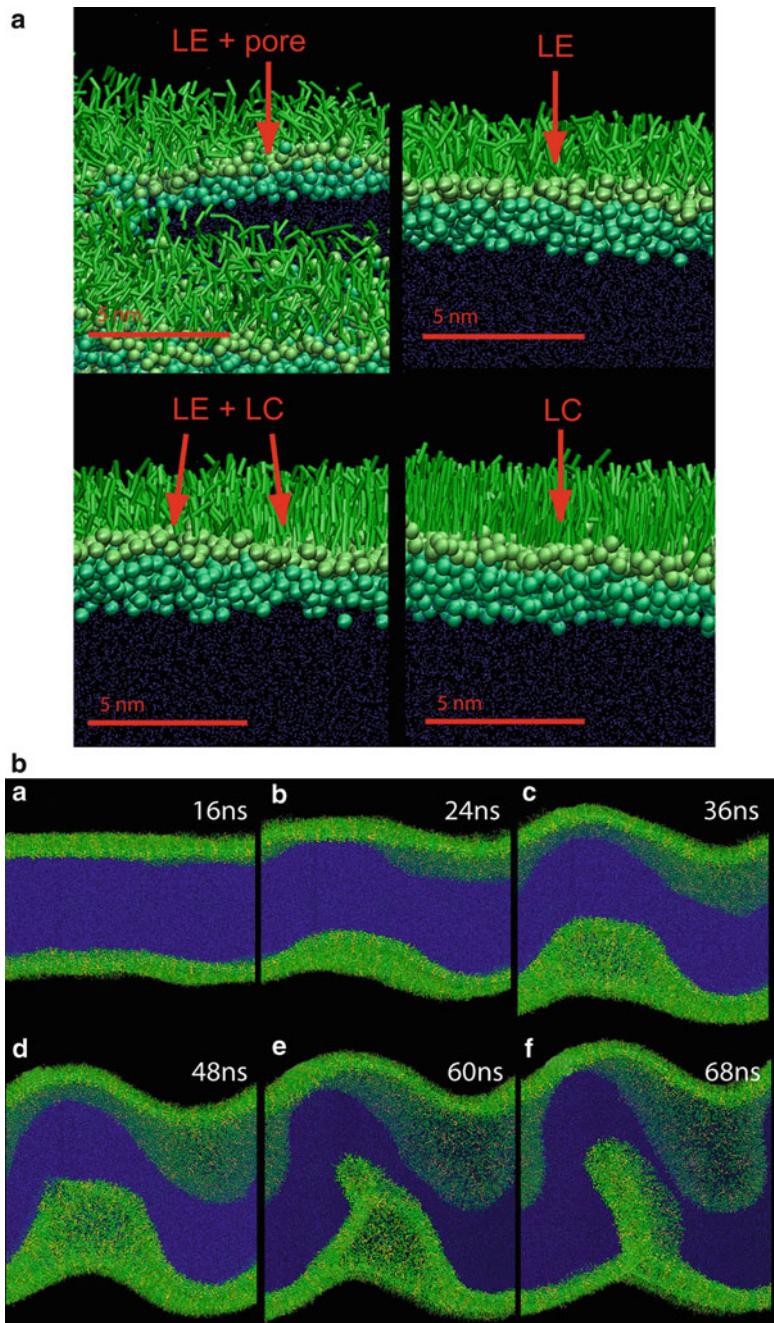


Fig. 6. Gas, liquid-expanded (LE), liquid-condensed (LC) phases and their coexistence (a), and monolayer collapse (b), reproduced using coarse-grained simulations. Reprinted with permission from (38, 67). Copyright 2012.

the same model, but requires calculating additional points on the isotherm when only experimental data is available. This is because the isotherms in simulations may be shifted with respect to experimental curves, as is the case if the surface tension at the pure

interface is offset (see Fig. 5b, c, Subheading 3.1 and Note 3). Other important factors are system size and simulation time, as they limit the critical size of the nucleus of the second phase and time of its formation, which diverge close to the phase transition point. Normally in the phase coexistence region the lateral box dimension is required to be at least larger than 10 nm, and transformation times take hundreds of nanoseconds. To simulate monolayer collapse (43, 48, 49, 67) (Fig. 6b), which manifests as 2D to 3D transformations between the interface and the subphase(s), low surface tension close to zero or negative values should be applied. The actual threshold value below which the monolayer can collapse depends on the force-field and on the system size, see Note 5. Similar to transformations between the 2D phases, there is a critical size of the nucleus: monolayer bending deformation called buckling or wrinkling, which is followed by folding. The energy barriers for collapse depend on the viscoelastic properties of the monolayer and the subphases (68, 69). These considerations are not relevant if squeeze-out of individual lipids rather than collective transformations is expected.

## 5. Notes

### 1. Target surface tension

- (a) *Problem:* Applied surface tension cannot be achieved. *Possible reason:* The monolayer is unstable at this surface tension. Instability can arise from ejection of molecules from the interface below the stability limit (e.g., at negative surface tension); in the phase coexistence region if phase separation cannot occur (see Subheading 4.4); monolayer rupture, if the applied surface tension is higher than the surface tension at the pure interface.
- (b) *Problem:* The calculated surface tension in a monolayer appears to be much higher than (the calculated value) of the pure interface. *Possible reason:* Monolayer is “too expanded” representing an unphysical state. *Solution:* Applying a surface tension (with a value in the range between zero and that of the pure interface in simulations (!)) instead of controlling monolayer area, which should noticeably reduce the area.
- (c) *Problem:* The calculated surface tension in the monolayer is negative. *Possible reason:* The monolayer is “too compressed.” *Solution:* Similar to (b), try applying a positive surface tension. Another possible reason: the surface tension is not calculated properly (e.g., the contribution

of a wall potential is not fully taken into account, etc., see Subheading 4.1).

## 2. Pores

*Problem:* The surface coverage is not homogeneous, pores appear in the monolayer. *Possible reasons:* Coexistence of the LE and gas phases, or unphysical behavior at the applied surface tension higher than that of the pure interface (e.g., when the surface tension at pure interface is underestimated). *Solution:* The latter case is similar to Note 1b and can be ruled out by calculating the surface tension at the pure interface and testing if the surface tension decreases with area near the considered point of the isotherm (38).

## 3. Surface pressure

If the surface tension at the pure interface reproduces experimental values, the relation between the surface tension and the surface pressure is straightforward according to formula (1). If it is underestimated, comparison to experimental data becomes less obvious. Previous simulations attempted to use the experimental values instead of calculated (39, 41), or to introduce an effective surface tension to compensate for the offset (38). None of these approaches is optimal as the isotherms become nonuniformly distorted in this case. However, the major inconsistencies arise at high surface tensions/low pressures in the expanded monolayer.

## 4. Phase transitions

*Problem:* The monolayer phase separation/transformation into a different phase are expected to occur (e.g., based on experimental data), but are not observed. *Possible reason:* The system size is too small, smaller than the critical nucleus of the second phase, and/or the simulation time is too short, shorter than the nucleation times. *Another possible reason:* Parameters of the force-field do not reproduce phase transition at these conditions. *Possible solution:* Changing the temperature and/or surface tension.

## 5. Monolayer collapse

(a) *Problem:* Collapse of a monolayer is observed only at high negative surface tensions. *Possible reason:* The system size is too small for monolayer buckling/wrinkling to develop and/or the monolayer/subphase(s) are too rigid.

(b) *Problem:* Only protrusions/squeeze-out of individual components are observed, neither bending nor folding occur. *Possible reason:* System size is too small making collective transformations geometrically unfeasible or energetically unfavorable. *Other possible reason:* Surfactants are soluble in the subphase(s) in the considered model.

## Acknowledgments

SB is supported by postdoctoral fellowship from the Alberta Heritage Foundation for Medical Research (AHFMR). DPT is an Alberta Ingenuity Health Solutions Scientist and Alberta Ingenuity Technology Futures Strategic Chair in (Bio)Molecular Simulation. This work was supported by the Natural Sciences and Engineering Research Council (Canada).

## References

1. Tanford C (1980) *The hydrophobic effect*. Wiley, New York, NY
2. Rosen MJ (2004) *Surfactants and interfacial phenomena*, 3rd edn. Wiley, Hoboken, NJ
3. Mouritsen OG (2005) Life—as a matter of fat. The frontiers collection. Springer, Heidelberg
4. McConnell HM (1991) Structures and transitions in lipid monolayers at the air–water-interface. *Annu Rev Phys Chem* 42:171–195
5. Knobler CM, Desai RC (1992) Phase-transitions in monolayers. *Annu Rev Phys Chem* 43:207–236
6. Mohwald H (1990) Phospholipid and phospholipid-protein monolayers at the air/water interface. *Annu Rev Phys Chem* 41:441–476
7. Knobler CM (1990) Seeing phenomena in flatland—studies of monolayers by fluorescence microscopy. *Science* 249(4971):870–874
8. Weis RM (1991) Fluorescence microscopy of phospholipid monolayer phase-transitions. *Chem Phys Lipids* 57(2–3):227–239
9. Meunier J (2000) Why a Brewster angle microscope? *Colloids Surf A: Physicochem Eng Aspects* 171(1–3):33–40
10. Binnig G, Quate CF, Gerber C (1986) Atomic force microscope. *Phys Rev Lett* 56(9):930–933
11. Schief WR, Dennis SR, Frey W, Vogel V (2000) Light scattering microscopy from monolayers and nanoparticles at the air/water interface. *Colloids Surf A: Physicochem Eng Aspects* 171(1–3):75–86
12. Ahlstrom P, Berendsen HJC (1993) A molecular-dynamics study of lecithin monolayers. *J Phys Chem* 97(51):13691–13702
13. Baoukina S, Marrink SJ, Tieleman DP (2009) Structure and dynamics of lipid monolayers: theory and applications. *Biomembr Front: Nanostruct Models Des Life* 2:75–99
14. McKinnon SJ, Whittenburg SL, Brooks B (1992) Nonequilibrium molecular-dynamics simulation of oxygen diffusion through hexadecane mono-layers with varying concentrations of cholesterol. *J Phys Chem* 96(25):10497–10506
15. Zhang YH, Feller SE, Brooks BR, Pastor RW (1995) Computer-simulation of liquid/liquid interfaces. I. Theory and application to octane/water. *J Chem Phys* 103(23):10252–10266
16. Feller SE, Zhang YH, Pastor RW (1995) Computer-simulation of liquid/liquid interfaces. 2. Surface-tension area dependence of a bilayer and monolayer. *J Chem Phys* 103(23):10267–10276
17. Kaznessis YN, Kim ST, Larson RG (2002) Simulations of zwitterionic and anionic phospholipid monolayers. *Biophys J* 82(4):1731–1742
18. Pohorille A, Benjamin I (1993) Structure and energetics of model amphiphilic molecules at the water liquid vapor interface—a molecular-dynamics study. *J Phys Chem* 97(11):2664–2670
19. Kaganer VM, Mohwald H, Dutta P (1999) Structure and phase transitions in Langmuir monolayers. *Rev Mod Phys* 71(3):779–819
20. Albrecht O, Gruler H, Sackmann E (1978) Polymorphism of phospholipid monolayers. *J Phys* 39(3):301–313
21. Ben-Shaul A (1995) *The structure and dynamics of membranes*. Elsevier, Amsterdam
22. Israelachvili JN (1985) *Intermolecular and surface forces*. Academic, London
23. Marsh D (1996) Lateral pressure in membranes. *Biochim Biophys Acta* 1286(3):183–223
24. Baoukina S, Marrink SJ, Tieleman DP (2010) Lateral pressure profiles in lipid monolayers. *Faraday Discuss* 144:393–409
25. Aveyard R, Haydon DA (1973) *An introduction to the principles of surface chemistry*. Cambridge University, New York, NY
26. Hermans J, Berendsen HJC, Vangunsteren WF, Postma JPM (1984) A consistent

- empirical potential for water–protein interactions. *Biopolymers* 23(8):1513–1518
27. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79(2):926–935
  28. Yesylevskyy SO, Schafer LV, Sengupta D, Marrink SJ (2010) Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput Biol* 6(6)
  29. Abascal JLF, Vega C (2005) A general purpose model for the condensed phases of water: TIP4P/2005. *J Chem Phys* 123(23)
  30. Wu Z, Cui QA, Yethiraj A (2010) A new coarse-grained model for water: the importance of electrostatic interactions. *J Phys Chem B* 114(32):10524–10529
  31. Shinoda W, DeVane R, Klein ML (2008) Coarse-grained molecular modeling of non-ionic surfactant self-assembly. *Soft Matter* 4(12):2454–2462
  32. Mohammad-Aghaie D, Mace E, Sennoga CA, Seddon JM, Bresme F (2010) Molecular dynamics simulations of liquid condensed to liquid expanded transitions in DPPC monolayers. *J Phys Chem B* 114(3):1325–1335
  33. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102(18):3586–3616
  34. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26(16):1668–1688
  35. van Gunsteren WF (1987) GROMOS. University of Groningen, Groningen
  36. Shushkov PG, Tsvetanov SA, Ivanova AN, Tadjer AV (2008) Dielectric properties tangential to the interface in model insoluble monolayers: theoretical assessment. *Langmuir* 24(9):4615–4624
  37. Veldhuizen R, Nag K, Orgeig S, Possmayer F (1998) The role of lipids in pulmonary surfactant. *Biochim Biophys Acta: Mol Basis Dis* 1408(2–3):90–108
  38. Baoukina S, Monticelli L, Marrink SJ, Tielemans DP (2007) Pressure-area isotherm of a lipid monolayer from molecular dynamics simulations. *Langmuir* 23(25):12617–12623
  39. Duncan SL, Larson RG (2008) Comparing experimental and simulated pressure-area isotherms for DPPC. *Biophys J.* doi: [doi:10.1529/biophysj.107.114215](https://doi.org/10.1529/biophysj.107.114215)
  40. Knecht V, Marrink SJ (2007) Molecular dynamics simulations of lipid vesicle fusion in atomic detail. *Biophys J* 92(12):4254–4261
  41. Adhangale PS, Gaver DP (2006) Equation of state for a coarse-grained DPPC monolayer at the air/water interface. *Mol Phys* 104(19):3011–3019
  42. Kandasamy SK, Larson RG (2005) Molecular dynamics study of the lung surfactant peptide SP-B1-25 with DPPC monolayers: insights into interactions and peptide position and orientation. *Biophys J* 88(3):1577–1592
  43. Shinoda W, DeVane R, Klein ML (2010) Zwitterionic lipid assemblies: molecular dynamics studies of monolayers, bilayers, and vesicles using a new coarse grain force field. *J Phys Chem B* 114(20):6836–6849
  44. Dominguez H, Smolyrev AM, Berkowitz ML (1999) Computer simulations of phosphatidylcholine monolayers at air/water and CCl<sub>4</sub>/water interfaces. *J Phys Chem B* 103(44):9582–9588
  45. Schneemilch M, Quirke N (2010) Molecular dynamics of nanoparticle translocation at lipid interfaces. *Mol Simul* 36(11):831–835
  46. Choe S, Chang R, Jeon J, Violi A (2008) Molecular dynamics simulation study of a pulmonary surfactant film interacting with a carbonaceous nanoparticle. *Biophys J* 95(9):4102–4114
  47. Kaznessis YN, Kim S, Larson RG (2002) Specific mode of interaction between components of model pulmonary surfactants using computer simulations. *J Mol Biol* 322(3):569–582
  48. Baoukina S, Monticelli L, Risselada HJ, Marrink SJ, Tielemans DP (2008) The molecular mechanism of lipid monolayer collapse. *Proc Natl Acad Sci U S A* 105(31):10803–10808
  49. Duncan SL, Larson RG (2010) Folding of lipid monolayers containing lung surfactant proteins SP-B1-25 and SP-C studied via coarse-grained molecular dynamics simulations. *Biochim Biophys Acta Biomembr* 1798(9):1632–1650
  50. Javanainen M, Monticelli L, de la Serna JB, Vattulainen I (2010) Free volume theory applied to lateral diffusion in Langmuir monolayers: atomistic simulations for a protein-free model of lung surfactant. *Langmuir* 26(19):15436–15444
  51. Rose D, Rendell J, Lee D, Nag K, Booth V (2008) Molecular dynamics simulations of lung surfactant lipid monolayers. *Biophys Chem* 138(3):67–77

52. Kulovesi P, Telenius J, Koivuniemi A, Brezesinski G, Rantamaki A, Viitala T, Puukilainen E, Ritala M, Wiedmer SK, Vattulainen I, Holopainen JM (2010) Molecular organization of the tear fluid lipid layer. *Biophys J* 99(8):2559–2567
53. Laing C, Baoukina S, Tieleman DP (2009) Molecular dynamics study of the effect of cholesterol on the properties of lipid monolayers at low surface tensions. *Phys Chem Chem Phys* 11:1916–1922
54. Berger O, Edholm O, Jahnig F (1997) Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* 72(5):2002–2013
55. Jorgensen WL, Tiradoreives J (1988) The Opls potential functions for proteins—energy minimizations for crystals of cyclic-peptides and crambin. *J Am Chem Soc* 110(6):1657–1666
56. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 111(27):7812–7824
57. Lopez CF, Nielsen SO, Moore PB, Shelley JC, Klein ML (2002) Self-assembly of a phospholipid Langmuir monolayer using coarse-grained molecular dynamics simulations. *J Phys Condens Matter* 14(40):9431–9444
58. Chanda J, Bandyopadhyay S (2006) Molecular dynamics study of surfactant monolayers adsorbed at the oil/water and air/water interfaces. *J Phys Chem B* 110(46):23482–23488
59. Gupta A, Chauhan A, Kopelevichc DI (2008) Molecular modeling of surfactant covered oil–water interfaces: dynamics, microstructure, and barrier for mass transport. *J Chem Phys* 128(23)
60. Nielsen SO, Lopez CF, Moore PB, Shelley JC, Klein ML (2003) Molecular dynamics investigations of lipid Langmuir monolayers using a coarse-grain model. *J Phys Chem B* 107(50):13911–13917
61. Knecht V, Muller M, Bonn M, Marrink SJ, Mark AE (2005) Simulation studies of pore and domain formation in a phospholipid monolayer. *J Chem Phys* 122(2):024704
62. Xing CY, Faller R (2009) Coarse-grained simulations of supported and unsupported lipid monolayers. *Soft Matter* 5(22):4526–4530
63. Nielsen SO, Srinivas G, Klein ML (2005) Incorporating a hydrophobic solid into a coarse grain liquid framework: graphite in an aqueous amphiphilic environment. *J Chem Phys* 123(12)
64. Yeh IC, Berkowitz ML (1999) Ewald summation for systems with slab geometry. *J Chem Phys* 111(7):3155–3162
65. Lague P, Pastor RW, Brooks BR (2004) Pressure-based long-range correction for Lennard-Jones interactions in molecular dynamics simulations: application to alkanes and interfaces. *J Phys Chem B* 108(1):363–368
66. den Otter WK, Shkulipa SA (2007) Intermolayer friction and surface shear viscosity of lipid bilayer membranes. *Biophys J* 93(2):423–433
67. Baoukina S, Monticelli L, Amrein M, Tieleman DP (2007) The molecular mechanism of monolayer-bilayer transformations of lung surfactant from molecular dynamics simulations. *Biophys J* 93(11):3775–3782
68. Lu WX, Knobler CM, Bruinsma RF, Twardos M, Dennin M (2002) Folding Langmuir monolayers. *Phys Rev Lett* 89(14):146107
69. Pocivavsek L, Dellsy R, Kern A, Johnson S, Lin BH, Lee KYC, Cerda E (2008) Stress and fold localization in thin elastic membranes. *Science* 320(5878):912–916
70. Lorenz CD, Travesset A (2006) Atomistic simulations of Langmuir monolayer collapse. *Langmuir* 22(24):10016–10024

# **Chapter 17**

## **Simulating DNA by Molecular Dynamics: Aims, Methods, and Validation**

**Nicolas Foloppe, Marc Guérout, and Brigitte Hartmann**

### **Abstract**

The structure and dynamics of the B-DNA double helix involves subtle sequence-dependent effects which are decisive for its function, but difficult to characterize. These structural and dynamic effects can be addressed by simulations of DNA sequences in explicit solvent. Here, we present and discuss the state-of-art of B-DNA molecular dynamics simulations with the major force fields in use today. We explain why a critical analysis of the MD trajectories is required to assess their reliability, and estimate the value and limitations of these models. Overall, simulations of DNA bear great promise towards deciphering the structural and physical subtleties of this biopolymer, where much remains to be understood.

**Key words:** DNA structure and dynamics, Force fields, Sequence-dependent DNA flexibility, Validation of simulations, Molecular dynamics

---

### **1. Introduction**

In physiological conditions, the DNA conformation remains globally close to the familiar so-called B-form. Yet, this apparently uniform helicoidal scaffold is very deceptive, and makes it particularly difficult to understand the many sequence-dependent subtle effects which are fundamental to the DNA function. Thus, elucidating the DNA behavior and function remains a challenging task in structural biology because it involves subtle and dynamic structural variations along the sequence. Understanding the relationship between DNA sequence and its structural properties remains a central but largely unresolved issue, frustrating for example attempts to decode the genetic messages in the numerous genomic sequences now available. Indeed, specific and nonspecific DNA interactions with ligands (protein, RNA, drugs, ions) are underpinned by the intrinsic structural and energetic properties of DNA.

These intrinsic properties influence the preferred DNA deformations in molecular assemblies, an effect termed indirect readout. The intrinsic sequence-dependent structure and dynamics of DNA should in principle be revealed when observing DNA free in solution.

Molecular dynamics (MD) simulations in explicit solvent have the potential to characterize DNA structure, dynamics, and energetics in atomic details. MD simulates molecular systems computationally along time, aiming to represent the motions and preferred states of a system in atomic details, using methods rooted in physical-chemistry. The last decade has witnessed impressive progress in this area, with the ability to simulate DNA in explicit solvent for hundreds of nanoseconds (1). One can simulate the DNA alone or complexed to other molecules such as proteins. For free regular double-helical DNA, any arbitrary sequence can be simulated, without the requirement for an input experimental structure (a major difference with proteins, making DNA simulations particularly attractive). From this, much has already been learned, especially regarding the interconversions between the DNA overall structural forms (2–7). However, the information derived from such simulations is only as reliable as the underlying force fields.

A force field provides the potential energy model underpinning molecular simulations. Such models have a direct bearing on the preferred conformations observed during a simulation (the lower energy states are more frequently populated). Force fields are devised to reproduce many experimental and physical properties, however such models are derived by the simultaneous fit of many empirical parameters. Thus, force fields are constructs of various qualities and may be at different stages of test and refinement. Major efforts have been devoted to the development of DNA force fields, in particular with the CHARMM27 potential (8) and the Amber family of force fields, Parm98 (9), Parm99 (10), and Parmbs0 (11). This has allowed to obtain stable simulations of double-stranded DNA in solution, to simulate the DNA preference for its A or B forms (12–14), to start addressing the role of counterions around DNA (15, 16), and to explore DNA flexibility in solution (17, 18). Comparison between simulations and solution NMR observables (in particular with respect to sugar puckles and glycosidic  $\chi$  angles) show that simulations provide a fair representation of features typical of the B-DNA form (19–21). These remarkable successes have reinvigorated the general interest in DNA conformations in solution, even fuelling expectations that they might now be obtained solely from first principles (22, 23).

Nevertheless, additional studies have unveiled limitations regarding the representation by the latest generation of force fields of DNA (24, 25) and its constituents (20) in solution. One study (24) found that order parameters for C–H vectors in unrestrained MDs agreed poorly with  $^{13}\text{C}$  NMR measurements.

There was a closer agreement between NMR and simulations when MDs were performed under NMR-derived restraints, suggesting that the unrestrained MDs suffered from inadequacies in the potentials (assuming an adequate degree of sampling). Similar conclusions were reached by comparing X-ray diffraction fingerprints of DNA in solution to their counterpart calculated from MD trajectories (25).

A stringent test (17) of three modern DNA force fields, Parm98 (9), Parmbsc0 (11), and CHARMM27 (8), was conducted by comparing simulated and NMR data on a 14 bp oligomer of special interest. Extensive unrestrained MD simulations were carried out with established protocols. The experimental reference data consisted in a large set of NMR data extracted from particularly well resolved spectra (17, 26, 27), providing information on the dynamical behavior of the backbone, sugars and bases, and their interdependence. The unrestrained MDs spontaneously retrieved the overall structural features that are common to any B-DNA sequence. So, the three force fields correctly represented the NMR observables representative of an overall B-form (intranucleotide distances, for instance) and the general mechanical coupling between the sugar conformations, the backbone states, and the spatial disposition of the bases within the double helix. The generality of these mechanical couplings is nowadays well documented (26, 28–35). However, there was poor agreement between NMR and unrestrained simulations regarding the sequence-dependent features of this DNA sequence.

Predicting the intrinsic properties of DNA sequences solely from MD simulations would require to account for the modulation of DNA structure and dynamics by its sequence, otherwise firmly established on large sets of X-ray DNA structures (28, 32) or NMR data (36). At this level of detail, in absence of experimental restraints, DNA simulations remain weak and unreliable. A visible shortcoming is the failure of unrestrained MDs to reproduce the sequence-dependent backbone populations (17, 37, 38). This is far from anecdotal because the sequence-dependent backbone motions are mechanically coupled to the overall DNA structure. Indeed, a recent study showed that MD simulations and experimental structures differed significantly regarding the magnitude and direction of global DNA curvature (39). Therefore, the force fields still need improvements before the intrinsic properties of DNA sequences can be predicted reliably based on simulations alone. CHARMM36 (106) was recently developed to improve backbone motions, but additional, systematic studies are still required to further validate this new potential.

Yet, MD simulations are now firmly part of the tool box to explore how DNA behaves in solution and to help interpret biophysical or biochemical experiments. For instance, NMR approaches are greatly strengthened when their output is analyzed

in the context of MD simulations. Indeed, deficiencies in the force fields may be compensated to a large extent by experimentally derived restraints (17, 24, 40). With improved hardware and software, running MD simulations is becoming straightforward. Nevertheless, care should be exercised when setting the simulation protocols, and the output needs to be interpreted rather than read literally. Thus, the aim of this chapter is to provide practical guidelines for MD simulations of DNA oligomers, concerning the MD protocols themselves and factors to take into account when assessing the validity of the trajectories.

We chose to emphasize regular DNA double helices, i.e., without special features such as mismatches or bulges. Other recent reviews dealt with molecular dynamics of RNA (41) and nucleic acid–protein complexes (42).

---

## 2. Materials

Several MD packages are available for biomolecular systems. The most popular packages for nucleic acids are Amber (review: (43)), CHARMM (review: (44)) and NAMD (45). NAMD uses the force fields developed either for Amber or CHARMM. Regardless of the selected package, a similar workflow of successive steps is applied to (1) prepare the system in its starting configuration, (2) equilibrate solvent and solute, (3) run a production simulation trajectory, and (4) analyze the output trajectory. A typical workflow used in plain MD is shown in Fig. 1. Several examples of Amber 10 input files are provided in Subheading 5. The Web sites <http://ambermd.org/> and <http://www.charmm.org/> give access to forums where there is helpful information. In particular, Forums/CHARMM Community/Script Archive in <http://www.charmm.org> contains useful examples of input files.

---

## 3. Methods

### 3.1. Building the System

#### 3.1.1. DNA Starting Structure

No DNA experimental structure is required for launching an MD of double-stranded DNA. This makes DNA simulations attractive, since any arbitrary DNA sequence is amenable to de novo simulation. An initial “canonical” DNA double-helical model can be used, and may be built using molecular modeling software or Web servers such NAB (<http://structure.usc.edu/make-na>) (46) or NUCGEN (<http://nucleix.mbu.iisc.ernet.in/nucgen/index.htm>) (47). These software can generate A-, B-, or Z-DNA helical forms. Building deformed DNA is also possible, for instance with NAB, which enables to integrate particular helicoidal parameter values.

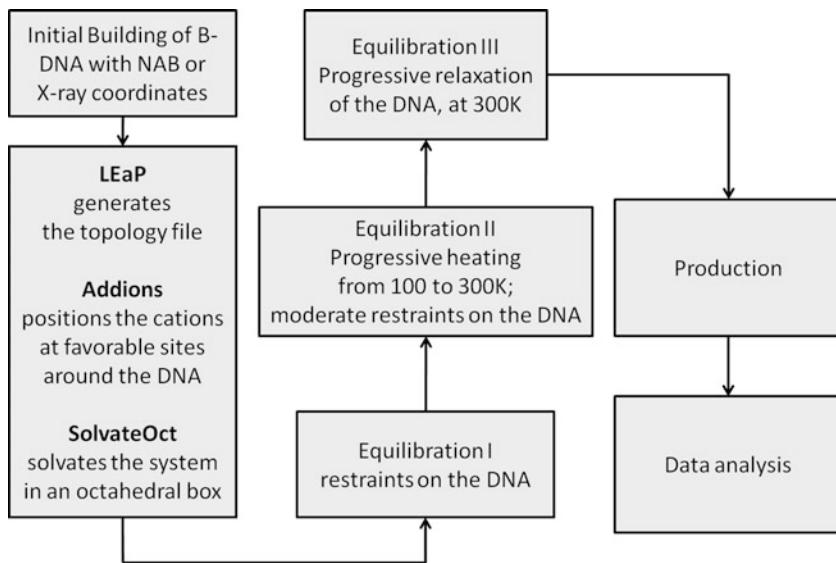


Fig. 1. Overall workflow for MD simulations of DNA. The workflow is illustrated with a protocol typically followed with Amber 10. The same main phases would be performed with CHARMM. In the building phase (*left*), several Amber LEaP modules (LEaP, Addions, SolvateOct) are used successively. LEaP also adds the hydrogen atoms, if necessary. The equilibration (*middle*) and production (*right*) phases are performed with the Sander module. The analysis frequently involves software specialized for the description of DNA structure, such as Curves (82), Curves+ (83) or 3DNA (84).

For DNAs containing special features such as mismatch, bulge, or abasic site, one may take advantage of relevant X-ray or NMR structures to initiate a simulation. These experimental structures are not free of biases. External forces (crystal packing, DNA–DNA intermolecular interactions, nonphysiological solvents or ionic strength, insufficient or inadequate restraints...) can impart structural artifacts influencing the subsequent MD trajectories. Thus, experimental-derived structures should be approached with caution, and running multiple MDs starting from different plausible conformations may be wise. Also, it provides useful tests for the convergence of the simulated properties.

Hydrogen atoms are missing in starting crystallographic structures, so they are added based on the topology and parameter files of the force field. These files are critical and should become familiar to the practitioner. They contain all the attributes of the atoms of the studied system (atom names, atom types, masses, force field parameters). Together with the molecular coordinates, they serve as input for the MD program.

### 3.1.2. Simulation Box Size and Shape

MD simulations use a primary central “box” which contains solute and solvent (water and ions). To avoid artifacts at the boundaries of this box, a procedure called periodic boundary conditions is usually used (see Subheading 3.2). In short, replicas of the primary box are positioned around this primary box, providing a well-solvated

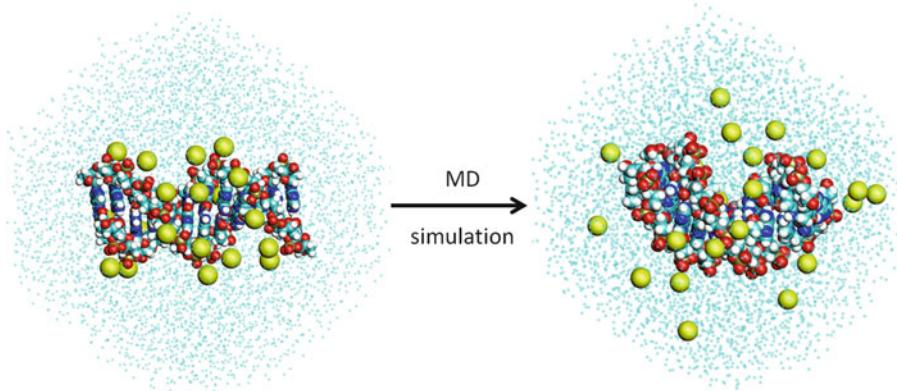


Fig. 2. Example of a 12 bp DNA oligomer in its simulation box. After the system preparation (*left*), the DNA conformation is a regular B-form double helix. The 22  $\text{Na}^+$  ions (minimum salt conditions) are close to the DNA. 4,495 water molecules (represented by *dots* for clarity) fill the truncated octahedral box. During the production phase (*right*), the DNA structure explores conformations which depart from its starting point. These new conformations should represent the intrinsic preferences of the investigated DNA sequence, which may differ significantly from a “canonical” B-form helix. Ions are free to diffuse in the whole box.

environment even if the DNA diffuses outside the primary box. An illustration of a primary box with DNA, ions and water molecules is shown on the left of Fig. 2.

For reasons of compute efficiency (amount of solvent), the primary box is usually not cubic. Instead, popular shapes include regular or truncated octahedrons, as well as dodecahedrons. These shapes minimize the amount of solvent relative to the size of the solute, hence their interest. A truncated octahedral box is frequently used, although a dodecahedron is even closer to a spherical shape (and even more economical regarding the amount of solvent).

The minimal size of the primary box depends on the length of the cutoffs truncating the nonbonded interactions. As a rule of thumb, the size of the box may extend by  $\sim 15 \text{ \AA}$  in all directions around the solute, ensuring four layers of water molecules around the DNA in the primary box. This is a reasonable compromise between a realistic DNA solvation and tractable computational times (48).

### 3.1.3. Ion Parameters and Initial Locations

Realistic simulations must be carried out in explicit solvent, water molecules but also ions. Most DNA simulations are still performed in “minimal salt” condition. This means a number of monovalent cations just sufficient for neutralizing the negative charges carried by the DNA backbone (one cation per phosphate group). Using such conditions is convenient, especially considering that some force field parameters for salts (49) lead to aggregates of  $\text{Na}^+$  or  $\text{K}^+$  with  $\text{Cl}^-$ , generating ion clusters or even  $\text{NaCl}/\text{KCl}$  crystals (50, 51), regardless the water model (52). One force field

parameterization (51) appears to completely avoid the formation of NaCl or KCl crystals, but cations tend to be anomalously trapped by the negatively charged phosphate groups (50). Other parameterizations for  $\text{Na}^+/\text{Cl}^-$  (51, 53–55) or  $\text{K}^+/\text{Cl}^-$  (53, 54) substantially decrease the propensity to form ion clusters and salt crystals (50). In particular, the Dang's parameters used for  $\text{K}^+$  with  $\text{Cl}^-$  in recent MDs of DNA result in appropriate level of KCl dissociation at physiological concentrations (38, 52). This opens the door to MD protocols going beyond the minimal salt conditions.

Ions can be positioned around the DNA at specific sites identified with a calculated electrostatic grid (56). Monovalent ions are thereby positioned at approximately 3 Å of the minima (cations) and maxima (anions) of the Coulombic potential on the DNA surface. Alternatively, the ions may be randomly distributed in the simulation box, at 5 Å or more from the DNA and 3.5 Å or more from one another (22). These two approaches did not lead to detectable differences in results in our investigations.

### 3.1.4. Water Models

The early TIP3P water model (57) is still classically associated to the Amber and CHARMM force fields. Other models are available, including SPC/E (58, 59). TIP3P and SPC/E parameters result in comparable DNA hydration patterns (50, 60), while SPC/E appears in better agreement with experiment on water diffusion rates (60). Use of the more recent TIP4PEW water model (61) did not significantly change the DNA structure compared to SPC/E (38).

## 3.2. Treatment of Electrostatics

The simulated structure and dynamics of highly charged molecules such nucleic acids depend crucially on the treatment of electrostatics (62–64). In particular, the long-range electrostatic interactions cannot be neglected without introducing severe artifacts altering the MD results. It means that electrostatic interactions need to be calculated between distant atoms, ideally between all nonbonded atom pairs. However, this would be too costly computationally; so, for every atom the electrostatic interactions are calculated only within a certain radius, or “cutoff.” Broadly speaking, there are two main approaches to treat the electrostatic interactions, spherical cutoff methods and lattice sum methods.

A widely adopted lattice sum method is the Particle Mesh Ewald (PME) method (65), a modified form of Ewald summation that efficiently calculates the infinite range Coulomb interactions under periodic boundary conditions (66). PME protocol involve (1) a real space cutoff (the distance cutoff used in Coulombic pairwise evaluations), (2) a cubic B-spline interpolation scheme, and (3) a charge grid spacing. Potentially, PME can introduce artifacts due to implicitly enforcing periodicity; however, these artifacts appear negligible when using large enough simulation boxes. Also, it should be noted

that the force fields and water models have not been developed under lattice sum conditions. Still, PME protocols perform reasonably well. In addition, PME does not imply excessive compute times. So, PME protocols have become very popular, maybe following a study (65) which suggested that PME is inherently superior to spherical cutoff methods. This view has now been revised (64), following tests (63, 67) and discussions. Thus, using PME is not a fundamental necessity.

Indeed, DNA simulations performed with the CHARMM force-shift atom-based spherical cutoff yield results similar to those obtained with PME (17, 63). The main shortcoming with spherical cutoff methods is the neglect of interactions beyond the cutoff. In addition, there are various kinds of spherical cutoffs, several of which are known to be inadequate. Published problems with spherical cutoff methods for biomolecules in aqueous solution can be attributed to inappropriately used spherical cutoffs, e.g., (1) no smoothing of the interactions at the cutoff, (2) use of a so-called switching function over a short range, or (3) using truncation based on “charge-groups.” These practices lead to well-documented artifacts, and the physical reasons for this are known. In contrast, the force-shift spherical cutoff performed adequately on a variety of systems sensitive to the treatment of electrostatics, including DNA (63, 64).

### **3.3. Equilibration and Production Phases**

#### *3.3.1. Equilibration*

The first phase of an MD simulation is an equilibration process, which distributes the kinetic and potential energies in accordance with the simulation conditions (e.g., force field, pressure, temperature). In practice, it allows the system to relax incrementally from its initial, somewhat arbitrary, configuration. Such arbitrary configuration includes the orientation of the water molecules around the solute, which was forcefully immersed in the middle of the simulation box. So, the solvent is usually allowed to relax first, while the DNA is kept fixed or restrained. Typically, this period includes a phase of heating up to 300 K. Then, one lets the DNA relax as well from its starting geometry, relieving possible internal strain in the covalent geometry and unfavorable steric contacts. The DNA restraints may be progressively relaxed by successive runs at 300 K. In the initial phases of the simulation, the NVT ensemble is sometimes preferred to the NPT ensemble ( $N$ : constant number of particles;  $V$ : constant volume;  $T$ : constant temperature;  $P$ : constant pressure). However, the NPT ensemble is usually preferred in the later production phase of the simulation, since it corresponds to most experimental conditions and allows a good control of the density of the system (68).

A numerical integration step of 2 fs for the equation of motion tends to be the norm, enabled by algorithms like SHAKE (69) which eliminate the fastest vibrational motions, by constraining

the bond lengths involving hydrogen atoms. The nonbonded pair-list is updated heuristically and the center-of-mass motion is removed at regular intervals (e.g., 10 ps). This latter adjustment overcomes the “flying ice cube syndrome,” an artifact draining high-frequency modes into low-frequency modes and resulting in excessive overall translational and rotational motions of the solute (70).

### 3.3.2. Production

The equilibration is followed by a production phase, which covers the section of trajectory during which one can in principle calculate physically sound properties. There should be continuity between the equilibration and the production phases. The production phase is continued with the same conditions used in the last period of equilibration (e.g., NPT, 300 K, 1 bar, SHAKE, regular updates of interacting atoms and center-of-mass motion removal). There is no hard and fast rule prescribing the adequate simulation duration, since the amount of sampling required depends on the properties of interest and on the system. However, for a small oligonucleotide in solution, it is typical to run the MD simulation for at least 15–20 ns (see Subheading 4.2), and one expects longer simulations to become the norm (1). Recording the system coordinates (“snapshot”) every picosecond is sufficient for a standard analysis of DNA structure and dynamics.

A snapshot extracted from an Amber trajectory is shown on the right-hand side of Fig. 2.

### 3.3.3. Thermostat

The temperature is a key parameter of a simulation, and it may be regulated in different manners, as recently reviewed (68). The Berendsen algorithm (71), implemented in Amber, may be employed. It does not strictly fix the temperature, but supplies exponential relaxation of instantaneous temperatures to a prescribed value. The temperature is adjusted by rescaling the atomic velocities in NVT and NPT simulations. The pressure is held constant by varying the simulation box volume. The Berendsen thermostat does not correspond to an exact thermodynamic ensemble and, thus, is not ideal for calculating quantities that depend on temperature variation. However, its use leads to reasonable results (see for instance (17, 22, 23, 38)), even in free energy calculations (72).

Alternative solutions probing exact thermodynamic ensembles are the Langevin (73) or Nosé-Hoover (74, 75, 105) schemes. Both methods introduce a friction term used to dynamically control the temperature. To our knowledge, data about Langevin or Nosé-Hoover thermostats used in DNA MD are very scarce (76), so evaluating their (dis)advantages with DNA is difficult. The results obtained with proteins (77, 78) may be transferable to other types of system.

## 4. Notes About the Validation of MD Trajectories

### 4.1. How to Describe the DNA Double Helix Structure

#### 4.1.1. Sugars, Phosphate Groups and Glycosidic Torsion Angles

Analyses of MDs must address the conformations of the different DNA components: sugar, phosphate groups and bases. The intrinsic DNA mechanics implies that the behaviors of sugar rings, backbone torsions and helical parameters—that describe the base positions within the double helix—are strongly interdependent (26, 28–35). So, deficiencies in the representation of one component affect the others.

We first recall how to describe the conformation of one DNA strand. Then, one has to describe the helical properties of double-stranded DNA.

The sugar ring conformations (“puckers”) are characterized by the pseudorotation phase angle, a function of the dihedral angles in the sugar ring. This pseudorotation angle allows a simple categorization of the sugar puckers, in *north* ( $300\text{--}50^\circ$ ), *east* ( $50\text{--}120^\circ$ ) and *south* ( $120\text{--}220^\circ$ ). The DNA backbone comprises two groups of torsions,  $\alpha/\beta/\gamma$  and  $\varepsilon/\zeta$  ( $\alpha$ :  $O_{3'}\text{--}P\text{--}O_5\text{--}C_{5'}$ ;  $\beta$ :  $P\text{--}O_5\text{--}C_5\text{--}C_{4'}$ ;  $\gamma$ :  $O_5\text{--}C_5\text{--}C_{4'}\text{--}C_{3'}$ ;  $\varepsilon$ :  $C_{4'}\text{--}C_3\text{--}O_3\text{--}P$ ;  $\zeta$ :  $C_3\text{--}O_3\text{--}P\text{--}O_{5'}$ ). These descriptors define the conformation of a DNA strand (Fig. 3).

The dihedral angles are reported according to the *g*– ( $240\text{--}0^\circ$ ), *g*+ ( $0\text{--}120^\circ$ ) and *trans* ( $120\text{--}240^\circ$ ) nomenclature. However,  $\varepsilon$  and  $\zeta$ , highly correlated, are generally expressed in terms of BI

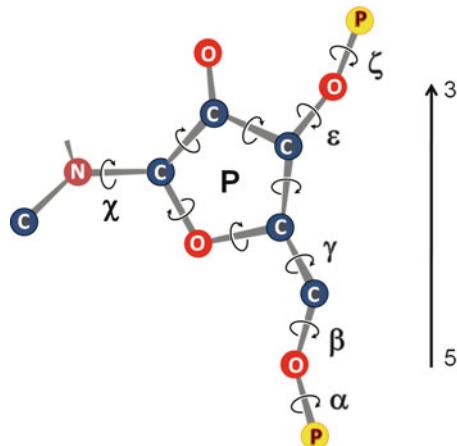
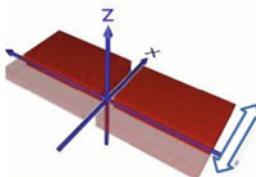


Fig. 3. Dihedral angles along a DNA chain. Dihedral angles are the key structural parameters to define the conformation of a molecule. Dihedral angles of the DNA backbone comprise  $\alpha$ :  $O_{3'}\text{--}P\text{--}O_5\text{--}C_{5'}$ ,  $\beta$ :  $P\text{--}O_5\text{--}C_5\text{--}C_{4'}$ ,  $\gamma$ :  $O_5\text{--}C_5\text{--}C_{4'}\text{--}C_{3'}$ ,  $\varepsilon$ :  $C_{4'}\text{--}C_3\text{--}O_3\text{--}P$  and  $\zeta$ :  $C_3\text{--}O_3\text{--}P\text{--}O_{5'}$ ; the glycosidic angle  $\chi$  is defined as  $O4'\text{--}C1'\text{--}N1\text{--}C2$  for purines and  $O4'\text{--}C1'\text{--}N9\text{--}C4$  for pyrimidines. Since the five torsions internal to the sugar ring are not independent, they are represented by the pseudorotation phase angle  $P$ , which allows to describe the ring conformations with one parameter.

### Base-pair parameter: X-disp



### Inter base-pair parameters:



Fig. 4. Schematic representation of three important DNA helicoidal descriptors. The translation of base-pairs towards either the minor or major grooves (X-disp, top) is related to the groove dimensions. The rotations between two successive base-pairs, twist (left, bottom) and roll (right, bottom), influence directly the double helix winding and curvature, respectively.

( $\varepsilon = \text{trans}$ ,  $\zeta = g-$ ) and BII ( $\varepsilon = g-$ ,  $\zeta = \text{trans}$ ) conformers, the two  $\varepsilon/\zeta$  combinations observed in DNA. These sub-states are ordinarily defined by the pseudo-angle ( $\varepsilon - \zeta$ ), with BI corresponding to  $(\varepsilon - \zeta) < 0^\circ$  and BII to  $(\varepsilon - \zeta) > 0^\circ$ . Sugar puckers backbone torsions should be analyzed in terms of percentages of their different conformations, not the corresponding average values.

The glycosidic torsion angle  $\chi$  (O4'-C1'-N1-C2 for pyrimidines and O4'-C1'-N9-C4 for purines) remains in *anti* in regular B-DNA (79), but this torsion should be scrutinized at mismatches, for instance.

#### 4.1.2. Helical Parameters

The spatial position of the bases with respect to the double helix axis is described by translational and rotational intra and inter base-pair parameters, following the Cambridge convention (80) and the “Tsukuba” reference frame (81). These structural parameters are referred to as helicoidal parameters. Among them, base-pair displacement (X-disp), twist and roll (Fig. 4) are particularly variable (81). Furthermore, these descriptors are related to groove dimensions (base-pair displacement), winding (twist) and curvature (roll).

The complete set of helical parameters, as well as groove dimensions and curvature directions and intensities, are usually calculated by Curves (82), Curves+ (83) or 3DNA (84).

#### 4.2. Convergence

The optimal duration of a simulation allowing a complete and converged sampling of B-DNA conformations remains unknown

and is likely to be system dependent. The adequate amount of sampling may also depend on input experimental restraints, e.g., from NMR.

Durations of 15–20 ns already provide a very helpful sampling of the sugar and backbone reconfigurations (17), but much longer simulations may ultimately be required. Due to the strong couplings between sugars, phosphate groups and the base orientations, one may hope that correctly exploring these local couplings results in a satisfactory representation of the overall DNA structure. Possible global rod-like motions could necessitate much longer trajectories (38), although an experimental evaluation of the characteristic times of such large scale motions remains difficult (85).

Below, we present routine tests that address the trajectory stability and consistency. Of course, they need to be applied with critical judgment and in the context of the scientific questions addressed.

#### *4.2.1. Simulated Structures Versus Reference Structure*

For oligomers, the RMSD (Root Mean Square Distance) calculated between each simulated snapshot and either the starting point or an ideal B-DNA should stabilize after the first nanosecond. Typical RMSDs between intact simulated oligomers of 12–15 bp and their idealized B-form counterpart are not expected to exceed 2–3 Å. However, these values can be significantly increased by curvatures of 20° or more. Opening (“fraying”) of the terminal base-pairs is not uncommon, but the simulation setup should be carefully examined (and probably corrected by hydrogen bond restraints) if this propagates to more central base-pairs.

#### *4.2.2. Structural Tests*

Additional tests rely on the structural analysis of the oligomer, excluding the first and last base-pairs to avoid end-effect issues. For production phases, the structural descriptors (Subheading 4.1) are obtained over nonoverlapping simulation windows (e.g., 5 ns or 10 ns) and compared. Obtaining similar results for different simulation blocks is a good indication of trajectory stability and convergence. This enables to refine the information given by the RMSDs and to determine which trajectory part should be emphasized in further analysis.

When they exist, one expects identical sequence fragments in different positions along an oligomer to give similar results, providing a convenient test of convergence. The relevance of such tests depends on the length of the sequence repeats. For instance, it should be kept in mind that the structure and dynamics of dinucleotides is modulated by their neighbors (36, 86).

#### ***4.3. Comparing Simulated DNA Structures to Experimental Features***

A particular DNA sequence may favor a different structure in solution compared to a crystal environment, because an oligomer X-ray structure can be affected by a number of factors, including forces exerted by the lattice (87–91). NMR-derived DNA oligomer structures should also be regarded with great caution, as they depend on refinement protocols and their force field.

Yet, when extracted from statistical analyses on large datasets of X-ray structures (28, 32), the structural trends regarding DNA and its components should be informative. Such analyses yielded results consonant with those obtained independently from NMR observables characterizing the intrinsic DNA malleability (26, 36). These general trends provide a helpful framework to approach the validity of simulated structures.

We summarize below some important trends in DNA structure and dynamics, from analyses of experimental data.

#### 4.3.1. Sugar Puckers

NMR strongly supports a dynamic equilibrium between *south* and *north* in B-DNA, with *south* being usually predominant (92–94). On average, purines and pyrimidines are associated with ~90 % and at least 70 % of *south* sugars, respectively. In simulations, the sugars attached to pyrimidines can significantly populate the *east* region, sometimes up to ~50 % (for instance with Parmbsc0 (17) and Parm99, in our experience). Unfortunately, NMR experiments do not distinguish between *south* and *east* conformers (94, 95), and the existence of *east* sugar populations is still an open question. A large amount (>30 %) of *east* pockers is not mirrored in the lower percentages observed in X-ray DNA structures (28), or their quantum-mechanical energetics (96). The sugar populations in simulations are clearly influenced by the underlying force field (20), which may deviate significantly from the intrinsic energetics calculated from first principles (20, 96). From the point of view of DNA mechanics, *north* and *east* sugars are not compatible with BII backbone conformers (2, 28, 31).

#### 4.3.2. Conformations of $\alpha/\beta/\gamma$ Backbone Angles

In free regular B-DNA,  $\alpha/\beta/\gamma$  are in *g*–/trans/*g*+ (28, 72). To our knowledge, no unusual conformations of  $\alpha/\beta/\gamma$  were observed on regular B-DNAs by NMR. However, it is notorious that the AMBER Parm94/98/99 force fields drive  $\alpha/\beta/\gamma$  in an unusual conformation (11, 23). The CHARMM27 force field is not prone to such unusual backbone conformations, maybe because each torsion was parameterized explicitly and very thoroughly (97, 98).

Artifactual backbone sub-states cause significant deformations of DNA, including in twist and roll (28, 72). Irreversible and frequent  $\alpha/\beta/\gamma$  flips are corrected with the Parmbsc0 force field (11), but long-lived unusual  $\alpha/\beta/\gamma$  conformations occasionally subsist (39). So, close attention needs to be paid to the  $\alpha/\beta/\gamma$  simulated conformations, together with an awareness of the force field. If unwanted unusual  $\alpha/\beta/\gamma$  are generated, one may filter out the undesirable snapshots before structural analysis. Indeed, there is no elegant solution to overcome this recurrent shortcoming of Amber force fields. However, the introduction of NMR distance restraints enables to restitute a realistic  $\alpha/\beta/\gamma$  behavior (17, 40).

### 4.3.3. Conformation of $\varepsilon/\zeta$ Backbone Angles

The torsions  $\varepsilon$  and  $\zeta$  are strongly correlated and define the two conformations explored by the phosphate groups in the B-form, BI ( $\varepsilon - \zeta \sim -90^\circ$ ) and BII ( $\varepsilon - \zeta \sim +90^\circ$ ). The DNA propensity to undergo these backbone transitions is highly sequence-dependent (26, 28, 29, 36) and was inferred at the dinucleotide level from a compilation of NMR  $^{31}\text{P}$  chemical shifts in solution (36). The resulting scale (Table 1) is currently the only purely experimental quantitative scale characterizing the sequence-dependent intrinsic flexibility of B-DNA in solution, for the ten complementary dinucleotide steps. This scale is essentially free from force field influences or crystal packing. Given the general importance of the BI–BII equilibrium for the overall DNA structure (next section), testing if simulated phosphate groups reproduce the BI/BII sequence effects is critical. In other words, the BI–BII equilibrium provides a structurally stringent and functionally relevant test of DNA force fields.

Unfortunately, the experimental BI/BII populations and their sequence dependence are not reproduced by current major force fields in unrestrained simulations (17, 38), although this is hopefully improved in the newly developed CHARMM36 force-field (106). Comparing to the experimental data, the simulated BII percentages of the GpA phosphates appears systematically

**Table 1**  
**Influence of the base sequence on the free B-DNA intrinsic flexibility**

Sequence	%BII
CpG•CpG	43•43
CpA•TpG	52•31
GpG•CpC	47•37
GpC•GpC	25•25
GpA•TpC	33•11
TpA•TpA	14•14
ApG•CpT	18•0
ApA•TpT	11•0
ApC•GpT	8•0
ApT•ApT	0•0

The DNA sequence is expressed in terms of the ten complementary dinucleotides base steps, summarized here from a previous study (36). The BII percentages (%BII) observed in solution by NMR are averaged over the two facing phosphates within each complementary dinucleotide. The average standard deviation of %BII is  $\pm 8$

overestimated, while those of CpA and TpG phosphates are largely underestimated, for example. Yet, it is reassuring that, under NMR restraints, CHARMM27 and Parm98 produced structures in close agreement with the experimental BII propensities (17).

#### 4.3.4. Helical Structure and Couplings

Average values and standard deviations of all the helical parameters describing the DNA double helix, extracted from a large dataset of X-ray structures, are available online (<http://rutchem.rutgers.edu/~olson/pdna.html>) (32). Although these data are not devoid of crystallization biases, they provide an estimate of the average structures and malleabilities of the ten possible dinucleotide steps, and thus provide a plausible reference to which simulated properties may be compared.

Inter base-pair parameters of roll and twist are intimately correlated (32, 99), and are associated to the three possible combination of phosphate conformations facing each other across the strands, BI•BI, BI•BII or BII•BII (26, 28, 29). BI•BI steps are associated to low twists/positive rolls, BII•BII steps to high twists/negative rolls, and BI•BII steps to intermediate values (Fig. 5 and Table 2).

In addition to these dinucleotidic effects, the base-pair displacement (X-disp) of DNA segments is sensitive to the density of BII steps (28, 36). Three or four phosphates together in BII in a tetramer segment (corresponding to a total of six phosphates) produce a shift of the bases towards the major groove, shaping a shallow major groove (depth around 3.0 Å) and a wide minor groove (width around 7.5 Å, Fig. 6) (100).

CHARMM27 reproduces particularly well the crystallographic values of twist and base-pair displacement (17). Generally, Parm98, Parmbsc0 and CHARMM27 correctly reflect the general DNA mechanics (17). This is, however, a delicate balance and, in our

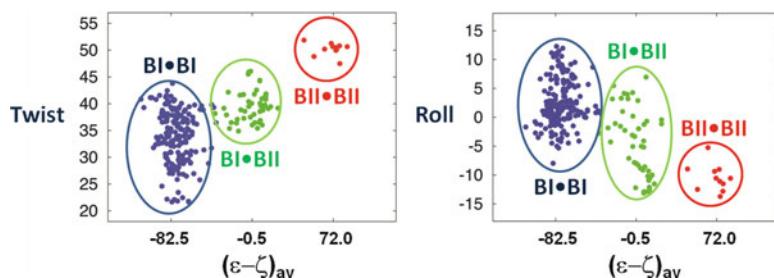


Fig. 5. Influence of the conformation of facing phosphate groups on helical parameters. The values (°) of the inter base-pair parameters twist and roll are plotted versus  $(\varepsilon - \zeta)_{av}$  values averaged on two facing phosphates. BI•BI (blue,  $(\varepsilon - \zeta)_{av} = -82.5^\circ$ ), BI•BII (green,  $(\varepsilon - \zeta)_{av} = -0.5^\circ$ ) and BII•BII (red,  $(\varepsilon - \zeta)_{av} = +72^\circ$ ) are the three possible conformational combinations of facing phosphates. The data were extracted from high resolution X-ray structures of free DNA (36).

**Table 2**  
**Influence of the backbone conformation on selected helicoidal parameters**

Backbone conformation	Twist	Roll
BI•BI	34.2 (5)	3.5 (5)
BI•BII	38.5 (5)	-0.6 (4)
BII•BII	49.4 (2)	-9.8 (5)

Representative average values of twist ( $^{\circ}$ ) and roll ( $^{\circ}$ ) for the three possible facing backbone conformations (BI•BI, BI•BII or BII•BII). The standard deviations are in brackets. The data were extracted from CpA•TpG and GpA•TpC steps in high resolution X-ray structures of free DNA (36). Twist and roll values were calculated with Curves (82)

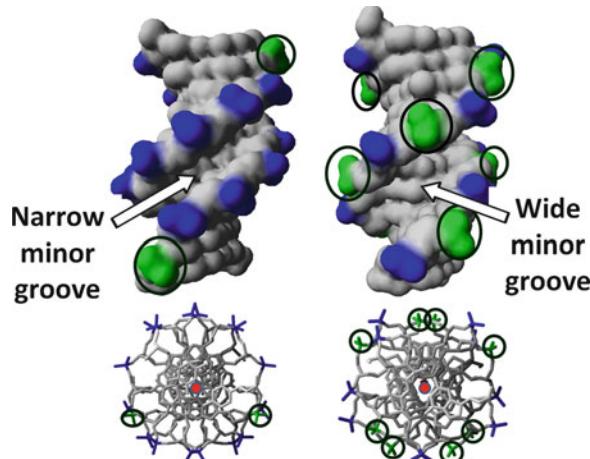


Fig. 6. Influence of the phosphate group conformations on the width of the DNA minor groove. The BI rich (left) and BII rich (right) DNA decamers are in the same orientations. BII phosphate groups are circled. The minor groove is considerably enlarged in the BII rich structure (top, right), due to a base displacement away from the helical axis (bottom, right).

experience, it can be degraded under restraints applied with excessive force-constants.

The MD analyses proposed here test whether the force fields capture the expected helical parameters in B-DNA, and thus the underlying DNA mechanics. Importantly, analyzing all these couplings also allows a fine evaluation of the conformational space explored by a given sequence. Having a detailed and dynamical view of the intrinsic properties of free DNA sequences is the prerequisite for understanding (rather than annotating) their biological functions (Fig. 7).

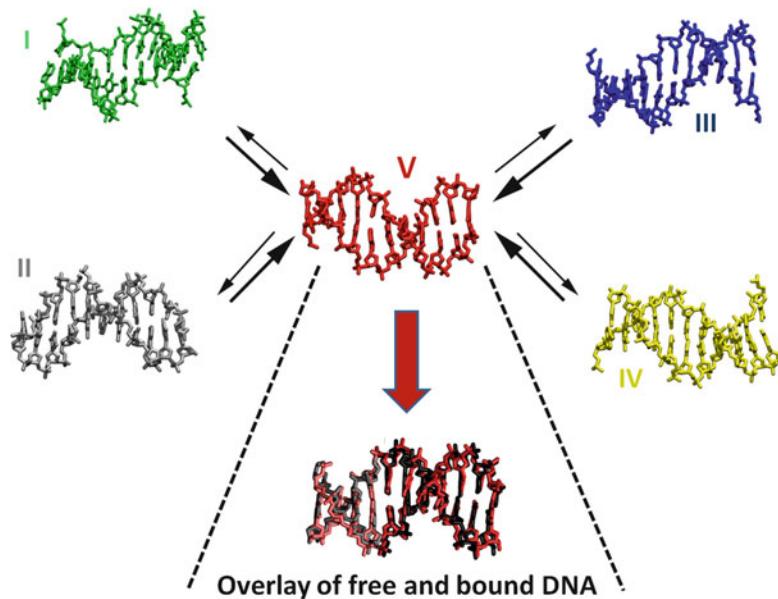


Fig. 7. DNA flexibility and recognition addressed by molecular dynamics and NMR in solution. The structure and dynamics of a DNA oligomer, which is selectively recognized by the Jun-Fos transcription factor, was characterized in solution by molecular dynamics and NMR experiments (17). The DNA oligomer was found to exist in solution as five predominant conformational families (I–V), subtly but significantly different. The most populated family V overlays closely with the X-ray DNA structure bound to Jun-Fos. This emphasizes the role of DNA pre-organization for recognition, and the importance of an accurate determination of DNA structures in solution.

#### 4.4. MD Under NMR Restraints

A common estimate of the agreement between experimental NMR measured data,  $d_{exp}$ , and their counterpart in restrained MD,  $d_{theor}$ , is to count the so-called violations between the two types of data. A violation is defined as the lack of overlap between  $d_{exp}$  and  $d_{theor}$ , taking into account the experimental errors and the MD standard deviations. This approach may be extended by calculating overall correlation coefficients between  $d_{exp}$  and  $d_{theor}$  values, to assess the overall similarity between the data profiles along the DNA sequence.

The following should be kept in mind when some NMR observables collected in routine experiments are compared to MD simulations to test the accuracy of a structural refinement under NMR restraints.

##### 4.4.1. Intra and Internucleotide Data

Classical restraints in NMR consist in intranucleotide (intrasugar and sugar-to-base) and internucleotide ( $base_i/base_{i+1}$  and  $sugar_i/base_{i+1}$ ) distances. The intranucleotide distances, are sensitive to the  $\chi$  torsions and to the sugar conformations (101), but they do not vary significantly within conformations under the B-form umbrella

(apart from disrupting features such as mismatches). Accordingly, the intranucleotide experimental distances are spontaneously respected in unrestrained MDs of regular B-DNA (17), since these simulations reproduce the gross structural features of the double helix. In contrast, the internucleotide distances reflect the subtleties of the relative positions of two successive bases (101, 102) and cover a rather wide range, with variations easily detectable by NMR. These internucleotide distances supplying data at the dinucleotide level are often complemented by residual dipolar coupling (RDC) measurements that give long-range informations (103). It is therefore much more relevant and insightful to assess the validity of the simulated structures based on the internucleotide distances and RDC, rather than on the somewhat insensitive intranucleotide distances.

In addition, the three internucleotide distances  $H2'_i-H6/8_{i+1}$ ,  $H2''_i-H6/8_{i+1}$ , and  $H6/8_i-H6/8_{i+1}$  were found (26, 27) cross-correlated and linearly coupled to  $^{31}\text{P}$  chemical shifts ( $\delta P$ ), that reflect the interconversion between the backbone BI and BII states. Importantly, these relationships provide a detailed check of the consistency of the NMR data and the possibility to extend the set of restraints for structural refinement through various extrapolations (17, 40).

#### 4.4.2. Backbone Conformations

The BI/BII ratios have been rarely considered until now in the NMR refinements, essentially due to difficulties in converting relevant NMR data into precise DNA backbone information. Recently, two approaches were developed to convert  $^{31}\text{P}$  chemical shifts into BI/BII ratios (26, 104). Thus,  $^{31}\text{P}$  chemical shifts contain important information about DNA backbone motions, of particular interest since these chemical shifts are measured very accurately. In absence of such data, the generic BII propensity scale (Table 1 and Subheading 4.3.3), may be used to relate the  $\text{BI} \leftrightarrow \text{BII}$  propensities to any sequence. Note that these strategies apply only to regular B-DNAs, devoid of unusual  $\alpha/\beta/\gamma$  torsion angles.

## 5. Examples of Typical Input Files for Amber 10

### 5.1. Construction of a B-DNA Duplex with NAB

```
molecule foo;
foo = fd_helix("abdna","tcgttagcaagct","dna");
putpdb("studied_DNA.pdb",foo,"-wwpdb");
```

The 12 bp DNA (“dna”) is in B-form (“abdna”). Its sequence is given for one 5’–3’ strand (“tcgttagcaagct”). The coordinates are written in the standard (“wwpdb”) output file “studied\_DNA.pdb.”

**5.2. Generation of the Simulation Box Containing the DNA, Ions and Water Molecules, with tLEaP or xLEaP**

```
toto = loadPDB "studied_DNA.pdb"
addions toto Na+ 0
solvateoct toto TIP3PBOX 15.0
saveamberparm toto studied_DNA_i_w.prmtop studied_
DNA_i_w.inpcrd
```

$\text{Na}^+$  cations are positioned around the DNA at specific sites identified with a calculated electrostatic grid (“addions”) in minimal salt condition (one  $\text{Na}^+$  per phosphate group, resultant net charge of 0 “0”). DNA and ions are solvated by TIP3P water molecules in a truncated octahedral box “solvateoct”, extending by  $\sim 15 \text{ \AA}$  in all directions around the solute.

The topology of the whole system (“studied\_DNA\_i\_w.prmtop”) and their initial coordinates (“studied\_DNA\_i\_w.inpcrd”) are then generated.

**5.3. Heating the System During the Equilibration Phase**

```
&cntrl
    imin=0,
    nstlim=5000,
    ntf=2, ntc=2, dt=0.002,
    nscom=5000,
    ntb=1, ntp=0, tempi=100.0, ntt=1,
    cut=9.0, nmropt=1
    ntr=1, restraint_wt=5.0, restraintmask=':1--24',
/
//Heating from 100 to 300 K
&wt type='TEMPO', istep1=0, istep2=5000, value1=
100.0, value2=300.0,
&end
&wt type='END', &end
```

The molecular dynamics (“imin=0”) is carried out with the Sander module during 10 ps (“nstlim=5000”). Using SHAKE (“ntf=2, ntc=2”) enables an integration step of 2 fs (“dt=0.002”). The nonbonded pair-list is updated heuristically by default when imin=0 or ntb>0. The center-of-mass motion is removed every 10 ps (“nscom=5000”). The NVT ensemble is used (“ntb=1, ntp=0”), and the temperature, initially of 100 K (“tempi=100.0”), is controlled with the Berendsen algorithm (“ntt=1”). The PME approach is applied by default when ntb>0, with a real space cutoff of 9 Å (“cut=9.0”).

The entire system is progressively heated (“nmropt=1”) from 100 K (“value1=100.0”) to 300 K (“value1=100.0”) in 10 ps (“istep1=0, istep2=5000”) with harmonic restraints of 5 kcal/mol/ $\text{\AA}^2$  (“restraint\_wt=5.0”) on the 12 bp DNA atoms (“restraintmask=':1--24'”).

### 5.4. Production Phase

```
&cntrl
  imin=0,
  nstlim=50000,
  ntf=2, ntc=2, dt=0.002,
  nscm=5000
  ntb=2, ntp=1, tempi=300.0, tautp=5.0, taup=5.0,
  cut=9.0,
  ntt=1,
  irest=1,
```

/

The molecular dynamics (“imin=0”) is carried out with the Sander module during 100 ps (“nstlim=50000”). Using SHAKE (“ntf=2, ntc=2”) enables an integration step of 2 fs (“dt=0.002”). The nonbonded pair-list is updated heuristically by default when imin=0 or ntb>0. The center-of-mass motion is removed every 10 ps (“nscm=5000”). The NPT ensemble is used (“ntb=2, ntp=1”), and the temperature, initially of 300 K (“tempi=300.0”), is controlled with the Berendsen algorithm (“ntt=1”) with coupling constants of 5 ps (“tautp=5.0, taup=5.0”). The PME approach is applied by default when ntb>0, with a real space cutoff of 9 Å (“cut=9.0”). The initial velocities are those of the preceding step (“irest=1”).

## References

- Perez A, Luque FJ, Orozco M (2007) Dynamics of B-DNA on the microsecond time scale. *J Am Chem Soc* 129:14739–14745
- Banavali NK, MacKerell AD Jr (2009) Characterizing structural transitions using localized free energy landscape analysis. *PLoS One* 4:e5525
- Banavali NK, Roux B (2005) Free energy landscape of A-DNA to B-DNA conversion in aqueous solution. *J Am Chem Soc* 127:6866–6876
- Cheatam TE 3rd, Kollman PA (1996) Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J Mol Biol* 259:434–444
- Knee KM, Dixit SB, Aitken CE, Ponomarev S, Beveridge DL, Mukerji I (2008) Spectroscopic and molecular dynamics evidence for a sequential mechanism for the A-to-B transition in DNA. *Biophys J* 95:257–272
- Langley DR (1998) Molecular dynamic simulations of environment and sequence dependent DNA conformations: the development of the BMS nucleic acid force field and comparison with experimental results. *J Biomol Struct Dyn* 16:487–509
- Noy A, Perez A, Laughton CA, Orozco M (2007) Theoretical study of large conformational transitions in DNA: the B ↔ A conformational change in water and ethanol/water. *Nucleic Acids Res* 35:3330–3338
- Foloppe N, MacKerell AD (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* 21:86–104
- Cheatam TE 3rd, Cieplak P, Kollman PA (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn* 16:845–862
- Wang J, Cieplak P, Kollman P (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074
- Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE 3rd, Laughton CA, Orozco M (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys J* 92:3817–3829

12. Cheatham TE 3rd (2004) Simulation and modeling of nucleic acid structure, dynamics and interactions. *Curr Opin Struct Biol* 14: 360–367
13. Mackerell AD Jr (2004) Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem* 25:1584–1604
14. Orozco M, Perez A, Noy A, Luque FJ (2003) Theoretical methods for the simulation of nucleic acids. *Chem Soc Rev* 32:350–364
15. Ponomarev SY, Thayer KM, Beveridge DL (2004) Ion motions in molecular dynamics simulations on DNA. *Proc Natl Acad Sci U S A* 101:14771–14775
16. Varnai P, Zakrzewska K (2004) DNA and its counterions: a molecular dynamics study. *Nucleic Acids Res* 32:4269–4280
17. Heddi B, Foloppe N, Oguey C, Hartmann B (2008) Importance of accurate DNA structures in solution: the Jun-Fos model. *J Mol Biol* 382:956–970
18. Perez A, Lankas F, Luque FJ, Orozco M (2008) Towards a molecular dynamics consensus view of B-DNA flexibility. *Nucleic Acids Res* 36:2379–2394
19. Arthanari H, McConnell KJ, Beger R, Young MA, Beveridge DL, Bolton PH (2003) Assessment of the molecular dynamics structure of DNA in solution based on calculated and observed NMR NOESY volumes and dihedral angles from scalar coupling constants. *Biopolymers* 68:3–15
20. Foloppe N, Nilsson L (2005) Toward a full characterization of nucleic acid components in aqueous solution: simulations of nucleosides. *J Phys Chem B* 109:9119–9131
21. Konerding DE, Cheatham TE 3rd, Kollman PA, James TL (1999) Restrained molecular dynamics of solvated duplex DNA using the particle mesh Ewald method. *J Biomol NMR* 13:119–131
22. Beveridge DL, Barreiro G, Byun KS, Case DA, Cheatham TE 3rd, Dixit SB, Giudice E, Lankas F, Lavery R, Maddocks JH, Osman R, Seibert E, Sklenar H, Stoll G, Thayer KM, Varnai P, Young MA (2004) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d (CpG) steps. *Biophys J* 87:3799–3813
23. Dixit SB, Beveridge DL, Case DA, Cheatham TE 3rd, Giudice E, Lankas F, Lavery R, Maddocks JH, Osman R, Sklenar H, Thayer KM, Varnai P (2005) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II: Sequence context effects on the dynamical structures of the 10 unique dinucleotide steps. *Biophys J* 89:3721–3740
24. Isaacs RJ, Spielmann HP (2004) Insight into G[bond]T mismatch recognition using molecular dynamics with time-averaged restraints derived from NMR spectroscopy. *J Am Chem Soc* 126:583–590
25. Zuo X, Cui G, Merz KM Jr, Zhang L, Lewis FD, Tiede DM (2006) X-ray diffraction “fingerprinting” of DNA structure in solution for quantitative evaluation of molecular dynamics simulation. *Proc Natl Acad Sci U S A* 103: 3534–3539
26. Heddi B, Foloppe N, Bouchemal N, Hantz E, Hartmann B (2006) Quantification of DNA BI/BII backbone states in solution implications for DNA overall structure and recognition. *J Am Chem Soc* 128:9170–9177
27. Heddi B, Foloppe N, Hantz E, Hartmann B (2007) The DNA structure responds differently to physiological concentrations of K(+) or Na(+). *J Mol Biol* 368:1403–1411
28. Djuranovic D, Hartmann B (2003) Conformational characteristics and correlations in crystal structures of nucleic acid oligonucleotides: evidence for sub-states. *J Biomol Struct Dyn* 20:771–788
29. Djuranovic D, Hartmann B (2004) DNA fine structure and dynamics in crystals and in solution: the impact of BI/BII backbone conformations. *Biopolymers* 73:356–368
30. Heddi B, Abi-Ghanem J, Lavigne M, Hartmann B (2010) Sequence-dependent DNA flexibility mediates DNase I cleavage. *J Mol Biol* 395:123–133
31. Isaacs RJ, Spielmann HP (2001) NMR evidence for mechanical coupling of phosphate B (I)–B(II) transitions with deoxyribose conformational exchange in DNA. *J Mol Biol* 311: 149–160
32. Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci U S A* 95:11163–11168
33. van Dam L, Levitt MH (2000) BII nucleotides in the B and C forms of natural-sequence polymeric DNA: a new model for the C form of DNA. *J Mol Biol* 304:541–561
34. Winger RH, Liedl KR, Pichler A, Hallbrucker A, Mayer E (1999) Helix morphology changes in B-DNA induced by spontaneous B(I) ↔ B(II) substrate interconversion. *J Biomol Struct Dyn* 17:223–235
35. Wu Z, Maderia M, Barchi JJ Jr, Marquez VE, Bax A (2005) Changes in DNA bending induced by restricting nucleotide ring pucker

- studied by weak alignment NMR spectroscopy. *Proc Natl Acad Sci U S A* 102:24–28
36. Heddi B, Oguey C, Lavelle C, Foloppe N, Hartmann B (2010) Intrinsic flexibility of B-DNA: the experimental TRX scale. *Nucleic Acids Res* 38:1034–1047
  37. Hart K, Nilsson L (2008) Investigation of transcription factor Ndt80 affinity differences for wild type and mutant DNA: a molecular dynamics study. *Proteins* 73:325–337
  38. Lavery R, Zakrzewska K, Beveridge D, Bishop TC, Case DA, Cheatham T 3rd, Dixit S, Jayaram B, Lankas F, Laughton C, Maddocks JH, Michon A, Osman R, Orozco M, Perez A, Singh T, Spackova N, Sponer J (2010) A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res* 38:299–313
  39. Lankas F, Spackova N, Moakher M, Enkhbayar P, Sponer J (2010) A measure of bending in nucleic acids structures applied to A-tract DNA. *Nucleic Acids Res* 38:3414–3422
  40. Abi-Ghanem J, Heddi B, Foloppe N, Hartmann B (2010) DNA structures from phosphate chemical shifts. *Nucleic Acids Res* 38:e18
  41. Hashem Y, Auffinger P (2009) A short guide for molecular dynamics simulations of RNA systems. *Methods* 47:187–197
  42. Mackerell AD Jr, Nilsson L (2008) Molecular dynamics simulations of nucleic acid–protein complexes. *Curr Opin Struct Biol* 18:194–199
  43. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
  44. Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30:1545–1614
  45. Phillips J, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
  46. Macke T, Case D (1998) Modeling unusual nucleic acid structures. ACS Publications, Washington, DC, pp 379–393
  47. Bansal M, Bhattacharyya D, Ravi B (1995) NUPARM and NUCGEN: software for analysis and generation of sequence dependent nucleic acid structures. *Bioinformatics* 11:281
  48. de Souza ON, Ornstein RL (1997) Effect of periodic box size on aqueous molecular dynamics simulation of a DNA dodecamer with particle-mesh Ewald method. *Biophys J* 72:2395–2397
  49. Aqvist J (1990) Ion–water interaction potentials derived from free energy perturbation simulations. *J Phys Chem* 94:8021–8024
  50. Noy A, Soteras I, Luque FJ, Orozco M (2009) The impact of monovalent ion force field model in nucleic acids simulations. *Phys Chem Chem Phys* 11:10596–10607
  51. Joung IS, Cheatham TE 3rd (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* 112: 9020–9041
  52. Auffinger P, Cheatham TE 3rd, Vaiana AC (2007) Spontaneous formation of KCl aggregates in biomolecular simulations: a force field issue? *J Chem Theor Comput* 3:1851–1859
  53. Beglov D, Roux B (1994) Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J Chem Phys* 100:9050–9063
  54. Dang L (1995) Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: a molecular dynamics study. *J Am Chem Soc* 117:6954–6960
  55. Jensen K, Jorgensen W (2006) Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions. *J Chem Theor Comput* 2:1499–1509
  56. Auffinger P, Westhof E (2000) Water and ion binding around RNA and DNA (C, G) oligomers. *J Mol Biol* 300:1113–1131
  57. Jorgensen WL, Chandrasekhar J, Madura JD (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
  58. Berendsen H, Grigera J, Straatsma T (1987) The missing term in effective pair potentials. *J Phys Chem* 91:6269–6271
  59. Glättli A, Daura X, Van Gunsteren W (2003) A novel approach for designing simple point charge models for liquid water with three interaction sites. *J Comput Chem* 24: 1087–1096
  60. Cheatham TE 3rd, Kollman PA (1998) Molecular dynamics simulation of nucleic acids in solution: How sensitive are the results to small perturbations in the force field and environment?" in *Structure, motion,*

- Interaction and Expression of Biological Macromolecules, Proceedings of the 10th Conversation. Eds: R.H. Sarma & M.H. Sarma, Adenine Press 99–116
61. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* 120:9665–9678
  62. Darden T, Perera L, Li L, Pedersen L (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* 7:R55–R60
  63. Norberg J, Nilsson L (2000) On the truncation of long-range electrostatic interactions in DNA. *Biophys J* 79:1537–1553
  64. Cheatham TE 3rd, Kollman PA (2000) Molecular dynamics simulation of nucleic acids. *Annu Rev Phys Chem* 51:435–471
  65. Cheatham T, Miller J, Fox T, Darden T, Kollman P (1995) Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA RNA, and proteins. *J Am Chem Soc* 117:4193–4194
  66. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys* 98: 10089–10092
  67. Beck DA, Armen RS, Daggett V (2005) Cut-off size need not strongly influence molecular dynamics results for solvated polypeptides. *Biochemistry* 44:609–616
  68. Hünenberger P (2005) Thermostat algorithms for molecular dynamics simulations. *Adv Comp Simul* 173:105–149
  69. Ryckaert J, Ciccotti G, Berendsen H (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341
  70. Harvey S, Robert K, Cheatham T III (1998) The flying ice cube: velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J Comput Chem* 19: 726–740
  71. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
  72. Varnai P, Djuranovic D, Lavery R, Hartmann B (2002) Alpha/gamma transitions in the B-DNA backbone. *Nucleic Acids Res* 30: 5398–5406
  73. Grest G, Kremer K (1986) Molecular dynamics simulation for polymers in the presence of a heat bath. *Phys Rev A* 33:3628–3631
  74. Nose J (1984) *J Chem Phys* 81:511
  75. Hoover W (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys Rev A* 31:1695–1697
  76. Luan B, Aksimentiev A (2008) Strain softening in stretched DNA. *Phys Rev Lett* 101:118101
  77. Cerutti DS, Duke R, Freddolino PL, Fan H, Lybrand TP (2008) Vulnerability in popular molecular dynamics packages concerning Langevin and Andersen dynamics. *J Chem Theor Comput* 4:1669–1680
  78. Khalili M, Liwo A, Jagielska A, Scheraga HA (2005) Molecular dynamics with the united-residue model of polypeptide chains. II. Langevin and Berendsen-bath dynamics and tests on model alpha-helical systems. *J Phys Chem B* 109:13798–13810
  79. Foloppe N, Hartmann B, Nilsson L, MacKerell AD Jr (2002) Intrinsic conformational energetics associated with the glycosyl torsion in DNA: a quantum mechanical study. *Biophys J* 82:1554–1569
  80. Dickerson RE (1989) Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res* 17:1797–1803
  81. Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakkeb Z, Sklenar H, Suzuki M, Tung CS, Westhof E, Wolberger C, Berman HM (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J Mol Biol* 313:229–237
  82. Lavery R, Sklenar H (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J Biomol Struct Dyn* 6:63–91
  83. Lavery R, Moakher M, Maddocks JH, Petkewichute D, Zakrzewska K (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res* 37:5917–5929
  84. Lu XJ, Olson WK (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* 31:5108–5121
  85. Peters JP 3rd, Maher LJ (2010) DNA curvature and flexibility in vitro and in vivo. *Quart Rev Biophys* 43:23–63
  86. Subirana JA, Faria T (1997) Influence of sequence on the conformation of the B-DNA helix. *Biophys J* 73:333–338
  87. Dickerson RE, Goodsell DS, Neidle S (1994) ...the tyranny of the lattice.... *Proc Natl Acad Sci U S A* 91:3579–3583

88. Hahn M, Heinemann U (1993) DNA helix structure and refinement algorithm: comparison of models for d(CCAGG<sub>m</sub>5CTGG) derived from NUCLSQ, TNT and X-PLOR. *Acta Crystallogr* 49:468–477
89. Heinemann U, Alings C, Hahn M (1994) Crystallographic studies of DNA helix structure. *Biophys Chem* 50:157–167
90. Lankas F (2004) DNA sequence-dependent deformability—insights from computer simulations. *Biopolymers* 73:327–339
91. Valls N, Wright G, Steiner RA, Murshudov GN, Subirana JA (2004) DNA variability in five crystal structures of d(CGCAATTGCG). *Acta Crystallogr Section D: Biol Crystallogr* 60:680–685
92. Tonelli M, James TL (1998) Insights into the dynamic nature of DNA duplex structure via analysis of nuclear Overhauser effect intensities. *Biochemistry* 37:11478–11487
93. van Wijk J, Huckriede BD, Ippel JH, Altona C (1992) Furanose sugar conformations in DNA from NMR coupling constants. *Methods Enzymol* 211:286–306
94. Wu Z, Delaglio F, Tjandra N, Zhurkin VB, Bax A (2003) Overall structure and sugar dynamics of a DNA dodecamer from homonuclear and heteronuclear dipolar couplings and <sup>31</sup>P chemical shift anisotropy. *J Biomol NMR* 26:297–315
95. Wijmenga SS, van Buuren BNM (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog Nucl Magn Reson Spectrosc* 32:287–387
96. Foloppe N, MacKerell AD Jr (1999) Intrinsic conformational properties of deoxyribonucleosides: implicated role for cytosine in the equilibrium among the A, B, and Z forms of DNA. *Biophys J* 76:3206–3218
97. Bosch D, Foloppe N, Pastor N, Pardo L, Campillo M (2001) Calibrating nucleic acids torsional energetics in force field: insights from model compounds. *J Mol Struct: Theorchem* 537:283–305
98. Foloppe N, MacKerell AD (1999) Contribution of the phosphodiester backbone and glycosyl linkage intrinsic torsional energetics to DNA structure and dynamics. *J Phys Chem B* 103:10955–10964
99. Hartmann B, Lavery R (1996) DNA structural forms. *Quart Rev Biophys* 29:309–368
100. Oguey C, Foloppe N, Hartmann B (2010) Understanding the sequence-dependence of DNA groove dimensions: implications for DNA interactions. *PLoS One* 5:e15931
101. Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York, NY
102. Lefebvre A, Fermandjian S, Hartmann B (1997) Sensitivity of NMR internucleotide distances to B-DNA conformation: underlying mechanics. *Nucleic Acids Res* 25:3855–3862
103. Mauffret O, Tevanian G, Fermandjian S (2002) Residual dipolar coupling constants and structure determination of large DNA duplexes. *J Biomol NMR* 24:317–328
104. Tian Y, Kayatta M, Shultis K, Gonzalez A, Mueller LJ, Hatcher ME (2009) (<sup>31</sup>)P NMR investigation of backbone dynamics in DNA binding sites (dagger). *J Phys Chem B* 113:2596–2603
105. Nose J (1984) *Mol Phys* 52:255
106. Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, Mackerell AD Jr. (2012) Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J. Chem. Theory Comput.* 8:348–362.

# **Chapter 18**

## **Simulation of Carbohydrates, from Molecular Docking to Dynamics in Water**

**Nicolas Sapay, Alessandra Nuriuso, and Anne Imberty**

### **Abstract**

Modeling of carbohydrates is particularly challenging because of the variety of structures resulting for the high number of monosaccharides and possible linkages and also because of their intrinsic flexibility. The development of carbohydrate parameters for molecular modeling is still an active field. Nowadays, main carbohydrates force fields are GLYCAMS06, CHARMM36, and GROMOS 45A4. GLYCAMS06 includes the largest choice of compounds and is compatible with the AMBER force fields and associated. Furthermore, AMBER includes tools for the implementation of new parameters.

When looking at protein–carbohydrate interaction, the choice of the starting structure is of importance. Such complex can be sometimes obtained from the Protein Data Bank—although the stereochemistry of sugars may require some corrections. When no experimental data is available, molecular docking simulation is generally used to obtain protein–carbohydrate complex coordinates. As molecular docking parameters are not specifically dedicated to carbohydrates, inaccuracies should be expected, especially for the docking of polysaccharides. This issue can be addressed at least partially by combining molecular docking with molecular dynamics simulation in water.

**Key words:** Molecular docking, Molecular dynamics, RESP charge derivation, GLYCAMS, AMBER, Autodock

### **List of Abbreviations**

CSD	Cambridge structural database
MD	Molecular dynamics
PDB	Protein data bank
RESP	Restricted electrostatic potential

---

### **1. Introduction**

Oligo- and polysaccharides have fundamental role in many biological processes. They are often exposed at the surface of cells, as glycoconjugates, and are involved in the communication

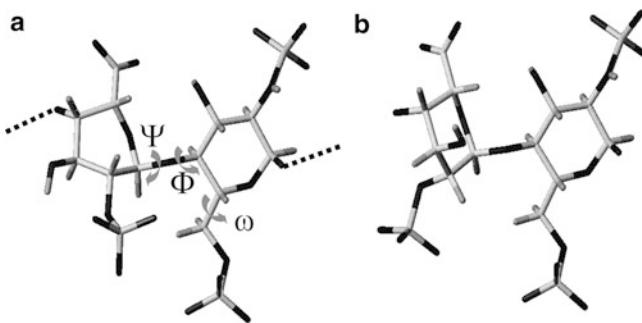


Fig. 1. Graphical representation of the disaccharide repeating unit of heparin (2-OS)- $\alpha$ -L-IdoA-1-4-(2-NS,6-OS)GlcNAc with indication of the torsion angles of interest. (a) IdoA ring in the  $^1C_4$  conformation. The dotted lines represent the extension of heparin sulfate chain. (b) IdoA ring in the  $^2S_0$  conformation.

with a large number of protein receptors. Specific interactions between protein and glycans are keys event in development and cancer. Glycans are also the targets of bacterial and viral receptors, as exemplified by the interaction between influenza virus hemagglutinins and human airway epithelia glycans. Establishing the shape of glycans, or more often the various shapes they can adopt, can therefore have application for drug development (1). However, the task is challenging.

The simulation of sugar is difficult because of the large amount of different monosaccharides, the variety of glycosidic linkages they can establish and the intrinsic flexibility of this class of compounds. The flexibility occurs at two levels: the flexibility of the glycosidic linkage between sugars and the flexibility of the furanose/pyranose rings. Furanose rings are naturally highly flexible while pyranose rings tend to have a more energetically favorable conformation (2) with some exceptions. The relative orientations of saccharide units are generally expressed in terms of the glycosidic linkage torsional angles  $\Phi$  and  $\Psi$  which are defined as  $\Phi = O5-C1-O-C'x$  and  $\Psi = C1-O-C'-C'(x - 1)$  for a  $(1-x)$  linkage (Fig. 1). Oligosaccharides are considered as flexible molecules, since they contain several glycosidic bonds about which there is a significant degree of conformational flexibility. These features make the conformational analysis of oligosaccharides a particular challenging exercise, both from the theoretical and the experimental point of view (3, 4). In addition of their high flexibility, carbohydrates are also considered difficult to model because of the occurrence of particular stereo-electronic effects, such as the anomeric and exoanomeric effects that are associated with configurational and conformational changes (5). Further complications arise for some families of sugars, such as glycosaminoglycans, because of their polyelectrolyte character where the influence of the large amount of acidic groups, in the form of carboxylate or sulfate, has to be taken into account during the modeling process.

The large variety of sugar molecules comes from the fact that each carbon on a sugar is asymmetric, each orientation producing a different compound. Additionally, each hydroxyl function can be involved in a glycosidic linkage. Finally, sugars are often modified by the cell enzymes (N-acetylation, sulfatation, etc.). As a consequence, it is very difficult, if not impossible, to obtain a comprehensive force field. To date, GLYCAMS06 is probably the most employed for sugar simulation. It is compatible with the AMBER protein force field, as well as the Generalized Amber force field. CHARMM and GROMOS also include parameters for sugars, although they are more limited.

---

## 2. Materials

### 2.1. Hardware

Partial charge derivation, molecular docking simulation or data analysis and visualization can be performed on any recent Linux workstation (see Note 1) equipped with a graphic card capable of 3D acceleration. A multicore workstation is certainly a plus. Molecular dynamics simulation should be performed on a cluster dedicated to parallel computation.

### 2.2. Choice of the Force Field

The initial choice of the force field requires great care as it will condition the simulation results, as well as the choice of the simulation software. To date, the main all atom force fields for carbohydrates are GLYCAMS06 (6), CHARMM version 36 (7–10) and GROMOS 45A4 (11). GLYCAMS06 offers the largest choice of carbohydrates—including iduronic acid, neuraminic acid and acetylated amino sugars—for a total of about 30 different compounds. In general, each compound is available with parameters for the  $\alpha$  and  $\beta$  anomeric configurations, the D and L enantiomeric forms and for all possible glycosidic linkage positions. CHARMM36 has a more limited choice of compounds. It only includes the main pyranosidic and furanosidic carbohydrates with parameters for the  $\alpha$  and  $\beta$  anomeric configurations, only the D enantiomeric form and for all possible glycosidic linkage positions. CHARMM36 also includes parameters for linear sugars and polyalcohols such as inositol. Both force fields include parameters for the methylation of the reducing end and for the O- and N-linkages with amino acids. Finally, GROMOS 45A4 only includes for unbranched hexopyranoses (11). It should be noted that GROMOS 45A4 has been optimized to enhance the stability of the  ${}^4C_1$  chair conformation, which might not be suited to carbohydrates like iduronic acid.

With the choice of the carbohydrates parameters comes the choice of the water model and the protein force field. GLYCAMS06 has been developed to be combined with the AMBER protein force fields (12–14) and the other AMBER-derived force fields,

including GAFF (15). CHARMM36 has been developed to be combined with the CHARMM22/CMAP protein force field, CMAP being a refinement specific for the protein backbone (16, 17), and the other CHARMM parameters. Both AMBER and CHARMM force fields have been developed with the TIP3P water model (18), although CHARMM use its own version of TIP3P with Lennard-Jones parameters for the hydrogen (19). The GROMOS 45A4 parameters are compatible with the other GROMOS force fields and the SPC water model (11).

As a first approach, we would recommend the combination GLYCAMS06 + AMBER99SB (14) + TIP3P. Firstly, GLYCAMS06 includes a very large choice of carbohydrates, which means the compounds of interest have a better chance to be ready for simulation. Secondly, you will benefit from the AmberTools software ([ambermd.org](http://ambermd.org)), which includes many facilities to build or modify carbohydrates from a PDB file or even from the sequence only. Finally, if your carbohydrate is available neither in GLYCAMS, nor in CHARMM, it will be easier to add new parameters to GLYCAMS06, as only the derivation of the partial charges will be needed, most of the time.

### **2.3. Choice of the Molecular Docking Software**

To date, the most popular molecular docking software are Glide ([www.schrodinger.com](http://www.schrodinger.com)), Dock ([dock.compbio.ucsf.edu](http://dock.compbio.ucsf.edu)), and Autodock ([autodock.scripps.edu](http://autodock.scripps.edu)). This latter one is probably the most often used for the docking of carbohydrates (20). Autodock includes several docking methods (a Lamarckian genetic algorithm, a simulated annealing algorithm and a Solis-Wets local search algorithm). Autodock also has the possibility to let flexible some side chains of the protein and to apply constraints on dihedrals, translation, or rotation. Finally, the Autodock Tools ([autodock.scripps.edu/resources/adt](http://autodock.scripps.edu/resources/adt)), released in parallel to Autodock, greatly facilitate the calculation setup and the analysis of the results.

### **2.4. Choice of the Molecular Dynamics Software**

All MD packages are released with a defined set of force fields because the file formats for the force field parameters, system topology and coordinates are specific to each package, as well as the implementation of the force field equations. Adding a new force field to a MD package is possible, but difficult (21, 22). To date, the combination GLYCAMS06 + AMBER99SB + TIP3P can be simulated with the AMBER (23) or the NAMD ([www.ks.uiuc.edu/Research/namd](http://www.ks.uiuc.edu/Research/namd)) (24) packages.

### **2.5. Obtaining Spatial Coordinates of Carbohydrate or Carbohydrate–Protein Complex**

The major databases containing experimental structures of carbohydrates are the Cambridge Structural Database (CSD; [www.ccdc.cam.ac.uk](http://www.ccdc.cam.ac.uk)) and the Protein Data Bank (PDB; [www.pdb.org](http://www.pdb.org)). There are only few academic databases specialized into carbohydrate structures and modeling, the main ones being Glyco3D (<http://glyco3d.cermav.cnrs.fr/glyco3d/>) and Glycosciences

([www.glycosciences.de](http://www.glycosciences.de)). In 2008, the PDB was containing about 3,500 carbohydrate structures (25). Most of them are N- or O-linked to a protein, or in complex with an enzyme, a lectin or an antibody. Isolated structures of carbohydrates remain scarce. The quality of those experimental sugar structures is still to be improved. The databank was primarily designed for protein structures. That is why the validation procedure is not well adapted to carbohydrates and a careful inspection of the coordinates is necessary.

Few tools exist to better assess the quality of sugar structures in PDB files. GlyTorsion ([www.glycosciences.de/tools/glytorsion](http://www.glycosciences.de/tools/glytorsion)) collects statistics on the torsion angles of the glycosidic rings and linkages. Carbohydrate Ramachandran Plot ([www.glycosciences.de/tools/carp](http://www.glycosciences.de/tools/carp)) calculates the conformational map of glycosidic linkages and compares it to the theoretical maps available in the GlycoMapsDB ([www.glycosciences.de/modeling/glycomapsdb](http://www.glycosciences.de/modeling/glycomapsdb)) (26).

Finally, few molecular modeling tools are available on line. Many models from experimental data as well as glycosidic linkages conformational maps are available on the Glyco3D Web site. Other sources of tridimensional models are the SWEET II ([www.glycosciences.de/modeling/sweet2](http://www.glycosciences.de/modeling/sweet2)) and the GLYCAM Web sites ([glycam.ccrc.uga.edu](http://glycam.ccrc.uga.edu)).

## **2.6. Obtaining Coordinates of Carbohydrate–Protein Complex**

The Protein Data Bank (PDB) is the main source of experimental molecular structures for proteins or carbohydrate–protein complexes. A subset of structures containing all lectins crystal structures with additional information is available in the lectin database (<http://www.cermav.cnrs.fr/lectines/>). Your protein might be available in a free state or bound to another sugar. If this sugar is very similar to yours, you can rebuild it by substituting the appropriate exocyclic group. If the protein is available with a completely different ligand or without one, molecular docking calculation will be needed first, in order to predict the coordinates of the complex. Note that sugars present in PDB are not necessarily well refined from the X-ray diffraction data, as the databank is mainly dedicated to proteins. It is strongly recommended to check the quality of the experimental data for the carbohydrate part.

## **2.7. Analyzing and Visualizing Data**

Many software products can be used to visualize and analyze structures from molecular docking and MD trajectories. Docking results obtained by using Autodock, for example, can be visualized and analyzed with the Autodock Tools. MD trajectories can be visualized with VMD (27) which also includes a large amount of analysis and file format conversion tools; it also possesses a TCL and Python interpreters for batch analysis. Additionally, all MD packages are delivered with at least some basic trajectory analysis tools (see Note 2).

### 3. Methods

In the present section, we describe the standard protocols used for preparing carbohydrate structure for docking calculations and MD simulations, using Autodock combined with the Amber molecular dynamics package. In particular, we describe in details how to obtain parameters compatible with GLYCAM force field and for this, the heparan sulfate is taken as an example (Fig. 1).

#### **3.1. Deriving Partial Charges with the RESP Method (12, 28)**

If the carbohydrate you want to model is not available in the GLYCAM06 force field, you will need to derive new parameters compatible with GLYCAM. This task is highly difficult and requires at least some basic knowledge in quantum chemistry. The simple method presented here is focused on deriving new charges for a carbohydrate compound, as other parameters have a good chance to be already present in GLYCAM06, AMBER99 or GAFF. In any case, the present method should not be considered as a black box able to generate any sugar parameters. This is only a quick and easy first approach. Eventually, you will have to check that the obtained parameters reliably reproduce the behavior of your carbohydrate. In our example, heparan sulfate is composed of monosaccharide units,  $\alpha$ -L-2-O-sulfate iduronic acid and  $\alpha$ -D-2-N-sulfate-6-O-sulfate-glucosamine bound by (1-4) linkages, each carrying a total charge of  $-2\text{ e}$  (Fig. 1).

1. Search for the atom types and bond parameters suited to your carbohydrate in the Glycam\_06x.dat file (x being the last version of the GLYCAM06 force field). If parameters are missing, search for them in gaff.dat file located in the \$AMBER-HOME/dat/LEaP/parm directory. In the later case, copy the needed parameters in a text file with the FRCMOD extension (see the AMBER documentation for a thorough description of the FRCMOD file format). For heparan sulfate, all bonded parameters are already defined in GLYCAM06.
2. Retrieve the atomic coordinates of your carbohydrate from the PDB or other databases. If they correspond to a crystal structure, do not forget to add the hydrogen atoms (see Note 3). For heparan sulfate, a NMR structure is available (PDB entry 1HPN (29)).
3. Split the carbohydrates into building blocks (see Note 4) as reported in the literature (30). The case of heparan sulfate is shown in Fig. 2. The main building blocks should correspond to the monosaccharide residues that must carry a null or integer charge. Other building blocks should correspond to the monosaccharide substituents. For example, heparan sulfate can be split into O-methyl glucosamine ( $0\text{ e}$ ), O-methyl iduronate ( $-1\text{ e}$ ), methyl sulfate ( $-1\text{ e}$ ) and methyl sulfamate ( $-1\text{ e}$ ).

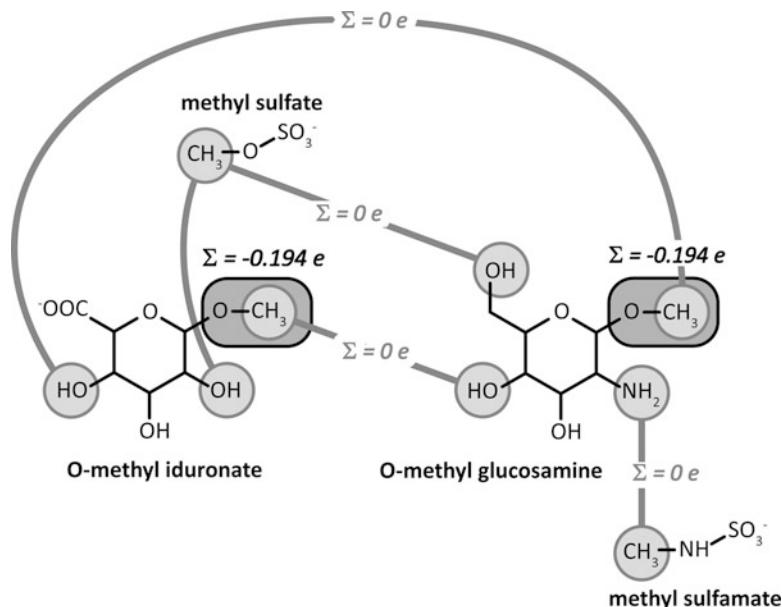


Fig. 2. Example of protocol for the derivation of RESP charges applied to the case of heparan sulfate. The polysaccharide is divided into four building blocks (*O*-methyl iduronate, *O*-methyl glucosamine, methyl sulfate, and methyl sulfamate). Intra- and intermolecular constraints applied during the RESP calculation are displayed in *black* and *gray*, respectively. The constraint of  $-0.194$  applied on the *O*-methyl groups correspond to the one used in the development of GLYCAM06. The intermolecular constraints of 0 enable to merge eventually the building blocks.

4. Model each block compound according to its conformational minimum of energy. This information can be obtained from theoretical, NMR or X-ray studies of the compound. If a compound can adopt several stable configurations, each one must be modeled. This is typically the case for glucosamine for which the group at O6, driven by  $\omega$  torsion angle, can be in a *gauche-gauche* or *gauche-trans* conformation. This is also the case for the  $\alpha$ -L-iduronate, which is a flexible ring that can adopt the  $^2S_0$ ,  $^1C_4$  or  $^4C_1$  puckering conformations (see Note 5). Save the coordinates of each conformation in PDB format.
5. Optimize each structure by quantum mechanics (QM) at the same level of theory than GLYCAM06, i.e., HF/6-31G\*. This can be done with QM calculation software like Gaussian ([www.gaussian.com](http://www.gaussian.com)). Save the optimized coordinates into PDB files. Then, the molecular electrostatic potential (MEP) should be calculated at the same level for each optimized structure.
6. Convert the PDB files into P2N files (see Note 6). A detailed description of the P2N file format is available on the R.E.D. server Web site (31) (<http://q4md-forcefieldtools.org/Tutorial/Tutorial-1.php#3>). A conversion tools is also released with the R.E.D. Tools (<http://q4md-forcefieldtools.org/RED/>). If a compound is associated to several conformations, they must all be merged into a single P2N file—the atom connectivity has

to be defined only for the first one. The MEP calculation outputs of each conformer must be merged in a single file as well, in the same order than in the P2N file.

7. Upload each P2N and QM files on the R.E.D. server and launch the calculation with a default constraint of 0.01, in agreement with the protocol employed for GLYCAM. The R.E.D. server actually automates the Restricted ElectroStatic Potential (RESP) calculation (12, 28). The same calculation can be performed with Amber and the Amber Tools, but may require more file manipulations to take into account multiple conformers. In the R.E.D. results, make sure that the standard error and the relative root mean square are within reasonable range. The actual threshold depends on the complexity of the molecule.
8. Now, the blocks created for the charge calculations have to be converted into GLYCAM units. A typical GLYCAM unit corresponds to a monosaccharide or to the capping of the reducing and nonreducing ends (see Notes 7 and 8). In the case of heparan sulfate, four new GLYCAM units must be defined. Two units will correspond to the central 2-sulfate iduronic acid (the *O*-methyl iduronate block merged with the methyl sulfate one) and to the 2,6-glucosaminosulfate (the *O*-methyl glucosamine block merged with the methyl sulfamate one) whereas the others two will be the nonreducing/reducing end units. In each monosaccharide, the *O*-methyl group on the carbon C1 and the hydrogen on the oxygen O4 must be removed as those atoms overlap when involved in the  $\alpha(1\rightarrow 4)$  glycosidic linkage. At this point, the GLYCAM units can be converted into the Amber PREP file format, i.e., an Amber topology file.

You have now an Amber PREP file where all the new GLYCAM units are defined with the proper partial charges. You may also need an Amber FRCMOD file where nonstandard bonded or Lennard-Jones parameters can be defined. More explanations about these files can be found in the Amber's manual or at the AMBER Web site (<http://ambermd.org/tutorials/>).

### **3.2. Docking a Sugar Compound to a Protein Structure**

Molecular docking simulation can predict the coordinates of a protein-ligand complex. Eventually, this result can be used as a starting point for further molecular modeling studies, especially MD simulations. In that latter case, it is important to keep some consistencies between the parameters employed for the docking simulation and those employed for the MD one. If possible, the same parameters should be used for the molecular docking and dynamics.

1. Extract the protein structure from a PDB file and prepare it as for a regular MD simulation. The preparation steps include the modeling of missing hydrogen atoms and protein side chains, the determination of the amino acids protonation state (especially

for histidine residues) and the removal of components not directly involved in the protein–ligand interaction, like crystallographic water molecules. If the ligand binding involves cofactors like calcium ions, they must be preserved within the protein structure. Ultimately, the protein coordinates should be saved in MOL2 format (see Note 8). Follow the same procedure for the ligand (see Note 9). It is recommended to center both protein and ligand structures on the origin and align them with their principal axes, for example with VMD. That will facilitate the set up of the docking grid latter. If atoms have been added to the protein or the ligand structures, their position should be minimized prior to the docking. All the steps described herein can be done with LEaP, available in the Amber Tools package, in combination with the Amber package, using the GLYCAM and AMBER99SB libraries.

2. Prepare the protein input files for Autodock with the Auto-Dock Tools. In particular, the MOL2 file should be converted to the Autodock PDBQT format (see Note 10). Define the grid within which the docking will be performed. It is generally centered on the geometrical center of the binding site and has to be large enough to accommodate rotations of the whole ligand (see Note 11). Define the atom types present in the ligand. Finally, run Autogrid to generate the docking grid that will be used by Autodock for the docking simulation. Those grids are specific to the conformation and the orientation of the protein. They have to be regenerated every time the protein coordinates are modified.
3. Load the ligand MOL2 file and save the ligand coordinates and charges in PDBQT format. Torsion angles, to be conformationally explored, have to be defined at that step (see Note 12). Autodock gives optimal results with approximately ten free torsions. More free torsions will significantly hinder the calculation speed and may decrease the accuracy of the results as well. If you are working with a polysaccharide, you should consider freezing or constraining the glycosidic linkages. This could also help in maintaining the  $\Phi$  angle in a correct conformation since Autodock does not contain specific potential energy function for the exoanomeric linkage.
4. Run Autodock using the Lamarckian genetic algorithm. Default parameters are generally satisfying, although you may need to increase the population size, the energy evaluations and the number of runs for large ligands. It is recommended to run the same simulation several times to estimate the precision of the method for the system of interest. You may also consider using the simulated annealing algorithm and compare the results obtained with the genetic algorithm.

5. Analyze the results with the AutoDock Tools. Autodock, as well as other docking software, has not been specifically developed for carbohydrates. That is why the Autodock score may not accurately match the experimental free energy of binding, especially for polysaccharides. However, the docking results should converge to a same result. This can be checked with the cluster analysis operated by Autodock after the docking calculation. The cluster with the lowest scores should include the majority of the poses. It may be worth minimizing the ligand coordinates, then recalculating the Autodock score and redoing the cluster analysis.
6. Extract the coordinates of the protein–ligand complex with the best Autodock score and store them in a PDB format (see Note 13).

### **3.3. Running a Molecular Dynamics Simulation of a Protein Sugar Complex**

The following steps can be useful for setting up a molecular dynamics simulation of carbohydrate–protein complexes by using the Amber package in combination with Ambertools.

1. Choose the starting coordinates of the complex, for example the result of a docking simulation, and save them in PDB format. This choice is crucial as it will entirely condition the MD results. A discussion with an experienced modeler will guide you through this choice.
2. Ensure that each residue of the simulated system is identified by a unique residue number. In LEaP, load the AMBER99SB (for the protein) and GLYCAMS06 (for the carbohydrate) parameters and prepare your system by adding missing hydrogen atoms, defining disulfide bridge, etc. (see Note 14). Neutralize the system by adding sodium or chloride counter ions and solvate with the TIP3P water model. The extension of the layer of water around the complex should be at least equal to the cut-off distance that will be used for the nonbonded interactions during the MD simulation. Save the topology and coordinate files for the simulation (see Note 15).
3. Set all important MD parameters in scripts representing the input files for running the simulation. A cut-off distance of at least 1 nm is recommended for nonbonded interaction. The electrostatics should be treated with an elaborated method, like the PME one (32, 33), not by a straight cut-off only (see Note 15).
4. Minimize the system in four successive steps. First, apply a restraint on all heavy atoms of the carbohydrate and the protein and minimize the position of the water molecules, the counter ions and the hydrogen atoms. Second, release the restraint on the carbohydrate and optimize its coordinates. Third, release the restraint on the protein side chains and run a new

optimization cycle. Finally, release all the restraints and minimize the energy of the whole system. The number of minimization steps depends on the convergence of the potential energy.

5. Slowly heat and pressurize the system to the target temperature and pressure, generally 298.15 K and 1 bar, (equilibration phase). Restraints can be applied on the protein and the carbohydrate, as previously. However, the whole system should be relaxed with no restraints for few nanoseconds at least before the next step. The temperature, the pressure and the potential energy should be equilibrated at that point.
6. Run the production part of the trajectory. The exact simulation time depends on the relaxation time of the phenomenon you are attempting to simulate. You may also consider running several simulations at the same time to improve the sampling. If the MD simulation follows a docking one, you should consider running simulations from different starting orientations of the ligand and check the consistency of the results. Finally, for rare or slow phenomena, classical simulations will not be sufficient. Sampling methods like umbrella sampling (34) or metadynamics (2, 35) should be considered instead.
7. Visualize and analyze the trajectory with appropriate software—VMD includes visualization tools dedicated to carbohydrates (36) (Fig. 3). In particular, verify that the energies, the temperature, the pressure, and the density of the system remain at equilibrium along the production phases. You should also plot the RMSD of the protein and the ligand backbones as a function of time to check the stability of the complex. For polysaccharides, check that the  $\Phi$  torsion angle remain within limits in agreement with the exoanomeric effect. For flexible glycosidic rings, check the ring puckering as well, for example by following variations of the Cremer–Pople parameters.

If you plan to compute the simulation with other MD packages, you should carefully check the correspondence with the Amber parameters. If you want to use NAMD, for example, detailed explanations are available in the user guide, in the section dedicated to Amber ([www.ks.uiuc.edu/Research/namd/2.7b3/ug/node14.html](http://www.ks.uiuc.edu/Research/namd/2.7b3/ug/node14.html)).

---

#### 4. Notes

1. Any Linux distribution can be used. For practical reason, it is probably better to choose one that favors stability over frequent update, like Red Hat Enterprise Linux or CentOS. Note that

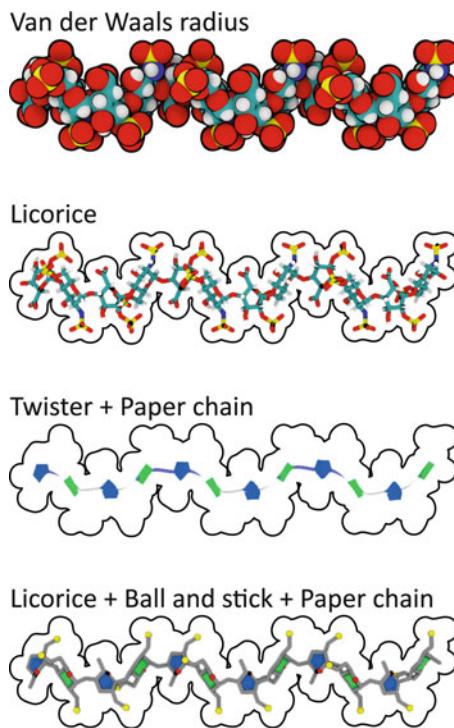


Fig. 3. Examples of carbohydrate representations available in VMD. The Twister and Paper Chain representation have been designed for cyclic and multibranched molecules, like polysaccharides. The van der Waals radius representation is outlined in *black* and drawn to scale.

recent Linux distributions tend to include GCC 4.1 as a default compiler. This version of GCC might be a source of difficulties for the compilation of docking or MD software. GCC 3 or 4.4 and newer should be preferred.

2. Trajectories generated with NAMD or AMBER can be analyzed with PTRAJ, released with the Amber Tools, or VMD. Note that with VMD, it is possible to save coordinates into many formats, which allows the use of many analysis tools. For example, it is possible to convert an NAMD trajectory file into a GROMACS one, then to use the GROMACS analysis tools—at least those which do not require a GROMACS topology file.
3. Hydrogen atoms can be added using molecular modeling software like Sybyl ([tripos.com](http://tripos.com)). Do not forget to minimize the hydrogen coordinates while keeping fixed the position of the heavy atoms.
4. Sugar polymers are characterized by having reducing and non-reducing ends. The nonreducing end is the saccharidic unit in which the anomeric carbon is linked, by a glycosidic bond to

another monomer or to a substituent (generally a methyl group), preventing opening of the ring to the aldehyde or keto form. The reducing end is the monomer in which the anomeric carbon is linked to an OH group. If the carbohydrate is a polysaccharide, the block monosaccharides must be methylated at the reducing end (i.e., oxygen O1) to take into account the fact it is involved in a glycosidic linkage.

5. Do not forget to check that the  $\Phi$  and  $\Psi$  dihedrals of the glycosidic linkage match the values present in the literature. The low energy conformations of glycosidic linkages vary with the nature of the monosaccharide involved.
6. The most important task is to define correctly the intra- and intermolecular constraints applied during the partial charge derivation. In agreement with the GLYCAM force field, the partial charge on hydrogens must be constrained to 0 e. Similarly, the charge of the O-methyl group at the reducing end of a monosaccharide must be constrained to  $-0.194\text{ e}$ . Finally, each linking group between blocks must have a charge of 0 e (Fig. 2).
7. Main nonreducing units are already defined in GLYCAM. The methyl and hydroxyl units correspond respectively to residues OME and ROH. Both units carry a charge of  $-0.194\text{ e}$ .
8. The conversion of R.E.D. blocks into GLYCAM units can be performed in three steps. First, create a simple PDB file of each GLYCAM units. Second, save each unit in MOL2 format and enter manually the RESP partial charges calculated previously with the R.E.D. server. Make sure the total charges of the monosaccharide units are integer and that the nonreducing unit has a total charge of  $+0.194\text{ e}$ . Third, convert the MOL2 file into an AMBER PREP topology file using the Amber Tools. MOL2 files can be directly generated from LEaP with the savemol2 command. It is recommended to also save the coordinates in PDB format using the savepdb command, as this can be useful latter for the analysis of the results. Note that with the MOL2 format, the atom connectivity and partial charges are explicitly described in addition to the coordinates. This is not the case with the PDB format, although the partial charges can be stored temporarily in the occupancy column, for example.
9. If you are working with a polysaccharide, it is important to remember that LEaP and GLYCAM list carbohydrate residues from the reducing to the nonreducing end (the name of GLYCAM reducing end residues stored in the GLYCAM library always begins by a 0). This is the opposite of what is generally observed in PDB structures. Each monosaccharide unit can be saved separately in PDB files. Then, the PDB files can be merged manually together in a unique file following the GLYCAM order (from the reducing to the nonreducing end).

10. If the MOL2 file is not correctly loaded by the Autodock Tools, try to convert it to the PQR format first. The conversion can be done with VMD.
11. Flexible protein residues can be defined at that step. They should be carefully chosen, however, since that will significantly slow down the calculation.
12. To date, Autodock does not allow changes of ring pucker. This is an issue for carbohydrates presenting flexible ring shapes, such as iduronic acid. In such cases, it is recommended to run several docking simulations with different ring conformations.
13. Autodock includes in its sources several useful Python scripts to manipulate Autodock results. Note that the ligand atoms extracted from the Autodock results may not be in the same order than in the input MOL2 file. This must be corrected prior to using those results in a MD simulation.
14. The protonation state of histidine residues can be checked with the WHAT IF software ([swift.cmbi.ru.nl/whatif](http://swift.cmbi.ru.nl/whatif)).
15. You can learn how to use Leap and how to prepare AMBER input files by following the tutorial available online (<http://ambermd.org/tutorials/basic/tutorial1/>), referring to the explicit water simulation section.

## References

1. Fadda E, Woods RJ (2010) Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects. *Drug Discov Today* 15:596–609
2. Spiwok V, Kralova B, Tvaroska I (2010) Modelling of beta-D-glucopyranose ring distortion in different force fields: a metadynamics study. *Carbohydr Res* 345:530–537
3. Imbert A, Pérez S (2000) Structure, conformation and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem Rev* 100:4567–4588
4. Woods RJ, Tessier MB (2010) Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Curr Opin Struct Biol* 20 (5):575–583
5. Tvaroska I, Bleha T (1989) Anomeric and exo-anomeric effect in carbohydrate chemistry. *Adv Carbohydr Chem Biochem* 47:45–123
6. Kirschner KN, Yongye AB, Tschaepel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL et al (2008) GLYCAMS06: a generalizable biomolecular force field for carbohydrates. *J Comput Chem* 29:622–655
7. Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW et al (2008) Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* 29:2543–2564
8. Hatcher E, Guvench O, MacKerell AD (2009) CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* 113:12466–12476
9. Hatcher ER, Guvench O, MacKerell AD (2009) CHARMM additive all-atom force field for acyclic polyalcohols, acyclic carbohydrates, and inositol. *J Chem Theory Comput* 5:1315–1327
10. Guvench O, Hatcher E, Venable RM, Pastor RW, MacKerell AD (2009) CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J Chem Theory Comput* 5:2353–2370
11. Lins RD, Hunenberger PH (2005) A new GROMOS force field for hexopyranose-based carbohydrates. *J Comput Chem* 26:1400–1412
12. Wang J, Cieplak P, Kollman P (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074
13. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W et al (2003) A point-charge force

- field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 24:1999–2012
14. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65:712–725
  15. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174
  16. Mackerell AD Jr, Feig M, Brooks CL 3rd (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25:1400–1415
  17. MacKerell AD Jr, Feig M, Brooks CL 3rd (2004) Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 126:698–699
  18. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
  19. Reiher WEI (1985) Theoretical studies of hydrogen bonding. Department of Chemistry, Harvard University, Cambridge
  20. Nurriso A, Kozmon S, Imbert A (2008) Comparison of docking methods for carbohydrate binding in calcium-dependent lectins and prediction of the carbohydrate binding mode to sea cucumber lectin CEL-III. *Mol Simul* 34:469–479
  21. DePaul AJ, Thompson EJ, Patel SS, Haldeman K, Sorin EJ (2010) Equilibrium conformational dynamics in an RNA tetraloop from massively parallel molecular dynamics. *Nucleic Acids Res* 38:4856–4867
  22. Sorin EJ, Pande VS (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys J* 88:2472–2493
  23. Case DA III, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM Jr et al (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
  24. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E et al (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
  25. Lutteke T (2009) Analysis and validation of carbohydrate three-dimensional structures. *Acta Crystallogr D* 65:156–168
  26. Frank M, Lutteke T, von der Lieth C-W (2007) GlycoMapsDB: a database of the accessible conformational space of glycosidic linkages. *Nucleic Acids Res* 35:287–290
  27. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:27–38
  28. Bayly CI, Cieplak P, Cornell W, Kollman PA (2002) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 97:10269–10280
  29. Mulloy B, Forster MJ, Jones C, Davies DB (1993) N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem J* 293:849–858
  30. Gouin SG, Vanquelef E, Fernandez JM, Mellet CO, Dupradeau FY, Kovensky J (2007) Multimannosides based on a carbohydrate scaffold: synthesis, force field development, molecular dynamics studies, and binding affinities for lectin Con A. *J Org Chem* 72:9032–9045
  31. Vanquelef E, Simon S, Marquant G, Garcia E, Klimerak G, Delepine JC, Cieplak P, Dupradeau FY (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res* 39: W511–517
  32. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an  $N[\text{center-dot}] \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
  33. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593
  34. Peric-Hassler L, Hansen HS, Baron R, Hunenberger PH (2010) Conformational properties of glucose-based disaccharides investigated using molecular dynamics simulations with local elevation umbrella sampling. *Carbohydr Res* 345:1781–1801
  35. Spiwok V, Tvaroska I (2009) Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides. *Carbohydr Res* 344:1575–1581
  36. Cross S, Kuttel MM, Stone JE, Gain JE (2009) Visualisation of cyclic and multi-branched molecules with VMD. *J Mol Graph Model* 28:131–139

# **Part III**

## **Mesoscopic Simulations and Coarse-Grained Models**

# Chapter 19

## Systematic Methods for Structurally Consistent Coarse-Grained Models

W.G. Noid

### Abstract

This chapter provides a primer on theories for coarse-grained (CG) modeling and, in particular, reviews several systematic methods for determining effective potentials for CG models. The chapter first reviews a statistical mechanics framework for relating atomistic and CG models. This framework naturally leads to a quantitative criterion for CG models that are “consistent” with a particular atomistic model for the same system. This consistency criterion is equivalent to minimizing the relative entropy between the two models. This criterion implies that a many-body PMF is the appropriate potential for a CG model that is consistent with a particular atomistic model. This chapter then presents a unified exposition of the theory and numerical methods for several approaches for approximating this many-body PMF. Finally, this chapter closes with a brief discussion of a few of the outstanding challenges facing the field of systematic coarse-graining.

**Key words:** Information theory, Relative entropy, Multiscale coarse-graining, Generalized Yvon-Born-Green theory, Boltzmann inversion, Variational principles

---

### 1. Introduction

As clearly demonstrated by the previous chapters, classical molecular dynamics (MD) and Monte Carlo simulations of atomically detailed models have contributed profound insight into many biological and condensed phase systems (1, 2). Classical force fields have been carefully parameterized to provide a reasonably realistic description of the atomistic structure and interactions for many molecular systems over a relatively wide range of thermodynamic conditions (3–5). With current hardware (6, 7) and software (8–12) resources, atomistic simulation methods allow exquisitely detailed investigations of phenomena that evolve on nanometer length scales and perhaps even microsecond time scales (13, 14). However, as a direct consequence of this detail, atomistic simulation methods remain limited in their capability

for efficiently and exhaustively investigating many processes of fundamental importance that evolve on significantly longer length and time scales, e.g., the folding of multidomain proteins or signal transduction events. These limitations have motivated tremendous interest in “coarse-grained” computational models that provide greater efficiency by representing systems in somewhat reduced detail.

The notion of “coarse-graining” high-resolution models to develop more tractable low-resolution models for studying emergent phenomena of interest is a recurring theme throughout science (15). Hydrodynamics (16) and thermodynamics (17) are two obvious examples of such coarse-grained theories. Similarly, classical atomistic models are frequently interpreted as a coarse-grained description of more detailed quantum mechanical models (18). Therefore, it is not surprising that pioneers such as Kirkwood made many seminal contributions to the statistical mechanics of coarse-graining long ago (19–21). Nor is it surprising that a staggering diversity of coarse-grained models have been applied to study a vast array of physical and biological phenomena. These models span a wide range from continuum and field theoretic models to particle-based models. Consequently, the present introduction to coarse-grained modeling will necessarily be of relatively limited scope.

This chapter will address “mesoscopic” particle-based coarse-grained (CG) models that retain sufficient detail to allow for chemically important distinctions. The fundamental particles in such models are interaction sites (sometimes called super-atoms) that represent groups of atoms, typically on the order of 2–10 “heavy” atoms (e.g., C, N, O, P) and associated hydrogen atoms, although this chapter certainly also applies for “bigger” sites. The interactions between these sites are then described by effective potentials that incorporate both energetic and entropic contributions resulting from averaging over (or, more formally, integrating out) uninteresting atomic details. Simulations with the resulting models can be three to four orders of magnitude more efficient than atomically detailed simulations for several reasons: Because CG models require fewer interacting sites to model the same system, fewer interactions must be calculated at each time step. Furthermore, these interactions are typically of shorter range than the interactions in atomistic models, e.g., long-ranged Coulombic interactions may be more effectively screened on the length scales relevant for CG interactions and can be subsumed into short-ranged effective potentials. These considerations are especially important for implicit solvent CG models that formally integrate out all of the solvent molecules and incorporate their effects into the interactions between the remaining sites (22). Moreover, as a consequence of averaging over atomistic configurations, the CG potential can implicitly include entropic effects and greatly reduce the need to explicitly sample a large number of degenerate atomically detailed configurations (23). Finally, by averaging

over interactions, the interactions between CG sites may be “softer,” (i.e., of lower frequency) which allows a larger integration time step for propagating dynamics (24). This effect is further enhanced because the friction and memory effects associated with fluctuations on an atomically detailed potential energy surface are dramatically reduced, allowing the CG model to rapidly sample larger regions of configuration space. However, as a side effect of this averaging, the CG effective potential may not be transferable to multiple systems or thermodynamic state points, and the relationship between the dynamics of atomistic and CG models becomes somewhat unclear.

Coarse-grained models have often been employed as intellectually appealing minimal models that are constructed to capture certain basic physical principles. These models then provide a powerful tool for assessing how generic design principles may govern phenomenon of interest, e.g., polymer configurations (25, 26), curvature-mediated aggregation of membrane-bound proteins (27), or protein folding (28). However, such minimal models may not provide a realistic description of any particular system (18). Moreover, these models may give the right answer for the wrong reason, as they only reflect the interactions that were initially specified in designing the model (29). While CG models may allow highly efficient investigations of complex phenomena, the resulting models may be misleading unless they incorporate “the right physics.” One possible way to quantify this statement would be to require that the model generates the correct distribution of CG microstates, i.e., the distribution of microstates that would be obtained if one used a “coarse-grained lens” to observe the configurations generated by an accurate high-resolution model. This chapter will consider this possibility in detail.

These considerations have motivated extensive recent effort in developing systematic approaches for determining “chemically specific” CG models that generate results that are consistent with an accurate high-resolution model for a particular system. These approaches typically focus on accurately reproducing either thermodynamic or structural properties. Klein and coworkers (30–36) and Marrink and coworkers (24, 37–41) have pioneered the development of CG models that reproduce thermodynamic properties determined from either experiment or simulation. In particular, surface tension, partitioning free energies, and density have frequently been employed to parameterize Lennard-Jones-type pair potentials between point sites. The resulting models have demonstrated remarkable transferability for various systems and a wide range of thermodynamic state points. In some cases, these models have also reproduced structural properties with semiquantitative accuracy. The recent developments by the Marrink group are reviewed in a later chapter.

“Bottom-up” strategies have also been developed for parameterizing CG potentials to accurately reproduce structural data

obtained from experiments or atomistic simulations. This chapter will focus on the theory and development of these approaches. The primary objectives of the present chapter are as follows:

1. Provide a readily accessible introduction to the statistical mechanical framework for relating atomistic and CG models of the same system.
2. Formulate and analyze quantitative consistency criterion for CG models that are “structurally consistent” with an underlying atomistic model.
3. Present a unified discussion of the computational and theoretical aspects of certain systematic methods for determining such models.
4. Outline some of the remaining outstanding challenges facing these approaches to systematic coarse-graining.

In particular, it is our hope that the present chapter will be distinguished from previous reviews by the accessibility of the introductory material and also by the unified approach to discussing a variety of systematic coarse-graining methods.

It is important to emphasize several major shortcomings of this chapter. As noted above, the present chapter is highly limited in scope and reflects the notation and personal perspective of the author. Moreover, this chapter does not attempt to address the staggering diversity in scope, application, and approach of coarse-grained models. For more information, the reader is referred to a several recent themed journal issues (42–45), and a recent book (46) devoted to the subject of CG models. In particular, many recent reviews (47–49) have addressed CG dynamics (50, 51) and models of, e.g., protein (18, 52–61), lipid (62–70), and polymer systems (71–74). Furthermore, the remaining chapters of this book provide additional insight into the great diversity and utility of CG models for a wide variety of applications.

---

## 2. Consistency Between Atomistic and Coarse-Grained Models

This chapter considers consistency between high-resolution and low-resolution mechanical models of a particular molecular system in an equilibrium canonical ensemble. The high-resolution model will be referred to as an “atomistic model,” and the corresponding interacting particles in the atomistic model will be referred to as “atoms.” The low-resolution model will be referred to as a “coarse-grained model,” and the corresponding interacting particles in the coarse-grained (CG) model will be referred to as “sites.” However, the following framework applies quite generally for parameterizing low-resolution particle-based

models from high-resolution particle-based models. It will be assumed that neither the atomistic nor the CG model includes rigid constraints. We shall also primarily address the case that the atoms and the sites correspond to structureless point particles, although it is possible to consider sites corresponding to asymmetric mass distributions (75–78) or that include internal degrees of freedom (79). The presentation and notation reflects the framework developed by Voth, Andersen, and coworkers (80).

## 2.1. Models and Notation

We begin by introducing notation and establishing the elementary statistical thermodynamics of atomistic and CG models (81). For convenience, atomistic quantities will be represented by lower case symbols, while CG quantities will be represented by capitalized symbols.

The microstate of an atomistic model with  $n$  atoms is specified by the set of  $n$  Cartesian coordinates  $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_n)$  and momenta  $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_n)$ . The Hamiltonian for the atomistic model is:

$$h(\mathbf{r}, \mathbf{p}) = \sum_i \frac{1}{2m_i} \mathbf{p}_i^2 + u(\mathbf{r}), \quad (1)$$

where  $m_i$  is the mass of atom  $i$  and  $u(\mathbf{r})$  is a classical interaction potential for the model. No assumptions will be made about the particular form of  $u$ . The canonical phase space distribution for the atomistic model factors into statistically independent configuration and momenta distributions:

$$p_r(\mathbf{r}) \propto \exp[-u(\mathbf{r})/k_B T] \quad (2)$$

$$p_p(\mathbf{p}) \propto \exp\left[-\sum_i \mathbf{p}_i^2 / 2m_i k_B T\right]. \quad (3)$$

In particular, the momenta distribution factors into independent Gaussian distributions for each  $\mathbf{p}_i$ .

The microstate of a CG model with  $N$  sites is specified by the set of  $N$  Cartesian coordinates  $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_N)$  and momenta  $\mathbf{P} = (\mathbf{P}_1, \dots, \mathbf{P}_N)$ . The Hamiltonian for the CG model is:

$$H(\mathbf{R}, \mathbf{P}) = \sum_I \frac{1}{2M_I} \mathbf{P}_I^2 + U(\mathbf{R}), \quad (4)$$

where  $M_I$  is the mass of site  $I$  and  $U(\mathbf{R})$  is a classical interaction potential for the model. The canonical phase space distribution for the CG model factors into statistically independent configuration and momenta distributions:

$$P_R(\mathbf{R}) \propto \exp[-U(\mathbf{R})/k_B T] \quad (5)$$

$$P_P(\mathbf{P}) \propto \exp\left[-\sum_I \mathbf{P}_I^2 / 2M_i k_B T\right]. \quad (6)$$

As above, the momenta distribution factors into independent Gaussian distributions for each  $\mathbf{P}_I$ .

## 2.2. Relating Atomistic and CG Models

The previous subsection introduced atomistic and CG models without establishing any formal connection between them. We relate the two models by introducing a set of mapping functions (80). We assume that each site physically represents a particular group of atoms in the atomistic model. We define a set of  $N$  linear mapping functions  $\mathbf{M}(\mathbf{r}) = (\mathbf{M}_1(\mathbf{r}), \dots, \mathbf{M}_N(\mathbf{r}))$  that define the Cartesian coordinates of each site as a linear function of atomic Cartesian coordinates with constant coefficients,  $c_{Ii}$ :

$$\mathbf{M}_I(\mathbf{r}) = \sum_i c_{Ii} \mathbf{r}_i, \quad (7)$$

for each site,  $I = 1, \dots, N$ , in the CG model. If all of the atoms are translated by a vector displacement, then the mapping should translate each site by the same vector. Consequently, it follows that the coefficients,  $c_{Ii}$ , should satisfy the following normalization:

$$\sum_i c_{Ii} = 1, \quad (8)$$

for each site. Because we treat the sites as independent point particles, it follows that the CG momenta are simply  $\mathbf{P}_I = M_I d\mathbf{R}_I/dt$ . This implies a set of linear mapping functions  $\mathbf{M}_P(\mathbf{p}) = (\mathbf{M}_{P1}(\mathbf{p}), \dots, \mathbf{M}_{PN}(\mathbf{p}))$  that define the Cartesian momenta of each site as a linear function of atomic Cartesian momenta:

$$\mathbf{M}_{PI}(\mathbf{p}) = M_I \sum_i c_{Ii} \mathbf{p}_i / m_i, \quad (9)$$

for each site.

For each site,  $I$ , we define the set of involved atoms,  $\mathcal{I}_I$ , as the set of atoms that provide a finite contribution to the sum in Eq. 7, i.e., the set for which  $c_{Ii} \neq 0$ . In principle, a single atom may be involved in multiple sites. In this case, it is also convenient to define a set of specific atoms that are involved in site  $I$  but are not involved in any other site. However, in the following, we shall assume that, for each site, all of the atoms involved in the site are also specific to the site. (This last assumption can be readily relaxed (80)).

The above assumptions are sufficiently general that the present framework applies to a wide range of particle-based CG models. In particular, the present framework is valid when sites are placed at, e.g., the center-of-mass (i.e.,  $c_{Ii} \propto m_i$ ) or the center-of-geometry (i.e.,  $c_{Ii} = 1/n_I$ , where  $n_I$  is the number of atoms involved in site  $I$ ) of a group of atoms. This framework also applies for implicit solvent CG models (23, 63), in which entire molecules have been eliminated from the model and for mixed resolution models in which some sites represent single atoms, while others represent groups of atoms (82). However, this framework does not apply for models,

e.g., the adaptive resolution models (83) discussed in a later chapter by Delle Site and Praprotnik, for which molecules dynamically change resolution during the course of a simulation (84). Nor does the following framework immediately apply for nonlinear coarse-graining maps.

### 2.3. Consistency Criteria

The mapping operators defined by Eqs. 7–9, in combination with the atomistic probability distributions defined by Eqs. 2 and 3, imply probability distributions for sampling atomistic phase space points that map to a point in the phase space for the CG model. Because the atomistic configuration and momenta are statistically independent random variables and because the mapping operators do not mix configuration and momenta variables, the implied CG phase space distribution factors into statistically independent configuration and momentum distributions. In particular, the probability of sampling an atomistic configuration  $\mathbf{r}$  that maps to a fixed CG configuration,  $\mathbf{R}$ , is given by

$$p_R(\mathbf{R}) = \langle \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) \rangle, \quad (10)$$

where the angular brackets denote canonical ensemble averages over the atomistic model and  $\delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) = \prod_I \delta(\mathbf{M}_I(\mathbf{r}) - \mathbf{R}_I)$ . Similarly,

$$p_P(\mathbf{P}) = \langle \delta(\mathbf{M}_{\mathbf{P}}(\mathbf{p}) - \mathbf{P}) \rangle \quad (11)$$

is the probability of sampling a set of atomistic momenta  $\mathbf{p}$  that maps to a fixed set of CG momenta,  $\mathbf{P}$ . Because the CG model considers fewer particles than the original atomistic model, the mapping operators map multiple atomistic configurations to the same CG configuration. The distribution,  $p_R(\mathbf{R})$  defined in Eq. 10 corresponds to the total Boltzmann weight of the volume element in the atomistic configuration space that maps to the specified point  $\mathbf{R}$  in the CG configuration space. Similarly,  $p_P(\mathbf{P})$  defined in Eq. 11 corresponds to the total Boltzmann weight of the volume element in the atomistic momentum space that maps to the specified point  $\mathbf{P}$  in the CG momentum space.

A natural, though certainly not unique, definition for consistency with an atomistic model requires that a consistent CG model reproduce the distributions of CG phase variables that are implied by the atomistic probability distributions and CG mappings (80):

$$P_R(\mathbf{R}) = p_R(\mathbf{R}), \quad (12)$$

$$P_P(\mathbf{P}) = p_P(\mathbf{P}). \quad (13)$$

These equations provide quantitative criterion for consistency between atomistic and CG models in configuration and momentum space, respectively.

The first criterion uniquely determines (to within an additive constant) the appropriate potential,  $U^0$ , for a CG model that is consistent with a particular atomistic model in configuration space:

$$U^0(\mathbf{R}) = -k_B T \ln p_R(\mathbf{R}) + \text{const} \quad (14)$$

$$= -k_B T \ln z(\mathbf{R}|u) + \text{const}, \quad (15)$$

in terms of a configuration-dependent partition function,

$$z(\mathbf{R}|u) = \int d\mathbf{r} \exp[-u(\mathbf{r})/k_B T] \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}). \quad (16)$$

The notation  $z(\mathbf{R}|u)$  in Eq. 15 implies that  $z$  is a function of the CG configuration  $\mathbf{R}$  and a functional of the atomistic potential  $u(\mathbf{r})$ , i.e., it depends upon the value of  $u(\mathbf{r})$  throughout the entire configuration space (85). For reasons that will be emphasized later,  $U^0$  is referred to as a many-body potential of mean force (PMF) (19, 86). The many-body PMF is not a conventional potential because it considers both entropic and energetic contributions to give appropriate Boltzmann weight to each CG configuration. Consequently, the many-body PMF depends upon the system and thermodynamic state point. The constants in Eqs. 14 and 15 are independent of configuration and, therefore, do not impact the configuration probability distribution. However, they are dependent upon the thermodynamic state point. This state point dependence may be difficult to determine but can contain important information that impacts the thermodynamic properties of the CG model.

The previous analysis indicates that, for any atomistic model and any CG mapping, there exists a potential,  $U^0$ , that is, sufficient to ensure that the CG model will be consistent with the underlying atomistic model in configuration space (80). However, in general, it remains impractical to evaluate the integral of Eq. 16. Consequently, much of the remaining chapter will discuss various systematic approaches for approximating  $U^0$ . The development of CG potentials that not only accurately approximate the many-body PMF at a single thermodynamic state point for a single system but are also transferable to multiple thermodynamic states and/or multiple systems remains a major challenge for the field.

While consistency in configuration space remains remarkably challenging, it is relatively easy to ensure consistency in momentum space between atomistic and CG models. A CG model will be consistent in momentum space with a particular atomistic model if and only if the following conditions are satisfied(80):

1. Each atom is involved in at most one site.
2. The mass of each site is given by

$$M_I = \left( \sum_{i \in I} \frac{c_{Ii}^2}{m_i} \right)^{-1}. \quad (17)$$

Note that this expression for the site mass depends upon both atomic masses and also mapping coefficients.

These conditions are obtained from the following considerations: Because each momentum variable  $\mathbf{P}_I$  is an independent Gaussian random variable in the CG model, in a consistent CG model the distribution  $p_P(\mathbf{P})$  in Eq. 11 must also factor into independent contributions for each  $\mathbf{P}_I$ . This is possible if and only if each atom is involved in at most one CG site. The condition upon the CG site mass then follows because (1) the mapping expresses each CG momentum variable as the sum of atomistic momenta, each of which is an independent Gaussian random variable; and because (2) the sum of independent Gaussian random variables is a Gaussian random variable with a variance given by the sum of the variances for the contributing random variables (87).

The consistency condition in momentum space has several implications. In the case that the mapping coefficients assign a site to the center-of-mass for an atomic group, i.e.,  $c_{Ii} \propto m_i$ , then Eq. 17 implies that the appropriate mass for that site is the total mass of the atoms involved in the site. Moreover, for a CG model that is consistent in momentum space, the center-of-mass mapping maximizes the mass of a site relative to any other mapping that maps the same atoms to the site. In particular, if a group of atoms with unequal mass are involved in a site, then Eq. 17 leads to a smaller site mass for a center-of-geometry mapping than for a center-of-mass mapping.

According to our earlier definition in Eq. 13, a CG model that is consistent in momentum space will reproduce the distribution of fluctuations in momentum space. However, this definition does not imply or ensure any particular relationship between the dynamical trajectory observed during a CG MD simulation and any dynamical trajectory of the underlying atomistic model. Standard Mori-Zwanzig theory (88, 89) demonstrates that any conservative and deterministic atomistic trajectory will appear to be nonconservative and nondeterministic when observed in the CG phase space. In particular, the initial conditions in the phase space for a single CG trajectory correspond to an ensemble of initial conditions in the atomistic phase space. Moreover, dynamical simulations of CG models with conservative interaction potentials necessarily neglect fluctuating atomistic interactions that perturb and retard the atomic dynamics. Consequently, dynamical information obtained from CG MD simulations that sample the canonical ensemble for a conservative potential should necessarily be carefully considered. The time scale of CG dynamics is frequently estimated empirically by comparing the observed dynamics with those that are known from either detailed simulations or experiment (90). However, the fundamental relationship between the dynamics of a CG model and the trajectories of an underlying atomistic model remains somewhat obscure. The development of CG models that accurately model the

dynamics of complex fluids and molecular systems remains a topic of active and extensive research that is beyond the scope of the present review. See, e.g., the later chapter by Lu.

## 2.4. Maximum Likelihood Approach

Recently, Shell proposed a maximum likelihood framework that provides a complementary perspective for considering consistency between atomistic and CG models (91). The maximum likelihood formalism assumes a mechanism for using the CG model to sample atomistic configurations. We shall use  $P_r(\mathbf{r})$  to denote the probability that the CG model samples a configuration  $\mathbf{R}$  that maps to an atomistic configuration  $\mathbf{r}$ . (We will consider this distribution in somewhat more detail momentarily.) The framework then considers the probability for the CG model to reproduce the distribution of configurations sampled by an atomistic model. In particular, if one discretizes the atomistic configuration space into distinguishable small volume elements of equal size and employs the CG model to perform  $n_s$  independent samples of the atomistic configuration space, then the probability that, during those  $n_s$  samples, each distinguishable configuration  $\mathbf{r}$  is sampled  $n(\mathbf{r})$  times is given by:

$$n_s! \prod_{\mathbf{r}} \frac{1}{n(\mathbf{r})!} P_r(\mathbf{r})^{n(\mathbf{r})}, \quad (18)$$

where  $\mathbf{r}$  indexes the discrete cells in the configuration space and  $W(\{n(\mathbf{r})\}) = n_s! / \prod_{\mathbf{r}} n(\mathbf{r})!$  is the multinomial factor corresponding to the number of distinguishable sequences of  $n_s = \sum_{\mathbf{r}} n(\mathbf{r})$  configurations in which each distinguishable configuration  $\mathbf{r}$  is represented  $n(\mathbf{r})$  times. If the  $n_s$  configurations were sampled from the underlying atomistic model with a configuration probability distribution,  $p_r(\mathbf{r})$ , then the expected number of samples for the configuration  $\mathbf{r}$  will be  $n(\mathbf{r}) = n_s p_r(\mathbf{r})$ . The probability that, during these  $n_s$  independent samples, the CG model samples each configuration  $\mathbf{r}$  with the same frequency as the atomistic model does (i.e., a total of  $n(\mathbf{r}) = n_s p_r(\mathbf{r})$  times in  $n_s$  samples) is given by the likelihood function:

$$L = n_s! \prod_{\mathbf{r}} \frac{1}{[n_s p_r(\mathbf{r})]!} P_r(\mathbf{r})^{[n_s p_r(\mathbf{r})]}. \quad (19)$$

The maximum likelihood approach parameterizes the CG model to maximize the probability of reproducing the distribution of atomistic configurations sampled by the atomistic model, i.e., the CG model should be parameterized to maximize the likelihood function  $L$  (91).

Considering  $n_s \rightarrow \infty$  samples, applying Stirling's approximation (92), passing to a continuous description of the configuration space, and defining a dimensioned relative entropy  $S_{rel} = -(1/n_s)k_B \ln L$  lead to

$$S_{rel} = k_B \int d\mathbf{r} p_r(\mathbf{r}) \ln \left[ \frac{p_r(\mathbf{r})}{P_r(\mathbf{r})} \right]. \quad (20)$$

The relative entropy corresponds to the Kullback–Leibler divergence from information theory (93) and quantitatively measures the overlap of the distributions of atomistic configurations generated by the atomistic and CG models, while weighting the overlap according to the atomistic distribution. Maximizing the likelihood function  $L$  corresponds to minimizing the relative entropy by definition. The Gibbs inequality (85), which will be discussed in greater detail below, implies that  $S_{rel} \geq 0$  and that  $S_{rel} = 0$  if and only if  $P_r(\mathbf{r}) = p_r(\mathbf{r})$  for all  $\mathbf{r}$ .

To proceed further in analyzing  $P_r$ , we employ the identity (87):

$$P_r(\mathbf{r}) = \int d\mathbf{R} P_{r|R}(\mathbf{r}|\mathbf{R}) P_R(\mathbf{R}), \quad (21)$$

where  $P_R(\mathbf{R})$  is the probability distribution for the CG model to sample a configuration  $\mathbf{R}$  and  $P_{r|R}(\mathbf{r}|\mathbf{R})$  is the conditional probability of obtaining the atomistic configuration  $\mathbf{r}$  given a fixed CG configuration  $\mathbf{R}$ . We assume that  $P_{r|R}(\mathbf{r}|\mathbf{R}) = 0$  unless  $\mathbf{R} = \mathbf{M}(\mathbf{r})$ . Consequently, we consider the following form for  $P_{r|R}$ :

$$P_{r|R}(\mathbf{r}|\mathbf{R}) = \frac{\mathcal{G}(\mathbf{r}) \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R})}{\Omega(\mathbf{R})}, \quad (22)$$

where  $\mathcal{G}(\mathbf{r})$  is a function of the atomistic configuration and  $\Omega(\mathbf{R})$  is a normalization that quantifies the degeneracy of atomistic configurations that map to the fixed configuration:

$$\Omega(\mathbf{R}) = \int d\mathbf{r} \mathcal{G}(\mathbf{r}) \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}). \quad (23)$$

Two natural choices for  $\mathcal{G}(\mathbf{r})$  are (1) a uniform distribution,  $\mathcal{G}(\mathbf{r}) = 1$  (91), and (2) a Boltzmann-weighted distribution,  $\mathcal{G}(\mathbf{r}) \propto \exp[-u(\mathbf{r})/k_B T]$ , which was earlier considered by Krishna et al. in proposing a reweighting scheme for developing temperature transferable potentials (94).

Employing Eqs. 21–23 in Eq. 20 and keeping track of dimensionless scaling, one obtains

$$S_{rel} = k_B \int d\mathbf{r} p_r(\mathbf{r}) \ln \left[ \frac{V^n}{V^N} \frac{p_r(\mathbf{r})}{P_R(\mathbf{M}(\mathbf{r}))} \right] + S_{map}, \quad (24)$$

where

$$S_{map} = k_B \left\langle \ln \left[ \frac{V^N}{V^n} \frac{1}{\mathcal{G}(\mathbf{r})} \Omega(\mathbf{M}(\mathbf{r})) \right] \right\rangle. \quad (25)$$

Equation 24 decomposes the relative entropy into two contributions: (1) a term which compares  $p_r$  and  $P_R$  and, therefore, explicitly depends upon the CG potential,  $U$ , according to Eq. 5; and (2) a

term,  $S_{\text{map}}$ , which is independent of  $U$  but reflects the degeneracy of the CG mapping. Further analysis of  $S_{\text{map}}$  may provide a useful guide for designing  $\mathbf{M}$ .

A straightforward calculation employing Lagrange multipliers then demonstrates that the relative entropy is minimized when

$$P_R(\mathbf{R}) = p_R(\mathbf{R}), \quad (26)$$

which is the same consistency criterion described above. Consequently, the earlier consistency criterion not only corresponds to generating the same distribution of CG configurations as the underlying atomistic model but also corresponds to minimizing the relative entropy between the two models, or equivalently, maximizing the overlap of the atomistic distributions generated by the two models. Recent work has expanded upon these considerations (95, 96).

### 3. Computational Algorithms for Consistency

The previous section demonstrated that the many-body PMF defined in Eq. 14 is the appropriate potential for a CG model that is consistent with a particular atomistic model in configuration space, i.e., a model that reproduces the distribution of CG configurations generated by an atomistic model (and an associated mapping). However, in general, it is not feasible to calculate, represent, or simulate this many-body potential. Consequently, the development of accurate approximations to this many-body PMF remains a major challenge for systematic coarse-graining. This section presents and discusses some of the most popular and successful methods for determining CG potentials as a systematic approximation to the many-body PMF. We note that Larini and Krishna have recently proposed a generalized mean field theory for unifying the various CG methods discussed below (and also others) (97).

We will employ consistent notation in an attempt to unify the presentation and to clarify the relationships between different methods. Simultaneously, we will attempt to maintain sufficient generality that the presentation will apply to a wide class of CG models. In particular, we will focus on approximate CG potentials that are of the following form:

$$U(\mathbf{R}) = \sum_{\zeta} \sum_{\lambda} U_{\zeta}(\psi_{\zeta}(\{\mathbf{R}\}_{\lambda})), \quad (27)$$

where  $\zeta$  indicates a particular interaction (e.g., a dihedral angle interaction) and  $U_{\zeta}$  is the corresponding potential (e.g., a dihedral angle potential) that is a function of a single scalar variable,  $\psi_{\zeta}$ , (e.g., a dihedral angle) that may be expressed as a function of the Cartesian coordinates,  $\{\mathbf{R}\}_{\lambda}$ , for a set of sites,  $\lambda$  (e.g., the four successively bonded sites that form a dihedral angle). This assumed form for the

potential applies, e.g., for molecular mechanics potentials that are frequently employed in CG models. The following analysis can also be reformulated in the case that the approximate potential includes terms that depend upon more than a single variable, e.g.,  $U_\zeta(\psi_{\zeta 1}(\{\mathbf{R}\}_\lambda), \psi_{\zeta 2}(\{\mathbf{R}\}_\lambda))$ , although that case will not be pursued here. Many CG methods bear striking similarities to classical density functional theories (98). To emphasize this similarity, we shall also find it convenient to represent the approximate potential as a sum of terms, each of which is a product of a potential,  $U_\zeta$ , and a conjugate field or density,  $\hat{\rho}_\zeta$

$$U(\mathbf{R}) = \sum_\zeta \int d\mathbf{x} U_\zeta(x) \hat{\rho}_\zeta(\mathbf{R}; x), \quad (28)$$

where

$$\hat{\rho}_\zeta(\mathbf{R}; x) = \sum_\lambda \delta(\psi_{\zeta\lambda}(\mathbf{R}) - x) \quad (29)$$

in terms of  $\psi_{\zeta\lambda}(\mathbf{R}) \equiv \psi_\zeta(\{\mathbf{R}\}_\lambda)$ . The quantity  $\hat{\rho}_\zeta$  defined by Eq. 29 is a density operator that can be interpreted as a counting device for determining, in a given configuration,  $\mathbf{R}$ , how many times the interaction potential  $U_\zeta(x)$  contributes to the total potential. Ensemble averages of  $\hat{\rho}_\zeta$  then determine structural correlation functions describing the interaction  $\zeta$ .

Several methods determine approximate potentials that reproduce a set of lower-order structural correlation functions, e.g., pair correlation functions, that are determined either from detailed simulations or experimental data. In particular, we shall consider atomistic (target) structural correlation functions of the form:

$$p_\zeta(x) = \langle \hat{\rho}_\zeta(\mathbf{M}(\mathbf{r}); x) \rangle = \int d\mathbf{r} p_r(\mathbf{r}) \sum_\lambda \delta(\psi_{\zeta\lambda}(\mathbf{M}(\mathbf{r})) - x), \quad (30)$$

which describes the distribution of CG structures obtained by mapping configurations sampled from an atomistic model with a probability distribution  $p_r(\mathbf{r})$ . We shall also consider corresponding CG correlation functions:

$$P_\zeta(x) = P_\zeta(x|U) = \langle \hat{\rho}_\zeta(\mathbf{R}; x) \rangle_U = \int d\mathbf{R} P_R(\mathbf{R}|U) \sum_\lambda \delta(\psi_{\zeta\lambda}(\mathbf{R}) - x), \quad (31)$$

which describes the distribution of structures sampled from a CG model with probability distribution  $P_R(\mathbf{R}|U)$ . The subscripted angular brackets in Eq. 31 indicate a canonical ensemble average evaluated for a CG model with potential  $U$ , which is further emphasized in the notation  $P_\zeta(x|U)$  and  $P_R(\mathbf{R}|U)$ . If  $\zeta$  corresponds to the nonbonded interaction between a pair of site types in distinct molecules, then  $p_\zeta(r)/r^2$  and  $P_\zeta(R)/R^2$  are proportional to the

corresponding site-site radial distribution functions (RDFs) generated from atomistic and CG simulations, respectively.

It should be emphasized that, even if a CG model reproduces a set of target lower-order structural correlation functions, i.e.,  $\{P_\zeta(x) = p_\zeta(x)\text{ for all } \zeta\}$ , it will not necessarily reproduce the higher-order structural correlations generated by the atomistic model (99). In particular, such a CG model is not guaranteed to reproduce the many-body distribution of sites generated by the atomistic model, i.e.,  $P_R(\mathbf{R}) \neq p_R(\mathbf{R})$ . Furthermore, such a model is not guaranteed to reproduce any particular thermodynamic properties of the atomistic model.

### 3.1. Direct Boltzmann Inversion

If the interactions considered in the CG model are statistically independent of one another in the atomistic model, then  $p_R$  factors into independent terms for each interaction:

$$p_R(\mathbf{R}) \propto \prod_{\zeta} \prod_{\lambda} p_{\zeta}(\psi_{\zeta\lambda}(\mathbf{R})). \quad (32)$$

In this case, each term in the CG potential can be determined from “direct Boltzmann inversion” of the corresponding distribution:

$$U_{\zeta}(x) = -k_B T \ln[p_{\zeta}(x)/J_{\zeta}(x)], \quad (33)$$

where  $J_{\zeta}$  corresponds to the Jacobian factor for the relevant order parameter, e.g.,  $J_{\zeta}(\theta) = \sin\theta$  for bond-angle interactions,  $J_{\zeta}(\phi) = 1$  for dihedral angle interactions, and  $J_{\zeta}(r) = r^2$  for central (radial) pair interactions (100). CG potentials determined from direct Boltzmann inversion have enjoyed considerable success in polymeric systems, as reviewed, e.g., in Refs. (90, 101). Direct Boltzmann inversion is often employed to parameterize bonded and other “hard” interactions that are only weakly coupled to the other interactions in the CG potential (102). In this case, the remaining interaction potentials, e.g., non-bonded pair potentials may be modeled with simple Lennard-Jones potentials to correctly reproduce excluded volume interactions or may be parameterized using iterative methods described below (103, 104).

However, for some biological and soft matter systems, the interactions in the CG potential may be highly coupled and direct Boltzmann inversion may not quantitatively reproduce the target distribution functions. The importance of the correlations between the different potential terms and the corresponding difficulty that this can create for CG modeling is very clearly indicated by considering a CG model for a molecular fluid in which each molecule is represented by a single point site and the CG potential only includes central pair potentials between each pair of sites. In this case,  $J_{\zeta}(r) \propto r^2$  so that  $p_{\zeta}(r)/J_{\zeta}(r)$  is proportional to the RDF between molecules generated by the atomistic model,  $g_{\zeta}(r)$ , and the CG potential  $U_{\zeta}(r)$  that is obtained from direct Boltzmann

inversion is equivalent (to within an additive constant) to the pair potential of mean force (ppmf),  $w_\zeta(r)$ :

$$U_\zeta(r) = w_\zeta(r) = -k_B T \ln g_\zeta(r). \quad (34)$$

Because the RDF has maxima resulting from the packing of molecules into solvation shells, the ppmf also oscillates with attractive minima at the distances corresponding to the maxima in the RDF. The force function obtained from the ppmf,  $-w'_\zeta(r)$ , is equal to the magnitude of the average force experienced by a molecule averaged over all configurations in which a second molecule is a fixed distance  $r$  away (85, 105). This force, which is referred to as the pair mean force, includes the average of the direct force between the pair, but also includes an environment-mediated force, i.e., the average correlated force from the surrounding molecules. In general, this contribution will not vanish; in fact, it is this contribution that generates the oscillations in the ppmf, which, in turn, result from the packing of the surrounding molecules into solvation shells.

If the ppmf is employed as a potential for simulating two sites in vacuum, then the environment-mediated forces included in the ppmf will ensure that the distribution of distances sampled between the two sites in vacuum equals the distribution for the pair in condensed phase. However, if the ppmf is employed as a potential for simulating the molecules in the condensed phase, then these environment-mediated interactions contributions to the ppmf will be included in the interaction potential between each pair of sites, and, consequently, the CG model will systematically overcount these interactions (106).

This problem is even more transparent in the trivial case in which both the atomistic and CG model correspond to the same simple Lennard-Jones fluid (i.e., no coarse-graining). In this case, the atomistic and CG pair potential should be the same, i.e., the original Lennard-Jones potential (106). However, the CG potential determined by direct Boltzmann inversion (i.e., the ppmf) will oscillate with maxima at the peaks that are present in the atomistic RDF due to simple packing of spheres.

It is important to recognize the distinction between the ppmf,  $w_\zeta(r)$ , which describes the average force between a pair and which is determined by a two-body correlation function, and the many-body PMF,  $U^0(\mathbf{R})$ , which describes the average force on  $N$  sites and which is determined by an  $N$ -body distribution (19, 86).

At present, there exists no direct (i.e., noniterative) method for determining CG potentials from structural correlation functions in such a way to ensure that the CG model will quantitatively reproduce a given set of correlation functions.

### 3.2. Iterative Boltzmann Inversion

The iterative Boltzmann inversion (IBI) method is a natural extension of direct Boltzmann inversion that addresses the correlations between different interactions by iteratively modifying the CG potentials to quantitatively reproduce a given set of structural distribution functions. The IBI method was pioneered for CG models of fluids and polymers by Muller-Plathe and coworkers (107). Schommers (108), Soper (109), and Reatto et al. (110) had earlier suggested somewhat similar approaches for determining empirical potentials for all-atom models of fluids. The IBI method also bears some similarity to the reverse Monte Carlo simulation method for iteratively determining structural ensembles from experimental scattering data (111, 112).

The motivation for the IBI method can be easily understood by further analyzing the simple model described above, in which each molecule is represented by a single site and the CG potential only includes central pair potentials (107, 109). Given a set of target site-site RDFs,  $\mathcal{g}_\zeta(r)$ , obtained from either experiment or all-atom simulation, one can determine a set of approximate potentials  $U_\zeta(r)$  from direct Boltzmann inversion. Simulations employing these potentials will then generate a set of CG site-site distribution functions,  $G_\zeta(r)$ . However, for the reasons described above, these CG RDFs may not quantitatively match the target site-site RDFs,  $\mathcal{g}_\zeta(r)$ . If, at some particular distance  $r_0$ , the RDF generated by the CG model is greater than the RDF generated by the atomistic model, i.e.,  $G_\zeta(r_0) > \mathcal{g}_\zeta(r_0)$ , this implies that the corresponding ppmf of the CG model,  $W_\zeta(r) = -k_B T \ln G_\zeta(r)$ , is more attractive than the ppmf of the atomistic model at the same distance, i.e.,  $W_\zeta(r_0) < w_\zeta(r_0)$ . Consequently, the CG model will more accurately reproduce the atomistic RDF at  $r_0$  if the CG ppmf is increased (made more repulsive) at that distance. Because the ppmf includes the direct interaction, the most straightforward way to increase the ppmf (and thus decrease the CG RDF) at  $r_0$  is to simply increase the pair potential at that distance. A similar argument suggests that the CG RDF can be increased at  $r_0$  by simply decreasing the pair potential at that distance. These ideas suggest that the CG potentials can be improved by modifying each term in the potential (107, 109):

$$\delta U_\zeta(x) = \varepsilon_\zeta (w_\zeta(x) - W_\zeta(x)) = -\varepsilon_\zeta k_B T \ln \left[ \frac{p_\zeta(x)}{P_\zeta(x|U)} \right], \quad (35)$$

where  $\varepsilon_\zeta$  is a factor that may be adapted for different interactions or iterations. After modifying the potentials according to Eq. 35, the CG model is simulated again to generate a new set of structural correlation functions. These functions are compared with the target correlation functions, the CG potentials are modified according to Eq. 35, and, assuming convergence, the iterative procedure is repeated until the CG model reproduces the target distribution functions with sufficient accuracy.

The IBI method has been successfully used for a range of molecular liquids and polymeric systems (113–121). The simplicity of the method is particularly appealing: Each term in the potential is modified by considering the conjugate distribution function. However, the general convergence properties of the method are not very well understood. As described above, the ppmf includes both the direct interaction and also an environment-mediated contribution. This latter contribution is only indirectly treated by the IBI method. In principle, because the different interactions may be highly coupled, modifying one particular potential can significantly impact the other distributions. Also, in some cases, the convergence of the IBI method has been sensitive to the order with which the different potentials are revised (104, 122). In particular, it has been suggested that high-frequency “hard” potentials should be optimized before low-frequency “soft” potentials (116, 123).

### 3.3. Inverse Monte Carlo

Lyubartsev and Laaksonen earlier introduced an inverse Monte Carlo (IMC) method for determining CG potentials that reproduce target distribution functions (124, 125). The IMC method has been employed to quantitatively model various molecular systems, including liquids (124), lipid bilayers (126–128), and ionic systems (129, 130). The IMC method has proven to be particularly amenable to theoretical characterization, and several other related approaches have provided additional insight into the IMC approach.

Similarly to the IBI method, the IMC method employs iterative CG simulations to determine potentials  $\{U_\zeta(x)\}$  that reproduce a set of target distributions,  $\{p_\zeta(x)\}$ . However, while the IBI method corrects each potential,  $U_\zeta$ , independently according to Eq. 35, the IMC method employs linear response theory to estimate how variations in the potential  $U_\zeta$  impact all of the distribution functions. A susceptibility matrix (124) describes how the distribution  $P_\zeta(x|U)$  varies with small changes in  $U_{\zeta'}(x)$ .

$$M_{\zeta\zeta'}(x, x') \equiv \frac{\delta P_\zeta(x|U)}{\delta U_{\zeta'}(x')} = -\frac{1}{k_B T} \langle \Delta \hat{\rho}_\zeta(\mathbf{R}; x) \Delta \hat{\rho}_{\zeta'}(\mathbf{R}; x') \rangle_U, \quad (36)$$

where the second equality follows from employing Eq. 28 and differentiating Eq. 31, with

$$\Delta \hat{\rho}_\zeta(\mathbf{R}; x) = \hat{\rho}_\zeta(\mathbf{R}; x) - \langle \hat{\rho}_\zeta(\mathbf{R}; x) \rangle_U. \quad (37)$$

The averages subscripted by  $U$  in Eqs. 36 and 37 are evaluated using the CG model with potential  $U$  according to Eq. 31, e.g.,  $\langle \hat{\rho}_\zeta(\mathbf{R}; x) \rangle_U = P_\zeta(x|U)$ . According to Eq. 36, the susceptibility matrix,  $M_{\zeta\zeta'}(x, x')$ , is proportional to the covariance matrix of fluctuations in the fields  $\{\hat{\rho}_\zeta\}$  sampled by CG simulations.

The linear analysis then relates variations in the CG potentials  $\delta U_\zeta$  to variations in the CG distributions according to:

$$\delta P_\zeta(x) = \sum_{\zeta'} \int dx' M_{\zeta\zeta'}(x, x') \delta U_{\zeta'}(x'). \quad (38)$$

In the IMC method,  $\delta P_\zeta(x)$  is defined by the difference between the target and CG simulated densities,  $p_\zeta(x) - P_\zeta(x|U)$ . Equation 38 is then solved for  $\delta U_{\zeta'}(x')$  to determine the corrections to the CG potential according to the Newton-Raphson method (125). In practice, Eq. 38 is discretized and numerically solved as a system of coupled linear algebraic equations (124). (Alternatively, as discussed further below, the potential may be expressed as a linear combination of discrete parameters which are refined using an analogous system of linear equations.) The CG model is then simulated with the updated potential; the resulting distributions are compared with the target distributions, and the updated potentials are again corrected according to Eq. 38. Assuming convergence, the process is iterated until the final CG potentials reproduce the target distributions with sufficient accuracy.

### **3.4. Fundamental Underpinnings of Structure-Focused CG Methods**

#### *3.4.1. Useful Tools*

The Gibbs inequality (85) and generating functions (17) are useful tools for theoretical analyses of CG models.

- 1. Gibbs inequality** Consider two positive functions (e.g., probability distributions),  $f$  and  $g$  that are defined on the same domain and have the same normalization, i.e.,

$$\int dx(f(x) - g(x)) = 0, \quad (39)$$

where  $x$  may be a high-dimensional variable. It then follows that

$$\int dx f(x) \ln\left(\frac{f(x)}{g(x)}\right) \geq 0 \quad (40)$$

and the equality is only obtained when  $f(x) = g(x)$  over the entire domain.

The Gibbs inequality implies the Gibbs-Bogoliubov-Feynman variational principle (85), which is closely related to several variational principles for CG modeling (49, 91). Within the canonical ensemble, the configuration probability distribution for a CG model with a potential  $U$  may be expressed:

$$P_R(\mathbf{R}|U) = V^{-N} \exp[-(U(\mathbf{R}) - F[U])/k_B T], \quad (41)$$

where  $F[U]$  is the excess (nonideal) configurational contribution to the Helmholtz free energy for a CG model with potential  $U$ . If  $f$  and  $g$  are canonical configuration distributions associated with different potentials,  $P_R(\mathbf{R}|U_A)$  and  $P_R(\mathbf{R}|U_B)$ , respectively, then Eq. 40 implies that:

$$F_B \leq F_A + \langle U_B(\mathbf{R}) - U_A(\mathbf{R}) \rangle_{U_A}, \quad (42)$$

where  $F_A = F[U_A]$ ,  $F_B = F[U_B]$ , and the angular brackets denote a canonical ensemble average weighted by  $P_R(\mathbf{R}|U_A)$ . The proof of the Gibbs inequality is straightforward (85, 131). The left hand side of Eq. 40 can be expressed:

$$\int dx g(x) \{z(x) \ln z(x) - z(x) + 1\}, \quad (43)$$

where  $z(x) = f(x)/g(x)$ , which is positive by hypothesis. The Gibbs inequality in Eq. 40 then immediately follows as a consequence of the fact that  $z \ln z - z + 1 \geq 0$  is a convex function for  $z > 0$  and achieves its minimum of 0 at  $z = 1$ .

- 2. Generating functions** The configurational integral is a prototypical example of a generating function (17).

$$Z[U] = V^N \exp[-F[U]/k_B T] \quad (44)$$

$$= \int d\mathbf{R} \exp \left[ -\frac{1}{k_B T} \sum_{\zeta} \int dx U_{\zeta}(x) \hat{\rho}_{\zeta}(\mathbf{R}; x) \right], \quad (45)$$

where the last line employs Eqs. 28 and 41. Differentiating  $-F[U]/k_B T = \ln(V^{-N}Z[U])$  with respect to potentials,  $U_{\zeta}$ , generates cumulants of the conjugate fields:

$$\frac{\delta}{\delta U_{\zeta}(x)} \ln(V^{-N}Z[U]) = -\frac{1}{k_B T} \langle \hat{\rho}_{\zeta}(\mathbf{R}; x) \rangle_U \quad (46)$$

$$= -\frac{1}{k_B T} P_{\zeta}(x|U) \quad (47)$$

$$\frac{\delta^2}{\delta U_{\zeta}(x) \delta U_{\zeta'}(x')} \ln(V^{-N}Z[U]) = \left( \frac{1}{k_B T} \right)^2 \langle \Delta \hat{\rho}_{\zeta}(\mathbf{R}; x) \Delta \hat{\rho}_{\zeta'}(\mathbf{R}; x') \rangle_U \quad (48)$$

$$= -\left( \frac{1}{k_B T} \right) M_{\zeta\zeta'}(x, x'). \quad (49)$$

Note that the second functional derivative is proportional to the susceptibility matrix considered by the IMC method.

### 3.4.2. Existence and Uniqueness

Given a set of target structural correlation functions,  $\{p_{\zeta}(x)\}$ , that correspond to a set of independent fields  $\{\hat{\rho}_{\zeta}(\mathbf{R}; x)\}$  in Eq. 30 and that are obtained from either atomistic simulation or experimental data, two important questions arise: (1) Does there exist a set of CG interaction potentials,  $\{U_{\zeta}(x)\}$ , that will reproduce the target distribution functions? and (2) Assuming that such a set exists, is this set of potentials unique?

The answer to the second question is to the affirmative: At least in principle and assuming that they exist, the set of potentials that reproduces a set of distributions is unique. The proof of this assertion follows from a straightforward extension of the celebrated Henderson uniqueness theorem (132), which is itself an extension of the Hohenberg-Kohn-Mermin theorem (133, 134) and follows from the Gibbs inequality. This result has been generalized for site-site intermolecular pair potentials (135). We have recently further generalized the proof for more complex potentials (96).

Consider two different CG potentials,  $U_A$  and  $U_B$ , that may be expressed in terms of two sets of potentials,  $\{U_{\zeta A}(x)\}$  and  $\{U_{\zeta B}(x)\}$ , according to Eq. 28. We assume that the set of corresponding conjugate fields,  $\{\hat{\rho}_{\zeta}(\mathbf{R}; x)\}$ , is linearly independent. These two potentials generate two sets of structural correlation functions  $\{P_{\zeta A}(x) = P_{\zeta}(x|U_A)\}$  and  $\{P_{\zeta B}(x) = P_{\zeta}(x|U_B)\}$ . Because the fields are independent, it follows that, if the two sets of potentials differ (by more than a constant) in any one term, then  $U_A(\mathbf{R}) - U_B(\mathbf{R})$  is not a constant. Consequently, the inequality is implied in the Gibbs inequality, Eq. 40. Employing the Gibbs-Bogoliubov-Feynman for both potentials according to Eq. 42 and combining the two relations leads to

$$\sum_{\zeta} \int dx \Delta U_{\zeta}(x) \Delta P_{\zeta}(x) < 0, \quad (50)$$

where  $\Delta U_{\zeta}(x) = U_{\zeta A}(x) - U_{\zeta B}(x)$  and  $\Delta P_{\zeta}(x) = P_{\zeta A}(x) - P_{\zeta B}(x)$ . The inequality can only be achieved if  $\Delta P_{\zeta}(x) \neq 0$  for some combination of  $\zeta$  and  $x$ . Therefore, if two CG models generate the same set of distribution functions,  $\{P_{\zeta}(x)\}$ , then the two sets of interaction potentials,  $\{U_{\zeta}(x)\}$ , employed by the models are the same to within an additive constant.

The answer to the first question is much less straightforward. Chayes and coworkers provided a variational proof demonstrating that, for any  $m$  – body distribution function that is the reduction of an equilibrium distribution (for a possibly more complex potential), i.e.,  $\rho^{(m)}(\mathbf{R}_1, \dots, \mathbf{R}_m) = \langle \delta(\mathbf{R}_1 - \mathbf{M}_1(\mathbf{r})) \cdots \delta(\mathbf{R}_m - \mathbf{M}_m(\mathbf{r})) \rangle$ , there does indeed exist an  $m$  – body potential that will reproduce the  $m$  – body distribution (136, 137). The proof is quite complicated and will not be further discussed here. Moreover, it is not clear to what extent this result can be generalized. In particular, it is not clear under what conditions there exist molecular mechanics potentials of the form given by Eq. 28 that will reproduce a set of corresponding distribution functions.

### 3.4.3. Variational Principles and Properties

Several variational principles have been identified for determining CG potentials that reproduce a given set of target distribution functions. These approaches typically construct a functional of the CG potential,  $\mathcal{F}[U]$ , that achieves a stationary point when the

target distribution functions are reproduced. Such a functional can be determined through two simple contributions:

1. The variational functional must couple each potential  $U_\zeta(x)$  to the target atomistic distribution function,  $p_\zeta(x)$ . The simplest way to do so is to include a term of the form:

$$\sum_\zeta \int dx p_\zeta(x) U_\zeta(x) = \int d\mathbf{r} p_r(\mathbf{r}) U(\mathbf{M}(\mathbf{r})) \\ = \int d\mathbf{R} p_R(\mathbf{R}) U(\mathbf{R}), \quad (51)$$

where we recall that  $p_R$  is the distribution of CG configurations implied by the atomistic distribution and CG mapping, so that Eq. 51 is the average of the CG potential weighted by the atomistic probability distribution.

2. Differentiation of the variational functional should also generate the corresponding CG distribution,  $P_\zeta(x|U)$ . The simplest way to do so is to include the generating function  $F[U] = -k_B T \ln(V^{-N} Z[U])$ , according to Eq. 46.

By combining these terms, one obtains

$$\mathcal{F}[U] = \int d\mathbf{R} p_R(\mathbf{R})(U(\mathbf{R}) - F[U]). \quad (52)$$

By construction,

$$\frac{\delta \mathcal{F}[U]}{\delta U_\zeta(x)} = p_\zeta(x) - P_\zeta(x|U), \quad (53)$$

so that  $\mathcal{F}[U]$  achieves its stationary point when each target distribution is reproduced by the CG model. The quantity  $\mathcal{F}[U]$  bears an intriguing similarity to an entropy function,  $S = (U - F)/T$ , for the CG model, but weights configurations according to the probability of sampling them from the **atomistic** model rather than from the CG model.

Murtola et al. (49) noted that  $\mathcal{F}[U]$  is closely related to the functional employed in the density functional analysis of Chayes and coworkers (136, 137) and also to the relative entropy  $S_{\text{rel}}$  considered by Shell (91). By employing the canonical configuration distribution for the CG model, the functional  $\mathcal{F}$  can also be expressed:

$$\mathcal{F}[U] = -k_B T \int d\mathbf{R} p_R(\mathbf{R}) \ln[V^N P_R(\mathbf{R}|U)] \quad (54)$$

$$= T(S_{\text{rel}}[U] + s[u] - S_{\text{map}}), \quad (55)$$

where  $s[u]$  is the configurational entropy of the underlying atomistic model and  $S_{map}$  was defined in Eq. 25. Because  $s[u]$  and  $S_{map}$  are both independent of  $U$ , it follows that the CG potential that minimizes the functional  $\mathcal{F}[U]$  also minimizes the relative entropy,  $S_{rel}$ , and thus maximizes the overlap of the atomistic and CG configuration distributions.

Because the target distributions,  $p_\zeta(x)$ , are independent of the CG potential, Eq. 53 implies that the curvature of  $\mathcal{F}[U]$ , i.e.,  $\delta^2 \mathcal{F}[U] / \delta U_\zeta(x) \delta U_{\zeta'}(x')$ , is proportional to the covariance matrix,  $\langle \Delta \hat{\rho}_\zeta(\mathbf{R}; x) \Delta \hat{\rho}_{\zeta'}(\mathbf{R}; x') \rangle_u$ , in Eq. (36) and is also proportional to the susceptibility matrix employed by the IMC method. Consequently, the IMC method corresponds to a Newton-Raphson method for finding the stationary point of  $\mathcal{F}[U]$  (49, 125). Therefore, according to Eq. 55, the IMC method is also equivalent to the relative entropy approach for determining effective potentials (49). This analysis also provides insight into the convergence properties of IMC and related approaches (49). Because the covariance is everywhere positive semi-definite, it follows that  $\mathcal{F}[u]$  has at most a single stationary point in the parameter space of potentials. However, these considerations are not sufficient to determine whether or not this stationary point exists.

Lyubartsev and Laaksonen also proved another intriguing variational property of the IMC method (124). They demonstrated that the pair potentials that reproduce a given set of target pair correlation functions also maximize the entropy of the CG model relative to any other model potential that reproduces the target correlation functions. This result also readily generalizes for molecular mechanics potentials (96): Given a set of target structural correlation functions  $\{p_\zeta(x)\}$  (and assuming such a potential exists), a molecular mechanics potential of the form given by Eq. 27 maximizes the entropy of the CG model relative to any other potential that reproduces the target correlation functions. This follows by applying the method of Lagrange multipliers to maximize the CG entropy with respect to variations in the CG distribution:

$$S[P_R] = -k_B \int d\mathbf{R} P_R(\mathbf{R}) \ln[V^N P_R(\mathbf{R})], \quad (56)$$

subject to the normalization condition:

$$\int d\mathbf{R} P_R(\mathbf{R}) - 1 = 0 \quad (57)$$

and the condition that each target correlation function is reproduced:

$$\int d\mathbf{R} P_R(\mathbf{R}) \hat{\rho}_\zeta(\mathbf{R}; x) - p_\zeta(x) = 0, \quad (58)$$

for each  $\zeta$  and  $x$ .

Papoian and coworkers (138, 139) have provided further insight by emphasizing the coarse-graining procedure as a molecular renormalization group calculation and emphasizing the utility of finding a compact representation of the CG potential:

$$U(\mathbf{R}) = \sum_{\alpha} K_{\alpha} S_{\alpha}(\mathbf{R}), \quad (59)$$

where the (few) discrete potential parameters,  $K_{\alpha}$ , are now interpreted as external fields,  $U_{\alpha}(x) \rightarrow K_{\alpha}$ . These fields then couple to (potentially many-body) structural observables  $\hat{\rho}_{\alpha}(\mathbf{R}; x) \rightarrow S_{\alpha}(\mathbf{R})$ , and determine a set of conjugate structural correlation functions,  $P_{\alpha}(x|U) \rightarrow \langle S_{\alpha}(\mathbf{R}) \rangle_U$ . The CG potentials are optimized in a linear iterative renormalization scheme similar to the IMC approach in order to reproduce a corresponding set of target correlation functions (124). The CG correlation functions correspond to derivatives of the partition function,  $Z[U]$ , with respect to the parameters,  $K_{\alpha}$ , according to Eq. 46. Consequently, the IMC and relative entropy approaches can also be interpreted as renormalization-based methods for reproducing derivatives of the partition function.

### 3.5. Multiscale Coarse-Graining

While the above methods parameterize potentials to reproduce lower-order structural correlation functions, Izvekov and Voth have suggested a multiscale coarse-graining (MS-CG) method for determining CG potentials (140–142). The MS-CG method employs a force-matching approach similar to that previously employed for ab initio simulations (143, 144). The MS-CG method has been successfully applied for modeling a wide range of condensed phase systems, including molecular and ionic liquids (106, 141, 145–147), lipid bilayers (140, 148, 149), nanoparticles (150), carbohydrates (151), peptides (152), and proteins (153).

Rather than directly or iteratively addressing structural correlation functions, the MS-CG method employs force information sampled from atomistic simulations and employs this information to directly approximate the many-body mean force field (80, 106), i.e., the force field derived from gradients of the many-body PMF:

$$\mathbf{F}_I^0(\mathbf{R}) = -\frac{\partial U^0(\mathbf{R})}{\partial \mathbf{R}_I} = \langle \mathbf{f}_I(\mathbf{r}) \rangle_{\mathbf{R}}, \quad (60)$$

where the second relation follows from direct differentiation of Eq. 15 and

$$\mathbf{f}_I(\mathbf{r}) = \sum_{i \in I} \mathbf{f}_i(\mathbf{r}) \quad (61)$$

is the net force on CG site  $I$  specified by the atomistic force field in configuration  $\mathbf{r}$ . The subscripted angular brackets in Eq. 60 denote a

conditioned canonical ensemble average evaluated over all the atomistic configurations  $\mathbf{r}$  that map to the fixed CG configuration  $\mathbf{R}$ :

$$\langle \alpha(\mathbf{r}) \rangle_{\mathbf{R}} = \langle \alpha(\mathbf{r}) \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) \rangle / \langle \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) \rangle, \quad (62)$$

for any quantity  $\alpha(\mathbf{r})$ . Equation 60 demonstrates that  $U^0$  is referred to as a “many-body potential of mean force” because the force field derived from its gradients is a many-body mean force field, i.e., a conditioned canonical ensemble average of the atomistic force  $\mathbf{f}_I(\mathbf{r})$  experienced by the site  $I$  in the atomistic configurations,  $\mathbf{r}$ , that map to the fixed CG configuration,  $\mathbf{R}$  (80, 86, 154). Because  $U^0$  is determined from a many-body distribution function, the mean force  $\mathbf{F}_I^0$  is generally a many-body function of the coordinates for all  $N$  sites. CG MD simulations that sample the canonical ensemble, while employing the many-body mean force field as a conservative force field, will generate a distribution of configurations that is consistent with the underlying atomistic model according to Eq. 12 (80).

To proceed in analyzing the MS-CG method, it is convenient to consider a CG force field  $\mathbf{F}$  as an object in an abstract vector space (80, 155–157). Each element in this space specifies a set of  $N$  vector-valued functions ( $\mathbf{F}_1(\mathbf{R}), \dots, \mathbf{F}_N(\mathbf{R})$ ) that define a Cartesian force on each site  $I$  in any configuration  $\mathbf{R}$ . An inner product between two CG force fields  $\mathbf{F}^{(1)}$  and  $\mathbf{F}^{(2)}$  is defined:

$$\mathbf{F}^{(1)} \odot \mathbf{F}^{(2)} = \left\langle \sum_I \mathbf{F}_I^{(1)}(\mathbf{M}(\mathbf{r})) \cdot \mathbf{F}_I^{(2)}(\mathbf{M}(\mathbf{r})) \right\rangle. \quad (63)$$

It is important to note that this inner product corresponds to an ensemble average weighted by the atomistic equilibrium configuration distribution,  $p_r(\mathbf{r})$ . A norm is defined according to  $\|\mathbf{F}\| = (\mathbf{F} \odot \mathbf{F})^{1/2}$ . Because the mean force field,  $\mathbf{F}^0$ , is a conditioned average of the atomistic force field,  $\mathbf{f}$ , the following identity holds for any CG force field,  $\mathbf{F}$  (80, 155):

$$\left\langle \sum_I \mathbf{f}_I(\mathbf{r}) \cdot \mathbf{F}_I(\mathbf{M}(\mathbf{r})) \right\rangle = \mathbf{F}^0 \odot \mathbf{F}. \quad (64)$$

The MS-CG force field is determined by minimizing a force-matching (143, 144) functional defined on the space of CG force fields (80, 140):

$$\chi^2[\mathbf{F}] = \frac{1}{3N} \left\langle \sum_I |\mathbf{f}_I(\mathbf{r}) - \mathbf{F}_I(\mathbf{M}(\mathbf{r}))|^2 \right\rangle. \quad (65)$$

As a direct consequence of Eq. 64, the MS-CG functional may be re-expressed:

$$\chi^2[\mathbf{F}] = \chi^2[\mathbf{F}^0] + \frac{1}{3N} \|\mathbf{F}^0 - \mathbf{F}\|^2, \quad (66)$$

where  $\mathbf{F}^0$  is the many-body mean force field. The second term  $\|\mathbf{F} - \mathbf{F}^0\|^2 \geq 0$  for all CG force fields,  $\mathbf{F}$ , and  $\|\mathbf{F} - \mathbf{F}^0\|^2 = 0$  if and only if  $\mathbf{F} = \mathbf{F}^0$ , i.e., if and only if  $\mathbf{F}_I(\mathbf{R}) = \mathbf{F}_I^0(\mathbf{R})$  for all  $I$  and  $\mathbf{R}$ . Consequently, the CG force field that provides the global minimum of  $\chi^2$  is the many-body mean force field, and the resulting CG model will be consistent with the underlying atomistic model from which the forces  $\mathbf{f}_I$  were sampled. Intriguingly, recent work has demonstrated that  $\chi^2$  can also be expressed as the gradient squared of the information function employed in the relative entropy approach (96).

Equations 65 and 66 also have an intuitive geometric interpretation. The vector space of CG force fields described above may be considered a subspace of a larger vector space of force fields that depend upon the atomistic configuration. The atomistic force field  $\mathbf{f} = (\mathbf{f}_1(\mathbf{r}), \dots, \mathbf{f}_N(\mathbf{r}))$  defined in Eq. 61 is a particular element in this larger vector space. The functional  $\chi^2[\mathbf{F}]$  then corresponds to the “distance” of a particular CG force field  $\mathbf{F}$  from the atomistic force field,  $\mathbf{f}$ . According to Eq. 60 the mean force field  $\mathbf{F}^0$  is a conditioned average of the atomistic force field, which corresponds to a geometric projection from the larger vector space into the vector subspace of CG force fields (88, 89, 154–159). Equivalently, Eq. 64 implies that the vector  $\mathbf{f} - \mathbf{F}^0$  is orthogonal to any vector in the subspace of CG force fields (155). By geometric arguments, the projection of a vector (i.e.,  $\mathbf{f}$ ) into a subspace determines the element (i.e.,  $\mathbf{F}^0$ ) that is minimum distance from the original vector. Thus, Eq. 66 corresponds to a simple statement of the Pythagorean theorem within the space of CG force fields: The square of the distance of the atomistic force field,  $\mathbf{f}$ , from a trial CG force field,  $\mathbf{F}$ , is equal to the sum of two terms that correspond to the orthogonal sides of a right triangle: (1) the square of the distance from the atomistic force field to the mean force field, i.e.,  $\chi^2[\mathbf{F}^0]$ ; and (2) the square of the distance from the mean force field to the trial CG force, i.e.,  $\|\mathbf{F} - \mathbf{F}^0\|^2$ . Because the first component is fixed by the original atomistic model and the CG mapping, minimizing the distance of a trial CG force field,  $\mathbf{F}$ , from the atomistic force field,  $\mathbf{f}$ , (i.e., minimizing  $\chi^2[\mathbf{F}]$ ) corresponds to minimizing the distance of the trial force field from the mean force field (i.e., minimizing  $\|\mathbf{F} - \mathbf{F}^0\|^2$ ). This distance reaches its global and unique minimum when the two force fields are equal.

As noted above for the many-body PMF, it is generally impossible to calculate, represent, or simulate the many-body mean force field,  $\mathbf{F}^0$ . In the context of the MS-CG theory, doing so would correspond to minimizing  $\chi^2$  with a complete basis set for the space of CG force fields, which is completely intractable for any complex molecular system. Consequently, it is important to identify practically useful incomplete basis sets for CG force fields (155). The following analysis will consider the basis set relevant for molecular mechanics potentials of the form given by Eq. 27, although

more complex potentials have also been used (160). Such approximate potentials are frequently employed and often provide a reasonably accurate approximation to the many-body mean force field. However, it is important to recognize that such a basis set is highly incomplete. At this time, it remains unclear under which circumstances such an incomplete basis set will allow an accurate approximation to the PMF.

We represent each term in the approximate potential with a discrete set of basis functions of a single variable,  $U_{\zeta d}(x)$ , with constant coefficients  $\phi_{\zeta d}$  that will serve as force field parameters (80, 155):

$$U_\zeta(x) = \sum_d \phi_{\zeta d} U_{\zeta d}(x). \quad (67)$$

A wide range of basis functions may be usefully used (96, 155, 157, 161–163). For instance, polynomial functions have been employed for harmonic and anharmonic bonded potentials, cosine series have been employed for torsional potentials, and discrete delta functions (“hat functions”) or flexible spline functions have been employed in representing bonded potentials and also nonbonded pair potentials. This basis expansion of the potentials determine a basis expansion of the corresponding force functions:

$$F_\zeta(x) = \sum_d \phi_{\zeta d} f_{\zeta d}(x), \quad (68)$$

where  $F_\zeta(x) = -dU_\zeta(x)/dx$  and  $f_{\zeta d}(x) = -dU_{\zeta d}(x)/dx$ . The force on site  $I$  may be expressed:

$$\mathbf{F}_I(\mathbf{R}) = \sum_\zeta \sum_d \phi_{\zeta d} \mathcal{G}_{I;\zeta d}(\mathbf{R}), \quad (69)$$

where

$$\mathcal{G}_{I;\zeta d}(\mathbf{R}) = \sum_\lambda \frac{\partial \psi_{\zeta \lambda}(\mathbf{R})}{\partial \mathbf{R}_I} f_{\zeta d}(\psi_{\zeta \lambda}(\mathbf{R})). \quad (70)$$

By using the force field representation introduced above and replacing the sums over  $\zeta$  and  $d$  with a single sum over a “super-index”  $D = \zeta d$ , Eq. 69 may be reexpressed (155):

$$\mathbf{F}(\phi) = \sum_D \phi_D \mathcal{G}_D. \quad (71)$$

In this representation, the CG force field,  $\mathbf{F}$ , has been expressed as a linear combination of force field vectors,  $\mathcal{G}_D$ , the components of which are given by Eq. 70. The set of vectors included in Eq. 71 defines an incomplete basis set that spans a vector subspace of CG force fields. The force field parameters  $\{\phi_D\}$  serve as constant coefficients that identify a particular force field within this subspace (80).

Employing this basis set expansion for the force field, the MS-CG functional becomes a simple quadratic function of the force field parameters:

$$\chi^2(\phi) = \chi^2[\mathbf{F}(\phi)] = \frac{1}{3N} \left\langle \sum_I \left| \mathbf{f}_I(\mathbf{r}) - \sum_D \phi_D \mathcal{G}_{I,D}(\mathbf{M}(\mathbf{r})) \right|^2 \right\rangle. \quad (72)$$

The same geometric arguments that determined the many-body mean force field as a projection of the atomistic force field now apply for the vector subspace spanned by a given basis set. The function  $\chi^2(\phi)$  may be re-expressed:

$$\chi^2(\phi) = \chi^2[\mathbf{F}^0] + \frac{1}{3N} \|\mathbf{F}^0 - \mathbf{F}(\phi)\|^2 \quad (73)$$

$$= \chi^2[\mathbf{F}^0] + \frac{1}{3N} \|\mathbf{F}^0 - \mathbf{F}(\phi^*)\|^2 + \frac{1}{3N} \|\mathbf{F}(\phi^*) - \mathbf{F}(\phi)\|^2, \quad (74)$$

where  $\mathbf{F}(\phi^*) = \sum_D \phi_D^* \mathcal{G}_D$  corresponds to the geometric projection of the mean force  $\mathbf{F}^0$  into the subspace spanned by the basis set  $\{\mathcal{G}_D\}$ . This force field then provides the unique global minimum of  $\chi^2(\phi)$  relative to any force field within this subspace (i.e., relative to any force field of the form given by Eq. 69). By minimizing  $\chi^2(\phi)$ ,  $\mathbf{F}(\phi^*)$  also minimizes the distance  $\|\mathbf{F}^0 - \mathbf{F}(\phi)\|^2$  and is, therefore, the force field within the subspace that is “closest” to the many-body mean force. Moreover, this force field is unique for a given basis set, although, if the basis vectors are not linearly independent, then  $\mathbf{F}(\phi^*)$  does not determine a unique set of parameters  $\phi^*$  (155). The present framework also applies in the case that a fixed “reference” contribution from, e.g., electrostatic interactions, is included in the CG potential (155).

This analysis forms the basis of the MS-CG variational principle for determining an optimized CG potential using forces sampled from atomistic MD simulations (80, 140, 141, 155, 163). If one minimizes  $\chi^2[\mathbf{F}]$  while using a complete force field basis set, then the resulting force field will be the many-body mean force field (106) and the associated CG model will be completely consistent with the underlying atomistic model (80). If one minimizes  $\chi^2(\phi)$  while using an incomplete force field basis set, then the resulting CG force field will be the projection of the mean force field onto the given basis (155). This force field will then be an “optimal approximation” to the mean force field in the sense that it minimizes the distance from the mean force field,  $\|\mathbf{F}^0 - \mathbf{F}(\phi)\|^2$ , relative to any other force field spanned by the basis (80, 159).

This corresponds to a different variational principle (or different metric) than is considered by the iterative structure-motivated CG methods. In general, given the same incomplete basis set, the MS-CG potential will differ from the potentials that are

obtained from iterative structure-motivated CG methods. Iterative structure-motivated methods (e.g., IBI or IMC) are designed to determine the CG potential that reproduces a set of target distribution functions. However, CG MD simulations that sample the canonical ensemble using the MS-CG potential are not guaranteed to quantitatively reproduce any particular distribution function. Rather, in the case of the MS-CG method, distribution functions are predictions of the method (106).

Importantly, Eq. 72 demonstrates that the MS-CG potential is determined from a linear least squares problem, for which powerful numerical techniques have been developed (164, 165). Two numerical approaches have been developed for determining the MS-CG potential (155, 163). In principle, both methods obtain identical potentials. Both methods employ a sequence of  $N_t$  sets of atomistic configurations and forces that have been sampled by either MD or MC simulation,  $\{\mathbf{r}_t, \mathbf{f}(\mathbf{r}_t)\}$ . The function  $\chi^2(\phi)$  may then be expressed:

$$\chi^2(\phi) = \frac{1}{3N_t N} \sum_t \sum_I \left| \mathbf{f}_I(\mathbf{r}_t) - \sum_D \phi_D \mathcal{G}_{I;D}(\mathbf{M}(\mathbf{r}_t)) \right|^2, \quad (75)$$

which may then be minimized directly by methods such as QR decomposition or singular value decomposition. If the calculation involves too many sites, parameters, and configurations for the minimization to be performed in memory, then  $\chi^2$  can be approximately solved via a block-averaging approach in which configurations are partitioned into smaller blocks, and  $\chi^2$  is expressed as a sum of residuals corresponding to each block. The optimal set of parameters are then approximated by averaging the parameters that minimize  $\chi^2$  for each block (141). In practice, the block-averaging approximation has provided an accurate and stable numerical method for determining the parameters for the MS-CG potential for a wide range of systems (163).

Alternatively, the sampled configurations and forces can be used to approximate equilibrium correlation functions and  $\chi^2$  can be expressed:

$$\chi^2(\phi) = \chi^2(0) - \frac{2}{3N} \sum_D b_D \phi_D + \frac{1}{3N} \sum_D \sum_{D'} \phi_D G_{DD'} \phi_{D'}, \quad (76)$$

where

$$b_D = \left\langle \sum_I \mathbf{f}_I(\mathbf{r}) \cdot \mathcal{G}_{I;D}(\mathbf{M}(\mathbf{r})) \right\rangle, \quad (77)$$

$$G_{DD'} = \left\langle \sum_I \mathcal{G}_{I;D}(\mathbf{M}(\mathbf{r})) \cdot \mathcal{G}_{I;D'}(\mathbf{M}(\mathbf{r})) \right\rangle. \quad (78)$$

Minimizing Eq. 76 with respect to  $\phi_D$  leads to the normal system of equations (106, 155, 163):

$$\sum_{D'} G_{DD'} \phi_{D'} = b_D. \quad (79)$$

The system of normal equations have the advantage that they require substantially less computational memory. However, the condition number of the matrix  $G_{DD}$  is the square of the condition number of the matrices involved in directly minimizing  $\chi^2$  using the approaches described in the previous paragraph. Nevertheless, the normal equations have also been successfully applied for a variety of different systems (106, 155–157, 162).

The normal equations formulated in Eqs. 77 further emphasize the geometric interpretation of the MS-CG method (156, 157). From the identity of Eq. 64, it follows that  $b_D$  corresponds to the inner product of the mean force field,  $\mathbf{F}^0$ , with the basis vector  $\mathcal{G}_D$ :

$$b_D = \mathbf{F}^0 \odot \mathcal{G}_D. \quad (80)$$

Similarly,  $G_{DD'}$  is a metric tensor that corresponds to the inner product between basis vectors:  $G_{DD'} = \mathcal{G}_D \odot \mathcal{G}_{D'}$ . The normal equations determine the MS-CG force field,  $\mathbf{F}$ , as a projection of the mean force field,  $\mathbf{F}^0$ , into the subspace spanned by  $\{\mathcal{G}_D\}$  by requiring that the MS-CG and mean force field have the same inner product with each basis vector:

$$b_D = \mathcal{G}_D \odot \mathbf{F}^0 = \mathcal{G}_D \odot \mathbf{F}, \quad (81)$$

which is equivalent to determining the vector  $\mathbf{F}$  that makes  $\mathbf{F}^0 - \mathbf{F}$  orthogonal to every force field spanned by the basis. Because the inner product in the space of force fields corresponds to an ensemble average and because the different basis vectors correspond to molecular interactions that are correlated in a condensed phase, the basis vectors are not orthogonal and  $G_{DD}$  is not diagonal. Rather, the off-diagonal elements of  $G_{DD}$  quantify the correlations between the different interactions included in the potential (156, 157). If the different terms in the potential were statistically independent, then the metric tensor  $G_{DD}$  would be diagonal and each force field parameter,  $\phi_D$ , could be directly determined as  $b_D/G_{DD}$  (106). This corresponds to the assumption that is implicit in direct Boltzmann inversion. However, in contrast to the IBI or IMC methods, which neglect these many-body correlations in determining a first trial potential and then address the correlations by iteratively correcting the potential, the MS-CG method employs the metric tensor to directly disentangle these correlations. These considerations are expanded in the following subsection.

### 3.6. Generalized YBG Theory

The above discussion naturally emphasized the importance of atomistic forces in determining the MS-CG potential. Therefore, it is somewhat surprising that the MS-CG potential can also be directly determined from structural correlation functions (106, 156).

The resulting analysis further clarifies the connection between the MS-CG method and other structure-motivated methods. This correspondence is most conveniently explored by employing a continuous basis set to represent CG force fields (157). In this case, we consider potentials of the form given by Eq. 28, and the corresponding force field may be expressed:

$$\mathbf{F} = \sum_{\zeta} \int dx \phi_{\zeta}(x) \mathcal{G}_{\zeta}(x), \quad (82)$$

where  $\phi_{\zeta}(x) = -dU_{\zeta}(x)/dx$  is a force function and  $\mathcal{G}_{\zeta}(x)$  is a corresponding force field vector with elements

$$\mathcal{G}_{I;\zeta}(\mathbf{R}; x) = \sum_{\lambda} \frac{\partial \psi_{\zeta\lambda}(\mathbf{R})}{\partial \mathbf{R}_I} \delta(\psi_{\zeta\lambda}(\mathbf{R}) - x). \quad (83)$$

We note in passing that the present argument can be generalized to the case that the CG potential includes a fixed reference term (166).

When expressed in the continuous basis set, the MS-CG functional becomes (157)

$$\chi^2[\{\phi_{\zeta}(x)\}] = \chi^2[0] - \frac{2}{3N} \sum_{\zeta} \int dx b_{\zeta}(x) \phi_{\zeta}(x) \quad (84)$$

$$+ \frac{1}{3N} \sum_{\zeta} \sum_{\zeta'} \int dx \int dx' \phi_{\zeta}(x) G_{\zeta\zeta'}(x, x') \phi_{\zeta'}(x'), \quad (85)$$

and the normal equations may be expressed:

$$\sum_{\zeta'} \int dx' G_{\zeta\zeta'}(x, x') \phi_{\zeta'}(x') = b_{\zeta}(x). \quad (86)$$

In terms of

$$b_{\zeta}(x) = \mathcal{G}_{\zeta}(x) \odot \mathbf{F}^0 = \left\langle \sum_I \mathbf{f}_I(\mathbf{r}) \cdot \mathcal{G}_{I;\zeta}(\mathbf{M}(\mathbf{r}); x) \right\rangle \quad (87)$$

$$G_{\zeta\zeta'}(x, x') = \mathcal{G}_{\zeta}(x) \odot \mathcal{G}_{\zeta'}(x') = \left\langle \sum_I \mathcal{G}_{I;\zeta}(\mathbf{M}(\mathbf{r}); x) \cdot \mathcal{G}_{I;\zeta'}(\mathbf{M}(\mathbf{r}); x') \right\rangle, \quad (88)$$

where  $b_{\zeta}(x)$  is the inner product of the many-body mean force field with the basis vector  $\mathcal{G}_{\zeta}(x)$  and  $G_{\zeta\zeta'}(x, x')$  is a many-body structural correlation function that corresponds to a metric tensor defined by the given basis set. The metric tensor  $G_{\zeta\zeta'}(x, x')$  should be compared with the susceptibility matrix,  $M_{\zeta\zeta'}(x, x')$ , employed in IMC and defined in Eq. 36. While  $M_{\zeta\zeta'}(x, x')$  considers the correlations between every set of interactions in the CG potential,  $G_{\zeta\zeta'}(x, x')$  trivially vanishes for any two potentials that do not act on a common site.

Equation 87 indicates that the MS-CG method employs atomistic force information as a surrogate for information regarding the many-body mean force field (156, 157). Consequently,  $b_\zeta(x)$  can be re-expressed:

$$b_\zeta(x) = \int d\mathbf{R} p_R(\mathbf{R}) \sum_I \mathcal{G}_{I;\zeta}(\mathbf{R}; x) \cdot \mathbf{F}_I^0(\mathbf{R}) \quad (89)$$

$$= k_B T \int d\mathbf{R} \sum_I \mathcal{G}_{I;\zeta}(\mathbf{R}; x) \cdot \frac{\partial}{\partial \mathbf{R}_I} p_R(\mathbf{R}), \quad (90)$$

where the second relation expresses the many-body mean force as gradients of the many-body PMF which is related to  $p_R$  according to Eq. 14. Performing integration by parts,  $b_\zeta(x)$  can be expressed entirely in terms of structural correlation functions (156, 157):

$$b_\zeta(x) = k_B T \left( \frac{d}{dx} \bar{\mathcal{G}}_\zeta(x) - L_\zeta(x) \right). \quad (91)$$

One then arrives at a system of linear integral equations for determining the MS-CG potential from structural correlation functions,  $G_{\zeta\zeta'}(x, x')$ ,  $\bar{\mathcal{G}}_\zeta(x)$ , and  $L_\zeta(x)$ :

$$\sum_{\zeta'} \int dx' G_{\zeta\zeta'}(x, x') \phi_{\zeta'}(x') = k_B T \left( \frac{d}{dx} \bar{\mathcal{G}}_\zeta(x) - L_\zeta(x) \right), \quad (92)$$

where

$$\bar{\mathcal{G}}_\zeta(x) = \left\langle \sum_\lambda |\nabla \psi_{\zeta\lambda}(\mathbf{M}(\mathbf{r}))|^2 \delta(\psi_{\zeta\lambda}(\mathbf{M}(\mathbf{r})) - x) \right\rangle, \quad (93)$$

$$L_\zeta(x) = \left\langle \sum_\lambda \nabla^2 \psi_{\zeta\lambda}(\mathbf{M}(\mathbf{r})) \delta(\psi_{\zeta\lambda}(\mathbf{M}(\mathbf{r})) - x) \right\rangle, \quad (94)$$

$$|\nabla \psi_{\zeta\lambda}(\mathbf{R})|^2 = \sum_I (\partial \psi_{\zeta\lambda}(\mathbf{R}) / \partial \mathbf{R}_I)^2, \quad \text{and} \quad \nabla^2 \psi_{\zeta\lambda}(\mathbf{R}) = \sum_I \partial^2 \psi_{\zeta\lambda}(\mathbf{R}) / \partial \mathbf{R}_I^2.$$

This system of linear integral equations, Eq. 92, is a generalization for molecular systems (156, 157) of the Yvon-Born-Green theory, which is a well-known integral equation theory for simple liquids (85, 105). If, for instance, the CG potential is composed of simple pair potentials,  $U(\mathbf{R}) = \sum_\lambda U^{(2)}(\psi_\lambda(\mathbf{R}))$ , where  $\lambda$  identifies a particular pair of sites  $\{I, J\}$ ,  $\psi_\lambda(\mathbf{R}) = |\mathbf{R}_I - \mathbf{R}_J|$  is the distance between the pair, and  $U^{(2)}(r)$  is a central pair potential. Then  $|\nabla \psi_\lambda(\mathbf{R})|^2 = 2$ ,  $\nabla^2 \psi_\lambda(\mathbf{R}) = 4/\psi_\lambda(\mathbf{R})$ ,  $\bar{\mathcal{G}}(r) = 2 \langle \sum_\lambda \delta(\psi_\lambda(\mathbf{R}) - r) \rangle$ , and  $L(r) = 2\bar{\mathcal{G}}(r)/r$ . In this case,  $G_{\zeta\zeta'}(r, r')$  is a spherically averaged three-body correlation function and  $\bar{\mathcal{G}}(r)$  is related to the RDF,  $\mathcal{G}(r) = \exp[-w(r)/k_B T]$ , according to  $\bar{\mathcal{G}}(r) = 2r^2 \mathcal{G}(r)/c$ , where  $c$  is a dimensioned normalization. It also follows that  $b(r)$  is closely related to the pair mean force,  $-w'(r)$ :

$$b(r) = (2r^2/c) k_B T \mathcal{G}'(r) = (2r^2/c) [-w'(r) \mathcal{G}(r)]. \quad (95)$$

After dividing through by the RDF,  $\mathcal{g}(r)$ , the generalized-YBG equation Eq. 92 may be re-expressed:

$$-\mathbf{w}'(\mathbf{r}) = \phi(\mathbf{r}) + \int d\mathbf{r}' \left( \frac{c}{2r^2} \right) \mathcal{g}^{-1}(\mathbf{r}) \bar{G}(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}'), \quad (96)$$

where  $\bar{G}(\mathbf{r}, \mathbf{r}')$  is the explicitly many-body contribution to the metric tensor  $G(\mathbf{r}, \mathbf{r}') = \bar{g}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') + \bar{G}(\mathbf{r}, \mathbf{r}')$ . Equation 96 is an explicit restatement of the conventional YBG theory that states that the pair mean force,  $-\mathbf{w}'(\mathbf{r})$ , which (as discussed earlier) is the average force on a pair of particles separated by a distance  $r$ , can be decomposed into two contributions: (1) the direct force between the pair at a fixed distance  $r$ ,  $\phi(\mathbf{r}) = -dU^{(2)}(\mathbf{r})/dr$ ; and (2) the average correlated force from the environment on a pair of particles separated by  $r$ . The integral term in Eq. 96 decomposes this environment-mediated force into contributions from shells of radius  $r'$  from the fixed particles.

The YBG theory of simple liquids had previously been considered to be an exact direct relation between a pair potential and the equilibrium two- and three-particle correlation functions generated by the pair potential (85, 105). This direct relationship had also been extended for rigid anisotropic molecules (167) and for polymers modeled with central pair potentials and/or rigid bonds (168–172). More recently, Weeks and coworkers have exploited the YBG theory to model long-ranged electrostatics (173–175). However, the theory derived above in Eq. 92 generalizes the YBG theory for arbitrarily complex molecules with flexible internal degrees of freedom modeled with molecular mechanics potentials (156). Moreover, the above discussion demonstrates that the generalized-YBG theory is also a powerful method for determining an optimal approximation of a given form to an unknown and arbitrarily complex potential directly (i.e., noniteratively) from structural correlation functions (156). In the case of CG modeling, the arbitrarily complex potential is the many-body PMF, the structural correlation functions reflect distributions of CG sites, the form of the approximate potential is determined from the force field basis set implicit in Eq. 92, and the optimal approximate potential corresponds to the MS-CG potential. Consequently, the g-YBG theory determines the MS-CG potential directly from structural correlation functions (106, 157), which can be obtained, perhaps, from a protein databank (176).

This analysis further clarifies the MS-CG method as a force-balance relation between the atomistic and CG model. The CG potentials,  $U_\zeta(x)$ , are determined to balance the average atomistic force along each basis vector. Higher-order structural correlations clearly play an important role in redistributing the many-body mean force among the various potentials. However, even if a CG model reproduces a set of lower-order structural correlation functions, in general, the model will not accurately reproduce the

higher-order many-body structural correlation functions of an atomistic model (99). Consequently, the MS-CG potential may not exactly reproduce the atomistic average force or, equivalently, the atomistic distribution for each degree of freedom (106).

There are, in fact, two relevant g-YBG equations when considering CG potentials (49, 106):

1. The set of structural correlation functions generated by the atomistic model determines one g-YBG equation. The MS-CG potential satisfies this (inverse) g-YBG equation and is the projection of the many-body mean force field onto an incomplete basis set.
2. The set of structural correlation functions generated by a given CG potential determines a second g-YBG equation. Even if the CG potential reproduces the lower-order structural correlation functions generated by the atomistic model, the higher-order structural correlations will be different. Consequently, the second g-YBG equation is distinct from the first g-YBG equation. The CG potential will be an exact (direct) solution of this second g-YBG equation.

Chu and coworkers have exploited this logic in developing an iterative-YBG approach that systematically corrects the potential so that, for each relevant degree of freedom, the direct and indirect forces generated by the CG model balance the corresponding mean force generated by the atomistic model (177).

## 4. Outstanding Challenges

The promised efficiency of CG models has attracted tremendous recent interest. Consequently, dramatic progress has been rapidly made in the development of computational algorithms and theoretical understanding for systematic coarse-graining methods that ensure consistency with high-resolution models. Nevertheless, systematic CG simulation methods are still in their infancy in comparison to the well-established atomistic simulation methods that have been intensely studied and optimized for decades. Extensive computational and theoretical study is still required before CG models might attain the accuracy, transferability, and predictive capability of all-atom models. The present section briefly discusses a few of the most significant outstanding challenges facing the field of systematic coarse-graining.

### 4.1. Transferability

One of the most important and well-known challenges facing CG models regards their transferability. This chapter focused on methods that attempt to systematically determine accurate approximations

to the many-body PMF. Equation 16 demonstrates that the PMF depends upon the system composition and volume by defining the configuration space that is integrated over and depends upon the temperature through the Boltzmann weight assigned to each configuration. Consequently, effective potentials that are designed as systematic approximations to the PMF should be expected to demonstrate similar state point dependence.

Considerable effort has been devoted to studying the temperature dependence of CG effective potentials. Several studies have demonstrated that, in some cases, CG potentials parameterized at a single temperature can provide an accurate description of atomistic structure over a wide range of temperatures (116, 147, 178). However, in other cases, CG potentials have demonstrated relatively little temperature transferability (179). Moreover, Ghosh and Faller warned that, even if the target atomistic distribution functions demonstrate little temperature dependence, effective potentials that are parameterized to reproduce the atomistic structure at one temperature may not accurately reproduce the atomistic structure at the second temperature (117). Several studies have also proposed systematic methods for modifying effective potentials parameterized at one temperature for simulations at a second temperature. For instance, Krishna et al. adapted the Ferrenberg-Swendsen reweighting algorithm (180) to use configurations sampled from the canonical ensemble at one temperature to determine the MS-CG potential at a second temperature (94). Alternatively, Müller-Plathe and coworkers employed an empirical analytic function to modify an effective potential for simulating ethylbenzene over a wide temperature range (118).

Somewhat less progress has been achieved in adapting systematic coarse-graining methods to determine effective potentials for modeling multiple molecular systems. Some studies have employed simple mixing rules to determine effective potentials for binary mixtures from potentials parameterized for one component systems (118, 181, 182). Allen and Rutledge have proposed a novel local-density-dependent adaptation of the IMC procedure to reproduce both the molecular structure and also the chemical potential of atomistic models for several different systems (183, 184). Mullinax and Noid generalized the MS-CG variational principle for “extended ensembles,” which are defined by combining equilibrium ensembles for multiple distinct systems. They then demonstrated that this approach systematically improved the transferability of potentials for modeling a variety of liquid mixtures over a wide range of compositions (162). Several other studies have attempted to minimize the “environment-mediated” contribution to effective pair potentials in order to obtain a more transferable direct potential. Wang, Voth, and coworkers have demonstrated that such potentials can provide surprisingly good transferability for ionic liquids (178). Villa, Peter, and van der Vegt have employed a

similar strategy to determine effective site-site pair potentials by subtracting the environment-mediated contribution from the corresponding ppmf (185, 186). Recently, they have applied this approach to determine implicit solvent benzene models for a range of benzene-water mixtures (187). They demonstrated that the resulting potentials correctly incorporate hydrophobic interactions over a range of concentrations and that, moreover, these potentials perform significantly better than potentials parameterized for neat benzene.

A third aspect of transferability addresses the utility of effective potentials in different ensembles. This chapter has been entirely developed within the canonical ensemble. However, effective potentials are frequently employed in simulations sampling, e.g., the isothermal-isobaric ensemble. Most commonly, CG simulations at constant NPT are performed by coupling the system to a barostat using the standard equations for the virial (85). The resulting pressure of the CG model is then typically much too large (176). This is not entirely surprising and is consistent with Weeks-Chandler-Andersen (WCA) theory, which implies that the local structure reported in, e.g., RDFs, primarily reflects the short-ranged, rapidly varying repulsive contributions to the potential and is relatively insensitive to the slowly varying long-ranged interactions that play an important role in stabilizing condensed phases (188). Consequently, CG potentials are frequently modified so that, when using standard (atomistic) expressions for the virial, the potential generates the correct average virial. This modification typically takes one of two forms:

1. CG potentials that are parameterized via iterative structure-motivated methods (e.g., IMC, IBI) to reproduce structural distribution functions are often modified by the addition of a linear ramp potential:  $U_{ramp}(x) = \alpha(x - x_0)$  (107, 123, 179). This modification introduces an additional term in the virial. The slope of the ramp potential can then be varied so that the CG model generates the correct average virial. Because the ramp potential varies slowly, it has relatively little (negative) impact upon the local structure of the CG model (as reported by, e.g., RDFs), in accord with WCA theory.
2. CG potentials can also be adapted for NPT simulations by employing an additional virial constraint in the potential calculation (128, 141, 142). This constraint then modifies the variational framework to determine the potential that best approximates the many-body PMF, subject to the constraint that the model generates the correct average virial when averaged over the configurations sampled from atomistic MD simulations.

Although, both approaches have enjoyed some success, neither approach rigorously addresses the statistical thermodynamics of

coarse-graining in the NPT ensemble. Recently Das and Andersen performed such an analysis (189). A key insight in this work was the formulation of a consistency criterion not only for sampling configurations, but also for sampling volume fluctuations in the NPT ensemble. They demonstrated the intuitive result that, in the thermodynamic limit, the optimal effective potential should be the same in the NVT and NPT ensembles. However, an additional volume-dependent contribution should be considered in the extended Lagrangian of the CG model. In some sense, this addresses the fact that the CG PMF is a function of volume and, consequently, the atomistic expression for the virial is no longer appropriate. This volume dependence is manifest not only in the configuration-dependent contribution of the PMF (i.e.,  $-k_B T \ln z(\mathbf{R})$  in Eq. 15) but also in the configuration-independent contribution to the PMF (i.e., *const* in the same equation). This analysis merits further consideration in future work.

A related issue is the issue of representability, i.e., the issue of which properties can possibly be accurately described with a CG model and an approximate potential. As detailed in this chapter, a variety of different avenues can be followed in deriving CG potentials, and no simple approximate potential can be expected to accurately reproduce every detail (e.g., thermodynamics, structure, and kinetics) of an atomistic model. For instance, it is readily shown that the (unique) pair potential reproducing a given RDF will not reproduce thermodynamic properties such as the energy (135). However, a more subtle consideration is the conclusion by Louis (190) that “an effective density-dependent pair potential  $\nu(r; \rho)$  cannot be properly interpreted without reference to the coarse-graining procedure by which it was derived.” This is clearly illustrated by the analysis of Das and Andersen, showing that volume-dependent contributions to the PMF should be addressed when attempting to reproduce the thermodynamic pressure with a CG model (189). Along similar lines, Voth and coworkers have recently suggested a mechanism for addressing the significance of missing atomic degrees of freedom when considering CG observables (160). Consequently, a major advance would be achieved if one could *a priori* determine which properties could possibly be reproduced by a CG model of a given form.

#### **4.2. Mappings and Basis Sets**

CG mappings are most frequently determined by physico-chemical intuition for a particular system. CG sites are typically assigned to provide adequate resolution for describing functional groups that are believed to be important in a particular process. This approach is often particularly useful in the case that the mapping is relatively “fine” and each site corresponds to a relatively small number of atoms. In contrast, this approach is less straightforward when the CG mapping is more “coarse” and each site corresponds to, e.g., dozens or hundreds of atoms. A number of studies have considered the design of CG mappings in this latter case. For instance, very

low-resolution mappings have been designed for proteins and biomolecules based upon consideration of rigidity analysis (191), preserving dynamic domains (192, 193), or retaining spatial mass distributions (194, 195).

Nevertheless, from the perspective of the author, the CG mapping remains quite underappreciated and poorly understood. The theory outlined above demonstrates that, subject to relatively few assumptions, for any atomistic model and associated mapping, there does exist a many-body PMF and a CG model with this potential will reproduce all of the structural properties of the underlying atomistic model. However, the PMF can never be attained and, instead, must be approximated with an incomplete basis set. The accuracy and transferability of the resulting potential then depend upon a subtle relationship between the mapping and the basis set. For a given mapping, the accuracy and transferability of the potential are fundamentally limited by the flexibility of the basis set. The accuracy and transferability of the potential can only be improved by expanding the basis set. Typically, molecular mechanics potentials are employed for their convenience and efficiency in CG simulations. It is not clear what will be the most useful potentials for expanding the basis set. Several recent studies have considered Stillinger-Weber (196) type potentials (77, 160), generalizations of Gay-Berne potentials (76, 197), or other non-central potentials (78, 198, 199). However, much work remains to be done in identifying a better basis that allow accurate and transferable potentials to be efficiently simulated. Conversely, for a given basis set, the accuracy and transferability of the resulting model strongly depends upon the CG mapping. For instance, Rühle et al. demonstrated that a simple molecular mechanics basis cannot accurately capture atomistic forces that appear as noncentral forces in the CG model (120). Moreover, they demonstrated that these errors led to a poor description of the conformers sampled by a CG hexane model. In addition, Mullinax and Noid recently demonstrated the somewhat counterintuitive result that the accuracy and transferability of CG potentials can actually decrease when increasing the resolution of the CG mapping (162).

#### **4.3. Predictive Capability**

Finally, a third major challenge addresses the capability to predict the accuracy and transferability of CG models. Very frequently, CG potentials are only validated when their results are compared with results from high-resolution models, most often for the system that was used in parameterizing the potential. Relatively little progress has been achieved in attempting to predict the accuracy of a CG model for a particular system. Recently, Chaimovich and Shell suggested that the relative entropy may allow predictions of the accuracy of a CG model, while using information extracted from the atomistic model, i.e., before actually performing CG

simulations (200). Alternatively, it may be possible to formulate self-consistency criterion when determining MS-CG potentials. The development of a priori estimates of the accuracy and transferability of CG models will represent a major advance.

---

## 5. Conclusions

In recent years, tremendous progress has been achieved in the development of new theories and computational methods for systematic coarse-graining. This chapter reviews a formal framework for considering these advances and then attempts to provide a unified discussion of both structure-motivated and also force-motivated approaches. It is shown that a generalization of the Yvon-Born-Green theory provides a conceptual bridge between the two approaches. This chapter also discusses some of the most outstanding challenges facing the field and reviews some of the progress made to date in addressing these challenges. It is anticipated that these challenges may well be overcome in the near future, and it is hoped that the unified framework presented in this chapter may play some small role in stimulating these advances.

---

## Acknowledgements

The author gratefully acknowledges the financial support of an NSF CAREER award (NSF Grant No. MCB 1053970) and also start-up funding from the Pennsylvania State University. The author gratefully acknowledges Dr. Navrotskaya for her contributions in developing the present unified approach to systematic coarse-graining and J.W.Mullinax for his contributions in developing the g-YBG theory. WGN also acknowledges the insight and instruction of G. A. Voth, H.C. Andersen, and G. Ayton in many helpful and formative discussions. Finally, WGN gratefully acknowledges Vinod Krishna, Qiang Du, M. Scott Shell, Garegin Papoian, Ard A. Louis, John Weeks, and Matej Praprotnik for many helpful and insightful conversations on the subjects of this work.

## References

1. Allen MP, Tildesley DP (1987) Computer simulation of liquids. Oxford Press, New York
2. Frenkel D, Smit B (2002) Understanding molecular simulation: from algorithms to applications, 2nd edn. Academic Press, New York
3. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J Am Chem Soc* 117:5179–5197

4. MacKerell AD et al (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
5. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
6. Mervis J (2001) Advanced computing—nsf launches teragrid for academic research. *Science* 293:1235–1236
7. Reed DA (2003) Grids, the teragrid, and beyond. *Computer* 36:62–68
8. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J Comp Chem* 4:187–217
9. Plimpton S (1995) Fast parallel algorithms for short-range molecular-dynamics. *J Comp Phys* 117:1–19
10. Case DA, Cheatham I, TE, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods R (2005) The Amber biomolecular simulation programs. *J Comp Chem* 26: 1668–1688
11. Christen M et al (2005) The gromos software for biomolecular simulation: Gromos05. *J Comp Chem* 26:1719–1751
12. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theor Comp* 4:435–447
13. Duan Y, Kollman PA (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 282:740–743
14. Gumbart J, Trabuco LG, Schreiner E, Villa E, Schulten K (2009) Regulation of the protein-conducting channel by a bound ribosome. *Structure* 17:1453–1464
15. Anderson PW (1972) More is different. *Science* 177:393–396
16. Boon J-P, Yip S (1992) Molecular hydrodynamics. Dover Publications, New York
17. Callen HB (1985) Thermodynamics and an introduction to thermostatics. Wiley, New York
18. Clementi C (2008) Coarse-grained models of protein folding: toy models or predictive tools? *Curr Opin Struc Biol* 18:10–15
19. Kirkwood JG (1935) Statistical mechanics of fluid mixtures. *J Chem Phys* 3:300–313
20. Kirkwood JG (1946) The statistical mechanical theory of transport processes. I. general theory. *J Chem Phys* 14:180–201
21. Irving JH, Kirkwood JG (1950) The statistical mechanical theory of transport processes. IV. the equations of hydrodynamics. *J Chem Phys* 18:817–829
22. Silbermann J, Klapp S, HL, Shoen M, Channamsetty N, Block H, Gubbins KE (2006) Mesoscale modeling of complex binary fluid mixtures: towards an atomistic foundation of effective potentials. *J Chem Phys* 124:074105
23. Levitt M, Warshel A (1975) Computer simulation of protein folding. *Nature* 253:694–698
24. Marrink SJ, de Vries AH, Mark AE (2004) Coarse grained model for semiquantitative lipid simulations. *J Phys Chem B* 108:750–760
25. ten Wolde PR, Chandler D (2002) Drying-induced hydrophobic polymer collapse. *Proc Natl Acad Sci USA* 99:6539–6543
26. Everaers R, Sukumaran SK, Grest GS, Svaneborg C, Sivasubramanian A, Kremer K (2004) Rheology and microscopic topology of entangled polymeric liquids. *Science* 303:823–826
27. Reynwar BJ, Illya G, Harmandaris VA, Muller MM, Kremer K, Deserno M (2007) Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature* 447:461–464
28. Thirumalai D, Klimov DK (1999) Deciphering the timescales and mechanisms of protein folding using minimal off-lattice models. *Curr Opin Struc Biol* 9:197–207
29. Hagler AT, Honig B (1978) On the formation of protein tertiary structure on a computer. *Proc Natl Acad Sci USA* 75:554–558
30. Shelley JC, Shelley MY, Reeder RC, Bandyopadhyay S, Klein ML (2001) A coarse grain model for phospholipid simulation. *J Phys Chem B* 105:4464–4470
31. Nielsen SO, Lopez CF, Srinivas G, Klein ML (2003) A coarse grain model for *n*-alkanes parameterized from surface tension data. *J Chem Phys* 119:7043–7049
32. Shinoda W, Devane R, Klein ML (2007) Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. *Mol Sim* 33:27–36
33. Khurana E, DeVane RH, Kohlmeyer A, Klein ML (2008) Probing peptide nanotube self-assembly at a liquid-liquid interface with coarse-grained molecular dynamics. *Nanoletters* 8:3626–3630

34. DeVane R, Shinoda W, Moore PB, Klein ML (2009) Transferable coarse grain nonbonded interaction model for amino acids. *J Chem Theor Comp* 5:2115–2124
35. DeVane R, Klein ML, Chiu CC, Nielsen SO, Shinoda W, Moore PB (2010) Coarse-grained potential models for phenyl-based molecules: I. parametrization using experimental data. *J Phys Chem B* 114:6386–6393
36. Chiu CC, DeVane R, Klein ML, Shinoda W, Moore PB, Nielsen SO (2010) Coarse-grained potential models for phenyl-based molecules: II. application to fullerenes. *J Phys Chem B* 114:6394–6400
37. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 111:7812–7824
38. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink SJ (2008) The MARTINI coarse-grained force field: extension to proteins. *J Chem Theor Comp* 4:819–834
39. Lopez CA, Rzepiela AJ, de Vries AH, Dijkhuizen L, Hunenberger PH, Marrink SJ (2009) Martini coarse-grained force field: extension to carbohydrates. *J Chem Theor Comp* 5:3195–3210
40. Periole X, Cavalli M, Marrink SJ, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theor Comp* 5:2531–2543
41. Lee H, de Vries AH, Marrink SJ, Pastor RW (2009) A coarse-grained model for polyethylene oxide and polyethylene glycol: conformation and hydrodynamics. *J Phys Chem B* 113:13186–13194
42. Partial special issue of *J Chem Theor Comp* 2 (3) May (2006)
43. (2009) Themed issue on coarse-grained modeling soft condensed matter. *Phys Chem Chem Phys* 11(12)
44. (2009) Soft matter themed issue on modeling soft matter. 2(22)
45. (2010) Faraday discussions on multiscale simulation of soft matter systems. 144
46. Voth GA (ed) (2009) Coarse-graining of condensed phase and biomolecular systems. CRC Press, USA
47. Ayton GS, Noid WG, Voth GA (2007) Multiscale modeling of biomolecular systems: in serial and in parallel. *Curr Opin Struc Biol* 17:192–8
48. Peter C, Kremer K (2009) Multiscale simulation of soft matter systems—from the atomistic to the coarse-grained level and back. *Soft Matter* 5:4357–4366
49. Murtola T, Bunker A, Vattulainen I, Deserno M, Karttunen M (2009) Multiscale modeling of emergent materials: biological and soft matter. *Phys Chem Chem Phys* 11:1869–1892
50. Kroger M (2004) Simple models for complex nonequilibrium fluids. *Phys Rep* 390:453–551
51. Engquist EWB, Li XT, Ren WQ, Vandeneijnden E (2007) Heterogeneous multiscale methods: a review. *Comm Comp Phys* 2:367–450
52. Head-Gordon T, Brown S (2003) Minimalist models for protein folding and design. *Curr Opin Struc Biol* 13:160–167
53. Bahar I, Rader A (2005) Coarse-grained normal mode analysis in structural biology. *Curr Opin Struc Biol* 15:586–592
54. Tozzini V (2005) Coarse-grained models for proteins. *Curr Opin Struc Biol* 15:144–150
55. Bond PJ, Holyoake J, Ivetac A, Khalid S, Sansom MSP (2007) Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *J Struct Biol* 157:593–605
56. Liwo A, Czaplewski C, Oldziej S, Scheraga HA (2008) Computational techniques for efficient conformational sampling of proteins. *Curr Opin Struc Biol* 18:134–139
57. Hills RD, Brooks CL (2009) Insights from coarse-grained go models for protein folding and dynamics. *Int J Mol Sci* 10:889–905
58. Kuznetsov IB (2009) Simplified computational methods for the analysis of protein flexibility. *Curr Prot Pept Sci* 10:607–613
59. Eom K, Yoon G, Kim JI, Na S (2010) Coarse-grained elastic models of protein structures for understanding their mechanics and dynamics. *J Comp Theor Nanosci* 7:1210–1226
60. Tozzini V (2010) Multiscale modeling of proteins. *Acc Chem Res* 43:220–230
61. Bahar I, Lezon TR, Yang LW, Eyal E (2010) Global dynamics of proteins: bridging between structure and function. *Ann Rev Biochem* 39:23–42
62. Scott HL (2002) Modeling the lipid component of membranes. *Curr Opin Struc Biol* 12:495–502
63. Brannigan G, Lin L. C-L, Brown F. LH (2006) Implicit solvent simulation models for biomembranes. *Eur Biophys J* 35:104–124
64. Muller M, Katsov K, Schick M (2006) Biological and synthetic membranes: what

- can be learned from a coarse-grained description? *Phys Rep* 434:113–176
65. Venturoli M, Sperotto MM, Kranenburg M, Smit B (2006) Mesoscopic models of biological membranes. *Phys Rep* 437:1–54
  66. Brown F, LH (2008) Elastic modeling of biomembranes and lipid bilayers. *Ann Rev Phys Chem* 59:685–712
  67. Bennun SV, Hoopes MI, Xing CY, Faller R (2009) Coarse-grained modeling of lipids. *Chem Phys Lipids* 159:59–66
  68. Noguchi H (2009) Membrane simulation models from nanometer to micrometer scale. *J Phys Soc Jpn* 78:041007
  69. Pandit SA, Scott HL (2009) Multiscale simulations of heterogeneous model membranes. *Biochim Biophys Acta-Biomembranes* 1788:136–148
  70. Deserno M (2009) Mesoscopic membrane physics: concepts, simulations, and selected applications. *Macromol Rapid Comm* 30:752–771
  71. Glotzer SC, Paul W (2002) Molecular and mesoscale simulation methods for polymer materials. *Ann Rev Mat Res* 32:401–436
  72. Paul W, Smith GD (2004) Structure and dynamics of amorphous polymers: computer simulations compared to experiment and theory. *Rep Prog Phys* 67:1117–1185
  73. Guenza MG (2008) Theoretical models for bridging timescales in polymer dynamics. *J Phys: Condens Matter* 20:033101–0331024
  74. Ganesan V, Ellison CJ, Pryamitsyn V (2010) Mean-field models of structure and dispersion of polymer-nanoparticle mixtures. *Soft Matter* 6:4010–4025
  75. Buchete NV, Straub JE, Thirumalai D (2004) Development of novel statistical potentials for protein fold recognition. *Curr Opin Struc Biol* 14:225–32
  76. Babadi M, Everaers R, Ejtehadi MR (2006) Coarse-grained interaction potentials for anisotropic molecules. *J Chem Phys* 124:174708
  77. Molinero V, Moore EB (2009) Water modeled as an intermediate element between carbon and silicon. *J Phys Chem B* 113:4008–4016
  78. Alemani D, Collu F, Cascella M, Dal Peraro M (2010) A nonradial coarse-grained potential for proteins produces naturally stable secondary structure elements. *J Chem Theor Comp* 6:315–324
  79. Murtola T, Karttunen M, Vattulainen I (2009) Systematic coarse graining from structure using internal states: application to phospholipid/cholesterol bilayer. *J Chem Phys* 131:055101
  80. Noid WG, Chu J-W, Ayton GS, Krishna V, Izvekov S, Voth GA, Das A, Andersen HC (2008) The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J Chem Phys* 128:244114
  81. Hill TL (1987) An introduction to statistical thermodynamics. Dover reprint, New York
  82. Shi Z, Chen K, Liu Z, Kallenbach NR (2006) Conformation of the backbone in unfolded proteins. *Chem Rev* 106:1877–1897
  83. Praprotnik M, Delle Site L, Kremer K (2008) Multiscale simulation of soft matter: from scale bridging to adaptive resolution. *Ann Rev Phys Chem* 59:545–571
  84. Ensing B, Nielsen SO, Moore PB, Klein ML, Parrinello M (2007) Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics. *J Chem Theor Comp* 3:1100–1105
  85. Hansen J-P, McDonald IR (1990) Theory of simple liquids, 2nd edn. Academic Press, London
  86. Liwo A, Oldziej S, Pincus MR, Wawak RJ, Rackovsky S, Scheraga HA (1997) A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J Comp Chem* 18:849–873
  87. van Kampen NG (2007) Stochastic processes in Physics and Chemistry, 3rd edn. North-Holland
  88. Zwanzig R (1960) Ensemble method in the theory of irreversibility. *J Chem Phys* 33:1338–1341
  89. Mori H (1965) Transport collective motion and brownian motion. *Prog Theor Phys* 33:423–455
  90. Baschnagel J et al (2000) Bridging the gap between atomistic and coarse-grained models of polymers: status and perspectives. *Adv Poly Sci* 152:41–156
  91. Shell MS (2008) The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *J Chem Phys* 129:144108
  92. Bender CM, Orszag SA (1999) Advanced mathematical methods for scientists and engineers: asymptotic methods and perturbation theory. Springer, Berlin Heidelberg
  93. Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Stat* 22:79–86
  94. Krishna V, Noid WG, Voth GA (2009) The multiscale coarse-graining method. iv.

- transferring coarse-grained potentials between temperatures. *J Chem Phys* 131:024103
95. Chaimovich A, Shell MS (2011) Coarse-graining errors and numerical optimization using a relative entropy framework. *J Chem Phys* 134:094112
  96. Rudzinski JF, Noid WG (2011) Coarse-graining entropy, forces, and structures. *J Chem Phys* 135:214101
  97. Larini L, Krishna V (2011) A generalized mean field theory of coarse-graining. *J Chem Phys* 135:124103
  98. Evans R (1979) Nature of the liquid-vapor interface and other topics in the statistical-mechanics of nonuniform, classical fluids. *Adv Phys* 28:143–200
  99. Evans R (1990) Comment on reverse monte carlo simulation. *Mol. Sim.* 4:409–411
  100. Tschop W, Kremer K, Batoulis J, Burger T, Hahn O (1998) Simulation of polymer melts. i. coarse-graining procedure for polycarbonates. *Acta Poly* 49:61–74
  101. Praprotnik M, Junghans C, Delle Site L, Kremer K (2008) Simulation approaches to soft matter: generic statistical properties vs. chemical details. *Comp Phys Comm* 179:51–60
  102. Peter C, Kremer K (2009) Multiscale simulation of soft matter systems—from the atomistic to the coarse-grained level and back. *Soft Matter* 5:4357–4366
  103. Harmandaris VA, Adhikari NP, van der Vegt N. FA, Kremer K (2006) Hierarchical modeling of polystyrene: from atomistic to coarse-grained simulations. *Macromolecules* 39:6708–6719
  104. Peter C, Delle Site L, Kremer K (2008) Classical simulations from the atomistic to the mesoscale and back: coarse graining an azobenzene liquid crystal. *Soft Matter* 4:859–869
  105. Hill TL (1987) Statistical mechanics: principles and selected applications. Dover reprint, New York
  106. Noid WG, Chu J-W, Ayton GS, Voth GA (2007) Multiscale coarse-graining and structural correlations: connections to liquid state theory. *J Phys Chem B* 111:4116–27
  107. Müller-Plathe F (2002) Coarse-graining in polymer simulation: from the atomistic to the mesoscopic scale and back. *Chem Phys Chem* 3:754–769
  108. Schommers W (1983) Pair potentials in disordered many-particle systems: a study of liquid gallium. *Phys Rev A* 28:3599–3605
  109. Soper AK (1996) Empirical potential monte carlo simulation of fluid structure. *Chem Phys* 202:295–306
  110. Reatto L, Levesque D, Weis JJ (1986) Iterative predictor-corrector method for extraction of the pair interaction from structural data for dense classical liquids. *Phys Rev A* 33:3451–3465
  111. McGreevy RL, Pusztai L (1988) Reverse Monte Carlo simulation: a new technique for the determination of disordered structures. *Mol Sim* 1:359–367
  112. Keen DA, McGreevy RL (1990) Structural modelling of glasses using reverse Monte Carlo simulation. *Nature* 344:423–425
  113. Reith D, Putz M, Müller-Plathe F (2003) Deriving effective mesoscale potentials from atomistic simulations. *J Comp Chem* 24:1624–1636
  114. Milano G, Müller-Plathe F (2005) Mapping atomistic simulations to mesoscopic models: a systematic coarse-graining procedure for vinyl polymer chains. *J Phys Chem B* 109:18609–18619
  115. Sun Q, Faller R (2006) Crossover from unentangled to entangled dynamics in a systematically coarse-grained polystyrene melt. *Macromolecules* 39:812–820
  116. Vettorel T, Meyer H (2006) Coarse graining of short polyethylene chains for studying polymer crystallization. *J Chem Theor Comp* 2:616–629
  117. Ghosh J, Faller R (2007) State point dependence of systematically coarse-grained potentials. *Mol Sim* 33:759–767
  118. Qian H-J, Carbone P, Chen X, Karimi-Varzaneh HA, Liew CC, Müller-Plathe F (2008) Temperature-transferable coarse-grained potentials for ethylbenzene, polystyrene and their mixtures. *Macromolecules* 41:9919–9929
  119. Fischer J, Paschek D, Gieger A, Sadowski G (2008) Modeling of aqueous poly(oxyethylene) solutions. 2. Mesoscale simulations. *J Phys Chem B* 112:13561–13571
  120. Ruhle V, Junghans C, Lukyanov A, Kremer K, Andrienko D (2009) Versatile object-oriented toolkit for coarse-graining applications. *J Chem Theor Comp* 5:3211–3223
  121. Wang H, Junghans C, Kremer K (2009) Comparative atomistic and coarse-grained study of water: what do we lose by coarse-graining? *Eur Phys J E* 28:221–229
  122. Megariotis G, Vyrkou A, Leygue A, Theodorou DN (2011) Systematic coarse graining of 4-cyano-4'-pentylbiphenyl. *Ind Eng Chem Res ASAP* 50:546
  123. Reith D, Pütz M, Müller-Plathe F (2003) Deriving effective mesoscale potentials from atomistic simulations. *J Comp Chem* 24:1624–1636

124. Lyubartsev AP, Laaksonen A (1995) Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach. *Phys Rev E* 52:3730–3737
125. Lyubartsev A, Mirzoev A, Chen LJ, Laaksonen A (2010) Systematic coarse-graining of molecular models by the newton inversion method. *Faraday Disc* 144:43–56
126. Murtola T, Falck E, Patra M, Karttunen M, Vattulainen I (2004) Coarse-grained model for phospholipid/cholesterol bilayer. *J Chem Phys* 121:9156–9165
127. Lyubartsev AP (2005) Multiscale modeling of lipids and lipid bilayers. *Eur Biophys J* 35:53–61
128. Murtola T, Falck E, Karttunen M, Vattulainen I (2007) Coarse-grained model for phospholipid/cholesterol bilayer employing inverse Monte Carlo with thermodynamic constraints. *J Chem Phys* 121:075101
129. Lyubartsev AP, Laaksonen A (1997) Osmotic and activity coefficients from effective potentials for hydrated ions. *Phys Rev E* 55:5689–5696
130. Lyubartsev AP, Karttunen M, Vattulainen I, Laaksonen A (2003) On coarse-graining by the inverse Monte Carlo method: dissipative particle dynamics simulations made to a precise tool in soft matter modeling. *Soft Matter* 1:121–137
131. Ishihara A (1968) Gibbs-bogoliubov inequality. *J Phys A: Math Nucl Gen* 1:539–548
132. Henderson RL (1974) A uniqueness theorem for fluid pair correlation functions. *Phys Lett A* 49:197–198
133. Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. *Phys Rev B* 136:864–871
134. Mermin ND (1965) Thermal properties of inhomogeneous electron gas. *Phys Rev A* 137:1441–1443
135. Johnson ME, Head-Gordon T, Louis AA (2007) Representability problems for coarse-grained water potentials. *J Chem Phys* 126:144509
136. Chayes JT, Chayes L, Lieb EH (1984) The inverse problem in classical statistical mechanics. *Comm Math Phys* 93:57–121
137. Chayes JT, Chayes L (1984) On the validity of the inverse conjecture in classical density functional theory. *J Stat Phys* 36:471–488
138. Savelyev A, Papoian GA (2009) Molecular renormalization group coarse-graining of electrolyte solutions: applications to aqueous NaCl and KCl. *J Phys Chem B* 113:7785–7793
139. Savelyev A, Papoian GA (2009) Molecular renormalization group coarse-graining of polymer chains: applications to double-stranded DNA. *Biophys J* 96:4044–4052
140. Izvekov S, Voth GA (2005) A multiscale coarse-graining method for biomolecular systems. *J Phys Chem B* 109:2469–2473
141. Izvekov S, Voth GA (2005) Multiscale coarse graining of liquid-state systems. *J Chem Phys* 123:134105
142. Izvekov S, Chung PW, Rice BM (2010) The multiscale coarse-graining method: assessing its accuracy and introducing density dependent coarse-grain potentials. *J Chem Phys* 133:064109
143. Ercolessi F, Adams JB (1994) Interatomic potentials from first-principles calculations: the force-matching method. *Europ Phys Lett* 26:583
144. Izvekov S, Parrinello M, Burnham CJ, Voth, GA (2004) Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: a new method for force-matching. *J Chem Phys* 120:10896–10913
145. Wang YT, Voth GA (2005) Unique spatial heterogeneity in ionic liquids. *J Am Chem Soc* 127:12192–12193
146. Wang YT, Izvekov S, Yan TY, Voth GA (2006) Multiscale coarse-graining of ionic liquids. *J Phys Chem B* 110:3564–3575
147. Wang YT, Feng SL, Voth GA (2009) Transferable coarse-grained models for ionic liquids. *J Chem Theor Comp* 5:1091–1098
148. Izvekov S, Voth GA (2006) Multiscale coarse-graining of mixed phospholipid/cholesterol bilayers. *J Chem Theor Comp* 2:637–648
149. Izvekov S, Voth GA (2009) Solvent-free lipid bilayer model using multiscale coarse graining. *J Phys Chem B* 113:4443–4455
150. Izvekov S, Violi A, Voth GA (2005) Systematic coarse-graining of nanoparticle interactions in molecular dynamics simulation. *J Phys Chem B* 109:17019–17024
151. Liu P, Izvekov S, Voth GA (2007) Multiscale coarse-graining of monosaccharides. *J Phys Chem B* 111:11566–11575
152. Wang YT, Voth GA (2010) Molecular dynamics simulations of polyglutamine aggregation using solvent-free multiscale coarse-grained models. *J Phys Chem B* 114:8735–8743
153. Hills RD, Lu LY, Voth GA (2010) Multiscale coarse-graining of the protein energy landscape. *PLoS Comput Biol* 6:e1000827
154. Chorin AJ, Hald OH, Kupferman R (2000) Optimal prediction and the Mori-Zwanzig representation of irreversible processes. *Proc Natl Acad Sci USA* 97:2968–2973
155. Noid WG, Liu P, Wang YT, Chu, J-W, Ayton, GS, Izvekov S, Andersen HC, Voth GA

- (2008) The multiscale coarse-graining method. II. Numerical implementation for molecular coarse-grained models. *J Chem Phys* 128:244115
156. Mullinax JW, Noid WG (2009) A generalized Yvon-Born-Green theory for molecular systems. *Phys Rev Lett* 103:198104
157. Mullinax JW, Noid WG (2010) A generalized Yvon-Born-Green theory for determining coarse-grained interaction potentials. *J Phys Chem C* 114:5661–5674
158. Chorin AJ (2003) Conditional expectations and renormalization. *Multiscale Model Simul* 1:105–118
159. Chorin AJ, Hald OH (2006) *Stochastic Tools in Mathematics and Science*. Springer, New York
160. Larini L, Lu LY, Voth GA (2010) The multiscale coarse-graining method. vi. implementation of three-body coarse-grained potentials. *J Chem Phys* 132:164107
161. Lu L, Voth GA (2009) Systematic coarse-graining of a multicomponent lipid bilayer. *J Phys Chem B* 113:1501–1510
162. Mullinax JW, Noid WG (2009) Extended ensemble approach for deriving transferable coarse-grained potentials. *J Chem Phys* 131:104110
163. Lu LY, Izvekov S, Das A, Andersen HC, Voth GA (2010) Efficient, regularized, and scalable algorithms for multiscale coarse-graining. *J Chem Theor Comp* 6:954–965
164. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) *Numerical recipes in FORTRAN: the art of scientific computing*. Cambridge University Press, New York
165. Demmel JW (1997) *Applied Numerical Linear Algebra*. SIAM, Philadelphia
166. Mullinax JW, Noid WG (2010) Reference state for the generalized Yvon-Born-Green theory: application for a coarse-grained model of hydrophobic hydration. *J Chem Phys* 133:124107
167. Gubbins KE (1980) Structure of nonuniform molecular fluids—integrodifferential equations for the density-orientation profile. *Chem Phys Lett* 76:329–332
168. Whittington SG, Dunfield LG (1973) A Born-Green-Yvon treatment of polymers with excluded volume. *J Phys A: Math Nucl Gen* 6:484
169. Attard P (1995) Polymer Born-Green-Yvon equation with proper triplet superposition approximation results for hard-sphere chains. *J Chem Phys* 102:5411
170. Taylor MP, Lipson J. EG (1993) A site-site Born-Green-Yvon equation for hard sphere dimers. *J Chem Phys* 100:518
171. Taylor MP, Lipson J. EG (1995) A Born-Green-Yvon equation for flexible chain-molecule fluids. I. General formalism and numerical results for short hard-sphere chains. *J Chem Phys* 102:2118
172. Taylor MP, Lipson J. EG (1995) A Born, Green,Yvon equation for flexible chain-molecule fluids. I. Applications to hard-sphere polymers. *J Chem Phys* 102:6272
173. Katsov K, Weeks JD (2001) On the mean field treatment of attractive interactions in nonuniform simple fluids. *J Phys Chem B* 105:6738–6744
174. Chen YG, Kaur C, Weeks JD (2004) Connecting systems with short and long ranged interactions: local molecular field theory for ionic fluids. *J Phys Chem B* 108:19874–19884
175. Rodgers JM, Kaur C, Chen YG, Weeks JD (2006) Attraction between like-charged walls: short-ranged simulations using local molecular field theory. *Phys Rev Lett* 97:097801
176. Mullinax JW, Noid WG (2010) Recovering physical potentials from a model protein database. *Proc Natl Acad Sci USA* 107:19867
177. Cho HM, Chu JW (2009) Inversion of radial distribution functions to pair forces by solving the yvon-born-green equation iteratively. *J Chem Phys* 131:134107
178. Wang YT, Noid WG, Liu P, Voth GA (2009) Effective force coarse-graining. *Phys Chem Chem Phys* 11:2002–2015
179. Carbone P, Varzaneh H. AK, Chen X, Müller-Plathe F (2008) Transferability of coarse-grained force fields: the polymer case. *J Chem Phys* 128:064904
180. Ferrenberg AM, Swendsen RH (1988) New monte carlo technique for studying phase transitions. *Phys Rev Lett* 61:2635–2638
181. Mognetti BM, Oettel M, Yelash L, Virnau P, Paul W, Binder, K (2008) Spherically averaged versus angle-dependent interactions in quadrupolar fluids. *Phys Rev E* 77:041506
182. Mognetti BM, Virnau P, Yelash L, Paul W, Binder K, Müller M, MacDowell LG (2009) Coarse-grained models for fluids and their mixtures: comparison of Monte Carlo studies of their phase behavior with perturbation theory and experiment. *J Chem Phys* 130:044101
183. Allen EC, Rutledge GC (2008) A novel algorithm for creating coarse-grained, density dependent implicit solvent models. *J Chem Phys* 128:154115
184. Allen EC, Rutledge GC (2009) Evaluating the transferability of coarse-grained, density-dependent implicit solvent models to mixtures and chains. *J Chem Phys* 130:034904

185. Villa A, van der Vegt N. FA, Peter C (2009) Self-assembling dipeptides: including solvent degrees of freedom in a coarse-grained model. *Phys Chem Chem Phys* 11:2068–2076
186. Villa A, Peter C, van der Vegt N. FA (2009) Self-assembling dipeptides: conformational sampling in solvent-free coarse-grained simulation. *Phys Chem Chem Phys* 11:2077–2086
187. Villa A, Peter C, van der Vegt N. FA (2010) Transferability of nonbonded interaction potentials for coarse-grained simulations: benzene in water. *J Chem Theor Comp* 6:2434–2444
188. Andersen HC, Chandler D, Weeks JD (1976) Roles of repulsive and attractive forces in liquids : the equilibrium theory of classical fluids. *Adv Chem Phys* 34:105
189. Das A, Andersen, HC (2010) The multiscale coarse-graining method. V. Isothermal-isobaric ensemble. *J Chem Phys* 132:164106
190. Louis AA (2002) Beware of density dependent pair potentials. *J Phys: Condens Matter* 14:9187–9206
191. Gohlke H, Thorpe MF (2006) A natural coarse graining for simulating large biomolecular motion. *Biophys J* 91:2115–2120
192. Zhang ZY, Lu LY, Noid WG, Krishna V, Pfaendtner J, Voth GA (2008) A systematic methodology for defining coarse-grained sites in large biomolecules. *Biophys J* 95:5073–5083
193. Zhang ZY, Pfaendtner J, Grafmuller A, Voth GA (2009) Defining coarse-grained representations of large biomolecules and biomolecular complexes from elastic network models. *Biophys J* 97:2327–2337
194. Arkhipov A, Freddolino P, Schulten, K (2006) Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure* 129:1767–1777
195. Zhang Z, Voth GA (2010) Coarse-grained representations of large biomolecular complexes from low-resolution structural data. *J Chem Theor Comp* 6:2990–3002
196. Stillinger FH, Weber TA (1985) Computer-simulation of local order in condensed phases of silicon. *Phys Rev B* 31:5262–5271
197. Morriss-Andrews A, Rottler J, Plotkin, SS (2010) A systematically coarse-grained model for dna and its predictions for persistence length, stacking, twist, and chirality. *J Chem Phys* 132:035105
198. Liwo A, Czaplewski C, Pillardy J, Scheraga HA (2001) Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field. *J Chem Phys* 115:2323–2347
199. Liwo A, Kazmierkiewicz R, Czaplewski C, Groth M, Oldziej S, Wawak RJ, Rackovsky S, Pincus MR, Scheraga HA (1998) United-residue force field for off-lattice protein-structure simulations: III. Origin of backbone hydrogen-bonding cooperativity in united-residue potentials. *J Comp Chem* 19:259–276
200. Chaimovich A, Shell MS (2010) Relative entropy as a universal metric for multiscale errors. *Phys Rev E* 81:060104

# Chapter 20

## The Martini Coarse-Grained Force Field

Xavier Periole and Siewert-Jan Marrink

### Abstract

The Martini force field is a coarse-grained force field suited for molecular dynamics simulations of biomolecular systems. The force field has been parameterized in a systematic way, based on the reproduction of partitioning free energies between polar and apolar phases of a large number of chemical compounds. In this chapter the methodology underlying the force field is presented together with details of its parameterization and limitations. Then currently available topologies are described with a short overview of the key elements of their parameterization. These include the new polarizable Martini water model. A set of three selected ongoing studies using the Martini force field is presented. Finally the latest lines of development are discussed.

**Key words:** Coarse-grained models, Martini force field, Molecular dynamics simulations, Biomolecular systems

---

### 1. Introduction

The use of coarse-grained (CG) models in a variety of simulation techniques has proven to be a valuable tool to probe the time and length scales of systems beyond what is feasible with traditional all-atom (AA) models. Applications to lipid systems in particular, pioneered by Smit et al. (1), have become widely used. A large diversity of coarse-graining approaches is available; they range from qualitative, solvent-free models, via more realistic models with explicit water, to models including chemical specificity (for recent reviews see refs. 2–4). Models within this latter category are typically parameterized based on comparison to atomistic simulations, using iterative Boltzmann schemes (5–7) or force matching (8) approaches. Our own model (9, 10), coined the Martini force field, has also been developed in close connection with atomistic models; however, the philosophy of our coarse-graining approach is different. Instead of focusing on an accurate reproduction of structural details at a particular state point for a specific system, we aim for

a broader range of applications without the need to reparameterize the model each time. We do so by extensive calibration of the chemical building blocks of the CG force field against thermodynamic data, in particular oil/water partitioning coefficients. This is similar in spirit to the recent development of the GROMOS force field (11). Processes such as lipid self-assembly, peptide membrane binding, and protein–protein recognition depend critically on the degree to which the constituents partition between polar and non-polar environments. The use of a consistent strategy for the development of compatible CG and atomic-level force fields is of additional importance for its intended use in multiscale applications (12). The overall aim of our coarse-graining approach is to provide a simple model that is computationally fast and easy to use, yet flexible enough to be applicable to a large range of biomolecular systems. Detailed tutorials have been made available for people to be able to play around with simple systems and thereby get their hands on the different aspects of coarse-graining and running different types of systems. These tutorials are downloadable free of charge from the Martini Web site: <http://cgmartini.nl>.

Currently, the Martini force field provides parameters for a variety of biomolecules, including many different lipids, cholesterol, amino acids, sugars, DNA, fullerene, collagen, dendrimers, and more. A protocol for simulating peptides and proteins is also available. Extensive comparison of the performance of the Martini model with respect to a variety of experimental properties has revealed that the model performs generally quite well (“semi-quantitatively”) for a broad range of systems and state points. Properties accurately reproduced include structural (e.g., liquid densities (9), area/lipid for many different lipid types (9), accessible lipid conformations (13), the tilt angle of membrane spanning helices (14), or helix-helix packing motifs (15, 16), elastic (e.g., bilayer bending modulus (9), rupture tension (10)), dynamic (e.g., diffusion rates of lipids (9, 10), peptides (17) and proteins (18), water transmembrane (TM) permeation rate (9), time scales for lipid aggregation (9)), and thermodynamic (e.g., bilayer phase transition temperatures (19, 20), propensity for interfacial versus TM peptide orientation (14), lipid desorption free energy (10), membrane domain formation (21, 22)) data.

The remainder of this chapter is organized as follows. A detailed description of the CG methodology is presented in the next section, discussing both its abilities and its limitations. Subsequently, a nonexhaustive list of currently available topologies is given and finally examples of three applications are presented, namely, the molecular face of lipid rafts, the gating of a membrane-embedded mechanosensitive channel, and evidence for binding sites of cardiolipins mediating mitochondrial supercomplex formation. A short look at the future prospects of the Martini force field concludes this chapter.

---

## 2. General Methodology

### 2.1. The Looks of Martini

*The name:* The name “Martini” of the force field was coined in 2007 with the release of version 2.0 (10). Martini is the nickname of the city of Groningen in the Netherlands where the force field was developed and where its development continues to date. A famous landmark in the city is the 100 m high Martini tower. The name also reflects the universality of the cocktail with the same name; how a few simple ingredients (chemical building blocks) can be endlessly varied to create a complex palette of taste.

*The mapping:* The Martini model is based on a four-to-one mapping (10), i.e., on average four heavy atoms are represented by a single interaction center, with an exception for ring-like molecules. To map the geometric specificity of small ring-like fragments or molecules (e.g., benzene, cholesterol, and several of the amino acids), the general four-to-one mapping rule is insufficient. Ring-like molecules are therefore mapped with higher resolution (up to two-to-one). The model considers four main types of interaction sites: polar (P), non-polar (N), apolar (C), and charged (Q). Within a main type, subtypes are distinguished either by a letter denoting the hydrogen-bonding capabilities (d = donor, a = acceptor, da = both, 0 = none) or by a number indicating the degree of polarity (from 1 = low polarity to 5 = high polarity). The mapping of representative biomolecules is shown in Fig. 1. For reasons of computational efficiency the mass of the CG beads is set to 72 amu (corresponding to four water molecules) for all beads, except for beads in ring structures, for which the mass is set to 45 amu.

*Nonbonded interaction potentials:* All particle pairs  $i$  and  $j$  at distance  $r_{ij}$  interact via a Lennard-Jones (LJ) potential:

$$V_{\text{LJ}} = 4\epsilon_{ij}[(\sigma_{ij})^{12} - (\sigma/r_{ij})^6] \quad (1)$$

The strength of the interaction, determined by the value of the well-depth  $\epsilon_{ij}$ , depends on the interacting particle types. The value of  $\epsilon$  ranges from  $\epsilon_{ij} = 5.6 \text{ kJ/mol}$  for interactions between strongly polar groups to  $\epsilon_{ij} = 2.0 \text{ kJ/mol}$  for interactions between polar and apolar groups mimicking the hydrophobic effect. The effective size of the particles is governed by the LJ parameter  $\sigma = 0.47 \text{ nm}$  for all normal particle types. For the special class of particles used for ring-like molecules, slightly reduced parameters are defined to model ring-ring interactions;  $\sigma = 0.43 \text{ nm}$ , and  $\epsilon_{ij}$  is scaled to 75% of the standard value. The full interaction matrix can be found in the original publication (10). In addition to the LJ interaction, charged groups (type Q) bearing a charge  $q$  interact via a Coulombic energy function with a relative dielectric constant  $\epsilon_{\text{rel}} = 15$  for explicit screening:

$$V_{\text{el}} = q_i q_j / 4\pi\epsilon_0\epsilon_{\text{rel}}r_{ij} \quad (2)$$

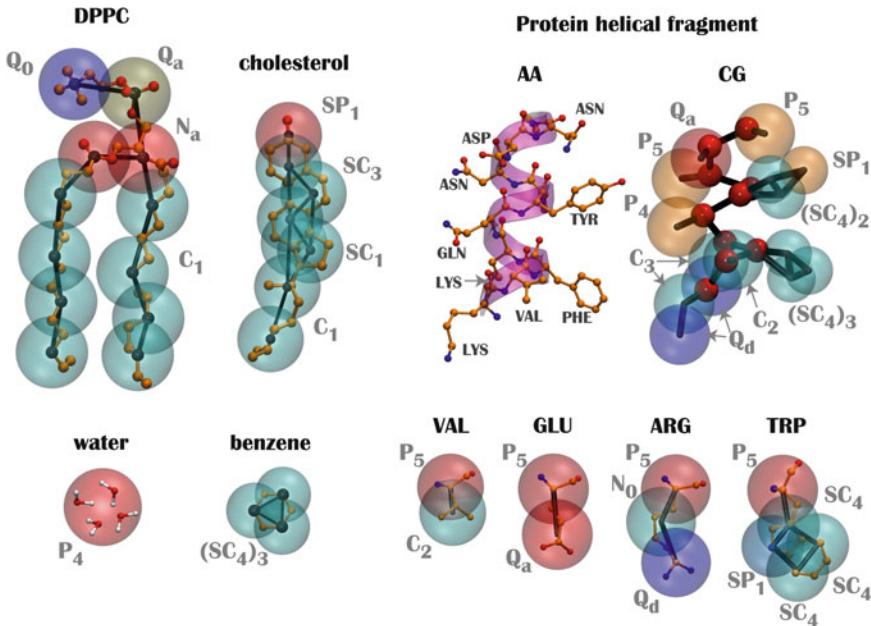


Fig. 1. The looks of Martini. Mapping between the chemical structure at the atomistic level (AA) with the coarse-grained (CG) Martini model for DPPC, cholesterol, water, benzene, a protein helical fragment, and a few amino acids (valine, glutamic acid, arginine, and tryptophan). The CG beads are shown as transparent vdW spheres. For clarity, in the case of the protein helical fragment the AA and CG representations are shown side-by-side and the CG backbone beads are represented by small spheres. Hydrogens are only shown for the atomistic water.

Note that the nonbonded potential energy functions are used in their shifted form. The nonbonded interactions are cut off at a distance  $r_{\text{cut}} = 1.2$  nm. The LJ potential is shifted from  $r_{\text{shift}} = 0.9$  nm to  $r_{\text{cut}}$ . The electrostatic potential is shifted from  $r_{\text{shift}} = 0.0$  nm to  $r_{\text{cut}}$ . Shifting of the electrostatic potential in this manner mimics the effect of a distance-dependent screening.

*Bonded interaction potentials:* Bonded interactions are described by the following set of potential energy functions:

$$V_b = 1/2 K_b (d_{ij} - d_b)^2, \quad (3)$$

$$V_a = 1/2 K_a [\cos(\varphi_{ijk}) - \cos(\varphi_a)]^2, \quad (4)$$

$$V_d = K_d [1 + \cos(\theta_{ijkl} - \theta_d)], \quad (5)$$

$$V_{id} = K_{id} (\theta_{ijkl} - \theta_{id})^2, \quad (6)$$

acting between bonded sites  $i, j, k, l$  with equilibrium distance  $d_b$ , angle  $\varphi_a$ , and dihedral angles  $\theta_d$  and  $\theta_{id}$ . The force constants  $K$  are generally weak, inducing flexibility of the molecule at the CG level resulting from the collective motions at the fine-grained level. The bonded potential  $V_b$  is used for chemically bonded sites, and the

angle potential  $V_a$  to represent chain stiffness. Proper dihedrals  $V_d$  are presently only used to impose secondary structure of the peptide backbone, and the improper dihedral angle potential  $V_{id}$  is used to prevent out-of-plane distortions of planar groups. LJ interactions between nearest neighbors are excluded.

*Implementation:* The functional form of the CG force field was originally developed for convenient use with the GROMACS simulation software (23). Example input files for many systems can be downloaded from <http://cgmartini.nl>. The current version of the Martini force field is denoted as version 2.0 (lipids only) or 2.1 (including proteins). The general form of the potential energy functions has allowed other groups to implement our CG model (with small modifications) also into other major simulation packages such as NAMD (24), GROMOS (13) and Desmond (25). Note that the groups of Schulten (26, 27) and Sansom (28, 29) have developed CG protein force fields compatible with the Martini lipid force field, but different from the Martini protein force field.

## 2.2. Parameterization of Nonbonded Interactions

In order to parameterize the nonbonded interactions of the CG model, a systematic comparison to experimental thermodynamic data has been performed. Specifically, the free energy of hydration, the free energy of vaporization, and the partitioning free energies between water and a number of organic phases were calculated for each of the 18 different CG particle types. Concerning the free energies of hydration and vaporization, the CG model reproduces the correct trend (10). The actual values are systematically too high, however, implying that the CG condensed phase is not as stable with respect to the vapor phase as it should be. The same is true with respect to the solid phase. This is a known consequence of using a LJ 12-6 interaction potential, which has a limited fluid range. Switching to a different nonbonded interaction potential could, in principle, improve the relative stability of the fluid phase (see Subheading 5). As long as its applications are aimed at studying the condensed phase and not at reproducing gas/fluid or solid/fluid coexistence regions, the most important thermodynamic property is the partitioning free energy. Importantly, the water/oil partitioning behavior of a wide variety of compounds can be accurately reproduced with the current parameterization of the Martini model. Table 1 shows results obtained for the partitioning between water and a range of organic phases of increasing polarity (hexadecane, chloroform, and octanol) for a selection of the 18 CG particle types. The free energy of partitioning between organic and aqueous phases,  $\Delta G_{\text{oil/aq}}$ , was obtained from the equilibrium densities  $\rho$  of CG particles in both phases:

$$\Delta G_{\text{oil/aq}} = kT \ln (\rho_{\text{oil}} / \rho_{\text{aq}}) \quad (7)$$

**Table 1**

**Oil (Hex), chloroform (CLF), and octanol (OCO)/water partitioning free energies for a selection of the 18 CG particle types, compared to experimental values of the corresponding chemical building blocks**

Building block	Type	Hex/water		CLF/water		OCO/water	
		CG	Exp	CG	Exp	CG	Exp
Acetamide	P <sub>5</sub>	-28	-27	-18	-20	-10	-8
Water	P <sub>4</sub>	-23	-25	-14	-	-9	-8
Propanol	P <sub>1</sub>	-11	-10	-2	-2	-1	0
Propylamine	N <sub>d</sub>	-7	-6	0	1	3	3
Methylformate	N <sub>a</sub>	-7	-6	0	4	3	0
Methoxyethane	N <sub>0</sub>	-2	1	6	-	5	3
Butadiene	C <sub>4</sub>	9	11	13	-	9	11
Chloropropane	C <sub>3</sub>	13	12	13	-	14	12
Butane	C <sub>1</sub>	18	18	18	-	17	16

The experimental data are compiled from various sources (see ref. 10), the simulation data are obtained using Eq. 7. All values are expressed in kJ/mol and obtained at T = 300 K.

The equilibrium densities can be obtained directly from a long MD simulation of the two-phase system in which small amounts (around 0.01 mole fraction proved sufficient to be in the limit of infinite dilution) of the target substance are dissolved. With the CG model, simulations can easily be extended into the multimicrosecond range, enough to obtain statistically reliable results to within 1 kJ/mol for most particle types. As can be judged from Table 1, comparison to experimental data for small molecules containing four heavy atoms (the basic mapping of the CG model) reveals a close agreement to within 2 kT for almost all compounds and phases; indeed, agreement is within 1 kT for many of them. Expecting more accuracy of a CG model might be unrealistic. Note that the multiple nonbonded interaction levels allow for discrimination between chemically similar building blocks, such as saturated versus unsaturated alkanes or propanol versus butanol (which would be modeled as Nda) or ethanol (P2). A more extensive table including all particle types and many more building blocks can be found in the original publication (10).

The thermodynamic integration (TI) approach is also routinely used to determine solute partitioning free energies. It has a few advantages over the equilibrated two-phase technique: (1) it avoids uncertainties due to finite concentration effect since it truly reflects infinite dilution, (2) it solves the sampling issue observed in cases

where the solute favors one phase much more than the other. The use of TI coupled with a thermodynamic cycle for the calculation of solvation free energy is well documented in the literature. One may consult the article from Villa and Mark for a detailed description (30). Technical aspects of the calculation using the Gromacs package are described in a tutorial available on the Martini Web site. A comparison of solvation free energies obtained with TI compared to the direct approach is given in (31).

### **2.3. Parameterization of Bonded Interactions**

To parameterize the bonded interactions, we use structural data that are either directly derived from the underlying atomistic structure (such as bond lengths of rigid structures) or obtained from comparison to fine-grained simulations. In the latter procedure, the fine-grained simulations are first converted into a “mapped” CG (MCG) simulation by identifying the center-of-mass of the corresponding atoms as the MCG bead. Second, the distribution functions are calculated for the mapped simulation and compared to those obtained from a true CG simulation. Subsequently the CG parameters are systematically changed until satisfactory overlap of the distribution functions is obtained. Using this procedure, simulations of bulk alkanes have been used to determine the optimal values of the “standard” equilibrium bond distance of 0.47 nm and force constant of  $K_b = 1,250 \text{ kJ/mol/nm}^2$ , and equilibrium angle of 180° with force constant of  $K_a = 25 \text{ kJ/mol}$ . Likewise, standard bonded parameters have been derived for unsaturated alkanes (10), and the phospholipid headgroup (9).

Although it is advised to try sticking to the use of COM of atom groups in the mapping, in cases of complex molecules it might be more convenient to use specific atoms instead. This would be fine and might for instance allow a better representation of the mechanistic of the molecule e.g., to represent the relative orientation of two bound planes (aromatics) it would be important to realistically incorporate the rotation around the bond into the model. This might be easier using specific sites in the atomistic model. In extreme cases it might even be necessary to use a “double bead” representation in which one bead carries the bonded terms while the other one carries the nonbonded terms. In this approach it is important to choose the link between the two beads adequately (virtual sites may be used). This solution should not be a standard.

### **2.4. Limitations of the Model**

The potential range of applications of the CG model is very broad. There are, however, certain important limitations that should be kept in mind.

*Limited stability of fluid phase:* First of all, the model has been parameterized for the fluid phase. Properties of solids, such as crystal packing, are not expected to be accurate. Both the gas and the solid phase appear too stable with respect to the fluid phase.

This is a known consequence of the use of the LJ 12-6 potential for the nonbonded interactions. The thermodynamic behavior of solid–fluid and gas–fluid interfaces should therefore be interpreted with care, at least at the quantitative level. In applications where such interfaces are formed (especially the water–vapor interface in, e.g., rupture of lipid monolayers) these limitations have to be kept in mind. In biomolecular simulations, a related problem is the potential freezing of the Martini water model. The LJ parameters for water ( $\epsilon = 5.0$  kJ/mol,  $\sigma = 0.47$  nm) bring it into the solid state region of the LJ phase diagram, although the use of a shift potential reduces the long-range attractive part and the CG water is more fluid compared to the standard LJ particle. We have previously determined the freezing temperature of the CG water as  $290 \pm 5$  K (9, 10, 32). While this is admittedly higher than it should be, in most applications freezing is not observed as long as no nucleation site is formed. Apart from simulations performed at lower temperatures, rapid freezing is a potential problem in systems where a nucleation site is already present (a solid surface, but also an ordered bilayer surface may act as one) or when periodicity enhances the long-range ordering (e.g., for small volumes of water). In those cases in which the freezing poses a problem, a simple pragmatic solution has been presented in the form of antifreeze particles (10). This works in some cases, but has apparently given problems in combination with solid supports (33). Freezing might also be more easily observed with Martini implementations in other packages, notably with Gromos which uses an extended LJ interaction (13).

*Entropy-enthalpy compensation:* Furthermore, the parameterization is based on free energies. The inherent entropy loss on coarse graining is necessarily compensated for by a reduced enthalpy term (13). The enthalpy/entropy balance of many processes may therefore be biased when modeled at the CG level and affect its temperature dependence, although not necessarily weakening it. For instance, the temperature-dependent hydration free energy for linear alkanes was found to be more pronounced in the CG representation compared to an AA representation (13). As is true for any force field, applications outside the temperature range used for parameterization ( $\sim 270$ – $330$  K) have to be considered with care. Although absolute entropies are clearly underestimated due to the loss of atomistic degrees of freedom, entropy differences can still be accurate. As an example, for the dimerization of WALP23 in DOPC, a direct comparison of enthalpies and entropies obtained from experiment and simulation is possible: Yano and Matsuzaki (34) measured values of  $\Delta H = -31$  kJ/mol and  $\Delta TS = +19$  kJ/mol, which can be compared to the respective values of  $-30$  kJ/mol and  $+15$  kJ/mol obtained by Ash and coworkers (35) using the Martini force field. Thus, not only is the dimerization free energy difference  $\Delta G$  obtained with the Martini model ( $-15$  kJ/mol) in good agreement with experiment ( $-13$  kJ/mol), but also the enthalpic and entropic contributions to it.

*Implicit screening:* Another difficulty encountered in our CG model, and perhaps in most coarse-graining approaches, is the correct modeling of the partitioning of polar and charged compounds into a low dielectric medium. Because of the implicit screening, the interaction strength of polar substances is underestimated in nonpolarizable solvents. Applications involving the formation of polar/charged complexes in a nonpolar environment are especially prone to be affected. The inability to form a TM water pore upon dragging a lipid across the membrane is an example (10). The recent development of a Martini water model that includes orientational polarization by the means of a dipole represented by two drude charges attached to each water bead (see below for a detailed description) allows to correct for some of these effects (31). Apart from the implicit screening in the CG model, the neglect of long-range electrostatic forces poses a further limitation. Pairwise interactions beyond 1.2 nm (between two and three CG beads away) are not taken into account. In principle long-range electrostatic interactions could be added to the CG model, in ways similar to those used in atomistic simulations. One has to realize that a modification of the electrostatic interaction scheme might affect other system properties.

*Effective time scale:* The CG dynamics is faster than the AA dynamics because the CG interactions are much smoother compared to atomistic interactions. The effective friction caused by the fine-grained degrees of freedom is missing. Based on comparison of diffusion constants for a range of systems (including simple solvents and lipids) in the CG model versus experimental data, the effective time sampled using CG interactions is three- to eightfold larger. When interpreting the simulation results with the CG model, a standard conversion factor of 4 is used, which is the effective speed-up factor in the diffusion dynamics of CG water compared to real water. The same order of acceleration of the overall dynamics is also observed for a number of other processes, including the permeation rate of water across a membrane (9), the sampling of the local configurational space of a lipid (13), the aggregation rate of lipids into bilayers (9) or vesicles (23), and the self-diffusion of lipids (9, 10), transmembrane peptides (17), and proteins (18). However, the speed-up factor might be quite different in other systems or for other processes. Particularly for protein systems, no extensive testing of the actual speed-up due to the CG dynamics has been performed, although protein translational and rotational diffusion was found to be in good agreement with experimental data in simulations of CG rhodopsin (18). In general, the time scale of the simulations has to be interpreted with care.

*Time step:* Martini has been parameterized using time steps in the range of 20–40 fs. Whether you can use 40 fs or have to settle for a somewhat smaller time step depends on your system, and on your

attitude toward coarse-grained modeling, as explained below. First, the Martini force field is not an atomistically detailed force field. Many assumptions underlie the model, the major one being the neglect of some of the atomistic degrees of freedom. As a result, the interactions between particles are effective ones and the energy landscape is highly simplified. This simplified energy landscape allows for a greatly increased sampling speed at the cost of a loss of detail. This makes CG models in general so powerful. The emphasis, therefore, should not be to sample the energy landscape as accurately as possible, but rather, as effectively as possible. This is in contrast to traditional all-atom models, for which the energy landscape is more realistic and an accurate integration scheme is more important. In practice, the inherent “fuzziness” of the Martini model makes the presence of small energy sinks or sources a less critical problem than in accurate atomistic simulations. Second and most importantly, structural properties are rather very robust with respect to time step; for a time step up to 40 fs, there are no noticeable effects on structural properties of the systems investigated. Moreover, thermodynamic properties such as the free energy of solvation also appear insensitive to the size of the time step. Thus, if the goal is to generate representative ensembles quickly, large time steps seem acceptable. Whereas one can debate the first argument (i.e., the “idealistic” versus “pragmatic” view of the power of CG simulations), the second argument (i.e., the insensitivity of both structural and thermodynamic properties to the magnitude of the time step) implies that a reduction of the time step to 10 fs or below, as recently suggested (36), is a waste of computer time (32). Nevertheless, time steps of 40 fs and beyond may be pushing the limits too far for certain systems. We therefore recommend a time step of 20–30 fs, in combination with an enlarged neighborlist cutoff (to 1.4 nm) to be on the safe side. Of course, one should always check whether or not results are biased by the choices made. Given that the largest simplifications are made at the level of the interaction potentials, this can best be done by comparing to results obtained using more detailed models.

*Fixed secondary structure:* Finally, in applications of peptides and proteins one has to be aware that secondary structure (SS) transformations are not modeled in the current parameterization (see Subheading 3). The secondary structure is essentially fixed by the use of a dihedral potential energy function. The backbone bead type is also function of the SS to take into account the fact that when involved in interactions stabilizing a given SS element the backbone is less prompted to engage in other interactions. The backbone interaction strength is therefore decreased when involved in a SS element. This approach allows discrimination between various secondary structure elements but prevents realistic transitions between them. Processes in which folding and unfolding are playing a substantial role are therefore not suitable for modeling

with our current CG force field. Movements of secondary structure elements with respect to each other are possible, however, and were shown to be quite realistic in a recent application of the gating of a membrane-embedded mechanosensitive channel (see Subheading 4).

---

### 3. Available Topologies

The transferability that the Martini model has been conceived on has been used by our group but also by other groups to make topologies for generic building blocks for several molecule types as lipids, proteins, sugars, but also for more specific ones such as fullerene, DNA, collagen, and polymers. Of special interest is also the recently developed topology for a polarizable water model. The topologies are available on the Martini Web site: <http://cgmartini.nl>, and discussed in some detail below. A general recipe for construction of your own model is also given.

#### 3.1. Basic Topologies for Lipids, Proteins, Sugars

*Lipids:* The Martini force field has been originally developed with a strong orientation toward lipid systems. The lipid models have been thoroughly tested on many types of systems covering not only the bilayer state but also micellar, monolayer, hexagonal, and cubic phases. An example topology for a dipalmitoyl-phosphatidylcholine (DPPC) molecule is shown in Fig. 1a. The hydrophobic tails consist of C1 type particles, the glycerol moiety of Na particles of intermediate polarity, and the head group of a negatively charged Qa particle for the phosphate group, and a positively charged Q0 particle for the choline group. Double bonds in the tail can be effectively modeled using slightly less hydrophobic beads (C2, C3) together with a change of the angle potential (Eq. 4) that governs the stiffness and orientation of the lipid tails. Likewise, changes in the lipid head group particle types allow one to discriminate between the most important lipid classes. Currently available topologies include PC (phosphatidylcholine), PE (phosphatidylethanolamine), PG (phosphatidylglycerol), and PS (phosphatidylserine), with all common tails such as hexanoyl, lauroyl, palmitoyl, oleoyl, stearoyl, linoleyl, and arachidonoyl tails. The properties that are reproduced with these models include structural (area per lipid and bilayer thickness), elastic (bending rigidity and area compressibility), thermodynamic (phase transition temperature and line tension) and dynamical (lipid diffusion and water permeation rate) data. In addition to two-tail lipids, the topologies of a variety of single-tail lipids are available, as well as for sphingolipid (SM). However, parameterization of the latter has proven problematic due to lack of clear experimental data and consistent atomistic models for SM bilayers. Furthermore, lipid topologies have been parameterized for specific lipids such

as cholesterol (37, 38), cardiolipins (39), triglycerides (40), and bolalipids (41).

*Proteins:* To select particle types for the amino acids, systematic comparison to experimental partitioning free energies between polar and apolar media was also used in line with the general Martini philosophy (see Subheading 2). Table 2 shows the resulting assignment of the amino acid side chains and the associated partitioning free energies. Where available, the experimental values are reproduced to within 2 kT, a level of accuracy that is difficult to obtain even with atomistic models. Most amino acids are mapped onto single standard particle types, similarly to the recent work of other groups (26, 29). The positively charged amino acids are modeled by a combination of a Q-type and an N- or C-type particle. The bulkier ring-based side chains are modeled by three or four beads of the special class of ring particles. The Gly and Ala residues are only represented by the backbone particle. Figure 1 shows the mapping of a few of them. To compensate for the lack of directionality in the backbone interactions (H-bonds) the type of the particle used for the backbone bead is made a function of its secondary structure; when free in solution or in a coil or bend, the backbone has a strong polar character (P type), while as part of a helix or beta strand the interbackbone hydrogen bonds reduce the polar character significantly (N type). The final bead type of each amino acid was refined to match data obtained from atomistic simulations: (1) PMF of side chain analogues when crossing a DOPC–water interface, (2) amino acid association constants, (3) partitioning and orientation of pentapeptides at the water–cyclohexane interface, (4) tilt, orientation and interaction of helical transmembrane peptides (14).

For the bonded interactions, distributions were derived directly from the protein databank, using the proper AA to CG mapping (14). These distributions reflect all possible configurations for a large number of different systems under a variety of conditions. Keeping the aim of our CG model in mind, namely, to be able to simulate many biomolecules with a single set of parameters, this is the least biased information. Using this procedure, bonded parameters were derived for the backbone (BB) potentials, namely, the BB-BB bonded potential, the BB-BB-BB angle potential, and the BB-BB-BB-BB dihedral potential. Furthermore, for each amino acid, side chain (SC) distributions were obtained for the BB-SC bonded potential, the BB-BB-SC angle potential, and for the intra-SC potentials for amino acids containing more than one CG particle. The complete set of bonded and nonbonded parameters for proteins can be found elsewhere together with more details of the parameterization (14). In the current version of Martini the BB-BB-BB angle potential, and the BB-BB-BB-BB dihedral potential are used to enforce the secondary structure (SS) of the backbone, which is therefore an input parameter in our CG model. Different

**Table 2**  
**Free energy based mapping of select amino acid side chain analogues**

Side chain	Particle type	Mapping <sup>a</sup>	Oil/water	
			CG	Exp
LEU/ILE	C1		22	22/22
VAL/PRO	C2		20	17/-
MET/CYS	C5		9	10/5
SER/THR	P1		-11	-14/-11
ASN	P5		<-25	-28
GLN	P4		-23	-25
ASP <sup>-1</sup>	Qa		<-25	-
ASP <sup>0</sup>	P3		-18	-19
GLU <sup>-1</sup>	Qa		<-25	-
GLU <sup>0</sup>	P1		-11	-11
ARG <sup>+1</sup>	N0-Qd	N0: C $\beta$ -C $\gamma$ -C $\delta$ -N $\varepsilon$	<-25	-
ARG <sup>0</sup>	N0-P4	Qd/P4:C $\zeta$ -N $\omega$ 1-N $\omega$ 2	-23	-25
LYS <sup>+1</sup>	C3-Qd	C3: C $\beta$ -C $\gamma$ -C $\delta$	<-25	-
LYS <sup>0</sup>	C3-P1	Qd/P1:C $\varepsilon$ -N $\omega$	-1	-2
HIS	SC4-SP1-SP1	SC4: C $\beta$ -C $\gamma$ SP1: C $\delta$ -N $\varepsilon$ SP1: N $\delta$ -C $\varepsilon$	-19	-20
PHE	SC4-SC4-SC4	SC4: C $\beta$ -C $\gamma$ -C $\delta$ 1 SC4: C $\delta$ 2-C $\varepsilon$ 2 SC4: C $\varepsilon$ 1-C $\zeta$	19	17
TYR	SC4-SC4-SP1	SC4: C $\beta$ -C $\gamma$ -C $\delta$ 1 SC4: C $\delta$ 2-C $\varepsilon$ 2 SP1: C $\varepsilon$ 1-C $\zeta$ -OH	-1	-2
TRP	SC4-SP1-SC4-SC4	SC4: C $\beta$ -C $\gamma$ -C $\delta$ 2 SP1: C $\delta$ 1-N $\varepsilon$ 1-C $\varepsilon$ 1 SC4: C $\varepsilon$ 2-C $\zeta$ 2 SC4: C $\varepsilon$ 1-C $\omega$	12	9

The experimental partitioning free energies are obtained for cyclohexane/water, the simulation results for butane/water, using Eq. 7. All values are expressed in kJ/mol and obtained at  $T = 300$  K

<sup>a</sup>The mapping is reported only for amino acid side chains composed of more than one bead

dihedral and angle parameters are used to distinguish a helix, a strand, or a random coil. It is therefore not possible to study realistic folding-unfolding events at this stage.

Furthermore the structural SS parameters were extracted from a statistical analysis of protein structures and thus describe ideal secondary structure elements, e.g.,  $\alpha$ -helices and  $\beta$ -sheets. Note that this may result in an unusually high RMSD of the model with respect to the experimental structure but does not mean it is unfolding. In cases where the specificity of the local deformations of the protein backbone (e.g., helix kinks) is of interest the use of the elastic network together with the Martini model (ElNeDyn) is recommended (16). The elastic network defines a network of springs between backbone beads. We have shown that with an appropriate set of values for the cutoff distance limiting the extend of the network and of the force constant of the springs this model is able to simultaneously reproduce global and local deviations, fluctuations, as well as collective motions of all-atom protein models (16). It is of importance to note that since in this version of the protein force field the backbone bead is placed on the  $C\alpha$  atom all bonded parameters (equilibrium values and force constants) have been reparameterized accordingly (16).

*Carbohydrates:* The parameterization of carbohydrates followed the same philosophy as was used previously for lipids and proteins, focusing on the reproduction of partitioning free energies of small compounds between polar and non polar phases (42). The carbohydrate building blocks considered are the monosaccharides glucose and fructose, and the disaccharides sucrose, trehalose, maltose, cellobiose, nigerose, laminarabiose, kojibiose, and sophorose. For a single sugar ring (consisting of 12 nonhydrogen atoms) three CG particles are used. This level of resolution preserves the general mapping procedure for the Martini force field (4/1 mapping), the geometrical shape of the rings, and allows for a distinction between different types of monosaccharides through variation in the bond lengths, angles, and CG particle types. The particle types that best reproduced experimental partitioning data were found to be the polar class (P1–P4) of particles. Disaccharides are modeled as two to three bead units connected by a single bond, which mimics the glycosidic linkage. This geometry allows the definition (and subsequent parameterization) of the important  $\varphi$  and  $\psi$  dihedral angles, which determine the relative orientation and flexibility of the two sugar residues. The set of fine-grained particles represented by the CG beads is chosen to be different for the monosaccharide and the corresponding disaccharide residue. This somewhat non-obvious choice is based on the ability to represent the typical polar/apolar character of the disaccharides, with the apolar part corresponding to the central part along the glycosidic linkage. Oligosaccharides are constructed adding disaccharide residues through additional bonds. Bonded parameters for these saccharides were optimized by comparison to conformations sampled with an atomistic force field. Special attention was given to the

representation of the rotameric states. Applications of the carbohydrate model to the oligosaccharides amylose and curdlan show a preservation of the main structural properties, with three orders of magnitude more efficient sampling than the atomistic counterpart. The investigation of the cryo- and anhydroprotective effect of glucose and trehalose on a lipid bilayer showed a strong decrease of the melting temperature, in good agreement with both experimental findings and atomistic simulation studies (42).

### **3.2. Polarizable Water Model**

The classic Martini water model does not bear charges and, consequently, is blind to electrostatic fields and polarization effects. A uniform screening is used instead. While this is a reasonable approximation for the bulk water, problems arise at interfaces between water and other phases and in the presence of charged particles. In our constant effort to improve the accuracy and the range of applicability of the Martini force field, a water model providing a better representation of the electrostatic interaction has been recently developed (31). It includes orientational polarizability by means of a dipole described by two net charges attached to the water bead, which carries the LJ potential (see Fig. 2a). The parameterization of the model followed three criteria: (1) The dielectric constant and (2) particle density of the model should be close to that of real water, and (3) the partitioning free energies between water and organic solvents, one of the cornerstones of the Martini model, should remain unaffected. The final polarizable model has a dielectric constant of 75.6, a density of  $1,043 \text{ kg/m}^3$  and adjustments of its LJ parameters allowed to reproduce the right solvation energies for most particles. The LJ self- and cross-interactions of charged particles were modified to compensate for the important change in their interaction with the water model. Details of the model can be found elsewhere (31). Note that a model similar in spirit was published around the same time by Wu et al. (43).

While preserving many important properties of the classical Martini water model, the polarizable version significantly improves others. To illustrate the benefit of incorporating orientational polarization in the CG water model electroporation events were studied (31). It was notably shown that both poration of an octane slab in water under an external electric field and the poration of a lipid bilayer due to the electric field created by an ionic imbalance across the membrane are phenomena described realistically by the model. In Fig. 2c an example of the electroporation due to an unbalance of ion concentration is shown. A large water pore is spontaneously formed allowing the ions to cross the membrane bilayer. When the ion concentrations are equivalent on both sides of the bilayer the pore closes. The effect of the polarizable water model is more clearly seen in the snapshot presented in Fig. 2d,

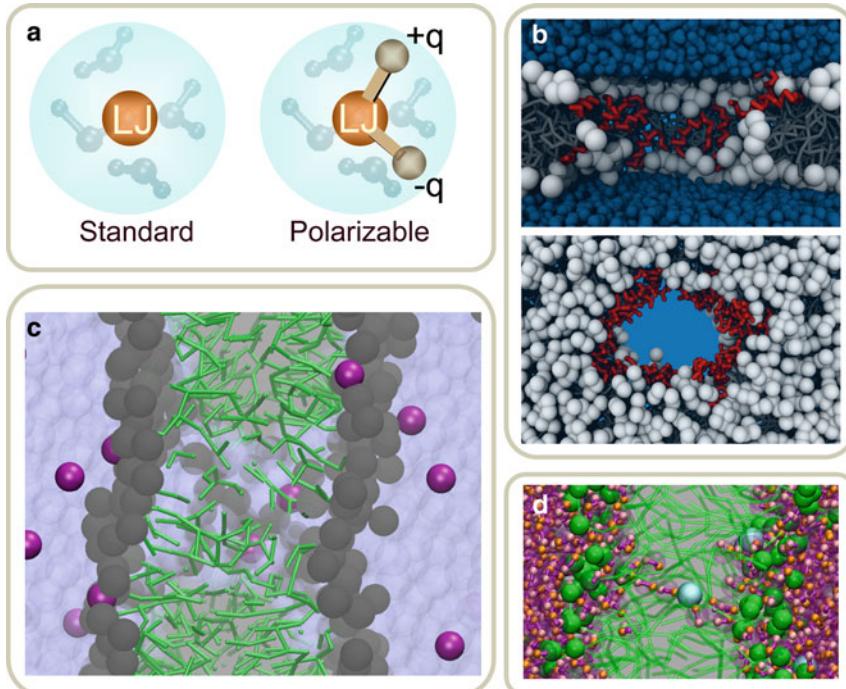


Fig. 2. New polarizable water model for Martini. The polarizable water model is shown in panel **a** together with the standard water model. In panel **b**, a transmembrane pore formed by magainin peptides is shown from the side (*top*) and the top (*bottom*) of the lipid bilayer. For clarity, in the side view the water is shown on both sides of the lipid bilayer and hidden in the transmembrane region; in the top view the water is not shown. In panel **c**, a water pore formed due to the presence of an ion gradient across the lipid bilayer is shown. The lipid head groups are depicted by *gray spheres* and the tail by *sticks*. The ions are the *solid spheres* and cross the membrane through the pore. In panel **d**, the dipoles of the polarizable water molecules are shown when solvating an ion upon crossing the membrane at low electric field. The water dipoles clearly orient due to the presence of the ion.

where a single ion permeated the membrane under a low electric field. During the permeation process, the ion keeps being hydrated to avoid the cost of desolvation in a low dielectric medium. The polarizable waters reorient in such a way that their dipoles stabilize the ionic charge. Another example of the potential range of applications with the polarizable water model is the formation of transmembrane pores by antimicrobial peptides. Although pores are formed with the nonpolarizable water model when self-assembly simulations are used (i.e., starting from randomly distributed systems), it was not possible to observe their formation on a preformed lipid bilayer. With the polarizable water model when a number of magainin peptides are added in the aqueous phase next to a DMPC bilayer, spontaneous pore formation is observed as is illustrated in Fig. 2b. The pore is of a disordered toroidal nature, in agreement with previous results from atomistic simulations (44, 45).

### 3.3. Specific Topologies

*Nanomaterials:* Fullerenes are molecules entirely composed of carbon atoms forming hollow spheres, ellipsoids, or tubes. They are products emerging from the development and use of nanomaterials and their effect on health and environment has become a concern. Monticelli and coworkers (46) have determined the topology of the C<sub>60</sub> molecules based on a large collection of experimental data including its partitioning behavior in different solvents. The fullerenes are described by 16 CG beads, staying close to a 4/1 mapping, based on the bead type for benzene (SC4). This topology was used to describe the thermodynamics and mechanism of interaction of fullerene aggregates—C<sub>60</sub> aggregates formed in aqueous solution—with a lipid bilayer (46). The authors primarily observed a complete dissolution of the fullerene aggregates into the lipid bilayer and noted to their surprise only relatively little mechanical damage of the bilayer. Carbon nanotubes (CNTs) are of great interest in a wide range of potential applications but still only little is known concerning their behavior in biological environment. Wallace and Sansom have built carbon nanotube (CNT) topologies using C-type particles from an earlier version of Martini (9). The model was used to study the interaction between CNT and several detergents with the aim of reducing CNT self-affinity and disperse them in an aqueous environment (47, 48). They also looked at the response of a lipid bilayer upon the penetration of CNTs of different radius and with different orientations relative to the membrane plane (49).

*Collagen:* Collagen molecules are abundant proteins in the human body where their mechanical properties provide stability, elasticity and strength to connecting tissues. A collagen topology (50) was built as an extension to the Martini protein force field. Collagen molecules are repeats of the GXY motif where X and Y may be any residue but often prolines and hydroxyprolines. They adopt a peculiar right-handed triple helix with a backbone conformation that was not yet available in the Martini force field. The authors defined a set of bonded terms (bond, angles and dihedrals) based on collagen experimental structures (for reference values) and atomistic simulations of small building blocks (for force constants). The hydroxyproline (hP) was first defined using the Black and Mould hydrophobicity scale to determine the most appropriate bead type (C5-P1) for its nonbonded interactions. The model was finally tested by calculating some mechanical features of the collagen molecule [(G-P-hP)<sub>10</sub>]<sub>3</sub>. The Young's modulus and persistence length of the model were found in good agreement with experimental and computational values in the literature.

*DNA:* Double stranded DNA may mix with lipids to form lipoplexes. These have become attractive as they present a very promising alternative to viral gene vectors for the intracellular gene delivery. Their primary advantage being to have a lower toxicity to their host, but they are still not as efficient. Computer

simulations might help improving this formalism. Khalid and coworkers (51) have developed a coarse-grained model of DNA compatible with the Martini force field. The authors use Nd, Na and Nda particle types to describe the bases; the phosphate and deoxyribose groups are described by a Qa and C-Na bead type, respectively. Bonded terms were not a particular issue since an elastic network was used to connect all particles within 0.7 nm by a spring with a 1,500 kJ/mol/nm<sup>2</sup> force constant. The persistence length of the model was tuned in agreement with experimental values. The model was not tested further against experimental or computational data (51). The authors studied the nanoarchitecture of lipoplexes (52), which has been suggested to influence their efficiency, and were able to observe phase transitions and stability of large lipoplexes that were compatible with results obtained by SAXS.

*Polymers:* The use of CG models to describe polymer dynamics antedates the use of such models for lipid dynamics (53). Yet there is a need also for Martini models for polymeric systems, especially since it would then be possible to simulate systems in which polymers interact with biomolecules, which are already defined in Martini. One example is the use of dendrimers as efficient drug carriers for transmembrane cargo delivery. A Martini model for dendrimers was developed by Lee and Larson (54) who studied pore formation by polyamidoamine (PAMAM) dendrimers in DMPC and DPPC bilayers. Third and 5th generation dendrimers interact with DPPC and may cause pore formation depending on the chemical details of the dendrimers and the salt concentration (54). In a follow-up study (55), 5th and 7th generation dendrimers were simulated and showed distinct behavior depending on their charges, in agreement with atomic force microscopy experiments. Relatively subtle differences in structure caused significant differences in pore formation and water permeation through the bilayer, suggesting that the CG simulations are accurate enough to help design and understand this type of molecules (55). Other polymers with a large range of bioapplications are polyethylene oxide (PEO) and polyethylene glycol (PEG). They have in particular been conjugated to an array of pharmaceuticals to overcome limitations of low solubility, short circulating lifetime, and immunogenicity. The Martini model was recently parameterized for PEO and PEG (56). Based on reproducing structural properties such as the swelling in aqueous solution, the particle type for the C-O-C repeat unit was set to SNa. However, in order to reproduce the appropriate density of the melt, the self-interaction was somewhat increased. Another polymer for which Martini parameters are now available is polystyrene (57).

### 3.4. How to Make Your Own Topology

Here we present a simple three-step recipe, or guide, on how to proceed in parameterizing new molecules using the Martini model. The first step consists of mapping the chemical structure to the CG representation, the second step is the selection of appropriate

bonded terms, and the third step is the optimization of the model by comparing to AA level simulations and/or experimental data.

*Step I, mapping onto CG representation:* The first step consists in dividing the molecule into small chemical building blocks, ideally of four heavy atoms each. The mapping of CG particle types to chemical building blocks, examples of which are presented in Table 1, subsequently serves as a guide towards the assignment of CG particle types. Because most molecules cannot be entirely mapped onto groups of four heavy atoms, some groups will represent a smaller or larger number of atoms. In fact, there is no reason to map onto an integer number of atoms, e.g., a pentadecane mapped onto four C1 particles implies that each CG bead represents 3.75 methyl(ene) groups. In case of more substantial deviations from the standard mapping scheme, small adjustments can be made to the standard assignment. For instance, a group of three methyl(ene)s is more accurately modeled by a C2 particle (propane) than the standard C1 particle for saturated alkanes. The same effect is illustrated by the alcohols: whereas the standard alcohol group is modeled by a P1 particle (propanol), a group representing one less carbon is more polar (P2, ethanol), whereas adding a carbon has the opposite effect (Nda, butanol). Similar strategies can be used for modulation of other building blocks. To model compounds containing rings, a more fine-grained mapping procedure can be used. In those cases, the special class of S-particles is appropriate.

*Step II, selecting bonded interactions:* For most molecules the use of a standard bond length (0.47 nm) and force constant of  $K_b = 1,250 \text{ kJ/mol/nm}^2$  appears to work well. In cases where the underlying chemical structure is better represented by using different values, there is no restriction in adjusting these values. Especially for ring structures, much smaller bond lengths are required. For rigid rings, the harmonic bond and angle potentials are replaced by constraints, as was done for benzene and cholesterol. For linear chain-like molecules, a standard force constant of  $K_a = 25 \text{ kJ/mol}$  with an equilibrium bond angle  $\varphi_a = 180^\circ$  best mimics distributions obtained from fine-grained simulations. The angle may be set to smaller values to model unsaturated *cis*-bonds (for a single *cis*-unsaturated bond  $K_a = 45 \text{ kJ/mol}$  and  $\varphi_a = 120^\circ$ ) or to mimic the underlying equilibrium structure more closely in general. In order to keep ring structures planar, improper dihedral angles should be added. For more complex molecules (e.g., cholesterol), multiple ways exist for defining the bonded interactions. Not all of the possible ways are likely to be stable with the preferred time step of  $\sim 30 \text{ fs}$ . Some trial-and-error testing is required to select the optimal set.

*Step III, refinement:* The coarse-graining procedure does not have to lead to a unique assignment of particle types and bonded interactions. A powerful way to improve the model is by comparison to

AA level simulations, analogous to the use of quantum calculations to improve atomistic models. Structural comparison is especially useful for optimization of the bonded interactions. For instance, the angle distribution function for a CG triplet can be directly compared to the distribution function obtained from the AA simulation, using the mapping procedure described earlier. The optimal value for the equilibrium angle and force constant can thus be extracted. Comparison of thermodynamic behavior is a crucial test for the assignment of particle types. Both AA level simulations (e.g., preferred position of a probe inside a membrane) and experimental data (e.g., the partitioning free energy of the molecule between different phases) are useful for a good assessment of the quality of the model. The balance of forces determining the partitioning behavior can be very subtle. A slightly alternative assignment of particle types may significantly improve the model. Once more, it is important to stress that Table 1 serves as a guide only; ultimately the comparison to AA simulations and experimental data should be the deciding factor in choosing parameters.

During the refinement of the bonded and nonbonded terms of the CG topology of a molecule, one should keep in mind that these terms are strongly interdependent. When associating two or more beads together the principle of additivity of solvation free energy (and therefore partitioning free energy) may be followed as a first approximation. But this will not hold in many situations. For instance in a molecule when the distance between two beads is set to one smaller than the standard 0.47 nm (close to the particle LJ diameter: 0.526 nm) their association will result in an overlap of their LJ spheres and thereby reduce the volume of the molecule as a whole when compared to the sum of the bead sizes. As a result, the solvation free energy of the molecule will be larger than the sum of the values for each bead and thus will be overestimated. This is due to the loss of the energy cost of generating a cavity to fit the overlapping section(s) of the molecule in the solvent. This effect may be as large as 10–15 kJ/mol when the distance is decreased from 0.47 to 0.25 nm.

---

## 4. Applications

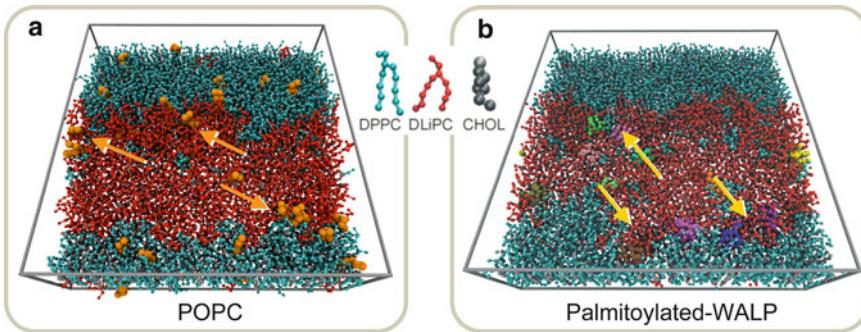
The list of applications of the Martini model to date is broad, reflecting the flexibility and transferability underlying our coarse-graining protocol. Some important applications include, among others, (protein mediated) vesicle fusion (58–61), lamellar phase transformations (19–21, 39), the collapse of monolayers (62–64), peptide-induced membrane modulation (65–68), the self-assembly of membrane-proteins (15, 18, 26, 27, 29, 69–71), gating of proteins (72–76), carbon nanotube-lipid interactions (46, 49),

confinement of copolymers (77), structure and dynamics of lipoprotein particles (26, 27, 40), pore formation by dendrimers (54, 55, 78), and formation of lipoplexes (52). In most of these studies the CG model, sometimes with small changes or extensions from the published version, performed well when compared to either experimental data or to more detailed atomistic models. Here we selected three applications of ongoing studies performed in our own lab, namely, simulations of raft-like lipid membranes, gating of mechanosensitive channels, and the involvement of cardiolipins in the formation of membrane protein supercomplexes. All simulations were performed with the GROMACS simulation software (23), versions 4.0. In the applications, we will use an effective time rather than the actual simulation time to account for the speed-up in coarse-grained dynamics (see Subheading 2.4).

#### **4.1. Molecular Face of Lipid Rafts**

The lateral heterogeneity of biological membranes has important implications for the function of cells. Nevertheless, to study the organization of biological membranes remains a challenge, because it is inherently difficult to characterize fluctuating nanoscale assemblies in living cells. Model membranes and isolated plasma membranes are more frequently studied, because large-scale phase separation can occur in these systems (79–81). In particular, ternary mixtures of saturated lipids, unsaturated lipids, and cholesterol can segregate into two coexisting fluid lipid domains, a liquid-ordered (Lo) and liquid-disordered (Ld) phase. The Martini model has proven very useful to assess the molecular nature of these domains at the nanoscale, information that has thus far eluded experimental determination. We have been able to show the spontaneous separation of a saturated phosphatidylcholine (PC)/unsaturated PC/cholesterol mixture into a liquid-ordered and a liquid-disordered phase with structural and dynamic properties closely matching experimental data (22). The lipid bilayer used in this study consisted of a ternary mixture of saturated di-C16:0PC (dipalmitoyl-phosphatidylcholine, DPPC), doubly unsaturated di-C18:2PC (dilinoleoyl-phosphatidylcholine, DLiPC), and cholesterol (molar ratio 0.42:0.28:0.3). The near-atomic resolution of the simulations reveals remarkable features of both domains, and of the boundary domain interface. Furthermore, we predicted the existence of a small surface tension between the monolayer leaflets, which drives registration of the domains. At the level of molecular detail, raft-like lipid mixtures show a surprising face with possible implications for many cell membrane processes (22).

It is intriguing to devise molecules that specifically bind at the boundary interface between the different lipid domains, thereby modifying the boundary properties while leaving the bulk regions unaltered. As they are supposed to reduce the line tension (or energetic cost) of the one-dimensional boundary interface, such molecules can be called linactants, analogous to surfactants (which



**Fig. 3.** Prediction of linactive behavior in multicomponent lipid membranes. Simulation snapshots of a ternary lipid mixture (DPPC/DLIPC/Cholesterol) to which a minor fraction (2 mol%) of a fourth component has been added that prefers to be at the domain boundary between the liquid-ordered and liquid-disordered phases. In panel **a**, the fourth component is POPC, a hybrid lipid with one saturated and one unsaturated tail. In panel **b**, the fourth component is a double-palmitoylated WALP peptide. Arrows point at a few individual molecules residing at the domain boundary, i.e., to be linactive.

modify the surface tension at an oil–water interface) (82). Possible line-active molecules could be, e.g., certain lipids, or lipid-anchored or transmembrane proteins. Recently we simulated the partitioning of other lipids as well as of transmembrane helices in the ternary mixture described above (65, 83). To this ternary mixture, small amounts of the fourth component were added. The idea was to introduce enough molecules to obtain proper statistics during the MD simulation, while perturbing the phase diagram of the ternary system as weakly as possible. Figure 3a shows typical line-active behavior of a palmitoyloleoyl-phosphatidylcholine (POPC) lipid. POPC, with both a saturated and an unsaturated tail, does not like either the Ld or the Lo phase, and has a tendency to accumulate at the domain boundary interface. In doing so, the line tension between the domains is reduced by about 30%, a strong effect given the low overall concentration of the POPC lipid (2 mol%). Similar line-active behavior was also observed for lyso-PC, a lipid with only one tail. In the case of TM peptides, most peptides studied, irrespective of hydrophobic matching conditions, have a strong preference for the Ld phase. Free energy calculations show that the enthalpic contribution due to the packing of the lipids drives the lateral sorting of the helices. In the Lo phase, the lipids are packed tightly together with cholesterol; this packing is distorted in the presence of a TM helix. The Ld phase, which is disordered, can accommodate the peptide with less impact on the lipid packing interactions. To drive peptides to the Lo phase, lipid anchors are required. Figure 3b shows the distribution of model TM peptides (WALPs) with multiple saturated lipid tails attached. The peptides cluster near the Lo/Ld domain boundary. Without the anchors, WALP is only found in the Ld phase.

#### 4.2. Mechanosensitive Channels in Action

Mechanosensitive channels form transmembrane pores to counteract pressure gradient buildup by balancing the osmotic conditions on either side of the cell membrane (84, 85). When activated, e.g., by increased membrane tension, the mechanosensitive channel of large conductance (MscL) forms a nonselective transmembrane channel capable of quickly transporting large amounts of solvent and solutes. MscLs are usually active no longer than a few hundreds of milliseconds with characteristic, rapidly flickering activation–deactivation cycles plainly visible in single-channel traces. By introducing ingenious point-mutations (86) at the channel walls the activation and deactivation of MscL can be controlled by ambient pH and/or light (87). This makes MscL a functional nanovalve with engineerable properties for a rapid, targeted drug release from a suitable nanocontainer (e.g., a stable liposome) acting as a drug delivery vehicle. MscL is a pentamer and each of the five subunits consists of a transmembrane and a cytoplasmic domain connected by a flexible linker (Fig. 4). The two transmembrane helices TM1 and TM2 and the N-terminal helix S1 are arranged in a crisscross manner with

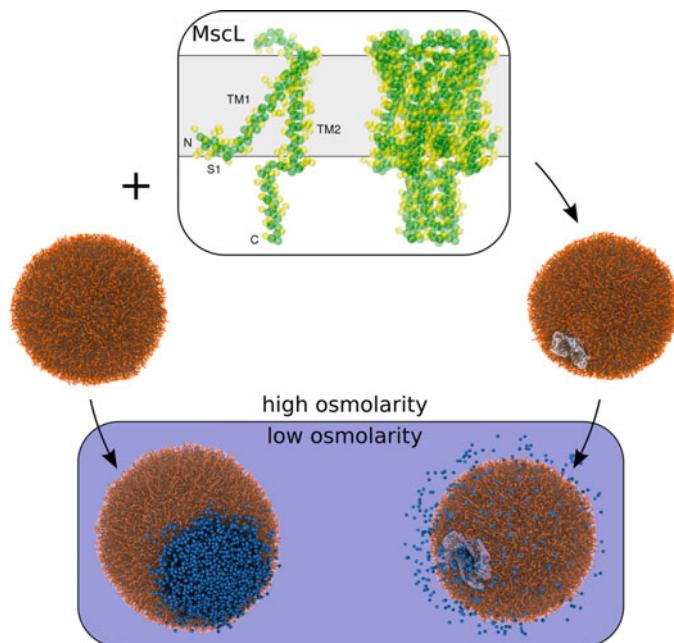


Fig. 4. Release of contents from pressurized liposomes. Response of a vesicle to a change in osmolarity with (*right*) and without (*left*) the presence of a mechanosensitive channel of large conductance (MscL). The MscL structure is shown in the *top panel*. One can appreciate the structure of a single subunit (*left*) and of the pentamer (*right*). The gray area represents the transmembrane region. *Left path:* an osmotic shock causes the vesicle to swell and ultimately to pop when its elastic limit is reached. *Right path:* under the same conditions but in the presence of an MscL the vesicle will survive by releasing the excess solvent (*small spheres*) through the opened channel.

TM1 and TM2 passing through the membrane, and S1 parallel to the membrane surface (Fig. 4). The transmembrane complex forms a ring-like structure with a hydrophobic lock region located slightly towards the cytoplasmic side from the center of the membrane. The cytoplasmic domains form a helical bundle at the mouth of the membrane channel.

The spontaneous gating of an MscL channel was studied using the Martini force field (10, 14). Originally, a lamellar setup was chosen (72), but more recently the gating of the channel in a small liposome was simulated (73). Simulations were carried out starting from the crystal structure of an MscL in its closed state (88) immersed in a small dioleoylphosphatidylcholine (DOPC) lipid vesicle measuring ~16 nm in diameter. Subsequently, a hypoosmotic shock condition was mimicked by gradually increasing the internal water content, and hence interior pressure, of the vesicle over a 0.5  $\mu$ s time window. Following the evolution of this system (over 10s of microseconds) under such a stress condition and comparing it to the case of a vesicle lacking the embedded MscL (cf. Fig. 4) it was found that (1) MscL activates at the limit of membrane elasticity, thereby releasing internal pressure and preventing membrane disruption; (2) The opening mechanism is asymmetric: the five subunits do open simultaneously but independently to accommodate the change in membrane thickness; (3) Flow of water through the channel is bi-directional; (4) Liposomal stress is relaxed on a submillisecond time scale. More details can be found elsewhere (73).

#### **4.3. Mitochondrial Supercomplex Formation**

Mitochondria are intracellular organelles, which are the power plants of most eukaryotic cells. The energy is produced in the oxidative phosphorylation (OxPhos) system and results from a series of electron transfer (oxidoreduction reactions) mainly carried out by four large membrane protein assemblies, the so-called respiratory chain complexes eukaryotic (I–IV) and by some small electron carriers (quinones and cytochrome c). The tortuous electron path through the respiratory chain triggers the transport of protons from the inside to the outside of the mitochondrial inner membrane, leading to an electrochemical gradient, which is used by the ATP synthase complex (complex V) for ATP synthesis. The respiratory chain complexes I–IV self-organize into supramolecular structures called respiratory “supercomplexes” (RSC) or “respirasomes” (89, 90). Examples are the supercomplexes consisting of complexes III and IV ( $\text{III}_2 + \text{IV}_1$  and  $\text{III}_2 + \text{IV}_2$ ) and even larger complexes additionally including complex I ( $\text{I}_1 + \text{III}_2 + \text{IV}_n$ ). While the existence and stoichiometry of supercomplexes has been established for a few organisms, the detailed structure (e.g., interfaces between complexes within supercomplexes) and most importantly the putative functional role and dynamics of those supercomplexes remains unclear. Notably, lipid membrane composition has been shown to be of great importance for the formation and stability of the

supercomplex  $\text{III}_2 + \text{IV}_2$ , and to have functional consequences. In that matter, cardiolipins (CLs), which compose 10–20% of the mitochondrial inner membrane, have been shown to be of primary importance (91–93). It is however not clear from the current experimental data by which mechanism CL might operate. Although cocrystallized CLs are found close to the surface of the protein in the transmembrane domain they are located deep inside a cavity and a helical subunit prevents them from having contact with bulk lipids and other embedded proteins.

Simulations of Martini models of complex  $\text{III}_2$  and IV embedded in a lipid membrane of POPC molecules containing ~20% cardiolipins were performed (94) and revealed preferential interfaces of CLs for both complexes  $\text{III}_2$  and IV (see Fig. 5). It is important to note that long simulations (10s of microseconds) were needed to reach acceptable convergence of the interfaces. Most interestingly, these actual *binding sites* differ significantly from the positions of CLs found in the crystal structures: they are at the surface of the proteins and not buried within the complexes. Furthermore self-assembly simulations of complexes  $\text{III}_2$  and IV in the presence of CLs showed that in the supercomplexes formed: (1) complexes conserve all the binding sites observed in the simulations when isolated; (2) supercomplexes make contact through those binding sites and therefore the binding sites are shared by the complexes. A typical example of a supercomplex formed is shown in Fig. 5d. These findings clearly show that bulk CLs strongly bind to the surface of the complexes reproducibly at specific locations and thereby provide anchors to form contacts with other complexes. Although this view contrasts with the earlier ideas that buried CLs were important for the formation of supercomplexes it allows the rationalization of the effect of the lack of cardiolipins on the formation of supercomplexes in a more intuitive and straightforward manner. Ongoing research in our laboratory aims at determining the precise role of the different sites and their degree of occupation in orienting the complexes relative to each other.

## 5. Outlook

In line with our constant efforts in improving the Martini coarse-grained force field and its range of applicability a few directions of development are currently being followed.

*New nonbonded form:* The use of a LJ 12-6 potential to describe the nonbonded interactions is, on hindsight, not the best choice. The steep repulsion leads to overstructuring of the Martini model in comparison to atomistic models, as evidenced for instance in the radial-distribution functions for simple alkanes (13). To reproduce

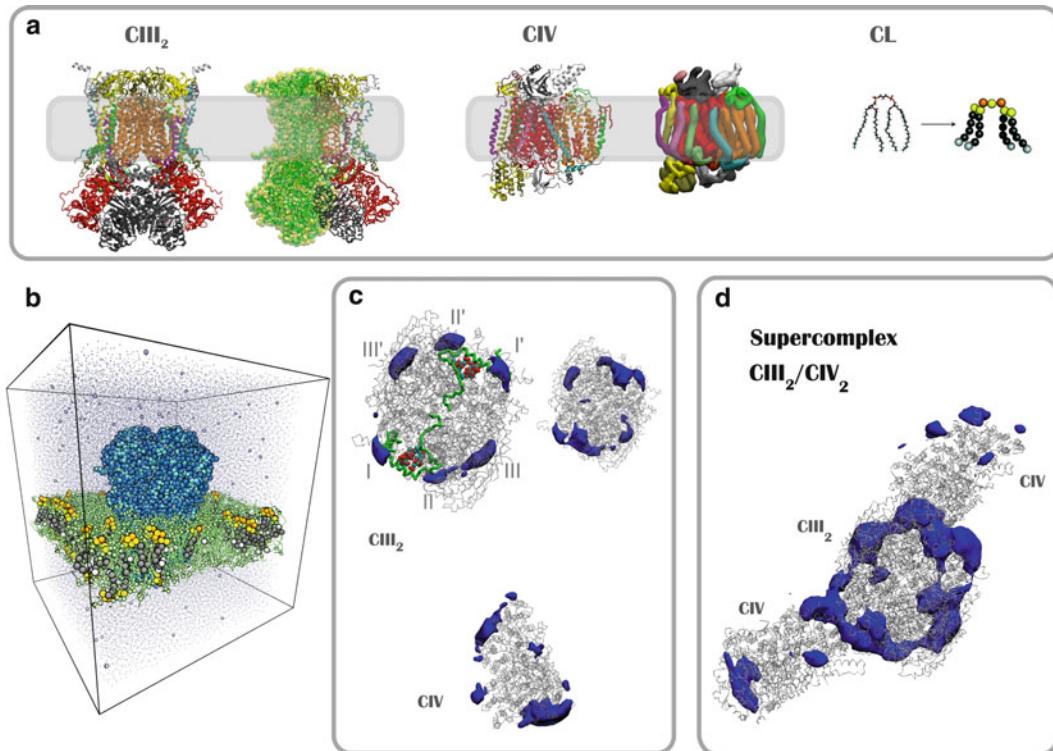


Fig. 5. Cardiolipins control mitochondrial supercomplex structure. The atomistic structures of complex-III homodimer ( $\text{CIII}_2$ ), complex-IV (CIV) and cardiolipins (CL) are shown in panel **a** together with their CG representation. In the case of  $\text{CIII}_2$  and CIV a few representations are used. One monomer of  $\text{CIII}_2$  is shown using vdW spheres. The second monomer is shown using a stick representation of the backbone ( $\text{C}_{\alpha}$ ) trace. The CIV CG model is represented by the density of the backbone beads obtained during a simulation. In panel **b** a typical system is shown. Here the  $\text{CIII}_2$  is embedded into a POPC membrane bilayer with 20 % of cardiolipins (*solid*). The aqueous phase is shown by small dots (*water beads*). Solvated ions are represented by spheres. In panel **c**,  $\text{CIII}_2$  (*top*) and CIV (*bottom*) are shown by gray transparent surfaces. The *solid* surfaces represent the regions of the system with high density of CLs. In panel **d**, an example is shown of the structure of the supercomplex  $\text{CIII}_2$ -CIV<sub>2</sub> as obtained by a self-assembly simulation. The solid surfaces represent the regions of the system with high density in CLs. The orientation of  $\text{CIII}_2$  in panels **c** and **d** is kept identical to ease the comparison of the CL high density regions in the individual complexes (panel **c**) and in the supercomplex.

the “softness” of real systems a power 9 or 10 repulsive term appears more appropriate, as in the recently reparameterized CG model of Klein and coworkers (95) that uses a LJ 9-6 form for most non-bonded interactions. The other drawback of the LJ 12-6 potential is the limited fluid range (see Subheading 2.4), preventing for instance the accurate reproduction of the experimental surface tension of hydrophilic compounds, including water. Possible solutions are given in the literature, such as switching to a power 4 attraction (96, 97), or to a different potential such as the Morse potential (98). For the Martini version 3.x we aim to change the nonbonded form to a softer potential with longer-ranged attraction.

*Foldable peptides and proteins:* The lack of structural flexibility of the current Martini protein backbone (see Subheadings 2.4 and 3) obviously limits the range of applicability of the model. While aiming at a coarse-grained model capable of folding a protein from its sequence may be overly ambitious, describing the backbone in a more realistic manner would be an appreciable improvement. For instance modeling the formation and realistic transition between secondary structure elements would be extremely beneficial. Available coarse-grained models for protein represent the backbone H-bonding in quite different manners and generally use a specific potential (74, 99–103). In contrast we want to keep a simple and generic. For instance, we have been exploring the possibility of using a fixed dipole (by the mean of two point charges of opposite sign) to represent the backbone polarity. The dipoles interact with other dipoles and particles through regular Coulomb and LJ potentials. Importantly, this description restores the directionality of backbone interactions. We have recently demonstrated proof of concept for generic polypeptides in the sense that secondary structure elements ( $\alpha$ -helices and  $\beta$ -sheets) may form, be stable and interconvert. The simplicity of the model allows easy tuning of the propensity of the backbone of each residue between extended and compact structures. This might be convenient if one wants to assign a certain preference for secondary structure at the residue level.

*Nucleotides:* One of the major classes of biomolecules that has not been thoroughly parameterized for Martini is the nucleotides, and the DNA and RNA molecules they constitute. Although there is a Martini model available for a small piece of DNA (51), a more generic approach is required to make the model applicable to all nucleic acids. We plan to perform such a parameterization analogous to our work on carbohydrates (42) described above. We will start with optimizing the description of the eight main nucleotides, focusing on thermodynamic partitioning behavior to select the best particle types, and comparing to atomistic simulations to select the bonded interactions. Subsequently we will consider conformational behavior of small oligonucleotides. Tuning of bonded parameters (angle and dihedral) will be used to result in mimicking the flexibility observed with atomistic models. The model can then be used to simulate large fragments of nucleic acids. Potentially an elastic network may be added (such as the ElNeDyn approach we developed for proteins) to assure a realistic behavior.

*Multiscale:* In combination, FG and CG models are even more powerful. One of the current challenges is to develop effective multiscale methods (104, 105), which combine the advantages of both levels of resolution. A useful, serial type of multiscale is the method of resolution transformation in which a CG configuration is back-mapped to an atomistic one. We have developed a

simple simulated-annealing based algorithm that can be used to perform resolution transformation for an arbitrary system (106). A subsequent simulation of the system with an atomistic force field allows one to assess, to some extent, the relevance of the configuration sampled at the CG level (107). A potentially powerful way to combine FG and CG force fields in parallel, is to define a specific region of interest to be modeled atomistically, whereas the rest of the system is described at the CG level. This so-called boundary approach, with a fixed separation between the two regions, is analogous to the idea of QM/MM methods in which quantum mechanics is combined with classical molecular mechanics. Exploratory work on boundary multiscaling methods combining FG and CG models has already been performed by several groups (24, 108, 109). We recently proposed a novel variation of the boundary method that is more general and requires only few additional parameters (110). Instead of deriving specific cross-interactions, we use virtual sites that couple the atomistic and CG degrees of freedom. These virtual sites are constructed from the center-of-mass of the underlying all-atom representation. The hybrid particles constituting the central molecule thus interact intramolecularly using FG forces, and intermolecularly according to the CG forces. An advantage is that no specific interactions between FG and CG beads need to be parameterized as is done in other boundary methods (24), making the method easily applicable to any system of interest. The method naturally combines the advantages of FG models (accurate description of the molecule of interest) and CG models (explicit treatment of the surrounding solvent at a speed 2–3 orders of magnitude larger compared to FG models). So far, the method has only been tested for simple systems consisting of pure butane, or dialanine peptides in either water or butane (110) as a proof of principle. Any type of CG (and FG) force field can be used, and results obtained with combining the Martini force field to the GROMOS force field have especially been encouraging. Within the field of biomolecules, the method appears ideally suited to study e.g., protein–ligand binding, where the active site and ligand are modeled in atomistic detail and the rest of the protein, together with the solvent, is coarse-grained. The combination of a flexible protein description and explicit solvent is envisioned to be important for molecular docking approaches (111, 112). Another area of potential applications is the complexation of proteins, with the surface of the protein in full detail and the protein core as well as the surrounding solvent at the CG level. Finally, our method offers an alternative approach to arrive at a “foldable” Martini model (see above), with the protein backbone in full atomic detail and the side chains and solvent at a CG level.

## Acknowledgments

The authors would like to thank the many people that have directly and indirectly contributed to the development of the Martini force field. In particular Alex de Vries and all the past and present members of the MD group in Groningen are acknowledged for their dynamism and enthusiasm in using, criticizing and improving Martini, as well as the groups of Peter Tieleman, Luca Monticelli, and Ilpo Vattulainen. Clement Arnarez, Martti Louhivuori, Lars Schafer, and Durba Sengupta provided data and images for the figures.

## References

- Smit B, Hilbers PAJ, Esselink K, Rupert LAM, Vanos NM, Schlijper AG (1990) Computer-simulations of a water oil interface in the presence of micelles. *Nature* 348:624–625
- Muller M, Katsov K, Schick M (2006) Biological and synthetic membranes: what can be learned from a coarse-grained description? *Phys Rep* 434:113–176
- Shillcock JC, Lipowsky R (2006) The computational route from bilayer membranes to vesicle fusion. *J Phys-Condens Mat* 18: S1191–S1219
- Venturoli M, Sperotto MM, Kranenburg M, Smit B (2006) Mesoscopic models of biological membranes. *Phys Rep* 437:1–54
- Elezgaray J, Laguerre M (2006) A systematic method to derive force fields for coarse-grained simulations of phospholipids. *Comput Phys Commun* 175:264–268
- Lyubartsev AP (2005) Multiscale modeling of lipids and lipid bilayers. *Eur Biophys J Biophys* 35:53–61
- Shelley JC, Shelley MY, Reeder RC, Bandyopadhyay S, Moore PB, Klein ML (2001) Simulations of phospholipids using a coarse grain model. *J Phys Chem B* 105:9785–9792
- Izvekov S, Violi A, Voth GA (2005) Systematic coarse-graining of nanoparticle interactions in molecular dynamics simulation. *J Phys Chem B* 109:17019–17024
- Marrink SJ, de Vries AH, Mark AE (2004) Coarse grained model for semiquantitative lipid simulations. *J Phys Chem B* 108:750–760
- Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 111:7812–7824
- Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676
- Chu JW, Ayton GS, Izvekov S, Voth GA (2007) Emerging methods for multiscale simulation of biomolecular systems. *Mol Phys* 105:167–175
- Baron R, Trzesniak D, de Vries AH, Elsener A, Marrink SJ, van Gunsteren WF (2007) Comparison of thermodynamic properties of coarse-grained and atomic-level simulation models. *Chem Phys Chem* 8:452–461
- Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink SJ (2008) The MARTINI coarse-grained force field: extension to proteins. *J Chem Theory Comput* 4:819–834
- Sengupta D, Marrink SJ (2010) Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes. *Phys Chem Chem Phys* 12:12987–12996
- Periole X, Cavalli M, Marrink SJ, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* 5:2531–2543
- Ramadurai S, Holt A, Schafer LV, Krasnikov VV, Rijkers DTS, Marrink SJ, Killian JA, Poolman B (2010) Influence of hydrophobic mismatch and amino acid composition on the lateral diffusion of transmembrane peptides. *Biophys J* 99:1447–1454
- Periole X, Huber T, Marrink SJ, Sakmar TP (2007) G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J Am Chem Soc* 129:10126–10132

19. Marrink SJ, Mark AE (2004) Molecular view of hexagonal phase formation in phospholipid membranes. *Biophys J* 87:3894–3900
20. Marrink SJ, Risselada J, Mark AE (2005) Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chem Phys Lipids* 135:223–244
21. Faller R, Marrink SJ (2004) Simulation of domain formation in DLPC-DSPC mixed bilayers. *Langmuir* 20:7686–7693
22. Risselada HJ, Marrink SJ (2008) The molecular face of lipid rafts in model membranes. *P Natl Acad Sci USA* 105:17367–17372
23. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) Gromacs: fast, flexible, and free. *J Comput Chem* 26:1701–1718
24. Shi Q, Izvekov S, Voth GA (2006) Mixed atomistic and coarse-grained molecular dynamics: simulation of a membrane-bound ion channel. *J Phys Chem B* 110:15045–15048
25. Bowers KJ, Chow E, Xu HF, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossaváry I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. Proceedings of the ACM/IEEE Conference on Supercomputing (SC06) Tampa, Florida
26. Shih AY, Arkhipov A, Freddolino PL, Schulten K (2006) Coarse grained protein-lipid model with application to lipoprotein particles. *J Phys Chem B* 110:3674–3684
27. Shih AY, Freddolino PL, Arkhipov A, Schulten K (2007) Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. *J Struct Biol* 157:579–592
28. Bond PJ, Holyoake J, Ivetac A, Khalid S, Sansom MSP (2007) Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *J Struct Biol* 157:593–605
29. Bond PJ, Sansom MSP (2006) Insertion and assembly of membrane proteins via simulation. *J Am Chem Soc* 128:2697–2704
30. Villa A, Mark AE (2002) Calculation of the free energy of solvation for neutral analogs of amino acid side chains. *J Comput Chem* 23:548–553
31. Yesylevskyy SO, Schafer LV, Sengupta D, Marrink SJ (2010) Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput Biol* 6:e1000810
32. Marrink SJ, Periole X, Tieleman DP, de Vries AH (2010) Comment on “On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models” by M. Winger, D. Trzesniak, R. Baron and W. F. van Gunsteren (2009) *Phys. Chem. Chem. Phys.*, 11: 1934, *Phys. Chem. Chem. Phys.* 12: 2254–2256
33. Xing CY, Faller R (2009) Coarse-grained simulations of supported and unsupported lipid monolayers. *Soft Matter* 5:4526–4530
34. Yano Y, Matsuzaki K (2006) Measurement of thermodynamic parameters for hydrophobic mismatch 1: Self-association of a transmembrane helix. *Biochem Us* 45:3370–3378
35. Ash WL (2009) Helix-helix interactions in membrane proteins probed with computer simulations. PhD Thesis (University of Calgary, Canada).
36. Winger M, Trzesniak D, Baron R, van Gunsteren WF (2009) On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models. *Phys. Chem. Chem. Phys.* 11:1934–1941
37. Marrink SJ, de Vries AH, Harroun TA, Katsaras J, Wassall SR (2008) Cholesterol shows preference for the interior of polyunsaturated lipid. *J Am Chem Soc* 130:10–11
38. Bennett WFD, MacCallum JL, Hinner MJ, Marrink SJ, Tieleman DP (2009) Molecular view of cholesterol flip-flop and chemical potential in different membrane environments. *J Am Chem Soc* 131:12714–12720
39. Dahlberg M (2007) Polymorphic phase behavior of cardiolipin derivatives studied by coarse-grained molecular dynamics. *J Phys Chem B* 111:7194–7200
40. Vuorela T, Catte A, Niemela PS, Hall A, Hyvonen MT, Marrink SJ, Karttunen M, Vattulainen I (2010) Role of lipids in spheroidal high density lipoproteins. *PLoS Comput Biol* 6:e1000964
41. Bulacu M, Periole X, Marrink SJ. (2012) In-silico design of robust boliplid membranes. *Biomacromol* 13:196–205
42. Lopez CA, Rzepiela AJ, de Vries AH, Dijkhuizen L, Hunenberger PH, Marrink SJ (2009) Martini coarse-grained force field: extension to carbohydrates. *J Chem Theory Comput* 5:3195–3210
43. Wu Z, Cui QA, Yethiraj A (2010) A New coarse-grained model for water: the importance of electrostatic interactions. *J Phys Chem B* 114:10524–10529
44. Leontiadou H, Mark AE, Marrink SJ (2006) Antimicrobial peptides in action. *J Am Chem Soc* 128:12156–12161
45. Sengupta D, Leontiadou H, Mark AE, Marrink SJ (2008) Toroidal pores formed by antimicrobial peptides show significant disorder. *BBA-Biomembranes* 1778:2308–2317

46. Wong-Ekkabut J, Baoukina S, Triampo W, Tang IM, Tieleman DP, Monticelli L (2008) Computer simulation study of fullerene translocation through lipid membranes. *Nat Nanotechnol* 3:363–368
47. Wallace EJ, Sansom MSP (2009) Carbon nanotube self-assembly with lipids and detergent: a molecular dynamics study. *Nanotechnology* 20:045101
48. Wallace EJ, Sansom MSP (2007) Carbon nanotube/detergent interactions via coarse-grained molecular dynamics. *Nano Lett* 7:1923–1928
49. Wallace EJ, Sansom MSP (2008) Blocking of carbon nanotube based nanoinjectors by lipids: A simulation study. *Nano Lett* 8:2751–2756
50. Gautieri A, Russo A, Vesentini S, Redaelli A, Buehler MJ (2010) Coarse-grained model of collagen molecules using an extended MARTINI force field. *J Chem Theory Comput* 6:1210–1218
51. Khalid S, Bond PJ, Holyoake J, Hawtin RW, Sansom MSP (2008) DNA and lipid bilayers: self-assembly and insertion. *J R Soc Interface* 5:S241–S250
52. Corsi J, Hawtin RW, Ces O, Attard GS, Khalid S (2010) DNA lipoplexes: formation of the inverse hexagonal phase observed by coarse-grained molecular dynamics simulation. *Langmuir* 26:12119–12125
53. Muller-Plathe F (2002) Coarse-graining in polymer simulation: from the atomistic to the mesoscopic scale and back. *Chem Phys Chem* 3:754–769
54. Lee H, Larson RG (2006) Molecular dynamics simulations of PAMAM dendrimer-induced pore formation in DPPC bilayers with a coarse-grained model. *J Phys Chem B* 110:18204–18211
55. Lee H, Larson RG (2008) Coarse-grained molecular dynamics studies of the concentration and size dependence of fifth- and seventh-generation PAMAM dendrimers on pore formation in DMPC bilayer. *J Phys Chem B* 112:7778–7784
56. Lee H, de Vries AH, Marrink SJ, Pastor RW (2009) A coarse-grained model for polyethylene oxide and polyethylene glycol: conformation and hydrodynamics. *J Phys Chem B* 113:13186–13194
57. Rossi G, Monticelli L, Puisto SR, Vattulainen I, Ala-Nissila T (2010) Coarse-graining polymers with the MARTINI force-field: polystyrene as a benchmark case. *Soft Matter* 7:698–708
58. Marrink SJ, Mark AE (2003) The mechanism of vesicle fusion as revealed by molecular dynamics simulations. *J Am Chem Soc* 125:11144–11145
59. Kasson PM, Kelley NW, Singhal N, Vrljic M, Brunger AT, Pande VS (2006) Ensemble molecular dynamics yields submillisecond kinetics and intermediates of membrane fusion. *P Natl Acad Sci USA* 103:11916–11921
60. Baoukina S, Tieleman DP (2010) Direct simulation of protein-mediated vesicle fusion: lung surfactant protein B. *Biophys J* 99:2134–2142
61. Smirnova YG, Marrink SJ, Lipowsky R, Knecht V (2010) Solvent-exposed tails as pre-stalk transition states for membrane fusion at low hydration. *J Am Chem Soc* 132:6710–6718
62. Baoukina S, Monticelli L, Marrink SJ, Tieleman DP (2007) Pressure-area isotherm of a lipid monolayer from molecular dynamics simulations. *Langmuir* 23:12617–12623
63. Duncan SL, Larson RG (2010) Folding of lipid monolayers containing lung surfactant proteins SP-B1-25 and SP-C studied via coarse-grained molecular dynamics simulations. *BBA-Biomembranes* 1798:1632–1650
64. Baoukina S, Monticelli L, Risselada HJ, Marrink SJ, Tieleman DP (2008) The molecular mechanism of lipid monolayer collapse. *P Natl Acad Sci USA* 105:10803–10808
65. Schäfer LV, Marrink SJ (2010) Partitioning of lipids at domain boundaries in model membranes. *Biophys J* 99(12):L91–L93
66. Fuhrmans M, Knecht V, Marrink SJ (2009) A single bicontinuous cubic phase induced by fusion peptides. *J Am Chem Soc* 131:9166–9167
67. Polyansky AA, Ramaswamy PE, Volynsky PE, Sbalzarini IF, Marrink SJ, Efremov RG (2010) Antimicrobial peptides induce growth of phosphatidylglycerol domains in a model bacterial membrane. *J Phys Chem Lett* 1:3108–3111
68. Khalfa A, Tarek M (2010) On the antibacterial action of cyclic peptides: insights from coarse-grained MD simulations. *J Phys Chem B* 114:2676–2684
69. Bond PJ, Wee CL, Sansom MSP (2008) Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. *Biochem-U*s 47:11321–11331
70. Psachoulia E, Fowler PW, Bond PJ, Sansom MSP (2008) Helix-helix interactions in membrane proteins: coarse-grained simulations of glycophorin a helix dimerization. *Biochemistry* 47:10503–10512

71. Sansom MSP, Scott KA, Bond PJ (2008) Coarse-grained simulation: a high-throughput computational approach to membrane proteins. *Biochem Soc T* 36:27–32
72. Yefimov S, van der Giessen E, Onck PR, Marrink SJ (2008) Mechanosensitive membrane channels in action. *Biophys J* 94:2994–3002
73. Louhivuori M, Risselada HJ, van de Rgiessen E, Marrink SJ (2010) Release of content through mechano-sensitive gates in pressurized liposomes. *Proc Natl Acad Sci USA* 107 (46):19856–19860. doi:[10.1073/pnas.1001316107](https://doi.org/10.1073/pnas.1001316107)
74. Wei GH, Mousseau N, Derreumaux P (2004) Complex folding pathways in a simple beta-hairpin. *Proteins* 56:464–474
75. Treptow W, Marrink SJ, Tarek M (2008) Gating motions in voltage-gated potassium channels revealed by coarse-grained molecular dynamics simulations. *J Phys Chem B* 112:3277–3282
76. Lycklama JA, Bulacu M, Marrink SJ, Driessens AJM (2010) Immobilization of the plug domain inside the SecY channel allows unrestricted protein translocation. *J Biol Chem* 285:23747–23754
77. Hatakeyama M, Faller R (2007) Coarse-grained simulations of ABA amphiphilic tri-block copolymer solutions in thin films. *Phys Chem Chem Phys* 9:4662–4672
78. Lee H, Larson RG (2008) Lipid bilayer curvature and pore formation induced by charged linear polymers and dendrimers: The effect of molecular shape. *J Phys Chem B* 112:12279–12285
79. Baumgart T, Hess ST, Webb WW (2003) Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Nature* 425:821–824
80. Veatch SL, Polozov IV, Gawrisch K, Keller SL (2004) Liquid domains in vesicles investigated by NMR and fluorescence microscopy. *Biophys J* 86:2910–2922
81. Lingwood D, Ries J, Schwille P, Simons K (2008) Plasma membranes are poised for activation of raft phase coalescence at physiological temperature. *P Natl Acad Sci USA* 105:10005–10010
82. Trabelsi S, Zhang S, Lee TR, Schwartz DK (2008) Linactants: Surfactant analogues in two dimensions. *Phys Rev Lett* 100:037802
83. Schäfer LV, de Jong DH, Holt A, Rzepiela AJ, de Vries AH, Poolman B, Killian JA, Marrink SJ (2010) Lipid packing drives the segregation of transmembrane helices into disordered lipid domains in model membranes. *Proc Natl Acad Sci USA* 108(4):1343–1348
84. Martinac B, Buechner M, Delcour AH, Adler J, Kung C (1987) Pressure-sensitive Ion channel in escherichia-colii. *P Natl Acad Sci USA* 84:2297–2301
85. Perozo E, Rees DC (2003) Structure and mechanism in prokaryotic mechanosensitive channels. *Curr Opin Struc Biol* 13:432–442
86. Kocer A, Walko M, Meijberg W, Feringa BL (2005) A light-actuated nanovalve derived from a channel protein. *Science* 309:755–758
87. Kocer A, Walko M, Feringa BL (2007) Synthesis and utilization of reversible and irreversible light-activated nanovalves derived from the channel protein MscL. *Nat Protoc* 2:1426–1437
88. Steinbacher S, Bass R, Strop P, Rees DC (2007) Structures of the prokaryotic mechanosensitive channels MscL and MscS. *Curr Top Membr* 58:1–24
89. Lenaz G, Genova ML (2009) Structural and functional organization of the mitochondrial respiratory chain: a dynamic super-assembly. *Int J Biochem Cell B* 41:1750–1772
90. Bultema JB, Braun HP, Boekema EJ, Kouril R (2009) Megacomplex organization of the oxidative phosphorylation system by structural analysis of respiratory supercomplexes from potato. *BBA-Bioenergetics* 1787:60–67
91. Zhang M, Mileykovskaya E, Dowhan W (2002) Gluing the respiratory chain together—Cardiolipin is required for supercomplex formation in the inner mitochondrial membrane. *J Biol Chem* 277:43553–43556
92. Zhang M, Mileykovskaya E, Dowhan W (2005) Cardiolipin is essential for organization of complexes III and IV into a supercomplex in intact yeast mitochondria. *J Biol Chem* 280:29403–29408
93. Pfeiffer K, Gohil V, Stuart RA, Hunte C, Brandt U, Greenberg ML, Schagger H (2003) Cardiolipin stabilizes respiratory chain supercomplexes. *J Biol Chem* 278:52873–52880
94. Arnarez C, Elezgaray J, Mazat JP, Marrink SJ, Periole X. Evidence for cardiolipin binding sites on the membrane exposed surface of the respiratory chain complex III. In preparation
95. Shinoda W, DeVane R, Klein ML (2010) Zwitterionic lipid assemblies: molecular dynamics studies of monolayers, bilayers, and vesicles using a new coarse grain force field. *J Phys Chem B* 114:6836–6849
96. Shinoda W, Devane R, Klein ML (2007) Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. *Mol Simulat* 33:27–36

97. DeVane R, Klein ML, Chiu CC, Nielsen SO, Shinoda W, Moore PB (2010) Coarse-grained potential models for phenyl-based molecules: I. Parametrization using experimental data. *J Phys Chem B* 114:6386–6393
98. Chiu CC, DeVane R, Klein ML, Shinoda W, Moore PB, Nielsen SO (2010) Coarse-grained potential models for phenyl-based molecules: II. Application to fullerenes. *J Phys Chem B* 114:6394–6400
99. Bereau T, Deserno M (2009) Generic coarse-grained model for protein folding and aggregation. *J Chem Phys* 130:235106
100. Liwo A, Pincus MR, Wawak RJ, Rackovsky S, Scheraga HA (1993) Prediction of protein conformation on the basis of a search for compact structures—test on avian pancreatic-polypeptide. *Protein Sci* 2:1715–1731
101. Maisuradze GG, Senet P, Czaplewski C, Liwo A, Scheraga HA (2010) Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *J Phys Chem A* 114:4471–4485
102. Ha-Duong T (2010) Protein backbone dynamics simulations using coarse-grained bonded potentials and simplified hydrogen bonds. *J Chem Theory Comput* 6:761–773
103. Alemani D, Collu F, Cascella M, Dal Peraro M (2010) A nonradial coarse-grained potential for proteins produces naturally stable secondary structure elements. *J Chem Theory Comput* 6:315–324
104. Ayton GS, Noid WG, Voth GA (2007) Multiscale modeling of biomolecular systems: in serial and in parallel. *Curr Opin Struc Biol* 17:192–198
105. Peter C, Kremer K (2009) Multiscale simulation of soft matter systems—from the atomistic to the coarse-grained level and back. *Soft Matter* 5:4357–4366
106. Rzepiela AJ, Schafer LV, Goga N, Risselada HJ, De Vries AH, Marrink SJ (2010) Software news and update reconstruction of atomistic details from coarse-grained structures. *J Comput Chem* 31:1333–1343
107. Rzepiela AJ, Sengupta D, Goga N, Marrink SJ (2010) Membrane poration by antimicrobial peptides combining atomistic and coarse-grained descriptions. *Faraday Discuss* 144:431–443
108. Neri M, Anselmi C, Cascella M, Maritan A, Carloni P (2005) Coarse-grained model of proteins incorporating atomistic detail of the active site. *Phys Rev Lett* 95:218102
109. Nielsen SO, Bulo RE, Moore PB, Ensing B (2010) Recent progress in adaptive multiscale molecular dynamics simulations of soft matter. *Phys Chem Chem Phys* 12:12401–12414
110. Rzepiela AJ, Louhivuori M, Peter C, Marrink SJ (2011) Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. *Phys Chem Chem Phys* 13:10437–10448
111. Zoete V, Grosdidier A, Michelin O (2009) Docking, virtual high throughput screening and in silico fragment-based drug design. *J Cell Mol Med* 13:238–248
112. Andrusier N, Mashiach E, Nussinov R, Wolfson HJ (2008) Principles of flexible protein-protein docking. *Proteins* 73:271–289

# Chapter 21

## Multiscale Molecular Modeling

Matej Praprotnik and Luigi Delle Site

### Abstract

We review the basic theoretical principles of the adaptive resolution simulation scheme (AdResS). This method allows to change molecular resolution on-the-fly during a simulation by changing the number of degrees of freedom in specific regions of space where the required resolution is higher than in the rest of the system. We also report about recent extensions of the method to the continuum and quantum regimes.

**Key words:** Multiscale modeling, Adaptive resolution simulation, AdResS

---

### 1. Introduction

Many problems in condensed matter, material science, and chemistry are multiscale in nature, meaning that the interplay between different scales plays a fundamental role. An exhaustive description of the related physical phenomena requires, from a theoretical or computational point of view, the simultaneous treatment of all the scales involved. This is a prohibitive task not only because of the large computational resources required but above all because the large amount of data produced would mostly contain information not essential to the problem analyzed; actually they may overshadow the underlying fundamental physics or chemistry of the system. A solution to this problem is that of treating the problems via multiscale approaches. In this case one simplifies the model of the physical system to the largest extent possible while keeping all the necessary details of the system where this is required. Multiscale methods have been developed and successfully applied to study solid state systems, where the atomistic models were either combined with the finite elements method (1–5) or linked to a quantum mechanical model (6, 7). In particular, for some systems (e.g., the solvation of a molecule) some regions (the first two or three solvation shells

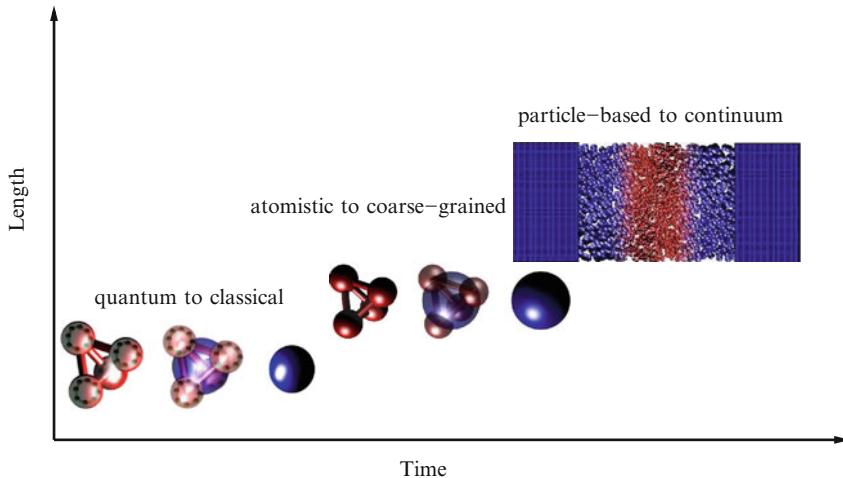


Fig. 1. Multiscale modeling: concurrent couplings. At the smallest scale the coupling is between the quantum and the classical (or coarse-grained) molecular resolution. The quantum resolution is here represented by the path-integral/polymer-ring representation of atoms (8). Going to a larger scale, one goes from the atomistic to the coarse-grained resolution(9) and finally at mesoscopic/macrosopic scales one couples particle-based representations to the continuum(10).

around the solute) require a higher resolution than the rest, but at the same time, because of the large fluctuations, an open boundary between the regions is required so that there is free exchange of particles. For that one requires adaptive resolution simulation schemes (see Fig. 1) which allow molecules to change resolution according to the region where they are instantaneously located. From a technical point of view, this means that the space is partitioned in regions characterized by different molecular resolutions where molecules can freely diffuse, changing their representation while keeping the overall thermodynamic equilibrium of the system (11, 12). The aim of this chapter is to review an adaptive simulation method, the adaptive resolution simulation (AdResS) method (9, 13), which has been shown to be particularly robust in coupling scales from the quantum to the classical atomistic up to the continuum (8, 14, 15).

## 2. Adaptive Resolution Simulation

### 2.1. Theoretical Framework

In this section we report the basic theoretical principles employed in the development of the AdResS. Let us consider a liquid of  $N$  molecules in a simulation box with a volume  $V$ , which is divided into two equally large domains  $A$  and  $B$ . In the domain  $A$ , we represent molecules on a low-resolution level, while the domain  $B$  is described using a higher resolution representation, for example, all-atom resolution. The number  $n$  of degrees of freedom (DOFs)

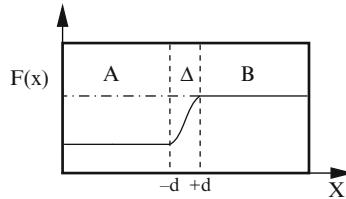


Fig. 2. Free energy density profile across the simulation box.

per molecule, which we explicitly treat in our description, is hence lower in the coarse-grained domain  $A$  and higher in the all-atom domain  $B$ . The two domains are in thermodynamic equilibrium and freely exchanging molecules, that is, the liquid is homogeneous across the whole simulation box. The necessary boundary condition for the thermodynamic equilibrium between the two representations is

$$\mu_A = \mu_B, \quad p_A = p_B, \quad T_A = T_B, \quad (1)$$

where  $\mu_A$ ,  $p_A$ , and  $T_A$ , and  $\mu_B$ ,  $p_B$ , and  $T_B$  are the chemical potentials, pressures, and temperatures of the liquid in the both domains, respectively (9, 11, 16). The question that arises at this point is how to achieve that the condition (1) is satisfied within a molecular dynamics simulation where the liquid has the same structure on both sides if analyzed on the lower resolution level. In an attempt to answer this question, we plot in Fig. 2 the free-energy density  $F = F(x)$  profile of the system associated to the DOFs that we explicitly consider in the simulation. Since the free energy is an extensive quantity, its value is lower in the domain  $A$  because the number of DOFs  $n_A$  is lower than in the domain  $B$ . For a smooth transition between the two resolutions, we introduce a transition regime  $\Delta$  (9) at the interface where the molecules slowly change their representation. In this regime they are in equilibrium with their actual surrounding and change continuously until the region of the new representation is reached. The molecules “arrive” fully equilibrated into the surrounding described by the new representation. The number of explicitly treated DOFs is  $n = n(x)$  with  $n_A = \text{const}_A$ ;  $n_B = \text{const}_B$ ; and  $n_A = n(x)$ . The system is in equilibrium which implies  $\lim_{x \rightarrow -d^-} \frac{\partial F_A(x)}{\partial x} = \lim_{x \rightarrow d^+} \frac{\partial F_B(x)}{\partial x} = 0$   $\Rightarrow \lim_{x \rightarrow -d^-} \frac{\partial n_A(x)}{\partial x} = \lim_{x \rightarrow d^+} \frac{\partial n_B(x)}{\partial x} = 0$ , and  $\partial F_A / \partial N_A + \phi = \mu_A$  and  $\partial F_B / \partial N_B = \mu_B$ , where  $\phi$  is the free energy per molecule associated with the integrated out DOFs. In the adaptive resolution scheme presented here, we do not book-keep the integrated out DOFs in the coarse-grained models. Rather, we reduce the many-body potential of the higher resolution representation into a reduced effective potential and keep only two-body terms. We do not treat explicitly the one-body terms, which depend only on the temperature (17), and hence do not contribute to the intermolecular forces.

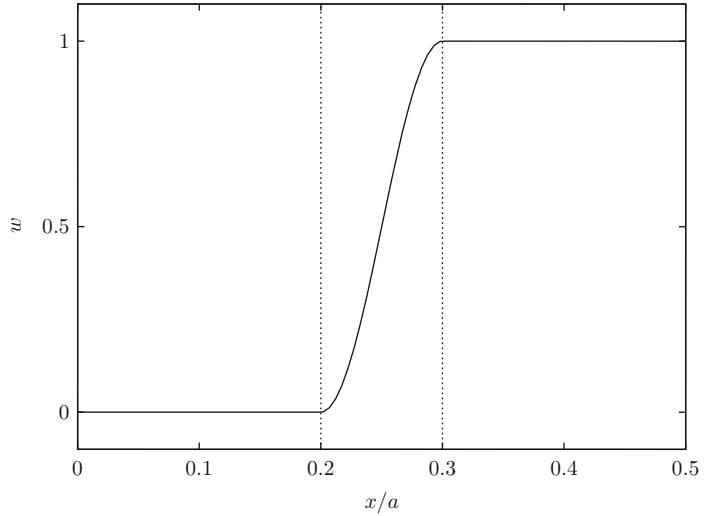


Fig. 3. The weighting function.

Instead, they can be viewed, in the adaptive scheme, as the equivalent of a latent heat. As explained later, in our model this part of the free energy is supplied or removed by a local thermostat. Of course, if we explicitly book-kept the one-body terms (three-body and higher terms are also omitted for numerical efficiency) in the effective potential, the free-energy profile would be flat across the whole simulation box (the dashed line in Fig. 2) (17) (P. Español, 2010, private communication). The free-energy density profile allows us to define a weighting function which has certain similarities with an order parameter (9). The weighting function  $w$  that determines the level of resolution across the system is presented in Fig. 3. It is introduced in such a way that  $w = 1$  and  $w = 0$  correspond to the high- and low-resolution regions, respectively, while the values  $0 < w < 1$  correspond to the transition regime. Thus, crossing the transition regime, the number of DOFs changes in a continuous manner. The change of resolution can be seen as a change in the dimensionality of the phase space of the switching DOF. If the particular DOF is fully considered ( $w = 1$ ), then the phase space has its full dimensionality. On the other hand, if it is completely switched off ( $w = 0$ ), it does not contribute to any statistical average, and thus, its dimensionality is zero. In the transition regime, the partially switched-on DOF contributes to statistical averages according to its weight  $w$ . To properly mathematically describe the continuous change of the phase space dimensionality and define thermodynamic properties, for example, temperature, in the transition regime, we resort to fractional calculus (18–22) and divide the transition regime into thin slabs, each with the different values  $w$  between 0 and 1.

For the fractional quadratic DOF  $p$  with the weight  $w$ , we can then write the partition function as

$$\begin{aligned}
\exp(-\beta F_p) &= C \int \exp(-\beta p^2/2) dV_w \\
&= 2C \int_0^\infty \exp(-\beta p^2/2) |p|^{w-1} \frac{dp}{\Gamma(w)} \\
&= \frac{2^{w/2} C \Gamma(w/2)}{\Gamma(w)} \beta^{-w/2} \sim \beta^{-w/2}, \tag{2}
\end{aligned}$$

where we introduced the infinitesimal volume element of the fractional configurational space defined as  $dV_w = |p|^{w-1} dp / \Gamma(w) = dp^w / (w\Gamma(w))$ . Here  $\Gamma$  is the gamma function (16). It directly follows  $\langle K_w \rangle = \frac{d\langle \beta F_p \rangle}{d\beta} = \frac{w}{2\beta} = \frac{w k_B T}{2}$ , where  $\langle K_w \rangle$  is the average kinetic energy per fractional quadratic DOF with the weight  $w$ . This is a generalization of the equipartition principle to non-integer quadratic DOFs (16, 23). In equilibrium  $T_A = T_B = T_\Delta = T$ , and thus,  $n_w \sim w$ . Using this framework the temperature in the transition region is correctly defined. This enables us to employ the local thermostat to control the free-energy difference between the two levels of resolution (24). The corresponding free energy for a generic quadratic switchable DOF  $p$  is (12)

$$F_p = \mu_p^{kin}(w) = -kT \log \left[ \int e^{-\beta p^2} d^w p \right], \tag{3}$$

and hence, the total explicit contribution of the entire set of switchable DOFs per molecule is

$$\mu^{kin}(w) = \sum_{DOF} \mu_p^{kin}(w). \tag{4}$$

The kinetic component to  $\phi$ , that is, the latent heat (in principle one could also include the quadratic intramolecular potential terms (17)), is<sup>1</sup>

$$\phi(w)^{kin} = \mu_B^{kin} - \mu^{kin}(w). \tag{5}$$

The analytical solution of Eq. 3 is:

$$\mu_p^{kin}(w) \sim C k T \left( \frac{w}{2} \right) \log(T) \tag{6}$$

where  $C$  is a constant,  $k$  the Boltzmann constant, and  $T$  is the temperature. Equation 6 is the ideal gas kinetic contribution to the chemical potential coming from the internal DOFs. Usually in a simulation with single-resolution representation of the molecules, this contribution to the chemical potential is ignored being only a trivial constant depending only on temperature (17). In our case, where the DOFs of interest might continuously change in going from

---

<sup>1</sup>As defined by Eqs. 3 and 4,  $\mu^{kin}(0) \neq \mu_A^{kin}$ . In fact,  $\mu_A^{kin} = \mu^{kin}(0) + \phi(0)^{kin} = \mu_B^{kin}$ .  $\mu^{kin}(w)$  represents only the contribution of the switched-on DOFs to the kinetic part of chemical potential. The rest is included in  $\phi(w)^{kin}$ .

one representation to another, each DOF in the transition region contributes differently according to the corresponding value of  $w(x)$ . Thus, the latent heat defined by Eq. 5 is crucial for keeping the thermodynamic equilibrium between two levels of resolution in the adaptive resolution simulations. The kinetic part of the free energy depends in the first order linearly on  $w$ , which determines the slope of the free-energy density profile in the transition regime in Fig. 2.

Next, we shall couple the two levels of resolution in a molecular dynamics approach: AdResS is built on a force-based approach (11). The interpolation formula for the pair force between molecules  $\alpha$  and  $\beta$  writes as

$$\mathbf{F}_{\alpha\beta} = w(X_\alpha)w(X_\beta)\mathbf{F}_{\alpha\beta}^B + [1 - w(X_\alpha)w(X_\beta)]\mathbf{F}_{\alpha\beta}^A \quad (7)$$

where  $\mathbf{F}_{\alpha\beta}^A$  is the force obtained from the potential of representation  $A$ , and  $\mathbf{F}_{\alpha\beta}^B$ , the one obtained from the potential of representation  $B$ ;  $w(X)$  is the switching function and depends on the center of mass positions  $X_\alpha$  and  $X_\beta$ , of the two interacting molecules. The basic idea of the scheme given by Eq. 7 is to enable the molecule to find its correct orientation in the liquid once it is given a random orientation at the low-resolution/transition regime boundary (9). Even though there might be some overlaps with neighboring molecules, the atomistic interactions are turned off ( $w = 0$ ) at that boundary. As the molecule approaches to the high-resolution regime, the atomistic interactions are gradually turned on, and the molecule on-the-fly finds its proper orientation. As it turns out, it is important to interpolate the forces and not the interaction potentials in order to satisfy Newton's third law (16, 23, 25). This is crucial for the local linear momentum conservation and proper diffusion of molecules across the transition regime. Each time a molecule leaves (or enters) the coarse-grained region, it gradually gains (or loses) its vibrational and rotational DOFs while retaining its linear momentum. The change in resolution carried out by AdResS is not time reversible as a given molecule in the low-resolution domain  $A$  corresponds to many orientations and configurations of the corresponding molecule in the high-resolution domain  $B$  (15). Since time reversibility is essential for energy conservation (26), AdResS does not conserve energy. In particular, the force in Eq. 7 is in general not conservative in the transition region (i.e., in general  $\oint \mathbf{F}_{\alpha\beta} \cdot d\mathbf{r} \neq 0$ ) (23, 25). Hence, to supply or remove the latent heat associated with the switch of resolution (see Eq. 5), we use a DPD thermostat (24, 27). The thermostat forces do not enter into the AdResS interpolating scheme, Eq. 7, instead they are added to the AdResS (9).

Using properly derived effective pair interactions  $\mathbf{F}_{\alpha\beta}^A$  (see Sect. 2), the boundary conditions as given by Eq. 1 are satisfied. The AdResS scheme (7) has been successfully applied to a liquid of tetrahedral molecules (9, 13), a generic macromolecule in solvent (28), and liquid water (14, 29). As already stated, to perform adaptive resolution simulations requires equilibrium between the

different regimes, which is facilitated by a transition zone. However, the necessary condition for thermodynamic equilibrium between two different representations as well as the transition zone, that is, the chemical potential, pressure, and temperature equivalence, can in some complex systems not be assured by a mere derivation of the effective pair interactions between coarse-grained molecules. Furthermore, one can also consider situations where we interface different molecular models at the same level of resolution (e.g., flexible and rigid atomistic water) with free, unhindered exchange of molecules among the regions of different molecular representation. A barrier free exchange of molecules over the border lines of molecular resolution is required to properly account for fluctuations. The number of DOFs in such cases remains the same, but forces acting on each atom are different in the different regions. To treat these scenarios, we recently introduced a generalization of AdResS (12), which allows to couple rather loosely connected molecular representations, that is, it maintains two different representations with, in general, different chemical potentials ( $\mu_A \neq \mu_B$ ) in thermodynamic equilibrium. In the generalized approach, we extend the original scheme, Eq. 7, by subtracting a thermodynamic force  $\mathbf{F}^{TD}$ . The total, force on molecule  $\alpha$  is after subtracting a thermodynamic force  $\mathbf{F}^{TD}$ ,

$$\mathbf{F}_\alpha = \sum_{\beta \neq \alpha} \left( w(X_\alpha) w(X_\beta) \mathbf{F}_{\alpha\beta}^B + [1 - w(X_\alpha) w(X_\beta)] \mathbf{F}_{\alpha\beta}^A \right) - \mathbf{F}^{TD}(X_\alpha) \quad (8)$$

where  $F_x^{TD} = -\frac{\partial \mu^{exc}}{\partial x}$  and  $\mu^{exc}$  is the excess chemical potential due to the intermolecular interactions (12). AdResS is a nonconservative scheme, and hence, the potential is not defined in the transition regime. Therefore, to calculate numerically the excess chemical potential, we proceed as follows. We divide the simulation box into regions of force fields  $A$  and  $B$  and the transition region in between. The region  $A$  is characterized by the value of the switching function  $w_0 = 0$ . The region  $B$  is characterized by the value of the switching function  $w_{N+1} = 1$ . In the transition regime, the value of  $w$  in the actual simulations varies continuously. Here we approximate this by discretizing  $w$  into  $N$  steps  $w_1, w_2, \dots, w_{N-1}, w_N$ . For any fixed value of  $w^2$ , the energy function is well-defined, and the excess chemical potential then is defined as  $\mu^{exc}(x_i) = \mu_{w_i}^{exc}$ , where the  $\mu_{w_i}^{exc}$  is the chemical potential of the molecules in a bulk system of the specific representation of  $w_i$ . To calculate numerically each  $\mu^{exc}(w_i)$ , one can use standard particle insertion methods. Repeating this procedure with all values of  $w_i$  leads to a position-dependent excess chemical potential  $\mu^{exc}(x)$ , which represent the second contribution to  $\phi$  (12).

---

<sup>2</sup>This is the value that one obtains by using the insertion methods in a hybrid system exclusively composed of hybrid molecules with a fixed level of resolution  $0 \leq w = w(x) = const. \leq 1$  corresponding to a fixed bulk value  $\mu_{w(x)}$ .

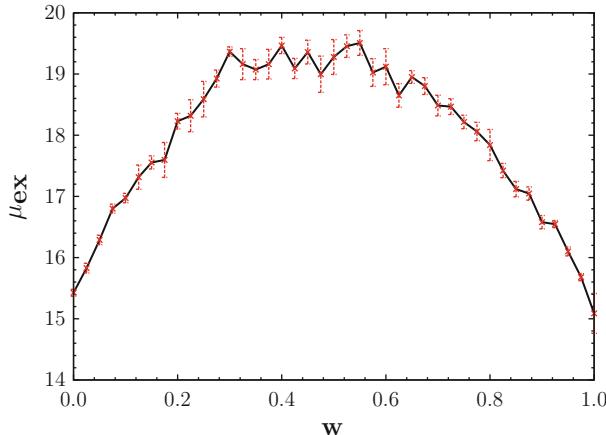


Fig. 4. The excess chemical potential. Reprinted with permission from (12). Copyright 2010, American Institute of Physics.

As a test case, we applied the above generalized scheme to a liquid of tetrahedral molecules (12). An intrinsic, though numerically negligible, problem in the original formulation of AdResS was the fact that we found some evident (but small) density variations through the transition regime due to the linear interpolation of forces in AdResS (13). We showed that employing the above derived scheme and introducing the corresponding thermodynamic force this problem can be solved. Figure 4 shows  $\mu^{exc}$ . The system is set up in such a way that the equation of state is the same in both the coarse-grained and all atom regimes at the temperature and density of the current simulation. Because of that,  $\mu^{exc}(x)$  is the same for  $w = 1$  and for  $w = 0$ . The resulting  $F_x^{TD}$  and the density profile are depicted in Fig. 5. The application of the thermodynamic force flattens the artificial density fluctuations by preserving the thermodynamic equilibrium between the two levels of resolution.

## 2.2. Mapping of Structural Properties

As discussed in the previous section, we have to map the structure of the low-resolution representation as close as possible to the reference high-resolution counterpart. This is necessary because we want the molecules to adapt, when they enter the high-resolution domain from the low-resolution region (through the transition one), as quickly as possible to the new environment. To this end, we reduce the many-body potential of the higher resolution representation into a reduced effective potential (30) using any of the standard methods, for example, iterative Boltzmann inversion (31–33), an iterative inverse statistical mechanics approach (34), force-matching scheme (35, 36), extended ensemble approach (37, 38). For more details of these methods, we refer the reader to the chapter by William G. Noid. Note that due to numerical efficiency, we consider only pair effective interactions and omit the higher or one-body

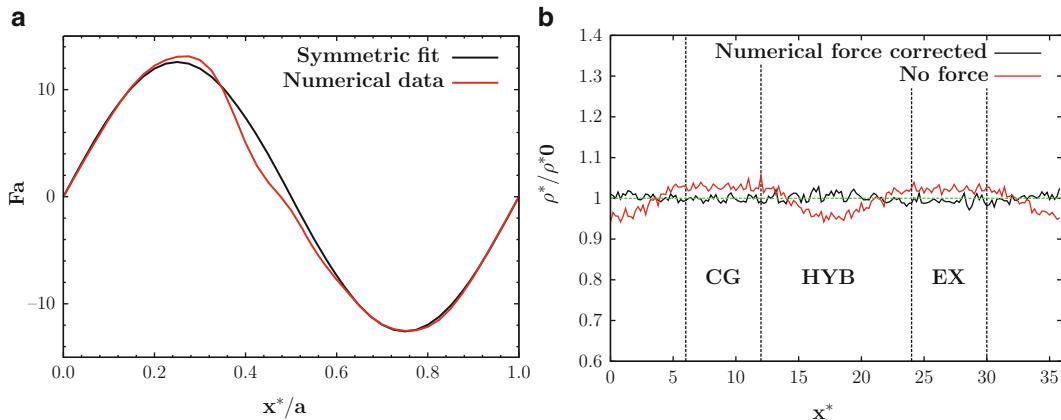


Fig. 5. Application of the thermodynamic force—reducing the density fluctuations in the transition regime. Reprinted with permission from (12). Copyright 2010, American Institute of Physics. (a) Thermodynamic force (b) Density across the simulation box.

terms (see the Section 1). Another issue is also whether to match the pressure or compressibility of the models. One can namely not match both (39). We need pressure equality to satisfy equilibrium condition (1), but on the other hand, the models should have the same compressibility in order to have the same thermodynamic fluctuations. To circumvent this problem, we resort here to the generalized AdResS scheme, Eq. 8, where via the effective potential we assure the same compressibility of the models while using the thermodynamic force we guarantee the equilibrium between the two levels of resolution.

As it turns out, AdResS is quite robust against the details of the low-resolution model, and we do not need to map the radial distribution functions exactly to the linethickness. This was demonstrated in the case of liquid water (15) where a non-perfect effective potential was deliberately used to study the robustness of the method, as depicted in Fig. 6.

### 2.3. Mapping of Dynamical Properties

Having mapped the structural properties via the effective potential derivation, one typically loses the control over dynamical properties of coarse-grained models. Due to soft effective interactions, the transport coefficient of the coarse-grained models differs quite substantially from the reference atomistic values, that is, diffusion constants and viscosities are too high/low, respectively. In our approach in order to match the dynamics of both levels of resolution, we slow down the dynamics of coarse-grained models using position-dependent local thermostats, as described below.

#### 2.3.1. Position Dependent Langevin Thermostat

If one is not concerned about hydrodynamics, one can use the stochastic Langevin thermostat where the transport coefficient is tuned by changing the strength of the coupling with the

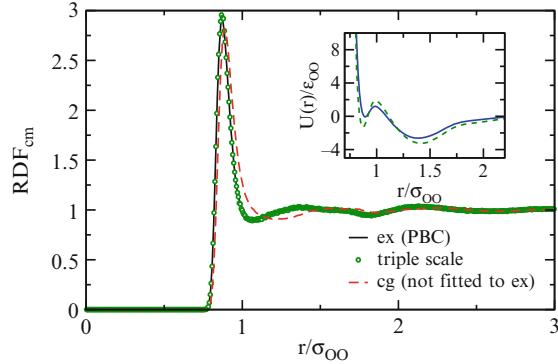


Fig. 6. RDFs and corresponding effective pair potential for water. Reprinted with permission from (15). Copyright 2009, American Institute of Physics.

thermostat. The Langevin equation with a position dependent coefficient  $\Gamma(x)$  can be written as (29)

$$m_i dv_i / dt = F_i - m_i \Gamma(x) v_i + R_i(x, t) \quad (9)$$

where  $R_i(x, t)$  is

$$\langle R_i(x, t) \rangle = 0, \quad (10)$$

$$\langle R_i(x, t_1) R_j(x, t_2) \rangle = 2\Gamma(x) m_i k T \delta(t_1 - t_2) \delta_{ij} \quad (11)$$

$$\Gamma(w) = \begin{cases} \Gamma_{cg} & \text{if } w \leq 0.6 \\ \alpha w + \beta & \text{if } 0.6 < w \leq 1.0. \end{cases} \quad (12)$$

This choice provides a simple interpolation between the two limit values of  $\Gamma(0.6) = \Gamma(0) = \Gamma_{cg} = 15 \text{ ps}^{-1}$  and  $\Gamma(1) = \Gamma_{all-atom} = 5 \text{ ps}^{-1}$ . The parameters  $\alpha$  and  $\beta$  are  $-25 \text{ ps}^{-1}$  and  $30 \text{ ps}^{-1}$ , respectively. In the adaptive resolution simulations, the force  $F_i$  is defined by the AdResS scheme. The results depicted in Fig. 7 clearly show that by employing our position dependent Langevin thermostat, we effectively slow down the dynamics of the coarse-grained models to be equal to the reference atomistic one.

### 2.3.2. Transverse DPD Thermostat

We have shown in the previous example that the coarse-grained dynamics can be slowed down by increasing the effective friction in the coarse-grained system using the position-dependent Langevin thermostat. However, it is well-known that the Langevin thermostat does not reproduce the correct hydrodynamics, that is, the hydrodynamic interactions are nonphysically screened. To correctly describe hydrodynamic interactions, one has to resort to the dissipative particle dynamics (DPD) thermostat. For tuning the transport coefficient of liquids, we extended the standard DPD thermostat (27) by including the damping of the transverse components of the relative velocity, yet keeping the advantages of conserving Galilei invariance and within our error bar also hydrodynamics (24).

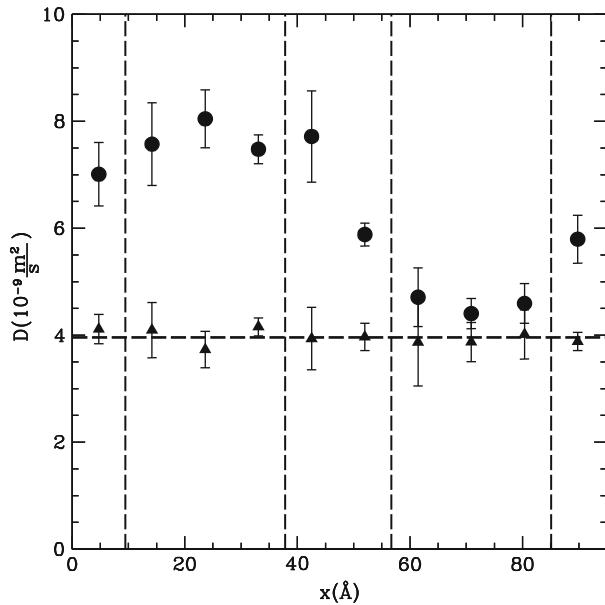


Fig. 7. Diffusion constant profile using the standard and position dependent Langevin thermostats. Reprinted with permission from (29). Copyright 2008, American Institute of Physics.

The Transverse DPD thermostat is introduced as follows (24):

$$\dot{\vec{p}}_i = \vec{F}_i^C + \vec{F}_i^D + \vec{F}_i^R, \quad (13)$$

where the first, second, and third term on the RHS denote conservative, damping, and random forces, respectively (40). The damping and random forces are expressed as:

$$\vec{F}_{ij}^D = -\zeta w^D(r_{ij}) \overset{\leftrightarrow}{P}_{ij}(\vec{r}_{ij}) \vec{v}_{ij}, \quad (14)$$

$$\vec{F}_{ij}^R = \sigma w^R(r_{ij}) \overset{\leftrightarrow}{P}_{ij}(\vec{r}_{ij}) \vec{\theta}_{ij}, \quad (15)$$

where  $\zeta$  and  $\sigma$  are the friction constant and the noise strength, respectively. Here  $\overset{\leftrightarrow}{P}_{ij}(\vec{r}_{ij})$  is a projection operator

$$\overset{\leftrightarrow}{P} = \overset{\leftrightarrow}{P}^T = \overset{\leftrightarrow}{P}^2, \quad (16)$$

which is symmetric in the particle indices ( $\overset{\leftrightarrow}{P}_{ij} = \overset{\leftrightarrow}{P}_{ji}$ ).

The noise vector  $\vec{\theta}_{ij}$

$$\langle \vec{\theta}_{ij}(t) \otimes \vec{\theta}_{kl}(t') \rangle = 2 \overset{\leftrightarrow}{I} (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) \delta(t - t') \quad (17)$$

is antisymmetric in the particle indices ( $\vec{\theta}_{ij} = -\vec{\theta}_{ji}$ ). The fluctuation-dissipation theorem is thus satisfied. If we choose the projector along the interatomic axis between particle  $i$  and  $j$   $\overset{\leftrightarrow}{P}_{ij}(\vec{r}_{ij}) = \hat{r}_{ij} \otimes \hat{r}_{ij}$ , we

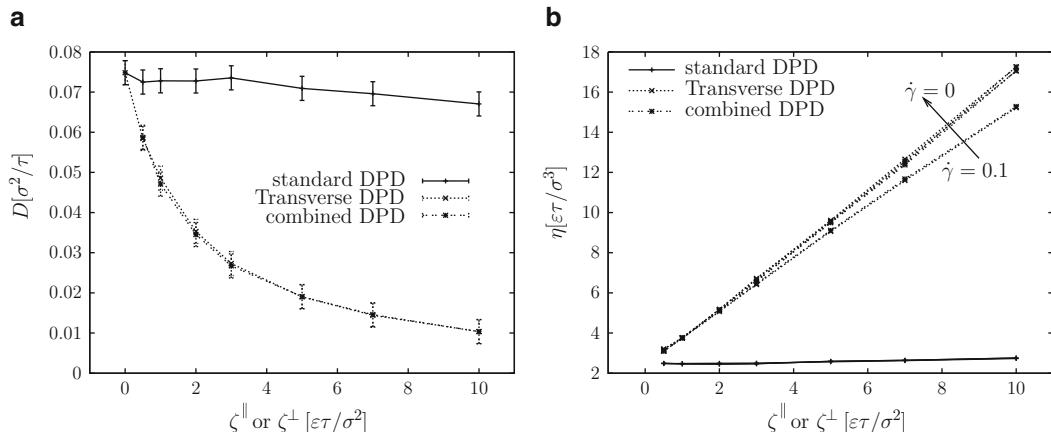


Fig. 8. Tuning the transport coefficients using transverse DPD thermostat. The figure is taken from (24). Reproduced by permission of The Royal Society of Chemistry. (a) Diffusion constant (b) Shear viscosity.

obtain the standard DPD thermostat, whereas  $\overset{\leftrightarrow}{P}_{ij}(\vec{r}_{ij}) = \overset{\leftrightarrow}{I} - \hat{r}_{ij} \otimes \hat{r}_{ij}$  yields the Transverse DPD thermostat.

As it turns out, the transport coefficient is rather insensitive to the value of the friction constant for damping the central relative velocities using the standard DPD thermostat. However, they are very sensitive to the value of the friction constant for damping the transverse relative velocities with the Transverse DPD thermostat as Fig. 8 shows (24).

### 3. AdResS: Technical Implementation

Very recently the AdResS scheme has been included into an open-source package ESPResSo for soft matter simulations (41, 42). This flexible implementation of AdResS will allow the simulation community to easily adapt the AdResS setup for their particular problems of interest. We hope and envisage that this will further boost the usage of AdResS.

### 4. Extension to Continuum Resolution

Recently, we have extended the AdResS scheme with the continuum description of a liquid modeled by the Navier-Stokes equation (10, 15). The triple-scale scheme was derived by combining two dual-scale schemes: AdResS, which couples the atomic and coarse-grained scales within a molecular dynamics (MD) simulation

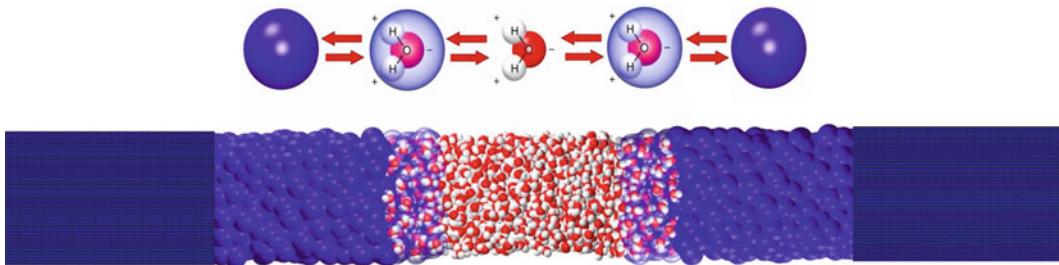


Fig. 9. The triple-scale model of liquid water (15).

framework, and a hybrid flux-exchange based continuum-MD scheme (HybridMD) developed by Delgado-Buscalioni et al. (43). The resulting triple-scale model consists of a particle-based micro-mesoscale MD region, which is divided into a central atomistic and a surrounding mesoscopic domain, and a macroscopic region modeled on the hydrodynamic continuum level as presented in Fig. 9 for liquid water. The central idea of the triple-scale method is to gradually increase the resolution as one approaches to the region of interest, that is, the atomistic region. The continuum and MD domains exchange information via mass and momentum fluxes. These fluxes are conserved across the interface between continuum and MD regions. The triple-scale approach is designed for molecular simulations of open domains with relatively large molecules, either in the grand canonical ensemble or under nonequilibrium conditions.

## 5. Extension to Quantum Level

A special treatment deserves the conceptual extension of AdResS to quantum problems. In general an adaptive approach that allows to pass from a quantum to a classical description and vice versa would require more than the mere change of number of DOFs. It requires the smooth passage from different kinds of physical principles. In fact while classical mechanics is governed by a deterministic evolution, quantum mechanics is characterized by the probabilistic character. For systems where electrons are explicitly treated, a classical-quantum adaptive scheme properly based on the Schrödinger equation would lead to the problem of variable number of particles, that is a varying particle normalization condition as the system evolves. So far schemes that treat electrons adaptively are based on practical solutions and not on a complete and consistent theoretical framework (6, 44). Instead if the quantum particles are the atoms (without considering explicitly the electrons), then the quantum problem, for some properties, can be mapped on an effective classical one. In this case, as a matter of fact, the adaptive coupling occurs between classical descriptions. The idea of

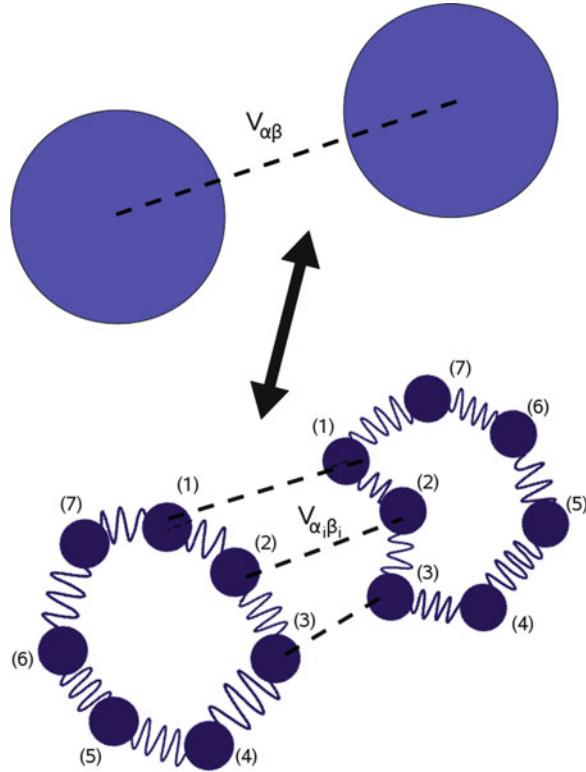


Fig. 10. Pictorial representation of the idea of path integral description of atoms. Two atoms,  $\alpha$  and  $\beta$  interacting in the classical approach as a rigid spheres become two polymer rings whose interatomic classical interaction is distributed over pairs of bead  $\alpha_i\beta_j$ . The neighboring beads  $i$ ,  $i + 1$  (and  $i$ ,  $i - 1$ ) are kept together by an elastic potential whose constant depends on the temperature of the system. The delocalization of the atoms onto several beads characterizes its quantum nature.

quantum atoms is based on the path-integral description (45); this latter describes the quantum atoms as classical polymer rings; in this context, the beads of a polymer ring are fictitious classical particles, as presented in Fig. 10. The simulation can then be performed adaptively so that one can have an atomistic or coarse-grained resolution in one region and a path-integral resolution, where each atom is represented by a polymer ring, in another with free exchange between the two regions. In this situation, the principles of AdResS apply straightforwardly as to the case of two classical regions characterized by different numbers of DOFs. The application to study the equilibrium statistical properties of a liquid of tetrahedral molecules has shown that indeed this idea is rather robust both conceptually and numerically (8) (see “quantum to classical” part of Fig. 1).

---

## 6. Conclusions and Outlook

The development of AdResSs is a subject of growing interest within the community of condensed matter, material science, and chemical physics. The interest in this kind of approach rises from the fact that, as discussed in this chapter, it may efficiently tackle the problem of interplay between different scales. This occurs by properly treating, simultaneously, all the relevant molecular DOF in each region of the simulation box and, because of the free diffusion of molecules from one region to another, properly accounting for density fluctuations. The practical consequences are an optimal employment of computational resources and an efficient analysis of the simulation data. In fact it allows for the reduction of computational costs by treating high-resolution models, which are computationally demanding, only in restricted regions and at the same time assures that details not relevant for the problem are not processed in regions where high resolution is not required. These last aspects allow in turn to derive a clear understanding of the basic physical features characterizing a given problem, avoiding that an excess of details overshadows the essential physics. In this sense, multiscale approaches as AdResS will allow not only to treat problems that before were prohibitive but also understand how the different scales, expressed by the molecular resolutions interfaced, are connected to each other. For example, in the solvation of a molecule, one may be interested to understand to which extend the local hydrogen bond network in the hydration shell is influenced by that of the bulk. To this aim one may employ AdResS as an analysis tool and perform simulations with varying sizes of the atomistic region interfaced with a coarse-grained region or even with the continuum where hydrogen bonds are not present. Depending on the minimum size of the atomistic region required to properly reproduce hydration properties, one can then comment on the extension of the influence of the hydrogen bond network on the hydration shell (46). In a similar spirit, one may think of the capture or release of a proton from a solvated molecule; in this case the quantum mechanical behavior of the proton passage would be described by, for example, the path-integral approach in a small region, while the large surrounding of bulk water can be represented with a coarser resolution. By further extending the AdResS methodology to continuum, one can go beyond equilibrium molecular dynamics simulations using periodic boundary conditions and study nonequilibrium processes, for example, fluid flows relevant for nanofluidics applications.

## Acknowledgements

Over the years we have collaborated with many colleagues and students on the topics described in this chapter. Out of them we would like to especially thank K. Kremer, R. Delgado-Buscalioni, C. Junghans, A. B. Poma, S. Poblete, C. Clementi, B. Lambeth, and S. Matysiak for fruitful collaboration and many discussions.

## References

1. Rafii-Tabar H, Hua L, Cross M (1998) A multi-scale atomistic-continuum modelling of crack propagation in a two-dimensional macroscopic plate. *J Phys: Condens Matter* 10:2375–2387
2. Broughton JQ, Abraham FF, Bernstein N, Kaxiras E (1999) Concurrent coupling of length scales: methodology and application. *Phys Rev B* 60:2391–2403
3. Hadjiconstantinou NG (1999) Combining atomistic and continuum simulations of contact-line motion. *Phys Rev E* 59:2475–2478
4. Smirnova JA, Zhigilei LV, Garrison BJ (1999) A combined molecular dynamics and finite element method technique applied to laser induced pressure wave propagation. *Comput Phys Commun* 118:11–16
5. Rottler J, Barsky S, Robbins MO (2002) Cracks and crazes: on calculating the macroscopic fracture energy of glassy polymers from molecular simulations. *Phys Rev Lett* 89:148304
6. Csanyi G, Albaret T, Payne MC, DeVita A (2004) 'Learn on the fly': a hybrid classical and quantum-mechanical molecular dynamics simulation. *Phys Rev Lett* 93:175503
7. Heyden A, Lin H, Truhlar DG (2007) Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations. *J Phys Chem B* 111:2231–2241
8. Poma AB, Delle Site L (2010) Classical to path-integral adaptive resolution in molecular simulation: towards a smooth quantum-classical coupling. *Phys Rev Lett* 104:250201
9. Praprotnik M, Delle Site L, Kremer K (2005) Adaptive resolution molecular dynamics simulation: changing the degrees of freedom on the fly. *J Chem Phys* 123:224106
10. Delgado-Buscalioni R, Kremer K, Praprotnik M (2008) Concurrent triple-scale simulation of molecular liquids. *J Chem Phys* 128:114110
11. Praprotnik M, Delle Site L, Kremer K (2008) Multiscale simulation of soft matter: from scale bridging to adaptive resolution. *Annu Rev Phys Chem* 59:545–571
12. Poblete S, Praprotnik M, Delle Site L, Kremer K (2010) Coupling different levels of resolution in molecular simulations. *J Chem Phys* 132:114101
13. Praprotnik M, Delle Site L, Kremer K (2006) Adaptive resolution scheme (adress) for efficient hybrid atomistic/mesoscale molecular dynamics simulations of dense liquids. *Phys Rev E* 73:066701
14. Praprotnik M, Matysiak S, Delle Site L, Kremer K, Clementi C (2007) Adaptive resolution simulation of liquid water. *J Phys: Condens Matter* 19:292201
15. Delgado-Buscalioni R, Kremer K, Praprotnik M (2009) Coupling atomistic and continuum hydrodynamics through a mesoscopic model: application to liquid water. *J Chem Phys* 131:244107
16. Praprotnik M, Kremer K, Delle Site L (2007) Adaptive molecular resolution via a continuous change of the phase space dimensionality. *Phys Rev E* 75:017701
17. Klapp SHL, Diestler DJ, Schoen M (2004) Why are effective potentials 'soft'? *J Phys: Condens Matter* 16:7331–7352
18. Nonnenmacher TF (1990) Fractional integral and differential equations for a class of Levy-type probability densities. *J Phys A: Math Gen* 23:L697S–L700S
19. Hilfer R (ed) (2000) Applications of fractional calculus in physics. World Scientific Publishing, Co. Pte. Ltd, Singapore
20. Cottrill-Shepherd K, Naber M (2001) Fractional differential forms. *J Math Phys* 42:2203–2212
21. Tarasov VE (2004) Fractional generalization of Liouville equations. *Chaos* 14:123–127
22. Tarasov VE (2005) Fractional systems and fractional Bogoliubov hierarchy equations. *Phys Rev E* 71:011102
23. Praprotnik M, Kremer K, Delle Site L (2007) Fractional dimensions of phase space variables:

- a tool for varying the degrees of freedom of a system in a multiscale treatment. *J Phys A: Math Theor* 40:F281–F288
24. Junghans C, Praprotnik M, Kremer K (2008) Transport properties controlled by a thermostat: an extended dissipative particle dynamics thermostat. *Soft Matter* 4:156–161
  25. Delle Site L (2007) Some fundamental problems for an energy-conserving adaptive-resolution molecular dynamics scheme. *Phys Rev E* 76:047701
  26. Janežič D, Praprotnik M, Merzel F (2005) Molecular dynamics integration and molecular vibrational theory: I. New symplectic integrators. *J Chem Phys* 122:174101
  27. Soddeemann T, Dünweg B, Kremer K (2003) Dissipative particle dynamics: a useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys Rev E* 68:046702
  28. Praprotnik M, Delle Site L, Kremer K (2007) A macromolecule in a solvent: adaptive resolution molecular dynamics simulation. *J Chem Phys* 126:134902
  29. Matysiak S, Clementi C, Praprotnik M, Kremer K, Delle Site L (2008) Modeling diffusive dynamics in adaptive resolution simulation of liquid water. *J Chem Phys* 128:024503
  30. Henderson RL (1974) A uniqueness theorem for fluid pair correlation functions. *Phys Lett* 49A:197–198
  31. Soper AK (1996) Empirical monte carlo simulation of fluid structure. *Chem Phys* 202:295–306
  32. Tschöp W, Kremer K, Hahn O, Batoulis J, Bürger T (1998) Simulation of polymer melts. II. From coarse-grained models back to atomistic description. *Acta Polym* 49:75–79
  33. Reith D, Pütz M, Müller-Plathe F (2003) Deriving effective mesoscale potentials from atomistic simulations. *J Comput Chem* 24:1624–1636
  34. Lyubartsev AP, Laaksonen A (1995) Calculation of effective interaction potentials from radial distribution functions: a reverse monte carlo approach. *Phys Rev E* 52:3730–3737
  35. Izvekov S, Parrinello M, Burnham CB, Voth GA (2004) Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: a new method for force-matching. *J Chem Phys* 120:10896–10913
  36. Izvekov S, Voth GA (2005) Multiscale coarse graining of liquid-state systems. *J Chem Phys* 123:134105
  37. Mullinax JW, Noid WG (2009) Generalized yvon-born-green theory for molecular systems. *Phys Rev Lett* 103:198104
  38. Mullinax JW, Noid WG (2009) Extended ensemble approach for deriving transferable coarse-grained potentials. *J Chem Phys* 131:104110
  39. Wang H, Junghans C, Kremer K (2009) Comparative atomistic and coarse-grained study of water: what do we lose by coarse-graining? *Eur Phys J E* 28:221
  40. Español P, Warren P (1995) Statistical mechanics of dissipative particle dynamics. *Europhys Lett* 30:191–196
  41. Limbach HJ, Arnold A, Mann BA, Holm C (2006) Espresso—an extensible simulation package for research on soft matter systems. *Comput Phys Comm* 174:704–727 <http://www.espresso.mpg.de>
  42. Junghans C, Poblete S (2010) A reference implementation of the adaptive resolution scheme in espresso. *Comput Phys Comm* 181:1449
  43. Fabritiis GD, Delgado-Buscalioni R, Coveney PV (2006) Multiscale modeling of liquids with molecular specificity. *Phys Rev Lett* 97:134501
  44. Bulo R, Ensing B, Vischer L (2009) Toward a practical method for adaptive qm/mm simulations. *J Chem Theor Comp* 5:2212
  45. Tuckermann M (2002) Path integration via molecular dynamics. Quantum simulations of complex many-body systems: from theory to algorithms, NIC Series, John von Neumann Institute for Computing, pp 269–298
  46. Bradley P, Lambeth C, Junghans C, Kremer K, Clementi C, Delle Site L (2010) On the Locality of Hydrogen Bond Networks at Hydrophobic Interfaces. *J Chem Phys* 133:221101

# Chapter 22

## Coarse-Grained Models for Protein Folding and Aggregation

Philippe Derreumaux

### Abstract

Coarse-grained models for protein folding and aggregation are used to explore large dimension scales and timescales that are inaccessible to all-atom models in explicit aqueous solution. Combined with enhanced configuration search methods, these simplified models with various levels of granularity offer the possibility to determine equilibrium structures, compare folding kinetics and thermodynamics with experiments for single proteins and understand the dynamic assembly of amyloid proteins leading to neurodegenerative diseases. I shall describe recent progress in developing such models, and discuss their potentials and limitations in probing the folding and misfolding of proteins with computer simulations.

**Key words:** Coarse-grained models, Proteins, Simulations, Structures, Thermodynamics, Dynamics, Kinetics, Misfolding, Aggregation

---

### 1. Introduction

Determining the lowest free energy (native) structure of a protein (1) and the mechanisms by which an amino acid sequence folds spontaneously into its native state are two theoretical challenges in modern structural biology. Under physiological conditions, most proteins display well-defined 3D structures, (2) though this paradigm is challenged by (1) intrinsically disordered proteins, which are expected to represent 30% of eukaryotic genomes encoded proteins with wholly or partially disordered structures, where folding is intimately coupled to binding to their biological targets (3), and (2) amyloid-proteins, free of any high sequence identity, which self-assemble into fibrils with a common cross- $\beta$  structure (4). In spite of constant efforts, structure determination is difficult for large size systems by NMR and X-ray diffraction, and identification of the successive conformational changes between the denatured and the native state is still limited by the time- and space-resolution of single-molecule biophysical experiments; e.g., a resolution of one picture per millisecond using

kinetic terahertz absorption spectroscopy (5). For amyloids, the situation is even more critical, since the insolubility of the end products and the transient character of the early formed oligomers preclude the use of the standard tools of structural biology.

In principle, all-atom molecular dynamics (MD) simulations (6) and generalized ensemble simulations such as replica-exchange MD (REMD) simulations (7, 8) in explicit aqueous solution are ideal methodologies for characterizing protein dynamics and thermodynamics, respectively. Irrespective of the accuracy of atomistic force fields, which approximate the Hamiltonians used in quantum chemistry that do not separate the electronic from the nuclear degrees of freedom, (9) generating dynamics on timescales beyond microseconds for fast folding or large-scale conformational change is still a tour de force. Typically, all-atom protein MD simulations in explicit solvent cover 0.1–1 μs (10–12). Obtaining equilibrium configurations and convergent thermodynamic properties such as the variation of the heat capacity as a function of temperature is even more difficult to reach, independently of the improved algorithms and the parameter used to overcome the multiple-minima problem (e.g., temperature or Hamiltonian) because the number of replicas scales with the square root of the number of degrees of freedom (13). Very often, the calculated free energy surface in explicit solvent is questioned in terms of accuracy and convergence (14). At the final level of complexity, we find binary or multicomponent protein interactions in solution or crowded environments, and notably the oligomerization/fibrillization of amyloid proteins occurring on the timescales of days in vitro and leading to neurodegenerative diseases (15), such as Alzheimer and Parkinson which constitute a major societal challenge. With increased life expectancy and an aging population, the number of patients with Alzheimer's disease is expected to reach 80 million in 2040, and there is no disease-modifying treatment thus far (16).

To push the boundaries in timescales and system sizes, it is essential to average out all fast degrees of motion and use coarse-grained (CG) models that group multiple atoms into beads via effective force fields. In their pioneered 1975 study, Levitt and Warshel offered the foundation of coarse-graining for protein folding. Using a two-bead representation for each amino acid, namely the C $\alpha$  and the centroid of the side chain, an effective implicit solvent force field and successive minimizations and normal mode thermalizations, they could fold the 58-residue BPTI protein within 6.5 Å from the experimental structure (17). Since then, various levels of granularity have been developed for protein folding. Figure 1 shows the all-atom, coarse-grained and C $\alpha$  representations for a tetrapeptide. The main difficulty in coarse-graining and reducing the number of particles is to design potential energy functions or force fields that retain the physics of the all-atom explicit solvent system in terms of structure, thermodynamics and dynamics (18, 19).

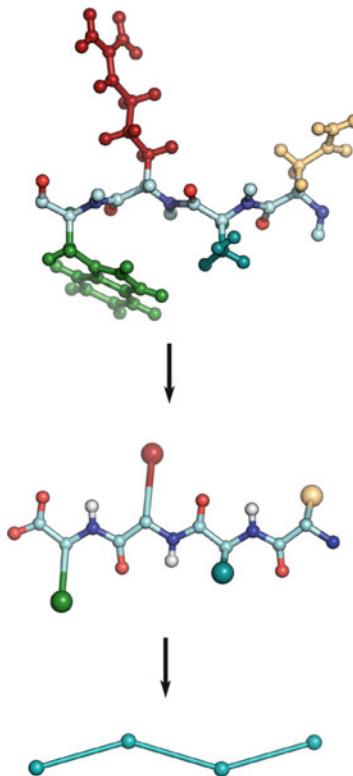


Fig. 1. The all-atom, OPEP coarse-grained and  $C\alpha$  representations for a tetrapeptide model.

This report presents a perspective on the state-of-the-art CG models used for globular protein folding. Elastic one-well (20) and double-well (21) network models and CGs for transmembrane protein, lipids, DNA, and RNA can be found in this book or elsewhere (22, 23). All the models vary substantially in their grainning levels and applications. Chapter 1 covers all the minimalist models represented by an ensemble of  $C\alpha$  beads. Most of them are biased towards the native structure by the simple or extended  $G_0$  models so as to examine protein thermodynamics, kinetics and structure ensembles in aqueous or crowded environments. Chapter 2 reports CG models including explicitly some side-chain information that can be used for folding one single or multiple proteins. Finally, I present the current force fields that can be used simultaneously for protein folding and aggregation.

---

## 2. Minimalist One Bead Models

The structure-based continuum  $G_0$ -like model was proposed for the simulation of protein folding where the protein conformation is represented by an ensemble of  $C\alpha$  atoms and the interactions

are specified to stabilize the native structure (24). The simpler analytical potential energy form is such that only attractive interactions following a Lennard-Jones potential are assigned to the native contacts whereas excluded volume repulsive interactions are assigned to the nonnative ones. In the native centric Gō-like model, all contacts are controlled by a unique parameter  $\varepsilon$ . A native contact is defined if the C $\alpha$  atoms are separated by at least three residues, and two heavy atoms, one from each residue, are less than a cutoff distance ranging from 4.5 to 6.5 Å apart in the experimental structure.

Using this structure-based Gō model with Langevin dynamics simulations combined with the WHAM technique, the dynamics and thermodynamics of any protein can be fully determined and compared with experimental data if available. Overall, the application of the minimalist Gō model on small proteins support the energy landscape theory where the folding landscape resembles a partially rough funnel riddled with traps where the protein can transiently resides (25–27). Employing a benchmark of 52 proteins ranging from 37 to 151 residues, Munoz et al. compared the experimental midpoint folding rates and stabilities in water at 298 K with model predictions and found that protein size is the primary factor determining folding rates and protein stability (28). More recently, a correlation coefficient of 0.92 was found between the simulated and experimentally determined folding rates using a total of one de novo and seven natural proteins (29). In this study, the simulated folding rates were estimated by calculating the mean first-passage time for folding transition from 400 randomly chosen unfolded states (with the fraction of native contacts  $Q < 0.05$ ) at the transition temperature,  $T_m$ , identified by a peak in the heat capacity profile. Folding cooperativity was characterized from the calorimetric criterion, i.e., the calculated ratio between the van't Hoff enthalpy (peak Cp value) and the calorimetric enthalpy (area under the Cp scan), the analysis of the transition of average energy or folding fraction ( $Q$ ) as a function of  $T$ , the conformation distributions as a function of  $T$  and the free energy profiles as a function of  $Q$ . Using the native centric Gō-model, DS119, a de novo protein with a  $\beta\alpha\beta$  fold was found to be an incipient downhill folder that folds more cooperatively than a downhill folder, but less cooperatively than a two-state folder (29). Overall, while the native centric Gō-model apparently works well, there are two main divergences between simulations and experiments.

First, in view of the critical role of water exclusion in folding, native centric CG models with elementary desolvation barriers have been developed (30–32). For a set of 13 proteins, folding rates simulated with desolvation barriers cover ~4.6 orders of magnitude, i.e., a range observed experimentally, whereas folding rates simulated without desolvation barriers cover only 2.2 orders of magnitude.

The transition-state ensemble (TSE) of a protein model with desolvation barriers generally has a higher average number of native contacts and is structurally less diffused than the TSE obtained without desolvation barriers (31). An increasing height of desolvation barriers leads to thermodynamically more cooperative folding/unfolding transitions and higher degrees of kinetic cooperativity as manifested by more linear rate-stability relationships under constant temperature (32). The interplay between search of the native structure and desolvation has also been explored on the SH3 protein and supports a folding mechanism where most of the structural formation of the protein is achieved before water is expelled from the hydrophobic core. After the folding transition, a near-native intermediate with partially solvated hydrophobic core is found and this is followed by a final step that squeezes out water molecules from the partially hydrated protein core (30).

Second, the standard Gō model lacks energetic frustration—compact nonnative states of very low energy cannot be explored—and energetic heterogeneity resulting from sequence details. One of early attempts in considering energetic frustration is the Thirumalai model consisting of hydrophobic (H), hydrophilic (P), and neutral (N) residues (33). The Hamiltonian is composed of Lennard-Jones between H residues, and a repulsive  $r^{-12}$  term otherwise, a bending potential and a dihedral angle potential adjusted so that the three minima, i.e., the *trans* and two *gauche* states, have the same energy and are separated by a reduced barrier. The bias towards the native state is provided by selecting a sequence with the appropriate pattern in the three-letter amino acid alphabet. Since the development of Thirumalai model for studying protein folding in aqueous solution, (33) in a chaperonin cavity ((34), with the finding of accelerated folding and an alternate folding pathway) or in the presence of crowding agents ((35, 36), with the prediction that crowding should enhance the stability of proteins and make the transition-state ensemble much broader), the dependence on the native state has been controlled by various means. These include maximization of the energy gap between the native and unfolded states (37), sequence periodicity and secondary structure propensity (38) or a priori knowledge of the native secondary structure (39). For instance, in the Head-Gordon model based on an alphabet of four bead types and therefore a reduced bead-bead interactions matrix compared to the Miyazawa-Jernigan matrix (40), each dihedral angle along the sequence are parameterized by favoring the formation of the assigned secondary structures over other competing local structures. Using this model, the known distinct thermodynamic and folding behaviors of proteins L and G, both sharing the same native topology, were reproduced (39).

More recently, the native centric Gō model was applied to study the folding of elongated proteins (41). Repeat proteins, composed of up to seven repeating units, can be viewed as simple cases of multidomain proteins in which the interface between the repeats acting as the domains is significantly larger than interfaces in common multidomain proteins. On the basis of experimental and MD-generated folding rates, Levy et al. have shown that, as the protein elongates, its folding kinetics slows down and deviates from the rate fitted on globular proteins based on the topological contact order. This observed slow kinetics results from a high population of stable intermediates composed of several consecutive repeats (41).

Escape from the target structures and thus the Gō approximation, while maintaining the C $\alpha$  representation, is clearly desirable to get the best agreement between experimental and predicted observables. One solution is provided by the REACH (Realistic Extension Algorithm via Covariance Hessian) coarse-graining method, in which the force constants of a residue-scale elastic network model are calculated from the variance–covariance matrix obtained from atomistic MD simulation (42). The local (up to 1–5) interactions and the inter- $\beta$ -strand interaction are modelled with a single force constant, while the nonbonded force constants are modelled as distance dependent with a double-exponential function. These terms can be learnt by averaging over different protein structures and the transferability of the force constants between different protein structural classes seems promising for exploring very rapidly the spatiotemporal motions of proteins. Whether the REACH dependence on the secondary-structural elements can be eliminated for protein folding remains to be determined nevertheless. Alternatively, it is possible to replace Coulomb interactions by dipole–dipole interactions and expand the bending angles from quadratic to quartic terms in the full space of ( $\phi$ ,  $\psi$ ) conformations (43) or to combine fragment-based two-body potential, fragment-based pseudodihedral angle potential,  $\beta$ -sheet stabilization potential, hydrophobic interaction and a steric exclusive repulsion so as to control the maximal number of contacts between residues (44). Simulations in the full configuration space are able to catch the formation of supersecondary structural motifs, like the transition of long single  $\alpha$ -helices to helix-turn-helix or  $\beta$ -hairpin assemblies, but are not yet mature for folding a real amino sequence (43). In contrast, Langevin dynamics simulations based on the assembly of 9-residue fragments enable the rapid structure-search for protein lengths up to 150 amino acids and the prediction can be improved when the candidate structures are cross-checked by an empirically derived score function (44). Such a fragment-based assembly approach, pivotal to the Rosetta structure prediction program, is not centered, however, on folding mechanisms.

### 3. Two- and Four-Beads Models for Folding

Going beyond the C $\alpha$  model by adding other particles for the backbone and the side chain has been followed by many groups in order to introduce energetic heterogeneity. Most multiple-beads models use implicit solvent representations, but attempts to couple a CG protein model with a CG solvent model are in progress (45, 46). It is possible to report the existing models, either by their energetic frustrations, biases (full vs. discrete conformational space), parameterization procedures or performances. For simplicity, I present the CG models by their levels of granularity ranging from two beads to six beads. Table 1 reports the most current approaches to explore configurational states with levels of frustration ranging from Gō-like (maximal frustration) to fully ab initio (free of any frustration).

Basdevant et al. developed a CG model with each amino acid represented by one to three coarse grains (one bead for the backbone and one or two beads for the side chains), with physical properties derived in a consistent bottom-up procedure from the higher-resolution all-atom AMBER force field (47). This CG description, very similar to the Zacharias model (48), was further modified by Ha-Duong to model flexibility (49). The bonded potentials are sequence-dependent and the secondary structure propensity of each amino acid is partially reflected in the Si-Bi-Bi + 1-Bi + 2 pseudotorsion angle, where Si and Bi denote the side-chain and backbone beads, respectively. To stabilize the secondary structures, the bonded force field is balanced by a simplified hydrogen bond (H-bond) potential based on dipole–dipole interactions. Tested on eight proteins with 17–98 amino acids, the model yields RMSD varying between 3 and 8 Å from the

**Table 1**  
**Current techniques to explore protein configurational states with their frustration levels ranging from Go-like (maximal) to fragment assembly (medium) and fully ab initio (free of any frustration)**

	Rosetta PEP-FOLD	Basin-hopping MC	MD LD	Meta-dynamics	Generalized ensemble
Ab initio	•		•	•	•
Fragment	•	•	•		•
Go-like		•	•		

MD and LD stand for molecular dynamics and Langevin dynamics, and MC for Monte Carlo

experimental structures using 200-ns MD simulations, indicating room for improvement (49).

Majek and Elber also proposed a model with a C $\alpha$  bead and one side-chain bead (50). Parameters were determined for distance dependent pair interactions, pseudo bonds, angles, torsions and a dipole–dipole interaction for H-bonds. The model was subject to MD simulations of 956 globular proteins starting from their native structures and discrimination between native and decoy structures for 2,470 proteins with 305,000 decoys and the “Decoys ‘R’ Us” dataset. Fifty-eight percent of tested proteins remain within 5 Å from their starting point within 20-ns MD simulations, but some proteins deviate by more than 9 Å. Overall, whether the Elber and Ha-Duong models work for protein folding remains to be determined.

De Mori et al. used as a first step a model where the protein is described by its C $\alpha$  trace and C $\beta$  centroids (51). The energy function incorporates effective pair-wise interactions among amino acids, knowledge-based constraints for backbone chirality, local propensities to form secondary motifs, and a term favoring tertiary packing. Using this simplified framework, the 36-residue villin subdomain consisting of three  $\alpha$ -helices was subject to several Monte Carlo simulations at distinct temperatures and the peak in the heat capacity,  $T_c$ , was identified. Seven different uncorrelated coarse-grained conformations at  $T_c$  were subsequently used for all-atom 50-ns MD simulations at 300 K in explicit solvent. Only one simulation approached the native NMR structure to within 2.4 Å for the core region spanning amino acids 9–32 and 3.7 Å for the whole protein (51). The application of this hybrid multiscale approach to more complex topologies has to be investigated.

The UNRES two-bead model (one unified bead for side chain and the peptide center, p, located in the centers of C $\alpha$ –C $\alpha$  bonds), refined over the years by Scheraga et al., is composed of complex local and nonbonded interaction analytical formulations, including anisotropic side chain—side-chain terms and backbone hydrogen bonding potential modelled by dipole interactions between the peptide centers. This model has been used to simulate microsecond protein folding dynamics using Langevin dynamics (52), and predict the equilibrium structure and free energy landscape of various proteins using various generalized ensemble techniques (53). Recently, the UNRES model has been coupled to meta-dynamics simulations by Cossio et al. to determine the optimal collective variables that enable folding a 48 amino acid protein, 1E0G, in the shortest possible time (54). One evident limitation of the UNRES model is that the calculated folding temperatures (~1,000 K) are much higher than the expected values (53).

Klein et al. employ a reduced representation in which roughly three to four heavy atoms (nonhydrogen) and adjacent hydrogens are mapped to a single CG (55). Most side chains are represented

by one CG, except Lys and Arg including a hydrophobic and a hydrophilic site and Tyr, Phe and Trp which use two, two and three particles, respectively. Intermolecular potentials based on Lennard-Jones style forms are parameterized using thermodynamic data including surface tension and density. The potential has been evaluated by comparing the solvent accessible surface area to all-atom representations and ranking of protein decoy data sets provided by Decoys “R” Us. Work on an intramolecular potential is in progress.

The MARTINI protein force field, based on partitioning free energy of amino acid side chains between water and oil phases to select the appropriate nonbonded interaction parameters, employs a four-to-one mapping; that is, on average, four heavy atoms are represented by a single interaction center, with an exception for ring-like molecules. A clear limitation of MARTINI is that it imposes secondary structure of the protein backbone, but many processes in which the folding and unfolding of secondary structures are not playing a substantial role can be explored on length and timescales that are 2–3 orders of magnitude larger compared to atomistic simulations (56).

SIMFOLD employs a four-bead representation including the N, C $\alpha$ , carbonyl C, and a centroid located at the center of mass of each side chain (57). The amide H atom and carbonyl O atom coordinates are not treated explicitly: they are derived from those of other main-chain atoms. The energy function consists of torsion angle potentials, van der Waals interactions, H-bond, hydrophobic, secondary structure propensity and pair-wise specific interactions. The functional forms and the parameters were optimized from the energy landscape theory by maximizing the energy gap between native and nonnative structures using a benchmark of 40 proteins. SIMFOLD was used on a structure prediction test of six proteins of lengths 47–76 with fragment assembly sampling method and a radius gyration bias. A native-like topology was obtained in one of the top five cluster centers of the predicted models for five proteins, and SIMFOLD-REMD simulations showed poorer performance than the fragment assembly method (57). Coupled to a multicanonical ensemble fragment assembly Monte Carlo method, a SIMFOLD study addressed what microscopic interactions make the protein energy landscape funnel-like. It was found that there is a limited number of compact conformations fully compatible with local structural preferences, and the native fold is one of them, and thus local structure preferences strongly shape up the folding funnel (58).

Rosetta incorporates a low-resolution five-bead representation that uses the backbone heavy atoms and a centroid located at the side-chain center of mass, followed by an all-atom representation (59). Rosetta uses a Monte Carlo fragment assembly of length 9 and then three residues to sample a series of scoring functions with increasing levels of chemical detail in order to identify native states.

Overall, the low-resolution energy function includes van der Waals hard sphere repulsion, environment, pair, side-chain packing density, secondary structure packing [helix–helix pairing, helix–strand pairing, strand–strand pairing, strand pair distance/register and strand arrangement into sheets], radius of gyration energetic contributions, contact order, and Ramachandran torsion angle filters. Additional hydrogen bonding (short- and long-range backbone–backbone hydrogen bond) energy terms are used during full-atom refinement. The Rosetta procedure enables structure prediction with atomic level accuracy for small proteins (60), but for larger proteins, the native state is virtually never explored. Rosetta’s coarse-grained potentials may actually impede accurate structure prediction at full-atom resolution in that the lowest-energy structures are not necessarily the most native-like (61, 62). Building on the Rosetta framework, Baker et al. have recently presented a general method to simultaneously model the folding and docking of multiple-chain interleaved homooligomers. Starting from fully extended monomer chains and using symmetrical constraints, for more than a third of the cases in a benchmark set of interleaved homooligomers, the method generates near-native models of large  $\alpha$ -helical bundles, interlocking  $\alpha$  sandwiches, and interleaved  $\alpha/\beta$  motifs with good accuracy (63).

Finally, Irback et al. developed a CG model containing three types of amino acids and five to six atoms per amino acid and with the only Ramachandran torsion angles as its degrees of freedom (64). The three backbone atoms N, C $\alpha$ , carbonyl C, and the H and O atoms are defined. Each side chain is represented by a single atom, C $\beta$ , which can be hydrophobic, polar, or absent (glycine). Using a threefold torsion term, a hard sphere potential, a 6–12 potential for hydrophobicity and an explicit two-body H-bond potential, the structures and thermodynamics of an helix, a helix-turn-helix and a bundle of three-helices were found in agreement with experiments (64). This CG has never been tested on all  $\beta$  and or mixed  $\alpha/\beta$  proteins, suggesting clear limitations in the parameters to discriminate native from nonnative structures for non-trivial topologies.

---

#### 4. Coarse-Grained Models for Folding and Amyloid Fibril Aggregation

There are two CG models that have been used extensively and simultaneously for protein folding and oligomerization. The first model by Dokhlyan et al. is specifically designed to be used with discontinuous MD (DMD) which requires square-well potentials between pairs of nonbonded beads and constraints between the neighboring beads (65). Along the backbone and the side chain, the model includes three particles (N, C $\alpha$ , carbonyl C) and up to

three CG beads, respectively depending on the properties of the individual amino acid. Because the carbonyl oxygen and the amide hydrogen are not explicitly present, an effective backbone H-bond is introduced between the nitrogen and carbonyl C atoms. The effective Hamiltonian also contains hydrophobic interactions, salt-bridge interactions and a dihedral angle term (65). In DMD, particles move with constant speeds along straight lines. When two particles reach a distance at which the potential is discontinuous, a collision occurs. The pair of particles with the shortest collision time is chosen as the next collision event and the new positions and velocities of the two particles involved are calculated based on conservation laws for the linear momentum, angular momentum, and total energy (66). The advantage of DMD is that the numerical integration of Newton's second law equations is avoided, resulting in a substantial decrease in computational burden and therefore the emergence of servers, and notably iFold which supports large-scale simulations of protein folding, thermal denaturation, thermodynamic scan, simulated annealing and *folding probability* analysis using CG protein model with a structure-based Gō-like potential (67). The DMD force field with up to three CG beads has been used to study the low molecular weight aggregates varying from dimers to hexamers of amyloid peptides (68), which are the most cytotoxic species in protein conformation diseases, including Alzheimer, prion and Parkinson (69, 70). The dynamics of transitions from monomers to larger oligomers has been explored using a four-bead DMD model on the two predominant A $\beta$  alloforms, A $\beta$ 1–40 and A $\beta$ 1–42, implicated in Alzheimer's disease, and some disease-modifying mutants (71, 72). Overall, although the four-bead DMD model cannot approach the accuracy of an all-atom model in explicit solvent, it enables the determination of possible structures and assembly mechanisms of proteins that can guide further experiments.

The second CG for protein folding and aggregation is OPEP (Optimized Potential for Efficient peptide structure Prediction) which employs five particles for the backbone (N, H, Ca, carbonyl C and O) and one bead for all side chains, except the proline amino acid which uses an all-atom representation (73, 74). The implicit solvent OPEP function is expressed as a sum of short-range (bond-lengths, bond angles, improper torsions of the side chains and the amide bonds, backbone torsions), van der Waals and two-body and four-body hydrogen-bonding interactions. OPEP has been coupled to Monte Carlo (73), the activation-relaxation or basin-hopping technique which explores the conformations through well-defined transition states (75), MD (76), and replica-exchange MD simulations (18). OPEP applications to amyloid oligomers in aqueous solution revealed unexpected transitions (such as the reptation move of  $\beta$ -strands in the late steps of aggregation that were confirmed experimentally) (75, 77) and structures (such as a

$\beta$ -barrel) (77, 78). REMD-OPEP simulations on amyloidogenic peptides clearly suggest that the nucleus prior to fibril formation goes beyond 20 chains for peptides up to eight amino acid lengths (unpublished data) and the  $\beta$ -strand-loop- $\beta$ -strand conformation, a signature of many proteins in their amyloid fibril states, is not encoded by the dimer (79, 80). REMD and MD simulations on a wide range of peptides and miniproteins are in progress to validate the force field at neutral pH, and a pH-dependent OPEP force field is also under development.

Although REMD-OPEP simulations provide lowest free energy conformations superposable on NMR structures and predicted melting temperatures in agreement with experiment for a couple of peptides, this protocol is computationally too time consuming for screening thousand peptides. Recently, we introduced a fast approach and server, PEP-FOLD, to the de novo prediction of peptides and miniproteins (81, 82). It first predicts, in the terms of a Hidden Markov Model-derived structural alphabet, a limited number of local conformations at each position of the structure. It then performs their assembly using a greedy procedure driven by the OPEP energy. On a benchmark of 52 peptides with 9–23 amino acids, PEP-FOLD generates lowest-energy conformations within 2.8 Å C $\alpha$  rmsd from the full NMR structures, outperforming previous approaches. For 13 miniproteins with 27–49 amino acids, PEP-FOLD reaches an accuracy of 3.6 and 4.6 Å for the most native and lowest-energy conformations, using the nonflexible regions identified by NMR. Overall, PEP-FOLD which turns the folding problem into a progressive assembly of rigid fragment conformations using a chain growth method, still poses questions, particularly for the analytical form of the Van der Waals interactions to be used since steric clashes occur much more frequently in a discrete space than in a continuous space. Despite this limitation, PEP-FOLD in its current state opens new perspectives for the design of novel peptide-based drugs (83).

---

## 5. Conclusions

This review has given a list of the CG models currently used for protein folding and aggregation in aqueous solution. These models varying in their level of granularity are permanently refined and optimized to make them as reliable and predictive as possible. While in theory, multibody potentials could be used, two-body interactions still dominate and all force fields are not transferable. In essence, protein CG force fields cannot reproduce exactly the structure, thermodynamics and dynamics inferred from experiments. Implicit water treatments accelerate folding by several orders of magnitude and impact folding mechanisms. The effects

of pH conditions on amyloid kinetics and structures (84) and of explicit solvent on recognition processes (85) are also well established. Rather, CG models are to be used for exploring dimensions, timescales, and rare events that are inaccessible to atomistic simulations so as to provide insights of the significant slow degrees of freedom of the macromolecular system and explanations of the folding, large-scale motion, recognition, and aggregation mechanisms of proteins. Clearly, CG simulations and results, which can bias and guide all-atom simulations in explicit solvent for further validation, constitute complementary tools to experimental studies and have the potential to significantly impact our knowledge and strategies in protein recognition and aggregation.

## References

1. Anfinsen CB (1973) Principles that govern the folding of proteins. *Science* 181:223–230
2. Andreeva A, Howorth D, Chandonia JM, Brenner SE, Hubbard TJ, Chothia C, Murzin A (2008) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res* 36:D419–D425, Database issue
3. Wright PE, Dyson HJ (2009) Linking binding and folding. *Curr Opin Struct Biol* 19:31–38
4. Goldschmidt L, Teng PK, Riek R, Eisenberg D (2010) Identifying the amyloome, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci U S A* 107(8):3487–3492
5. Kim SJ, Born B, Havenith M, Gruebele M (2008) Real-time detection of protein-water dynamics upon protein folding by Terahertz absorption. *Angew Chem Int Ed Engl* 47 (34):6486–6489
6. Karplus M, McCammon JA (2002) Molecular dynamics simulations of biomolecules. *Nat Struct Biol* 9(9):646–652
7. Sugita Y, Okamoto Y (1999) Replica exchange molecular dynamics method for protein folding. *Chem Phys Lett* 329:261–270
8. Berg BA, Neuhaus T (1992) Multicanonical ensemble: a new approach to simulate first-order phase transitions. *Phys Rev Lett* 68:9–12
9. Born M, Oppenheimer R (1927) Zur quantentheorie der molekülen. *Ann Phys Leipzig* 84:457
10. Duan Y, Kollman PA (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 282:740–744
11. Jayachandran G, Vishal V, Pande VS (2006) Using massively parallel simulation and Markovian models to study protein folding: examining the dynamics of the villin headpiece. *J Chem Phys* 124:164902
12. Wriggers W, Stafford KA, Shan Y, Piana S, Maragakis P, Lindorff-Larsen K, Miller PJ, Gullingsrud J, Rendleman CA, Eastwood MP, Dror RO, Shaw DE (2009) Automated event detection and activity monitoring in long molecular dynamics simulations. *J Chem Theory Comput* 5:2595–2605
13. Fukunishi H, Watanabe O, Takada S (2002) On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure. *J Chem Phys* 116(20):9058–9067
14. Sgourakis NG, Yan Y, McCallum SA, Wang C, Garcia AE (2007) The Alzheimer's peptides Abeta40 and 42 adopt distinct conformations in water: a combined MD/NMR study. *J Mol Biol* 368(5):1448–1457
15. De Simone A, Derreumaux P (2010) Low molecular weight oligomers of amyloid peptides display beta-barrel conformations: a replica exchange molecular dynamics study in explicit solvent. *J Chem Phys* 132(16): 165103
16. Mucke L (2009) Neuroscience: Alzheimer's disease. *Nature* 461(7266):895–897
17. Levitt M, Warshel A (1975) Computer simulation of protein folding. *Nature* 253:694–698
18. Chebaro Y, Dong X, Laghaei R, Derreumaux P, Mousseau N (2009) Replica exchange molecular dynamics simulations of coarse-grained proteins in implicit solvent. *J Phys Chem B* 113(1):267–274
19. Clementi C (2007) Coarse-grained models of protein folding: toy models or predictive tools? *Curr Opin Struct Biol* 17:1–6
20. Tozzini V (2005) Coarse-grained models for proteins. *Curr Opin Struct Biol* 15:144–150
21. Chu J-W, Voth GA (2007) Coarse-grained free energy functions for studying protein

- conformational changes: a double-well network model. *Biophys J* 93:3860–3871
22. Khalid S, Bond PJ, Holyoake J, Hawtin RW, Sansom MS (2008) DNA and lipid bilayers: self-assembly and insertion. *J R Soc Interface* 5(Suppl 3):S241–S250
  23. Pasquali S, Derreumaux P (2010) HiRE-RNA: a high-resolution coarse-grained energy model for RNA. *J Phys Chem B* 114(37):11957–66
  24. Ueda Y, Taketomi H, Go N (1978) Studies on protein folding, unfolding and fluctuations by computer simulation II. A three-dimensional lattice model of lysozyme. *Biopolymers* 17:1531–1548
  25. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4:10–19
  26. Bryngelson JD, Onuchic JN, Socci ND, Wolynes JD (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* 21:167–195
  27. Shakhnovich E (2006) Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. *Chem Rev* 106:1559–1588
  28. De Sancho D, Doshi U, Munoz V (2009) Protein folding rates and stability: how much is there beyond size? *J Am Chem Soc* 131:2074–2075
  29. Qi Y, Huang Y, Liang H, Liu Z, Lai L (2010) Folding simulations of a de novo designed protein with a betaalphabeta fold. *Biophys J* 98 (2):321–329
  30. Cheung MS, Garcia AE, Onuchic JE (2002) Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc Natl Acad Sci U S A* 99:685–690
  31. Ferguson A, Liu Z, Chan HS (2009) Desolvation barrier effects are a likely contributor to the remarkable diversity in the folding rates of small proteins. *J Mol Biol* 389(3):619–636
  32. Liu Z, Chan HS (2005) Solvation and desolvation effects in protein folding: native flexibility, kinetic cooperativity and enthalpic barriers under isostability conditions. *Phys Biol* 2(4): S75–S85
  33. Honeycutt JD, Thirumalai D (1990) Metastability of the folded states of globular proteins. *Proc Natl Acad Sci U S A* 87:3526–3529
  34. Jewett AI, Baumketner A, Shea JE (2004) Accelerated folding in the weak hydrophobic environment of a chaperonin cavity: creation of an alternate fast folding pathway. *Proc Natl Acad Sci U S A* 101(36):13192–13197
  35. Cheung MS, Klimov D, Thirumalai D (2005) Molecular crowding enhances native state sta-
  - bility and refolding rates. *Proc Natl Acad Sci U S A* 102:4753–4758
  36. Cheung MS, Thirumalai D (2007) Effects of crowding and confinement on the structures of the transition state ensemble in proteins. *J Phys Chem B* 111:8250–8257
  37. Das P, Matysiak S, Clementi C (2005) Balancing energy and entropy: a minimalist model for the characterization of protein folding landscapes. *Proc Natl Acad Sci U S A* 102:10141–10146
  38. Bellesia G, Jewett AI, Shea J-E (2010) Sequence periodicity and secondary structure propensity in model proteins. *Protein Sci* 19:141–154
  39. Yap E-H, Fawzi NL, Head-Gordon T (2008) A coarse-grained  $\alpha$ -carbon protein model with anisotropic hydrogen-bonding. *Proteins* 70:626–638
  40. Miyazawa S, Jernigan RL (1999) Self-consistent estimation of inter-residue protein contact energies based on an equilibrium mixture approximation of residues. *Proteins* 34:49–68
  41. Hagai T, Levy Y (2008) Folding of elongated proteins: conventional or anomalous? *J Am Chem Soc* 130:14253–14262
  42. Moritsugu K, Smith JC (2008) REACH Coarse-grained biomolecular simulation: transferability between different protein structural classes. *Biophys J* 95:1639–1648
  43. Alemani D, Collu F, Casella M, Dal Peraro M (2010) A nonradial coarse-grained potential for proteins produces naturally stable secondary structure elements. *J Chem Theory Comput* 6:315–324
  44. Sasaki TN, Cetin H, Sasai M (2008) A coarse-grained Langevin molecular dynamics approach to de novo protein structure prediction. *Biochem Biophys Res Commun* 369 (2):500–506
  45. Han W, Wan C, Wu Y (2008) Toward a coarse-grained protein model coupled with a coarse-grained solvent model: solvation free energies of amino acid side chains. *J Chem Theory Comput* 4:1891–1901
  46. Thorpe IF, Zhou J, Voth GA (2008) Peptide folding using multiscale coarse-grained models. *J Phys Chem B* 112:13079–13090
  47. Basdevant N, Borgis D, Ha-Duong T (2007) A coarse-grained protein-protein potential derived from an all-atom force field. *J Phys Chem B* 111:9390–9399
  48. Zacharias M (2003) Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Sci* 12:1271–1282

49. Ha-Duong T (2010) Protein backbone dynamics simulations using coarse-grained bonded potentials and simplified hydrogen bonds. *J Chem Theory Comput* 6:761–773
50. Majek P, Elber R (2009) A coarse-grained potential for fold recognition and molecular dynamics simulations of proteins. *Proteins* 76:822–836
51. De Mori G, Colombo G, Micheletti C (2005) Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics. *Proteins* 58:459–471
52. Maisuradze GG, Senet P, Czaplewski C, Liwo A, Scheraga HA (2010) Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *J Phys Chem A* 114:4471–4485
53. Czaplewski C, Kalinowski S, Liwo A, Scheraga HA (2009) Application of multiplexed replica exchange molecular dynamics to the UNRES force field: tests with  $\alpha$  and  $\alpha + \beta$  proteins. *J Chem Theory Comput* 5:627–640
54. Cossio P, Marinelli F, Laio A, Pietrucci F (2010) Optimizing the performance of bias-exchange metadynamics: folding a 48-residue LysM domain using a coarse-grained model. *J Phys Chem B* 114:3259–3265
55. DeVane R, Shinoda W, Moore PB, Klein ML (2009) Transferable coarse grain non-bonded interaction model for amino acids. *J Chem Theory Comput* 5:2115–2124
56. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tielemans DP, Marrink S (2008) The MARTINI coarse-grained force field: extension to proteins. *J Chem Theory Comput* 4:819–834
57. Fujitsuka Y, Takada S, Luthey-Schulten ZA, Wolynes PG (2004) Optimizing physical energy functions for protein folding. *Proteins* 54:88–103
58. Chikenji G, Fujitsuka Y, Takada S (2006) Shaping up the protein folding funnel by local interaction: lesson from a structure prediction study. *Proc Natl Acad Sci U S A* 103:3141–3146
59. Rohl CA, Strauss CE, Chivian D, Baker D (2004) Protein structure prediction using Rosetta. *Methods Enzymol* 383:66–93
60. Bradley P, Misura KM, Baker D (2005) Towards high-resolution de novo structure prediction for small proteins. *Science* 309:1868–1871
61. Bowman GR, Pande VS (2009) Simulated tempering yields insight into the low resolution Rosetta scoring functions. *Proteins* 74:777–788
62. Shmygelska A, Levitt M (2009) Generalized ensemble methods for de novo structure prediction. *Proc Natl Acad Sci U S A* 106:1415–1420
63. Das R, Andre I, Shen Y, Wu Y, Lemak A, Bansal S, Arrowsmith CH, Szyperski T, Baker D (2009) Simultaneous prediction of protein folding and docking at high resolution. *Proc Natl Acad Sci U S A* 106:18978–18983
64. Irback A, Sjunnesson F, Wallin S (2000) Three-helix-bundle protein in a Ramachandran model. *Proc Natl Acad Sci U S A* 97:13614–13618
65. Ding F, Buldyrev SV, Dokholyan NV (2005) Folding Trp-cage to NMR resolution native structure using a coarse-grained protein model. *Biophys J* 88:147–155
66. Smith AV, Hall CK (2001)  $\alpha$ -Helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model. *Proteins* 44:344–360
67. Sharma S, Ding F, Nie H, Watson D, Unnithan A, Lopp J, Pozefsky D, Dokholyan NV (2006) iFold: a platform for interactive folding simulations of proteins. *Bioinformatics* 22 (21):2693–2694
68. Ding F, LaRocque JJ, Dokholyan NV (2005) Direct observation of protein folding, aggregation, and a prion-like conformational conversion. *J Biol Chem* 280(48):40235–40240
69. Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ, Selkoe DJ (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* 416:483–484
70. Lesne S, Teng Koh M, Kotilinek L, Kayed R, Glabe CG, Yang A, Gallaher M, Ashe K (2006) A specific amyloid- $\beta$  protein assembly in the brain impairs memory. *Nature* 440:352–357
71. Urbanc B, Cruz L, Yun S, Buldyrev SV, Bitan G, Teplow DB, Stanley HE (2004) In silico study of amyloid beta-protein folding and oligomerization. *Proc Natl Acad Sci U S A* 101:17345–17350
72. Urbanc B, Betnel M, Cruz L, Bitan G, Teplow DB (2010) Elucidation of amyloid beta-protein oligomerization mechanisms: discrete molecular dynamics study. *J Am Chem Soc* 132:4266–4280
73. Derreumaux P (2000) Generating ensemble averages for small proteins from extended conformations by Monte Carlo simulations. *Phys Rev Lett* 85:206–209
74. Maupetit J, Tuffery P, Derreumaux P (2007) A coarse-grained protein force field for folding and structure prediction. *Proteins* 69 (2):394–408
75. Santini S, Mousseau N, Derreumaux P (2004) In silico assembly of Alzheimer's A $\beta$ 16–22

- peptide into  $\beta$ -sheets. *J Am Chem Soc* 126:11509–11516
76. Derreumaux P, Mousseau N (2007) Coarse-grained protein molecular dynamics simulations. *J Chem Phys* 126:025101–025106
77. Mousseau N, Derreumaux P (2005) Exploring the early steps of amyloid peptide aggregation by computers. *Acc Chem Res* 38:885–891
78. Lu Y, Derreumaux P, Guo Z, Mousseau N, Wei G (2009) Thermodynamics and dynamics of amyloid peptide oligomerisation is sequence-dependent. *Proteins* 5(4):954–963
79. Melquiond A, Dong X, Mousseau N, Derreumaux P (2008) Role of the region 23–28 in A $\beta$  fibril formation: insights from simulations of the monomers and dimers of Alzheimer's peptides A $\beta$ 40 and A $\beta$ 42. *Curr Alzheimer Res* 5 (3):244–250
80. Chebaro Y, Mousseau N, Derreumaux P (2009) Structures and thermodynamics of Alzheimer's amyloid- $\beta$  A $\beta$ (16–35) monomer and dimer by replica exchange molecular dynamics simulations: implication for full-length A $\beta$  fibrillation. *J Phys Chem B* 113:7668–7675
81. Maupetit J, Derreumaux P, Tuffery P (2009) PEP-FOLD: an online resource for de novo peptide structure prediction. *Nucleic Acids Res* 37:W498–W503, Web Server issue
82. Maupetit J, Derreumaux P, Tuffery P (2010) A fast method for large-scale de novo peptide and miniprotein structure prediction. *J Comput Chem* 31(4):726–738
83. Audie J, Boyd C (2010) The synergistic use of computation, chemistry and biology to discover novel peptide-based drugs: the time is right. *Curr Pharm Des* 16(5):567–582
84. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75:333–366
85. De Beer SB, Vermeulen NP, Oostenbrink C (2010) The role of water molecules in computational drug design. *Curr Top Med Chem* 10 (1):55–66

# Chapter 23

## Elastic Network Models: Theoretical and Empirical Foundations

Yves-Henri Sanejouand

### Abstract

Fifteen years ago, Monique Tirion showed that the low-frequency normal modes of a protein are not significantly altered when nonbonded interactions are replaced by Hookean springs, for all atom pairs whose distance is smaller than a given cutoff value. Since then, it has been shown that coarse-grained versions of Tirion's model are able to provide fair insights on many dynamical properties of biological macromolecules. In this chapter, theoretical tools required for studying these so-called Elastic Network Models are described, focusing on practical issues and, in particular, on possible artifacts. Then, an overview of some typical results that have been obtained by studying such models is given.

**Key words:** Protein, Normal mode analysis, Anisotropic network model, Gaussian network model, Low-frequency modes, B-factors, Thermal motion, Conformational change, Functional motion.

---

### 1. Introduction

In 1996, Monique Tirion showed that the low-frequency normal modes of a protein (see Section 3.1) are not significantly altered when Lennard-Jones and electrostatic interactions are replaced by Hookean (harmonic) springs, for all atom pairs whose distance is smaller than a given cutoff value (1). In the case of biological macromolecules, this seminal work happened to be the first study of an elastic network model (ENM). The ENM considered was an all-atom one, chemical bonds and angles being kept fixed through the use of internal coordinates, as often done in previous standard normal mode studies of proteins (2–4).

Soon afterwards, several coarse-grained versions of Tirion's ENM were proposed, in which each protein amino acid residue is usually represented as a single bead and where most, if not all, chemical "details" are disregarded (5, 6), including atom types and amino acid masses.

Since then, it has been shown that such highly simplified protein models are able to provide fair insights on the dynamical properties of biological macromolecules (5, 7–9), including those involved in their largest amplitude functional motions (10, 11), even in the case of large assemblies like RNA polymerase II (12), transmembrane channels (13, 14), whole virus capsids (15), or even the ribosome (16). As a consequence, numerous applications have been proposed, noteworthy for exploiting fiber diffraction data (17), solving difficult molecular-replacement problems (18, 19), or for fitting atomic structures into low-resolution electron density maps (19–23).

However, the idea that simple models can prove enough for capturing major properties of objects as complex as proteins had been put forward well before Tirion's introduction of ENMs in the realm of molecular biophysics. In the following, after a brief account of previous results supporting this claim (Section 2), theoretical tools required for studying an ENM are described (Section 3), focusing on practical issues and, in particular, on possible artifacts. Then, an overview of typical results that have been obtained by studying protein ENMs is given (Section 4).

## 2. Background

Indeed, coarse-grained models of proteins had been considered 20 years before M. Tirion's work, for studying what may well be the most complex phenomenon known at the molecular scale, namely, protein folding. Indeed, as soon as 1975, Michael Levitt and Arieh Warshel proposed to model a protein as a chain of beads, each bead corresponding to the  $C_\alpha$  atom of an amino acid residue, the centroid of each amino acid side chain being taken into account with another bead grafted onto the chain (24). That same year, Nobuhiro Go and his collaborators proposed an even simpler model in which the chain of beads is mounted on a two-dimensional lattice, each bead corresponding either to a single residue or, more likely, to a secondary structure element (e.g., an  $\alpha$ -helix) of a protein (25). Moreover, while the Levitt-Warshel model had been designed so as to study a specific protein, that is, a polypeptidic chain with a given sequence of amino acid residues, the Go model focuses on the conformation of the chain, more precisely, on the set of pairs of amino acids that are interacting together in the chosen (native) structure.

So, it is fair to view protein ENMs as off-lattice versions of the Go model.

Lattice models of proteins have been studied extensively since then so as to gain, for instance, a better understanding of the

sequence-structure relationship. Noteworthy, if the chain is short enough, all possible conformations on the lattice can be enumerated, allowing for accurate calculations of thermodynamic quantities and univoqual determination of the free energy minimum. Moreover, if the number of different amino acids is small enough, then the whole sequence space can also be addressed. For instance, in the case of the tridimensional cubic lattice, a 27-mer chain has 103,346 self-avoiding compact (i.e., cubic) conformations (26). On the other hand, if only two kinds of amino acids are retained, that is, if only their hydrophobic or hydrophilic nature is assumed to be relevant for the understanding of protein stability, then a 27-mer has  $2^{27}$  different possible sequences. This is a large number, but it remains small enough so that for each sequence the lowest-energy compact conformation can be determined and, when a nearly additive interaction energy is considered (27), the conclusion of such a systematic study happens to be an amazing one. Indeed, it was found that a few conformations (1% of them) are “preferred” by large sets of sequences (28). Moreover, although each of these sets forms a neutral net in the sequence space, it is often possible to “jump” from a preferred conformation to another, as a consequence of single-point mutations (29).

While the former property is indeed expected to be a protein-like one, allowing to understand why proteins are able to accommodate so many different single-point mutations without significant loss of both their structure and function, it is only during the last few years that the latter one has been exhibited. In particular, using sequence design techniques, a pair of proteins with 95% sequence identity, but different folds and functions, was recently obtained (30). If generic enough, such a property would help to understand how the various protein folds nowadays found on earth may have been “discovered” during the earliest phases of life evolution (e.g., prebiotic ones), since discovering a first fold could have proved enough for having access to many other ones, a single-point mutation after another.

In any case, this example shows how the study of simple models can help to think about, and maybe to understand better, major protein properties, in particular because such models can be studied on a much larger scale than actual proteins.

---

### 3. Theoretical Foundations

The vast majority of protein ENM studies rely on normal mode analysis (NMA) (9). Moreover, the hypotheses underlying this kind of analysis probably inspired the design of the first ENM. Actually, in her seminal work, M. Tirion performed NMA in order

to show that similar results can be obtained by studying an ENM or a protein described at a standard, semiempirical, level (1). So, hereafter, the principles of NMA are briefly recalled (more details can be found in classic textbooks (31, 32)). Next, the close relationship between NMA and the different types of ENMs is underlined.

### 3.1. Normal Mode Analysis

Newton's equations of motion for a set of  $N$  atoms cannot be solved analytically when  $N$  is large (namely,  $N > 2$ ), except in rare instances like the following, rather general, one. Indeed, for small enough displacements of the atoms in the vicinity of their equilibrium positions,  $V$ , the potential energy of the studied system can be approximated by the first terms of a Taylor series:

$$V = V_0 + \sum_{i=1}^{3N} \left( \frac{\partial V}{\partial r_i} \right)_0 (r_i - r_i^0) + \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} \left( \frac{\partial^2 V}{\partial r_i \partial r_j} \right)_0 (r_i - r_i^0)(r_j - r_j^0) \quad (1)$$

where  $r_i$  is the  $i$  th coordinate;  $r_i^0$ , its equilibrium value; and  $V_0$ , the potential energy of the system at equilibrium.

Since, within the frame of classical physics, the exact value of  $V$  is meaningless (only potential energy differences are expected to play a physical role),  $V_0$  can be zeroed. Moreover, since  $V_0$  is a minimum of  $V$ , for each coordinate:

$$\left( \frac{\partial V}{\partial r_i} \right)_0 = 0$$

This yields:

$$V = \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} \left( \frac{\partial^2 V}{\partial r_i \partial r_j} \right)_0 (r_i - r_i^0)(r_j - r_j^0) \quad (2)$$

In other words, if the atomic displacements around an equilibrium configuration are small enough, then the potential energy of a system can be approximated by a quadratic form.

On the other hand, if the system is *not* under any constraint with an *explicit* time dependence, then its kinetic energy can also be written as a quadratic form (31) and it is straightforward to show that when both potential and kinetic energy functions are quadratic forms, then the equations of atomic motion have the following, analytical, solutions (31–33):

$$r_i(t) = r_i^0 + \frac{1}{\sqrt{m_i}} \sum_{k=1}^{3N} C_k \alpha_{ik} \cos(2\pi\nu_k t + \phi_k) \quad (3)$$

where  $m_i$  is the atomic mass and where  $C_k$  and  $\Phi_k$ , the amplitude and phases of the so-called normal mode of vibration  $k$ , depend upon the initial conditions, that is, upon atomic positions and velocities at time  $t = 0$ . Noteworthy,  $C_k$  is a simple function of

$E_k$ , the total energy of mode  $k$ . In particular, if all modes have identical total energies, then

$$C_k = \frac{\sqrt{2k_B T}}{2\pi v_k} \quad (4)$$

where  $T$  is the temperature and  $k_B$  the Boltzmann constant. This means that the amplitude of mode  $k$  goes as the inverse of its frequency,  $v_k$ . As a matter of fact, when NMA is performed in the case of proteins, using standard all-atom force fields, it can be shown that modes with frequencies below 30–100 cm<sup>-1</sup> are responsible for 90–95% of the atomic displacements (34).

Note that such analytical solutions can provide various thermodynamic quantities like entropy, enthalpy, etc., and this, even at a quantum mechanical level of description (34).

In practice, the  $a_{ik}$ 's involved in Eq. 3, which give the coordinate contributions to mode  $k$ , are obtained as the  $k$ th eigenvector of  $\mathbf{H}$ , the mass-weighted Hessian of the potential energy, that is, the matrix whose element  $ij$  is

$$\left( \frac{\partial^2 V}{\sqrt{m_i m_j} \partial r_i \partial r_j} \right)_0 \quad (5)$$

By definition, the  $3N$  eigenvectors of a matrix like  $\mathbf{H}$  form an orthogonal basis set. This means that when  $k \neq l$ ,

$$\left( \frac{\partial^2 V}{\partial q_k \partial q_l} \right)_0 = 0$$

where  $q_k$  is the so-called normal coordinate, obtained by projecting the  $3N$  mass-weighted Cartesian coordinates onto eigenvector  $k$ , namely:

$$q_k = \sum_i^{3N} a_{ik} \sqrt{m_i} (r_i - r_i^0) \quad (6)$$

Moreover, the eigenvalues of  $\mathbf{H}$ , that is, the diagonal elements of the matrix obtained by expressing  $\mathbf{H}$  in this new basis set, provide the  $3N$  frequencies of the system since, for each mode  $k$ ,

$$\left( \frac{\partial^2 V}{\partial^2 q_k} \right)_0 = (2\pi v_k)^2$$

The eigenvalues and eigenvectors of a matrix are obtained by an operation called a diagonalization. In principle, for a real and symmetrical matrix like  $\mathbf{H}$ , such an operation is always possible. At a practical level, when the matrix size is not too large, that is, if

the matrix can be stored in the computer memory, algorithms and methods available in standard mathematical packages allow to get its eigenvalues and eigenvectors at a CPU cost raising as  $nN^2$ , where  $n$  is the number of requested eigensolutions. In other words, it is rather straightforward to obtain analytical solutions for the atomic motions, as long as small-amplitude displacements around a given, well-defined, equilibrium configuration are considered. Note that for a tridimensional system at equilibrium, at least six zero eigenvalues have to be obtained (except if the system is linear, in which case there are five of them), corresponding to the six possible rigid-body motions (translations or rotations) of the entire system. However, if the system is *not* at equilibrium, negative eigenvalues are usually observed. Moreover, significant mixing between rotation modes and some others can occur, leaving three zero eigenvalues only, that is, those corresponding to the three translation modes of the system (33).

The main drawback of NMA is obvious: the actual dynamics of a protein is much more complicated than assumed above. As a matter of fact, even on the short timescales considered within the frame of standard molecular dynamics simulations, a protein is able to jump from the attraction basin of an equilibrium configuration to another (35), and the number of these equilibrium configurations is so huge that it is unlikely for a nanosecond trajectory to visit one of them twice. In other words, while NMA focuses on protein dynamics at the level of a single minimum of the potential energy surface (PES), it is well-known that for proteins at room temperature, the relevant PES is a highly complex, multi-minima, one.

NMA has several other drawbacks. For instance, starting from a given protein structure, for example, as found in the Protein Databank (PDB), an equilibrium configuration has to be reached. This is usually done using energy-minimization techniques. As a consequence, the structure studied with NMA and a standard force field is always a distorted one, the  $C_\alpha$  root-mean-square deviation ( $C_\alpha$ -r.m.s.d) from the initial structure being typically of 1–2 Å (9).

More importantly, within the frame of NMA, it is not obvious to take solvent effects into account, as the meaning of an equilibrium configuration in the case of an ensemble of molecules in the liquid state is unclear. As a matter of fact, the first NMA studies of proteins were performed in vacuo (2–4, 36). Note that, nowadays, the availability of implicit solvent models, like EEF1 (37), offers a more satisfactory alternative.

However, as shown below, the main idea underlying the design of protein ENMs is not only to ignore the well-known drawbacks of NMA but, building upon its empirical successes, to add a few more on top of them.

### 3.2. The Elastic Network Model

In essence, there are two different types of ENMs, which differ by their dimensionality. The Gaussian network model (GNM), proposed by Ivet Bahar, Burak Erman, and Turkana Haliloglu in 1997 (5, 38), is a one-dimension model, while Tirion's model, later called the anisotropic network model (39) (ANM), is a tridimensional one.

#### 3.2.1. The Anisotropic Network Model

Although Eq. 2 may look simple, it relies on a large number of parameters, namely, the elements of the Hessian matrix Eq. 5. In order to make it even simpler, M. Tirion proposed to replace Eq. 2 by another quadratic form, namely:

$$V = \frac{1}{2} k_{\text{enm}} \sum_{d_{ij}^0 < R_c} (d_{ij} - d_{ij}^0)^2 \quad (7)$$

where  $d_{ij}$  is the actual distance between atoms  $i$  and  $j$ ,  $d_{ij}^0$  being their distance in the studied structure (1). This amounts to set Hookean springs between all pairs of atoms less than  $R_c$  Ångströms away from each other. Note that in Tirion's work, as well as in most ANM studies (there are notable exceptions (40)),  $k_{\text{enm}}$ , the spring force constant, is the same for all atom pairs. When it is so, the role of  $k_{\text{enm}}$  is just to specify which system of units is used,  $R_c$  being the only physically relevant parameter of the model. In other words, when studying an ENM, the major drawback added with respect to standard NMA is that most atomic details are simply ignored.

However, considering Eq. 7 instead of Eq. 2 has several practical advantages. First, an energy minimization is not required any more, since the configuration whose energy is the absolute minimum one ( $V = 0$ ) is known: it is the studied one. As a corollary, results obtained by studying ENMs are easier to reproduce. Indeed, an energy minimization not only introduces unwanted distortions in a structure, but it does it in a way that strongly depends upon the most tiny details of the protocol used, also as a consequence of the huge number of minima of a realistic PES for a biological macromolecule. Last but not least, as a straightforward consequence of Eq. 7, the elements of the Hessian matrix (see Eq. 5) are as simple as

$$h_{ij} = -k_{\text{enm}} \frac{(x_i - x_j)(y_i - y_j)}{\sqrt{m_i m_j} d_{ij}^2} \quad (8)$$

where  $h_{ij}$  is the element corresponding to the  $x$  and  $y$  coordinates of atoms  $i$  and  $j$ .

#### 3.2.2. The Gaussian Network Model

Because  $R_c$ , the cutoff value of an ANM, is usually rather small (see Section 3.2.3), the corresponding Hessian matrix is sparse, that is, most of its elements, Eq. 8, are zeroes. So, as proposed by I. Bahar, B. Erman, and T. Haliloglu (5), it is tempting to go another step further into the simplification process and to consider

the corresponding adjacency matrix, that is, the matrix whose elements are:

$$h_{ij} = -k_{enm} \quad (9)$$

when residues  $i$  and  $j$  are interacting ( $h_{ij} = 0$  otherwise). Note that in the case of an adjacency matrix, as well as for the Hessian matrix of an ANM,  $h_{ii}$ , the diagonal element  $i$ , is

$$h_{ii} = -\sum_{i \neq j} h_{ij} \quad (10)$$

Of course, with an adjacency matrix, information about directionality is missing. This is a major drawback of GNMs since this means that studying a GNM can only provide informations about motion amplitudes.

Note that GNMs are usually, if not always, set up at the residue level, while ANMs are sometimes studied at the atomic level, like in the seminal study of Tirion (1). From now on, to underline such (not so common) cases, these latter models will be coined “all-atom ANMs.”

### 3.2.3. The Cutoff Issue

The main, if not the only, parameter of an ENM is  $R_c$ . Although several studies have tried to justify the choice of a particular value for this parameter, typically by comparing calculated and experimental quantities, cutoff values over a wide range are still of common use, varying between 7 (41) and 16 Å (8).

For the most part, this probably reflects the fact the lowest-frequency modes of an ENM are usually “robust” (42), that is, little sensitive to the way the model is built. However, it is obvious that to be meaningful, the value of  $R_c$  has to be on the small side. Putting it to an extreme, in the case of a GNM (see Section 3.2.2), if  $R_c$  is so large that the adjacency matrix is completely filled with nonzero elements, its eigenvalues and eigenvectors, apart from being degenerate, will only depend upon  $N$ , the size of the system, and not upon its topology or its shape. As a consequence, they can for sure not provide any useful information. On the other hand, if  $R_c$  is too small, then the network of interacting residues is split into subnetworks, either free to rotate with respect to another one (in the case of an ANM) or completely independent from each other (in both ANM and GNM cases). Such dynamical properties are certainly not among those expected for a macromolecule, and this is why, in ANM studies, the smallest cutoff values used are of the order of 8–10 Å (10, 12), that is, larger than the typical distance between two interacting amino acid residues in a protein, namely, 6–7 Å (43, 44).

In practice, choosing a too small value for  $R_c$  yields additional zero eigenvalues.

So, if more than one (for a GNM) or six (for an ANM) zero eigenvalues are obtained, then it is highly recommended to increase  $R_c$ . Note that GNMs allow for the use of smaller values of  $R_c$  (a value of 7.3 Å is often chosen (41)) since in the case of a mono-dimensional model, a single connection is enough for avoiding any free translation of a group of atoms with respect to another. As a consequence, when a GNM is built with  $C_\alpha$  atoms picked from a single protein chain, that is, when all amino acid residues are chemically bonded to each other through peptidic bonds, a value of  $R_c$  as low as 4 Å (the typical distance between two consecutive  $C_\alpha$  atoms) can be used.

At first sight, it may seem that problems with small cutoff values could be solved with a distance-dependent spring force constant, as early proposed by Konrad Hinsen (6). However, it is clear that an exponential term, for instance, introduces a typical length which, when too small, yields similar artifacts. Indeed, in such a case, the additional free rigid-body motions obtained with a too small value for  $R_c$  are expected to be replaced by low-frequency motions involving the same too little-connected groups of atoms.

Note that with ENMs other kinds of spurious low-frequency motions can be observed. For instance, in crystal structures, protein N- and C-terminal ends are often found to extend away from the rest of the structure. As a consequence, large amplitude, usually meaningless, motions of these (almost) free ends can be found among the lowest-frequency modes. So, in order to obtain significant and clear-cut results, it is highly recommended to begin an ENM study by “cleaning” the studied structure, namely, by removing such free ends.

A similar kind of spurious low-frequency motion can be observed with all-atom ANMs, in which groups of little-connected atoms are involved, typically those at the end of long side chains (45). Note that an elegant way to cure such artifacts is to use the RTB approximation (46, 47), which allows to remove from the Hessian matrix all contributions associated to motions occurring inside each “block” the system is split into (RTB stands for Rotation Translation of Blocks). In most cases, a block corresponds to a given amino acid residue but, while atom-atom interactions are taken into account when the atoms belong to different blocks, each block can also correspond to a whole protein subunit, allowing for the study of systems as large as entire virus capsids (15).

## 4. Empirical Foundations

As illustrated above, ENMs and NMA are closely related. As a consequence, the theoretical foundations of ENMs are for the most part those of NMA. However, when applied to complex

molecular systems, NMA is known to have obvious drawbacks (see Section 3.1). So, if NMA is still widely performed, it is because of its empirical, sometimes unexpected, successes. As recalled below, most of these successes can also be achieved by studying ENMs.

#### 4.1. B-Factors

From Eq. 3 and Eq. 4, it is straightforward to show that  $\langle \Delta r_i^2 \rangle$ , the fluctuation of coordinate  $i$  with respect to its equilibrium value, is

$$\langle \Delta r_i^2 \rangle = \frac{k_B T}{m_i} \sum_{k=1}^{n_{nz}} \frac{\alpha_{ik}^2}{4\pi^2 v_k^2} \quad (11)$$

$n_{nz}$  being the number of nonzero frequency normal modes of the system, namely,  $n_{nz} = N - 1$  when a GNM is considered and  $n_{nz} = 3N - 6$  when it is an ANM. However, in practice, since such fluctuations scale as the inverse of  $v_k$ , the  $k$ th mode frequency, a sum over the lowest-frequency normal modes of the system is usually enough for obtaining a fair approximation (34).

On the other hand,  $B_i$ , the crystallographic Debye-Waller factor (the so-called isotropic B-factors) of atom  $i$ , is expected to be related to the fluctuations of its atomic coordinates through

$$B_i = \frac{8\pi^2}{3} \langle \Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2 \rangle \quad (12)$$

Although other physical factors are involved, like crystal disorder or lattice phonons, as well as nonphysical ones, like the number of water molecules included in the structure refinement process by crystallographers, significant correlations between B-factor values predicted using Eq. 11 and 12 and experimentally obtained ones have been reported in numerous cases.

For instance, in a study of 30 protein GNMs ( $R_c = 7.5\text{\AA}$ ), a mean value of  $0.62 \pm 0.13$  for this correlation coefficient was found (7). Interestingly, in the same study, 26 other proteins were considered, for which accurate relaxation measurements had been measured by NMR, and the mean correlation between the corresponding fluctuations and those obtained using Eq. 11 was found to be significantly higher, namely,  $0.76 \pm 0.04$ , a remarkable agreement with the experimental data being achieved in several cases, with a correlation coefficient over 0.9 for four of them (7). Amazingly, ANMs do not perform significantly better. For instance, in a study of 83 proteins ( $R_c = 16\text{\AA}$ ), a mean value for the correlation coefficient of  $0.68 \pm 0.11$  between predicted and isotropic B-factors was obtained (8) while, using the all-atom ANM ( $R_c = 5\text{\AA}$ ) implemented in the Elnémo webserver (14), which makes use of the RTB approximation (46, 47), a very similar value of  $0.68 \pm 0.13$  was found (8).

Note that in both studies mentioned above, when Eq. 11 was used, overall translations or rotations of the entire protein within the crystal cell were excluded from the calculation, while it is well

known that such motions are able to provide by themselves good correlations with experimental values (48). In other words, much better correlations with experimental B-factors can be obtained by mixing NMA predictions with protein rigid-body motions, the latter accounting partly for crystal disorder, but mostly for the phonon modes of the whole crystal. Interestingly, these latter modes can be taken into account within the frame of ENM studies, simply by including all crystal cell symmetries in the model (49–51).

Of course, such significant correlations with experimental data can only be obtained because the amplitude of atomic thermal fluctuations scales as the inverse of mode frequencies (see Eq. 14). Indeed, with crude models like ENMs, the actual high-frequency modes of a protein cannot be predicted, because such modes strongly depend upon the chemical details of the structure, only a few neighboring atoms (e.g., covalently bonded ones) being involved in the highest-frequency modes. This does not mean, though, that the high-frequency modes of an ENM cannot bring any useful information. Indeed, they correspond to local motions occurring within the parts of the structure whose density is the highest (38). Moreover, such regions are often found nearby enzyme active sites (52, 53).

On the other hand, the B-factor values themselves cannot directly be obtained by studying ENMs, since their average is proportional to  $k_{enm}$ . Indeed, it is customary to choose  $k_{enm}$  so as to match average experimental B-factor values (9). Another common way is to choose  $k_{enm}$  so as to reproduce the lowest-frequency of the system, as obtained using all-atom force fields (52).

#### 4.2. The Relationship with Protein Functional Motions

The seminal paper of Tirion ends with the statement that (1)

Tests performed on a periplasmic maltodextrin binding protein (MBP) indicate that the slowest modes do indeed closely map the open form into the closed form. (Tirion, in preparation)

The next paper of M. Tirion never came out, but her result was confirmed a few years later, as part of a study of 20 protein ENMs ( $R_c = 8\text{\AA}$ ) in both their ligand-free (open) and ligand-bound (closed) forms (10). Indeed, for MBP (see Fig. 1), it was found

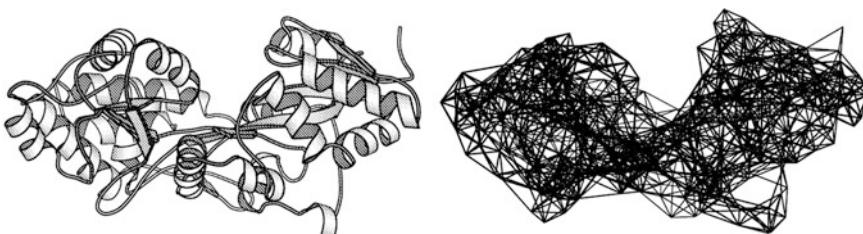


Fig. 1. Left: the open (ligand-free) form of maltodextrin binding protein (PDB identifier 1OMP). Right: the corresponding elastic network model. Pairs of  $C_x$  atoms are linked by springs (plain lines) when they are less than  $8\text{\AA}$  from each other. Drawn with molscript (54).

that the overlap between its second lowest-frequency mode and its functional conformational change is close to 0.9.

This means that 80% of the functional motion of MBP can be described by varying the normal coordinate associated to a single of its modes. Indeed,  $O_k$ , the overlap with mode  $k$ , is given by

$$O_k = \frac{\sum_i \Delta r_i \alpha_{ik}}{\sqrt{\sum_i \Delta r_i^2}} \quad (13)$$

where  $\Delta r_i$  is the variation of coordinate  $i$  between the open and the closed form after both structures have been superimposed (55). On the other hand, since the modes of MBP form an orthogonal basis set, the following property holds:

$$\sum_{k=1}^{n_{nz}} O_k^2 = 1 \quad (14)$$

More generally, it was found that when the conformational change of a protein upon ligand binding happens to be highly collective, one of its low-frequency normal modes often compares well with the experimental motion (overlap over 0.5 (10)). Since then, a study of nearly 4,000 cases has confirmed this result (11), while another study of a set of proteins with similar functions and shapes, but various folds, namely, DNA-dependent polymerases (12), has shown that the low-frequency modes of a protein, and hence the nature of its large-amplitude motions, are likely to be determined by its shape (10, 56, 57).

Indeed, this latter point has recently been confirmed in a rather direct way, by considering ENMs built in such a way that each amino acid interacts with a given number of neighbors (the closest ones). Then, at variance with cutoff-based ENMs, the rigidity of the system is fairly constant from a site to another. However, the relationship between the lowest-frequency modes of a protein and its functional motion is preserved. Specifically, it was found that the subspace defined by up to the 10–12 lowest-frequency modes of a protein is conserved, whatever model is used. Moreover, when no such, so-called robust, subspace exists, the functional motion of the protein is found to be either localized and/or of small amplitude (typically: less than 2–3 Å of C<sub>α</sub>-r.m.s.d.) (42).

In retrospect, these results make sense. First, a strong relationship between low-frequency modes and protein functional motions was first observed within the frame of NMA studies performed at a highly detailed, atomic level of description, noteworthy in the cases of lysozyme (58), hexokinase (59), citrate synthase (55), and hemoglobin (60). Since, as recalled above, it was later found that such a relationship also holds when most chemical details are removed, it is clear that the property captured by NMA has to be a very general one. On the other hand, K.

Hinsen has convincingly shown that the low-frequency modes of a protein can be used to split its structure into well-defined domains (6), with the additional advantage of a smooth, almost continuous, description of their boundaries. So, since it is well known that most large-amplitude protein functional motions can be well described as combinations of almost rigid-body motions of entire structural domains (61, 62), the relationship found between these motions and the low-frequency modes of ENMs is just another demonstration that whole quasi-rigid domain motions are involved in such modes. On the other hand, it is not that difficult to admit that the spatial clustering of amino acids into domains can be revealed by studying protein dynamical properties, even at a crude level of description. A corollary of this line of thought is that ENMs should perform better, as far as low-frequency and large-amplitude motions are concerned, in the case of large, multi-domain systems.

#### **4.3. Applications**

As illustrated above, NMA of ENMs seems to have a clear predictive power. So, given both the simplicity of these models and their coarse-grained nature, many applications have been proposed. For instance, as early suggested, being able to guess the pattern of atomic fluctuations through Eq. 11 may prove useful for refining crystal structures (63, 64).

However, most applications take advantage of the possibility to predict atomic displacements through the reciprocal of Eq. 6, namely:

$$r_i = r_i^0 + \frac{1}{\sqrt{m_i}} \sum_k^{n_{\text{sub}}} \alpha_{ik} q_k \quad (15)$$

where  $n_{\text{sub}}$  is the number of low-frequency modes considered to be enough for performing an accurate prediction. In the simplest case, mode amplitudes can be varied arbitrarily, one mode after the other. Indeed, in the light of enough experimental data, the analysis of such trajectories can prove enough for getting insights about the nature of the functional motion of a protein (13, 65). Some of the conformations thus obtained can also allow for solving difficult molecular-replacement problems, although it is often necessary to explore at least a couple of modes in order to reach a useful conformation (18). More generally, Eq. 15 can be used so as to reduce the dimensionality of the system and, thus, to find more easily protein conformations fulfilling a given set of constraints. For instance, it has been used for fitting known structures into low-resolution electron density maps (19–21, 23) providing, for instance, more detailed structural data for systems of major interest, like the ribosome (22).

Note that Eq. 15 is linear. As a consequence, atom motions follow straight lines and local distortions (of most chemical bonds, valence angles, etc.) cannot be avoided. So, for many applications, as well as for obtaining well-behaved normal mode trajectories, the conformations thus generated need to be “regularized” (18), using for instance a detailed all-atom force field and standard energy-minimization techniques.

## 5. Conclusion

Fifteen years after their introduction in the realm of molecular biophysics (1), thanks to their simplicity as well as to their coarse-grained nature, ENMs are becoming more and more popular. Indeed, many applications have been proposed, noteworthy within the frame of various structural biology techniques.

From a theoretical point of view, their relationship with NMA is obvious, since both approaches rely on a quadratic form for the energy function, the former, *par définition*, the latter, as a consequence of a small displacement, so-called harmonic (or linear) approximation.

From an empirical point of view, it has been extensively shown that normal mode studies of ENMs yield low-frequency, large-amplitude, and collective, motions which prove often similar to those obtained with an all-atom model and a standard empirical force field.

This is likely to be a consequence of the robustness of these motions (42). Moreover, such motions often provide fair predictions for the pattern of thermal atomic fluctuations (e.g., the crystallographic B-factors) or for the kind of functional motion a given protein can perform (e.g., its conformational change upon ligand binding).

## References

1. Tirion M (1996) Low-amplitude elastic motions in proteins from a single-parameter atomic analysis. *Phys Rev Lett* 77:1905–1908
2. Noguti T, Go N (1982) Collective variable description of small-amplitude conformational fluctuations in a globular protein. *Nature* 296:776–778
3. Go N, Noguti T, Nishikawa T (1983) Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc Natl Acad Sci USA* 80:3696–3700
4. Levitt M, Sander C, Stern P (1983) Normal-mode dynamics of a protein: bovine pancreatic trypsin inhibitor. *Int J Quant Chem* 10:181–199
5. Bahar I, Atilgan AR, Erman B (1997) Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* 2:173–181
6. Hinsen K (1998) Analysis of domain motions by approximate normal mode calculations. *Proteins* 33:417–429
7. Micheletti C, Lattanzi G, Maritan A (2002) Elastic properties of proteins: insight on the folding process and evolutionary selection of native structures. *J Mol Biol* 321:909–921
8. Kondrashov D, Van Wynsberghe A, Bannen R, Cui Q, Phillips G (2007) Protein structural variation in computational models and crystallographic data. *Structure* 15:169–177

9. Bahar I, Cui Q (eds) (2005) Normal mode analysis: theory and applications to biological and chemical systems. C&H/CRC Mathematical and Computational Biology Series, vol. 9. CRC press, Boca Raton
10. Tama F, Sanejouand YH (2001) Conformational change of proteins arising from normal mode calculations. *Prot Engineering* 14:1–6
11. Krebs WG, Alexandrov V, Wilson CA, Echols N, Yu H, Gerstein M (2002) Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. *Proteins* 48:682–695
12. Delarue M, Sanejouand Y-H (2002) Simplified normal modes analysis of conformational transitions in DNA-dependant polymerases: the elastic network model. *J Mol Biol* 320:1011–1024
13. Valadie H, Lacapere J-J, Sanejouand Y-H, Etchebest C (2003) Dynamical properties of the MscL of Escherichia coli: a normal mode analysis. *J Mol Biol* 332:657–674
14. Suhre K, Sanejouand Y-H (2004) Elnémo: a normal mode server for protein movement analysis and the generation of templates for molecular replacement. *Nucl Ac Res* 32: W610–W614
15. Tama F, Brooks III C (2002) The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus. *J Mol Biol* 318:733–747
16. Tama F, Valle M, Frank J, Brooks III CL (2003) Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc Natl Acad Sci USA* 100:9319–9323
17. Tirion M, ben Avraham D, Lorenz M, Holmes K (1995) Normal modes as refinement parameters for the F-actin model. *Biophys J* 68:5–12
18. Suhre K, Sanejouand Y-H (2004) On the potential of normal mode analysis for solving difficult molecular replacement problems. *Act Cryst D* 60:796–799
19. Delarue M, Dumas P (2004) On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. *Proc Natl Acad Sci USA* 101:6957–6962
20. Tama F, Miyashita O, Brooks III CL (2004) Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. *J Mol Biol* 337:985–999
21. Hinsen K, Reuter N, Navaza J, Stokes DL, Lacapere JJ (2005) Normal mode-base fitting of atomic structure into electron density maps: application to sarcoplasmic reticulum Ca-ATPase. *Biophys J* 88:818–827
22. Mitra K, Schaffitzel C, Shaikh T, Tama F, Jenni S, Brooks 3rd C, Ban N, Frank J (2005) Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* 438:318–324
23. Suhre K, Navaza J, Sanejouand Y-H (2006) Norma: a tool for flexible fitting of high resolution protein structures into low resolution electron microscopy derived density maps. *Act Cryst D* 62:1098–1100
24. Levitt M, Warshel A (1975) Computer simulation of protein folding. *Nature* 253:694–698
25. Taketomi H, Ueda Y, Go N (1975) Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effect of specific amino acid sequence represented by specific inter-unit interactions. *Int J Pept Prot Res* 7:445–459
26. Shakhnovich E, Gutin A (1990) Enumeration of all compact conformations of copolymers with random sequence of links. *J Chem Phys* 93:5967
27. Li H, Tang C, Wingreen N (1997) Nature of driving force for protein folding: a result from analyzing the statistical potential. *Phys Rev lett* 79:765–768
28. Li H, Helling R, Tang C, Wingreen N (1996) Emergence of preferred structures in a simple model of protein folding. *Science* 273:666–669
29. Trinquier G, Sanejouand YH (1999) New protein-like properties of cubic lattice models. *Phys Rev E* 59:942–946
30. He Y, Chen Y, Alexander P, Bryan P, Orban J (2008) NMR structures of two designed proteins with high sequence identity but different fold and function. *Proc Natl Acad Sci USA* 105:14412
31. Goldstein H (1950) Classical mechanics. Addison-Wesley, Reading, MA
32. Wilson E, Decius J, Cross P (1955) Cross molecular vibrations. McGraw-Hill, New York
33. Sanejouand Y-H (1990) Ph.D. Thesis. Université de Paris XI, Orsay, France
34. Levy R, Perahia D, Karplus M (1982) Molecular dynamics of an alpha-helical polypeptide: temperature dependance and deviation from harmonic behavior. *Proc Natl Acad Sci USA* 79:1346–1350
35. Elber R, Karplus M (1987) Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. *Science* 235:318–321
36. Brooks B, Karplus M (1983) Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc Natl Acad Sci USA* 80:6571–6575

37. Lazaridis T, Karplus M (1999) Effective energy function for proteins in solution. *Proteins* 35:133–152
38. Haliloglu T, Bahar I, Erman B (1997) Gaussian dynamics of folded proteins. *Phys Rev Lett* 79:3090–3093
39. Atilgan A, Durell S, Jernigan R, Demirel M, Keskin O, Bahar I (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 80:505–515
40. Hinsen K, Kneller G (1999) A simplified force field for describing vibrational protein dynamics over the whole frequency range. *J Chem Phys* 111:10766
41. Kundu S, Melton J, Sorensen D, Phillips Jr G (2002) Dynamics of proteins in crystals: comparison of experiment with simple models. *Biophys J* 83:723–732
42. Nicolay S, Sanejouand Y-H (2006) Functional modes of proteins are among the most robust. *Phys Rev Lett* 96:078104
43. Miyazawa S, Jernigan R (1985) Estimation of effective inter-residue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules* 18:534–552
44. Miyazawa S, Jernigan R (1996) Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term for simulation and threading. *J Mol Biol* 256:623–644
45. Tama F (2000) Ph.D. Thesis. Université Paul Sabatier, Toulouse, France
46. Durand P, Trinquier G, Sanejouand YH (1994) A new approach for determining low-frequency normal modes in macromolecules. *Biopolymers* 34:759–771
47. Tama F, Gadea F-X, Marques O, Sanejouand Y-H (2000) Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins* 41:1–7
48. Kuriyan J, Weis W (1991) Rigid protein motion as a model for crystallographic temperature factors. *Proc Natl Acad Sci USA* 88:2773
49. Simonson T, Perahia D (1992) Normal modes of symmetric protein assemblies. Application to the tobacco mosaic virus protein disk. *Biophys J* 61:410–427
50. Hinsen K (2008) Structural flexibility in proteins: impact of the crystal environment. *Bioinformatics* 24:521
51. Riccardi D, Cui Q, Phillips Jr G (2009) Application of elastic network models to proteins in the crystalline state. *Biophys J* 96:464–475
52. Juanico B, Sanejouand Y-H, Piazza F, De Los Rios P (2007) Discrete breathers in nonlinear network models of proteins. *Phys Rev Lett* 99:238104
53. Sacquin-Mora S, Laforet E, Lavery R (2007) Locating the active sites of enzymes using mechanical properties. *Proteins* 67:350–359
54. Kraulis P (1991) Molscript: a program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 24:946–950
55. Marques O, Sanejouand Y-H (1995) Hinge-bending motion in citrate synthase arising from normal mode calculations. *Proteins* 23:557–560
56. Lu M, Ma J (2005) The role of shape in determining molecular motions. *Biophys J* 89: 2395–2401
57. Tama F, Brooks C (2006) Symmetry, form, and shape: guiding principles for robustness in macromolecular machines. *Annu Rev Biophys Biomol Struct* 35:115–133
58. McCammon JA, Gelin BR, Karplus M, Wolynes P (1976) The hinge-bending mode in lysozyme. *Nature* 262:325–326
59. Harrison R (1984) Variational calculation of the normal modes of a large macromolecules: methods and some initial results. *Biopolymers* 23:2943–2949
60. Perahia D, Mouawad L (1995) Computation of low-frequency normal modes in macromolecules: improvements to the method of diagonalization in a mixed basis and application to hemoglobin. *Comput Chem* 19:241–246
61. Branden C, Tooze J et al (1991) Introduction to protein structure. Garland Publishing, New York
62. Gerstein M, Krebs W (1998) A database of macromolecular motions. *Nucl Acid Res* 26:4280–4290
63. Diamond R (1990) On the use of normal modes in thermal parameter refinement: theory and application to the bovine pancreatic trypsin inhibitor. *Acta Cryst A* 46:425–435
64. Kidera A, Go N (1990) Refinement of protein dynamic structure: normal mode refinement. *Proc Natl Acad Sci USA* 87:3718–3722
65. Sanejouand Y-H (1996) Normal-mode analysis suggests important flexibility between the two N-terminal domains of CD4 and supports the hypothesis of a conformational change in CD4 upon HIV binding. *Prot Eng* 9:671–677

# Chapter 24

## An Introduction to Dissipative Particle Dynamics

Zhong-Yuan Lu and Yong-Lei Wang

### Abstract

Dissipative particle dynamics (DPD) is a particle-based mesoscopic simulation method, which facilitates the studies of thermodynamic and dynamic properties of soft matter systems at physically interesting length and time scales. In this method, molecule groups are clustered into the dissipative beads, and this coarse-graining procedure is a very important aspect of DPD as it allows significant computational speed-up. In this chapter, we introduce the DPD methodology, including its theoretical foundation and its parameterization. With this simulation technique, we can study complex behaviors of biological systems, such as the formation of vesicles and their fusion and fission processes, and the phase behavior of lipid membranes.

**Key words:** Dissipative particle dynamics, Parameterization, Biomacromolecules, Vesicle, Lipid membrane

---

### 1. Introduction

The way scientists looking at the biological world is changing dramatically. One reason is that experimentalists can determine structures of functional proteins and enzymes with high-resolution microscopy and explore how these biomolecules are expressed in the human body. The other factor contributing to present state of scientific research is the advent of computer simulation techniques. Computer simulation, making connection between microscopic details of a model and its macroscopic properties of interest, has emerged as a field of research with unique capabilities, most notably the ability to analyze and predict various phenomena.

Biomolecular systems, such as lipid membranes, are extremely complex. Typically, lipid molecules undergo motions like bond and angle fluctuations, trans-gauche conformation transitions, and rotation around the molecular axes. Moreover, lipids exhibit collective phenomena, including thermally activated dynamic processes, such as vesicle fusion and fission, as well as thermodynamic

phase transitions. Generally speaking, such systems are inherently hierachic: atomistic details of the lipid molecules are essential, to some extent, for their functioning; however, the global thermodynamic and structural properties of the lipid membrane are often regulated by larger length and time-scale motions.

For computer simulations, a single method cannot be able to cover the full spatiotemporal range. Actually, different simulation approaches are needed to describe molecular motions at different scales. The main practical limitation comes from available computational capacity: detailed models can only be studied over limited length and time scales. For example, electronic structure calculations are only practical up to tens or hundreds of atoms. Molecular dynamics (MD) is a powerful simulation technique proven to produce highly realistic results in a wide variety of applications. However, the computational costs of detailed interaction model in this paradigm severely limit its applicability beyond extremely small spatiotemporal scales. Such restrictions have driven the development of coarse-grained model and mesoscopic simulation techniques. By attempting to preserve essential physical information, mesoscopic simulations could be able to represent thermodynamic properties and dynamic behaviors of the underlying microscopic system. This provides fascinating alternative to investigate complex behaviors of biological systems with simplified model.

A wide variety of simulation techniques have been developed to model phenomena occurring at the mesoscale. Among these methods, dissipative particle dynamics (DPD) is an approach originally invented for carrying out particle-based simulations reflecting the correct hydrodynamic behavior of complex fluids. In the past 20 years, DPD had a great development and turned to be one of the most promising simulation methods to investigate mesoscopic properties of soft matter. In this chapter, we give a survey about the theoretical background of this method.

---

## 2. Dissipative Particle Dynamics

DPD was originally proposed by Hoogerbrugge and Koelman as a novel scheme for mesoscopic simulations of complex fluids (1, 2). It combines the advantages of MD and lattice gas automata. By integrating out the internal degrees of freedom, fast motions in the system are averaged out and the remaining structures are represented by a set of “beads” with given masses and sizes. A bead represents a small region of fluids or molecular groups that is large on molecular scale, but is still macroscopically small. Beads interact with each other via soft potentials and their time evolutions are assumed to be governed by the Newton’s equations of motion. Español and Warren reinterpreted the DPD model and derived the correct fluctuation–dissipation relation for the friction

and noise terms in DPD (3). Formulating the Fokker-Planck equation for the DPD model, they obtained the condition under which the steady state solution to the stochastic differential equation would satisfy the canonical distribution with a well-defined temperature.

Generally speaking, the regular DPD method is described by

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i, \quad m_i \frac{d\vec{v}_i}{dt} = \vec{f}_i, \quad (1)$$

where  $\vec{r}_i$ ,  $\vec{v}_i$ , and  $m_i$  are position, velocity, and mass of bead  $i$ , respectively. For simplicity, the bead masses are taken as the unit in most of the DPD simulations. The total force  $\vec{f}_i$  acting on each bead consists of three parts:

$$\vec{f}_i = \sum_{i \neq j} (\vec{F}_{ij}^C + \vec{F}_{ij}^D + \vec{F}_{ij}^R), \quad (2)$$

where  $\vec{F}_{ij}^C$ ,  $\vec{F}_{ij}^D$ , and  $\vec{F}_{ij}^R$  represent conservative, dissipative, and random forces between bead  $i$  and  $j$ , respectively.

One important advantage of the DPD model is that all interactions are pair-wise additive and satisfy Newton's third law, so that linear and angular momentums are conserved.

## 2.1. Conservative Force

The thermodynamic properties of the system are determined by the conservative force. It is directed along the line connecting the centers of interacting beads and typically takes the repulsion form, as shown in (3), which is much softer than that commonly used in MD simulations.

$$\vec{F}_{ij}^C = \begin{cases} \alpha_{ij} \left(1 - \frac{r_{ij}}{r_c}\right) \vec{e}_{ij}, & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases}. \quad (3)$$

Here  $\alpha_{ij}$  is the maximum repulsion strength, and  $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ ,  $r_{ij} = |\vec{r}_{ij}|$ ,  $\vec{e}_{ij} = \vec{r}_{ij}/r_{ij}$ . The corresponding pair conservative potential is shown in (4).

$$U_{ij} = \begin{cases} \frac{\alpha_{ij}}{2} \left(1 - \frac{r_{ij}}{r_c}\right)^2, & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases}. \quad (4)$$

For comparison, the Lennard-Jones (LJ) potential, Weeks-Chandler-Anderson (WCA) potential (also known as the truncated and shifted LJ potential at its minimum value) and DPD potential are shown in Fig. 1.

The conservative parameter  $\alpha$  [in (3)] is one of the most important parameters in DPD simulations. Based on the assumption that all dissipative beads are spheres with equal size and mass, Groot and Warren (4) mapped the DPD repulsion parameter  $\alpha_{ii}$  for one-component system to the compressibility of fluid and determined  $\alpha_{ii}$  as a function of number density  $\rho$  and system temperature.

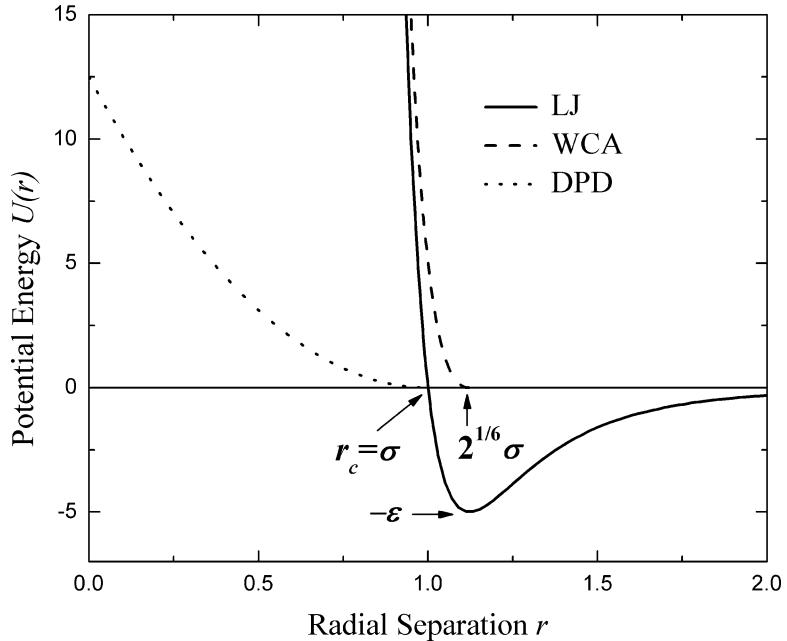


Fig. 1. The comparison between the LJ potential, the purely repulsive WCA potential and the soft potential in DPD.

In order to consider interactions between components with different chemical identities, they related  $\alpha_{ij}$  between different components to the solubilities, usually specified by Flory-Huggins parameter  $\chi_{ij}$  in polymer chemistry. This link is an effective mapping procedure between DPD model and Flory-Huggins theory at fixed density, and the repulsion parameter  $\alpha$  can be readily defined by experimental thermodynamic parameters.

Maiti and McGrother (5) reexamined the Groot-Warren mapping scheme and presented another procedure to coarse-grain complex fluid from its underlying chemistry. Taking the DPD repulsion parameter as a function of bead size, they found that the (experimental) interfacial tension in a segregated binary mixture is an ideal physical property to define the DPD repulsion parameter. Hence, using interfacial tension as the fitting target would be another choice to determine DPD interaction parameter  $\alpha$ .

Flekkøy and Coveney proposed that the dissipative beads could also be defined as cells on the Voronoi lattice with variable masses and sizes (6, 7). This Voronoi lattice arises naturally from the coarse-graining procedure, which may be applied iteratively and thus represents a form of renormalization group mapping. It enables not only to select any desired local scale for mesoscopic description of a given problem but also to deal with situations in which several different length scales are simultaneously present.

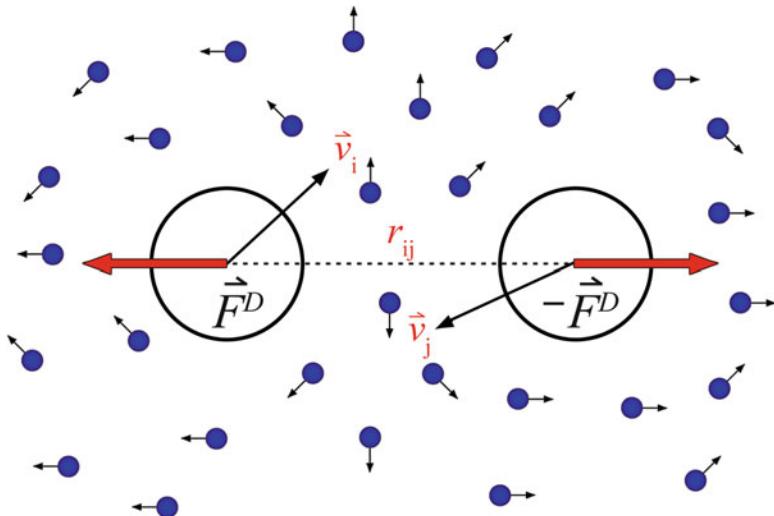


Fig. 2. Dissipative force  $\vec{F}_{ij}^D$ .

## 2.2. Dissipative and Random Forces

Both dissipative and random forces between all pairs of interacting beads are subjected to the same short-range cut-off distance. These two forces are given, respectively, by

$$\vec{F}_{ij}^D = -\gamma w^D(r_{ij})(\vec{e}_{ij} \cdot \vec{v}_{ij})\vec{e}_{ij}, \quad (5)$$

and

$$\vec{F}_{ij}^R = \sigma w^R(r_{ij})\theta_{ij}\vec{e}_{ij}, \quad (6)$$

where  $\gamma$  is the dissipation strength,  $\sigma$  is the noise strength, and  $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$  is the relative velocity between two interacting beads.  $\theta_{ij}$  is a Gaussian white noise function with symmetry property to ensure conservation of momentum and has the following stochastic properties.

$$\langle \theta_{ij}(t) \rangle = 0 \quad \text{and} \quad \langle \theta_{ij}(t)\theta_{kl}(t') \rangle = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\delta(t - t'), \quad (7)$$

$w^D$  and  $w^R$  are  $r$ -dependent weight functions, which are not independent with each other but must be coupled together through a fluctuation-dissipation relation, as shown in (8).

$$w^D(r_{ij}) = [w^R(r_{ij})]^2, \quad (8)$$

$$\sigma^2 = 2\gamma k_B T.$$

These conditions are necessary to ensure the system under consideration in a canonical ensemble (NVT) at thermodynamic equilibrium state. From (5) we can see that dissipative force is proportional to the relative velocity so as to remove energy by decreasing velocities of the interacting beads. As shown in Fig. 2, two beads interacting and moving toward or away from

each other will be slowed down by the dissipative forces. This inherently reflects the friction between the coarse grained beads on the mesoscale.

The random force acts as a heat pump and inputs energy back into the system to counteract the energy sink due to the dissipative force. The two forces are coupled together serving as a momentum-conserving thermostat, which is, by construction, local and Galilean invariant, and therefore can recover the correct hydrodynamic behavior on sufficiently large length and time scales.

Schmidt number  $Sc$  is a dimensionless number defined as the ratio between the kinetic viscosity for the momentum transfer,  $\nu$ , and the diffusion coefficient  $D$  for mass transfer:  $Sc = \nu/D$ . It is used to characterize typical fluid flows in which simultaneous momentum and mass transfers are present. In classical DPD simulations, the ultra soft potential leads to slower momentum transfer so that  $Sc$  is close to 1. This value corresponds to a situation often found in gases, while in real liquids  $Sc$  is of the order  $10^3$  or even larger. In order to manifest the physical behavior of complex fluid accurately, an appropriate scheme to capture the correct  $Sc$  value is needed.

Lowe formulated a new approach for DPD based on Andersen thermostat (called Lowe-Andersen or LA method), which replaces the temperature controlling formed by dissipative and random forces, and meanwhile conserves momentum (8). The relative velocities of interacting beads are taken from the Maxwellian distribution which maintains satisfactory temperature control. It was found that by using LA thermostat, correct radial distribution function (RDF) and high Schmidt number corresponding to polymeric systems can be achieved (9, 10).

### **2.3. Integration Scheme**

In DPD, the pair-wise coupling of the beads through random and dissipative forces makes the integration of the equations of motion a nontrivial task. One reason is that the random forces make the DPD equations of motion formally stochastic and there is no guarantee for time reversibility during integration process. The other difficulty arises from the dissipative forces, which depend explicitly on the relative velocities between interacting beads, while the bead velocities in turn are partially governed by the dissipative forces. In order to represent the stochastic motions of DPD beads, various finite time step implementations based on analogy to higher-order solvers for conservative systems have been proposed. Some of them have shown good performance (4, 8, 11–17).

#### *2.3.1. Verlet-Based Algorithm*

Groot and Warren adopted a modified velocity-Verlet algorithm for DPD simulations (GW-VV) (4). In respect that the dissipative forces depend on velocities, they empirically introduced a variable  $\lambda$  to predict new velocities, which can be used to calculate the dissipative forces. The value of  $\lambda$  depends on the system of interest

and practically  $\lambda = 0.65$  has been overwhelmingly used in typical DPD simulations. Detailed algorithm is shown in (9).

$$\begin{aligned}
 \vec{v}_i &\leftarrow \vec{v}_i + \frac{1}{2m} (\vec{F}_i^C \delta t + \vec{F}_i^D \delta t + \vec{F}_i^R \sqrt{\delta t}), \\
 \vec{v}_i^0 &\leftarrow \vec{v}_i + \lambda \frac{1}{m} (\vec{F}_i^C \delta t + \vec{F}_i^D \delta t + \vec{F}_i^R \sqrt{\delta t}), \\
 \vec{r}_i &\leftarrow \vec{r}_i + \vec{v}_i \delta t, \\
 \text{Calculate } \vec{F}_i^C(\vec{r}), \vec{F}_i^D(\vec{r}_i, \vec{v}_i^0), \vec{F}_i^R(\vec{r}), \\
 \vec{v}_i &\leftarrow \vec{v}_i + \frac{1}{2m} (\vec{F}_i^C \delta t + \vec{F}_i^D \delta t + \vec{F}_i^R \sqrt{\delta t}). \tag{9}
 \end{aligned}$$

Besold et al. considered using traditional velocity-Verlet algorithm directly for DPD integration (DPD-VV) (13). DPD-VV algorithm, described in (10), accounts for the self-consistent relationship between dissipative forces and velocities in an approximate fashion by updating dissipative forces for a second time at the end of each integration step. This integration scheme performed well if the conservative interactions are sufficiently strong and dominate the behavior of the system in which the dissipative and random forces act only as a weak thermostat.

$$\begin{aligned}
 \vec{v}_i &\leftarrow \vec{v}_i + \frac{1}{2m} (\vec{F}_i^C \delta t + \vec{F}_i^D \delta t + \vec{F}_i^R \sqrt{\delta t}), \\
 \vec{r}_i &\leftarrow \vec{r}_i + \vec{v}_i \delta t, \\
 \text{Calculate } \vec{F}_i^C(\vec{r}), \vec{F}_i^D(\vec{r}_i, \vec{v}_i), \vec{F}_i^R(\vec{r}), \\
 \vec{v}_i &\leftarrow \vec{v}_i + \frac{1}{2m} (\vec{F}_i^C \delta t + \vec{F}_i^D \delta t + \vec{F}_i^R \sqrt{\delta t}), \\
 \text{Calculate } \vec{F}_i^D(\vec{r}_i, \vec{v}_i). \tag{10}
 \end{aligned}$$

By minimizing the mutual dependence between dissipative forces and velocities, den Otter and Clarke (17) proposed a new approach (OC integrator), which adopts leap-frog algorithm with predefined variables  $\alpha_{\text{OC}}$  and  $\beta_{\text{OC}}$ . These two parameters, describing the relative weights of random forces with respect to dissipative and conservative forces, are determined prior to the actual DPD simulations by calculating the averages from the system in which both the kinetic and the configurational temperatures equal to the desired temperature. It had been shown that OC integrator performed well in both ideal gas and softly interacting DPD fluid. Details of the OC integrator are shown in (11).

$$\begin{aligned}
 \vec{v}_i &\leftarrow \vec{v}_i + \alpha_{\text{OC}} \frac{1}{m} (\vec{F}_i^C \delta t + \vec{F}_i^D \delta t) + \beta_{\text{OC}} \frac{1}{m} \vec{F}_i^R \sqrt{\delta t}, \\
 \vec{r}_i &\leftarrow \vec{r}_i + \vec{v}_i \delta t, \\
 \text{Calculate } \vec{F}_i^C(\vec{r}), \vec{F}_i^D(\vec{r}_i, \vec{v}_i), \vec{F}_i^R(\vec{r}). \tag{11}
 \end{aligned}$$

### 2.3.2. The LA Algorithm

The LA method obeys the same conservation laws and is similar to DPD as it is aimed for studies of coarse-grained models in terms of soft interactions (8). In the integration of LA method, one firstly integrates Newton's equations of motion with time step  $\delta t$ , as expressed in (12), and then thermalizes the system with Andersen-type thermostat. For all pairs of beads maintaining  $r_{ij} < r_c$ , they exchange velocities with the bath via a valid Monte Carlo scheme. This LA algorithm can also produce canonical distribution and is by construction local and Galilean invariant. It therefore preserves hydrodynamics, just as regular DPD method. In LA, there is no dissipative or random force so that the system will not suffer from the self-consistent relation between dissipative forces and velocities. Another advantage is that the rate of how often bead velocities are thermalized can be varied over a wide range, which implies that the dynamic properties of the system can be tuned in a controlled fashion.

$$\vec{v}_i \leftarrow \vec{v}_i + \frac{1}{2m} \vec{F}_i^C \delta t,$$

$$\vec{r}_i \leftarrow \vec{r}_i + \vec{v}_i \delta t,$$

$$\text{Calculate } \vec{F}_i^C(\vec{r}),$$

For all pairs of particles for which  $r_{ij} \leq r_c$ :

Generate  $\vec{v}_{ij}^0 \cdot \vec{e}_{ij}$  from a Maxwell distribution  $\xi_{ij} \sqrt{2k_B T^*/m}$ ,

$$2\Delta_{ij} = \vec{e}_{ij}(\vec{v}_{ij}^0 - \vec{v}_{ij}) \cdot \vec{e}_{ij},$$

$$\vec{v}_i \leftarrow \vec{v}_i + \Delta_{ij},$$

$$\vec{v}_j \leftarrow \vec{v}_j - \Delta_{ij},$$

with probability  $\Gamma \Delta t$ .

(12)

As compared to MD simulations, the main advantage of DPD method is its soft interaction potential, which allows for much larger time step. However, in order to keep equilibrium state of the system in canonical ensemble, time step  $\delta t$  should not be too large, which had been confirmed in the simulations by Marsh and Yeomans (18). In practice, one should therefore carefully check the dependence of equilibrium properties on time step, which may vary if different integration algorithms are adopted. With GW-VV integrator, Groot and Warren found that time step  $\delta t = 0.06$  will be the acceptable upper limit without significant loss of temperature control (4). Den Otter and Clarke found that GW-VV scheme shows a good agreement between the kinetic and the configurational temperatures at small time steps but is pessimistic at larger ones (17). Other studies had shown that besides integration algorithm, the choice of time step is indeed a crucial aspect for DPD simulations (19). Allen proposed a straightforward approach to estimate the optimum time step for desired configurational

temperature from hypervirial relation (20). He found that  $\delta t = 0.02$  may be the largest time step of the choice.

Practically speaking, the choice of integration scheme depends on the system of interest. If the structural properties of the system are the main focus, the integration time step can take a larger value for accelerating the evolution of the system to equilibrium. However, if one focuses on the dynamic behaviors, both integrator and time step should be chosen carefully.

### 3. Parameterization and Simulation Tutorial

The basic characteristic of a mesoscopic simulation method is that it could correctly reproduce physical properties of a real system beyond certain length and time scales. As mentioned above, DPD beads do not represent atoms but liquid elements or a group of molecules. This requires specified mapping procedure between microscopic system and the parameters for coarse-grained model.

For single component fluid, such as water, some of the important thermodynamic properties should be described correctly by DPD model (4). This requires that the pressure, the liquid compressibility, and so on can be mapped onto a real fluid. We may consider, for example, the dimensionless reciprocal compressibility, defined by

$$\kappa^{-1} = \frac{1}{\rho k_B T \kappa_T} = \frac{1}{k_B T} \left( \frac{\partial P}{\partial \rho} \right)_T, \quad (13)$$

where  $\kappa_T$  is the isothermal compressibility of the liquid with  $\kappa_T = -\frac{1}{V}(\partial V / \partial p)_{T,N}$ . For DPD single component system at sufficient large densities, Groot and Warren (4) found that the equation of state of DPD liquid should be described by

$$P = \rho k_B T + 0.1 \alpha k_B T \rho^2 \quad (14)$$

This means that the dimensionless compressibility, as introduced in (13), is  $K^{-1} = 1 + 0.2 \alpha \rho / k_B T$ . For water at room temperature (298 K), the dimensionless compressibility takes the value  $K^{-1} = 15.98$ . Therefore the repulsion parameter for coarse-grained DPD water bead can be expressed by  $\alpha_{ii} = 75 k_B T / \rho$ .  $\rho = 3$  is usually adopted in DPD simulations for maximizing the computation efficiency, thus the repulsion parameter between coarse-grained DPD water beads should be  $\alpha_{ii} = 25 k_B T$ . Similarly, the repulsion parameters for other types of DPD fluid beads can be estimated. But in practice, this  $\alpha_{ii} = 25 k_B T$  for  $\rho = 3$  is widely used in DPD simulations, since water can represent a large group of incompressible liquids.

The repulsion parameter  $\alpha_{ij}$  between different types of beads in a fluid mixture should be further defined for DPD simulations. Groot and Warren mapped  $\alpha_{ij}$  to the  $\chi_{ij}$  parameter in Flory-Huggins (FH) theory, so that DPD can be viewed, to some extent, as a continuous version of lattice FH model.

FH theory presents a generic thermodynamic model for polymer solutions. In this theory, the beads, representing either polymer segments or solvents, are restricted on the lattices. The internal energy is described as a perturbation from ideal mixing, i.e., only the excess energy over pure components is taken into account. For binary system, the mixing free energy per lattice site can be expressed by

$$\frac{F_{\text{mix}}}{k_B T} = \frac{\phi_A}{N_A} \ln \phi_A + \frac{\phi_B}{N_B} \ln \phi_B + \chi \phi_A \phi_B, \quad (15)$$

where  $N_A$  and  $N_B$  are number of segments per A and B polymer, respectively,  $\chi$  is the Flory-Huggins interaction parameter, and  $\phi_A$  and  $\phi_B$  are volume fractions of A and B components with the relation  $\phi_A + \phi_B = 1$ . In (15), the first two terms capture the combinatorial entropy of mixing, which always favors mixing but is not too large. The parameter  $\chi$  represents the interaction between A and B segments. When  $\chi$  is positive and larger than the critical value, mixture will separate into two phases.

From the equation of state,  $P = \rho k_B T + 0.1\alpha k_B T \rho^2$ , for DPD fluid, one can get the following equation

$$\rho k_B T + 0.1\alpha k_B T \rho^2 = P = -\left(\frac{\partial F}{\partial V}\right)_T = -V \frac{\partial f_v}{\partial V} - f_v = \rho \frac{\partial f_v}{\partial \rho} - f_v, \quad (16)$$

i.e.,

$$\rho k_B T + 0.1\alpha k_B T \rho^2 = \rho \frac{\partial f_v}{\partial \rho} - f_v, \quad (17)$$

where  $f_v$  is the free energy per volume. Dividing both sides by  $\rho^2$  and follow an integration, we can obtain the free energy per volume for single component DPD fluid as

$$\frac{f_v}{k_B T} = \rho \ln \rho + 0.1\alpha \rho^2 + c_1 \rho \quad (18)$$

For a binary system,  $f_v$  can be similarly written as

$$\begin{aligned} \frac{f_v}{k_B T} = & \rho_A \ln \rho_A + \rho_B \ln \rho_B + 0.1\alpha_{AA} \rho_A^2 + 0.1\alpha_{BB} \rho_B^2 \\ & + c_1 \rho_A + c_1 \rho_B + 0.2\alpha_{AB} \rho_A \rho_B \end{aligned} \quad (19)$$

Assuming that  $\alpha_{AA} = \alpha_{BB} = \alpha$  and  $\rho_A + \rho_B = \rho$  is constant, (19) can be expressed by

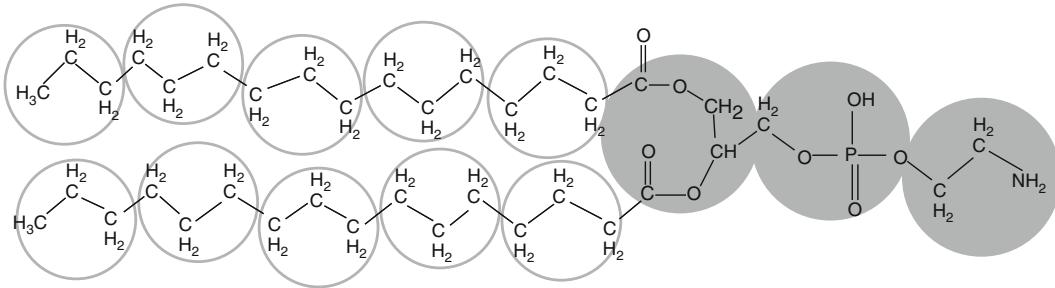


Fig. 3. A phospholipid molecule and its mapping on DPD model. Each bead represents roughly the same volume.

$$\begin{aligned} \frac{f_v}{\rho k_B T} &= \phi_A \ln \phi_A + \phi_B \ln \phi_B + \ln \rho (\phi_A + \phi_B) \\ &\quad + c_1 + 0.1 \left( \alpha \rho - \frac{2\alpha \rho_A \rho_B}{\rho} + \frac{2\alpha_{AB} \rho_A \rho_B}{\rho} \right) \\ &= \phi_A \ln \phi_A + \phi_B \ln \phi_B + 0.2(\alpha_{AB} - \alpha)\phi_A \phi_B \rho + c \end{aligned} \quad (20)$$

in which  $\phi_A = \rho_A / \rho$ . By defining  $\chi = 0.2\rho(\alpha_{AB} - \alpha)$ , (20) turns to be

$$\frac{f_v}{\rho k_B T} = \phi_A \ln \phi_A + \phi_B \ln \phi_B + \chi \phi_A \phi_B + c \quad (21)$$

One can see from (21) that the free energy per volume is very similar to the Flory-Huggins type free energy per site as shown in (15). Groot and Warren (4) determined a linear relation between  $\chi$  and  $\alpha$ , as shown in (22), by fitting to the interfacial properties for binary fluid mixtures.

$$\begin{aligned} \chi &= (0.286 \pm 0.002)(\alpha_{AB} - \alpha) \quad \text{for } \rho = 3, \\ \chi &= (0.689 \pm 0.002)(\alpha_{AB} - \alpha) \quad \text{for } \rho = 5. \end{aligned} \quad (22)$$

The mapping scheme between FH  $\chi_{ij}$  parameter and DPD repulsion parameter  $\alpha_{ij}$  establishes a solid foundation of DPD simulations. For a complex fluid system, once FH  $\chi$  parameter between different components is known, one can easily derive the corresponding DPD repulsion parameter  $\alpha_{ij}$ . In practice, the FH  $\chi$ -parameter can be obtained via different methods, for example from the solubility parameters, which are readily obtained in experiments or from MD simulations.

A good example for the successful application of DPD methods on biomolecule and surfactant systems was illustrated by Groot and Rabone, who parameterized the conservative force-field for DPD to represent phosphatidylethanolamine (PE) lipids and polyethylene oxide (PEO) surfactant (21). For the lipid molecule, triplet of methyl groups was coarse-grained into one tail bead, whereas the glycerol head group was described by three beads, as show in Fig. 3. Coarse-grained lipid molecule was constructed by linking beads together using Hookean springs with potential  $U_{i,i+1}(r) = 0.5k(r_{i,i+1} - l_0)^2$ , where  $i$  and  $i + 1$  label

adjacent beads in the coarse-grained molecule. Spring constant  $k$  and equilibrium bond length  $l_0$  were chosen so as to fix the average bond length to desired value. For PEO, the volume for three EO groups is the same as that for six  $\text{CH}_2$  groups, which means one DPD bead may represent 1.5 EO groups. Interaction parameters between DPD beads can be derived from experimental compressibility and solubility data. To produce the correct solubility of hexane, heptanes, and octane in water, Groot and Rabone found  $\chi_{\text{hydrocarbon}-\text{water}} \approx 6.0$ , which appears to be relatively independent of temperature. The second  $\chi$ -parameter to match is the one between PEO and water. Based on the fact that PEO and water mix in almost all ratios at room temperature but not mix ideally at higher temperatures, we can start from their demixing temperature, then extrapolate the temperature dependent  $\chi$ -parameter back to room temperature. Following this routine, Groot and Rabone estimated that the interaction parameter between PEO and water beads should be  $\chi_{\text{PEO}-\text{water}} \approx 0.30 \pm 0.04$ . The third FH parameter to fit is  $\chi_{\text{PEO}-\text{hydrocarbon}}$ . Because an EO group consists of  $2/3$  of hydrocarbon and  $1/3$  of oxygen, Groot and Rabone intuitively assumed that the  $\chi$  parameter between C and EO is  $1/3$  of that between C and water, that is  $\chi_{\text{PEO}-\text{hydrocarbon}} = 2.0$ . The lipid head groups contain more oxygen than EO does and also possess partial charges, therefore they can be treated as waters. With these parameters, Groot and Rabone constructed tensionless lipid membrane and found that the area per lipid molecule from simulation matches very well with that from experiment. Furthermore, they investigated pore formation in a lipid bilayer induced by the presence of nonionic surfactant.

Another good example was shown by Smit and coworkers, who had considered a model of dimyristoyl phosphatidyl choline (DMPC) with different coarse-grained levels (22). By matching the area per lipid and the hydrophobic thickness of DMPC bilayer with corresponding experimental values, they refined the interaction parameters between different species. Based on this model, they investigated the lipid membrane phase behavior as a function of temperature, head group interaction, and tail length (23, 24). At low values of head-head repulsion, the bilayer undergoes transitions from the sub gel phase  $L_c$  via the flat gel phase  $L_\beta$  to the fluid phase  $L_\alpha$  with increasing temperature (see Fig. 4). For higher values of head-head repulsion, the transition from the  $L_c$  to the  $L_\alpha$  phase occurred via the tilted gel phase  $L_{\beta'}$ , and sometimes the rippled phase  $P_{\beta'}$  appears during transition, depending on the tail length. Moreover, by introducing small molecules in the membrane, such as alcohol (25, 26) and cholesterol (27), the small molecules can be absorbed inside the membrane and this adsorption may induce structure changes of the membrane, which subsequently affects the functions of intrinsic membrane proteins.

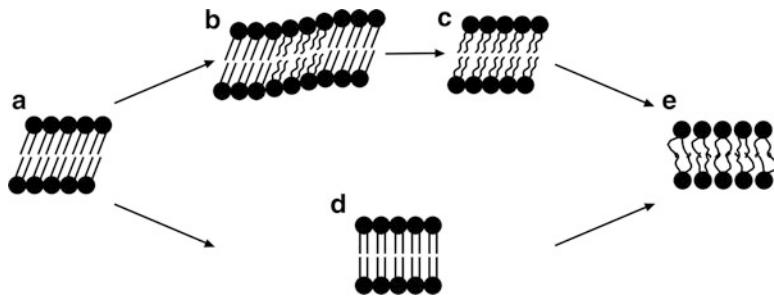


Fig. 4. Schematic drawing of the various lipid bilayer phases: (a)  $L_c$  phase, (b)  $P_\beta$  phase, (c)  $L_\beta$  phase, (d)  $L_\beta$  phase, and (e)  $L_x$  phase.

Shillcock, Lipowsky, and coworkers introduced another phenomenological coarse-grained model for lipids (28, 29). Lipid bilayers are composed of amphiphiles with no specific chemical identity. The repulsion parameters between DPD beads can be estimated from experimental thermodynamic properties, such as surface tension, membrane area stretch modulus, and bending rigidity. With this general lipid model, they studied tension-induced (30–33) and transmembrane protein-induced (34) vesicle fusions. Refer to the chapter 26 by Shillcock of this book for more details.

In practice, the procedure for determining the coarse-grained potential for biomolecular simulations is by no means unique. With inverse techniques, such as reverse Monte Carlo (RMC) and iterative Boltzmann inversion (IBI) methods, one can retrieve numerical coarse-grained potentials from the structures of the complex liquids. By transferring numerical function forms from RMC into DPD, Lyubartsev et al. found that the RDF computed from coarse-grained DPD simulations agree well with those obtained from all-atomistic MD simulations (35). Chen et al. adopted IBI method to coarse-grain the polyethylene systems (36). The numerical potentials were derived through an automatic updating program by mapping the RDF from LA method onto the ones from MD simulations. The so-obtained coarse-grained numerical potentials can be applied in larger systems under the same thermodynamic conditions. Meantime, they also introduced a possible scheme, based on the Rosenbluth sampling, to fine-grain the equilibrated coarse-grained system which may be further used in atomistic MD simulations. This coarse- and fine-graining procedure allows us to achieve the equilibrium fast enough on a mesoscopic length and time scale.

Basically, in DPD, the thermodynamic properties of the system are mainly controlled by conservative force. While the dynamic properties are intimately tied to the trajectories as they depend on the autocorrelation of the velocities and forces through the Green-Kubo relations and consequently related to thermostat.

With effective thermostat, the system of interest should produce correct dynamic properties. The dissipative and random forces are virtually coupled together to form a DPD thermostat, in which, as we have mentioned above,  $w^D$  and  $w^R$  cannot be chosen arbitrarily. Practically, in most of DPD simulations, (23) is widely adopted due to its simple form.

$$w^D(r_{ij}) = [w^R(r_{ij})]^2 = \begin{cases} \left(1 - \frac{r_{ij}}{r_c}\right)^2, & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases}. \quad (23)$$

Kremer and coworkers adopted a more computationally efficient form given by (24) (37, 38). Other forms of  $w^D$  and  $w^R$  are also possible.

$$w^D(r_{ij}) = [w^R(r_{ij})]^2 = \begin{cases} 1, & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases}. \quad (24)$$

For complex fluid systems, in order to obtain the precise dynamic properties,  $w^D$  and  $w^R$  should be gauged in detail. Eriks-son et al. illustrated that the thermostat should normally be considered as an integrated part of the effective dynamics, not only an added mechanism to ensure correct relaxation toward equilibrium (39). By tuning the magnitude and functional form of the dissipative and random forces, the diffusion coefficient and viscosity of the original system can be approximately matched by the coarse-grained model. For the numerical potential obtained from IBI method, Qian et al. showed that by selecting proper noise strength, the dynamic properties for the full-size system and for the coarse-grained model can be in consistent with each other (40).

One of the most attractive features of DPD method is its enormous versatility in order to construct simple models for complex fluids. In DPD, the Newtonian fluid is made “complex” by adding additional interactions between fluid particles. Just by tuning the conservative interactions between fluid particles, one can easily construct models for biomacromolecules, polyelectrolyte, and lipid molecules. Given the simple implementation of this method, DPD has been proven to be a powerful technique to study the behavior of complex fluids, such as the formation of micelles and vesicles (41) and their fusion and fission behavior (42–45). Another fascinating characteristics of DPD method is its thermostat, which can capture the hydrodynamic behavior of the fluid on sufficiently larger scales. With this merit, the DPD thermostat can also be adopted independently to simulate the dynamic behavior of, such as red blood cells (46) and vesicles (47), under flow fields.

Although the classical DPD model can successfully describe the complex fluid systems at constant temperatures, it is not able to conserve the total energy. Espa  ol, Avalos, and Mackie independently (48–50) proposed a generalized DPD algorithm that

incorporates the conservation of the total energy (so called EDPD) in particle–particle interactions, such that the thermal conduction and temperature gradients could be modeled on the hydrodynamic scale. Español (51) combined DPD and smoothed particle hydrodynamics (SPH) together and proposed the smoothed dissipative particle dynamics (SDPD) model, which improves some inherent deficiencies in original DPD method. To obtain the correct description of nonideal fluids, Pagonabarraga and Frenkel (52) introduced many-body DPD method (MDPD). In this method, the amplitude of the soft repulsions depends on the local excess free energy, which provides a wider range of possibilities for the equation of state. Following this routine, Trofimov et al. (53) refined this model for strongly nonideal systems and generalized it to multicomponent mixtures, and later they extended this model to constant pressure ensemble (54). With these extended DPD models, one could investigate the vapor–liquid coexistence phenomena (55).

## 4. Conclusions

DPD has been proved to be a versatile simulation technique, which facilitates the studies of thermodynamic and dynamic properties of soft matter systems at physically interesting length and time scales. Two key innovations in DPD are soft repulsions and pair-wise additive dissipative and random forces. The conservative force is extremely soft. After fitting to the compressibility or solubility of real fluid, DPD is able to capture many essential properties at mesoscopic level. The dissipative and random forces combine together to form as a momentum-conserving thermostat, therefore DPD method can preserve the hydrodynamic modes which are important to correctly describe complex fluids. With the original DPD method and its extended model, one can handle variety of problems such as phase behavior of biological membranes, biomacromolecules subjected to different flows and geometries, and morphology transitions of self-assembly entities. These simulation results could provide a solid validation to theoretical characterization and a natural complement to experimental observation.

## References

1. Hoogerbrugge PJ, Koelman JMVA (1992) Simulation microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys Lett* 19:155–160
2. Koelman JMVA, Hoogerbrugge PJ (1993) Dynamic simulation of hard-sphere suspensions under steady shear. *Europhys Lett* 21:363–368
3. Español P, Warren PB (1995) Statistical mechanics of dissipative particle dynamics. *Europhys Lett* 30:191–196
4. Groot RD, Warren PB (1997) Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J Chem Phys* 107:4423–4435

5. Maiti A, McGrother S (2004) Bead-bead interaction parameters in dissipative particle dynamics: relation to bead-size, solubility parameter, and surface tension. *J Chem Phys* 120:1594–1601
6. Flekkøy EG, Coveney PV (1999) From molecular dynamics to dissipative particle dynamics. *Phys Rev Lett* 83:1775–1778
7. Flekkøy EG, Coveney PV, Fabritiis GD (2000) Foundation of dissipative particle dynamics. *Phys Rev E* 62:2140–2157
8. Lowe CP (1999) An alternative approach to dissipative particle dynamics. *Europhys Lett* 47:145–151
9. Chen LJ, Lu ZY, Qian HJ et al (2005) The effect of Lowe-Andersen temperature controlling method on the polymer properties in mesoscopic simulation. *J Chem Phys* 122:104907
10. Koopman EA, Lowe CP (2006) Advantages of a Lowe-Andersen thermostat in molecular dynamics simulations. *J Chem Phys* 124:204103
11. Novik KE, Coveney PV (1998) Finite-difference methods for simulation models incorporating nonconservative forces. *J Chem Phys* 109:7667–7677
12. Pagonabarrage I, Hagen MHJ, Frenkel D (1998) Self-consistent dissipative particle dynamics algorithm. *Europhys Lett* 42:377–382
13. Besold G, Vattulainen I, Karttunen M et al (2000) Towards better integrators for dissipative particle dynamics simulations. *Phys Rev E* 62:R7611–R7614
14. Vattulainen I, Karttunen M, Besold G et al (2002) Integration schemes for dissipative particle dynamics simulation: from softly interacting systems towards hybrid models. *J Chem Phys* 116:3967–3979
15. Nikunen P, Karttunen M, Vattulainen I (2003) How would you integrate the equations of motion in dissipative particle dynamics simulations. *Comput Phys Commun* 153:407–423
16. Symeonidis V, Karniadakis GE (2006) A family of time-staggered schemes for integrating hybrid DPD models for polymers: algorithms and applications. *J Comput Phys* 218:82–101
17. Den Otter WK, Clarke JHR (2001) A new algorithm for dissipative particle dynamics. *Europhys Lett* 53:426–431
18. Marsh CA, Yeomans JM (1997) Dissipative particle dynamics: the equilibrium for finite time steps. *Europhys Lett* 37:511–516
19. Jakobsen AF, Mouritsen OG, Besold G (2005) Artifacts in dynamical simulations of coarse-grained model lipid bilayers. *J Chem Phys* 122:204901
20. Allen MP (2006) Configurational temperature in membrane simulations using dissipative particle dynamics. *J Phys Chem B* 110:3823–3830
21. Groot RD, Rabone KL (2001) Mesoscopic simulation of cell membrane damage, morphology changes and rupture by nonionic surfactant. *Biophys J* 81:725–736
22. Kranenburg M, Nicolas JP, Smit B (2004) Comparision of mesoscopic phospholipids-water models. *Phys Chem Chem Phys* 6:4142–4151
23. Kranenburg M, Laforgue C, Smit B (2004) Mesoscopic simulations of phase transitions in lipid bilayers. *Phys Chem Chem Phys* 6:4531–4534
24. Kranenburg M, Smit B (2005) Phase behavior of model lipid bilayers. *J Phys Chem B* 109:6553–6563
25. Kranenburg M, Vlaar M, Smit B (2004) Simulating induced integration in membranes. *Biophys J* 87:1596–1605
26. Patra M, Salonen E, Terama E et al (2006) Under the influence of alcohol: the effect of ethanol and methanol on lipid bilayers. *Biophys J* 90:1121–1135
27. De Meyer F, Smit B (2009) Effect of cholesterol on the structure of a phospholipids bilayer. *Proc Natl Acad Sci USA* 106:3654–3658
28. Shillcock JC, Lipowsky R (2002) Equilibrium structure and lateral stress distribution of amphiphilic bilayers from dissipative particle dynamics simulations. *J Chem Phys* 117:5048–5061
29. Gao LH, Shillcock JC, Lipowsky R (2007) Improved dissipative particle dynamics simulations of lipid bilayers. *J Chem Phys* 126:015101
30. Shillcock JC, Lipowsky R (2005) Tension-induced fusion of bilayer membranes and vesicles. *Mater* 4:225–228
31. Grafmüller A, Shillcock JC, Lipowsky R (2007) Pathway of membrane fusion with two tension-dependent energy barriers. *Phys Rev Lett* 98:218101
32. Gao LH, Lipowsky R, Shillcock JC (2008) Tension-induced vesicle fusion: pathway and pore dynamics. *Soft Matter* 4:1208–1214
33. Grafmüller A, Shillcock JC, Lipowsky R (2009) The fusion of membranes and vesicles: pathway and energy barriers from dissipative particle dynamics. *Biophys J* 96:2658–2675
34. Shillcock JC, Lipowsky R (2006) The computational route from bilayer membranes to

- vesicle fusion. *J Phys Condens Matter* 18: S1191–S1219
35. Lyubartsev AP, Karttunen M, Vattulainen I et al (2003) On coarse-graining by the inverse Monte Carlo method: dissipative particle dynamics simulations made to a precise tool in soft matter modeling. *Soft Mater* 1:121–137
  36. Chen LJ, Qian HJ, Lu ZY et al (2006) An automatic coarse-graining and fine-graining simulation method: application on polyethylene. *J Phys Chem B* 110:24093–24100
  37. Soddeleman T, Dünweg B, Kremer K (2003) Dissipative particle dynamics: a useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys Rev E* 68:046702
  38. Junghans C, Praprotnik M, Kremer K (2008) Transport properties controlled by a thermostat: an extended dissipative particle dynamics thermostat. *Soft Matter* 4:156–161
  39. Eriksson A, Jacobi MN, Nystrom J et al (2008) Effective thermostat induced by coarse-graining of simple point charge water. *J Chem Phys* 129:024106
  40. Qian HJ, Liew CC, Müller-Plathe F (2009) Effective control of the transport coefficients of a coarse-grained liquid and polymer models using the dissipative particle dynamics and Lowe-Andersen equations of motion. *Phys Chem Chem Phys* 11:1962–1969
  41. Wu SG, Guo HX (2009) Simulation study of protein-mediated vesicle fusion. *J Phys Chem B* 113:589–591
  42. Yamamoto S, Hyodo S (2003) Budding and fission dynamics of two-component vesicles. *J Chem Phys* 118:7937–7943
  43. Hong BB, Qiu F, Zhang HD et al (2007) Budding dynamics of individual domains in multicomponent membranes simulated by N-varied dissipative particle dynamics. *J Phys Chem B* 111:5837–5849
  44. Liu YT, Zhao Y, Liu H et al (2009) Spontaneous fusion between the vesicle formed by  $A_{2n}(B_2)_n$  type comb-like block copolymers with a semiflexible hydrophobic backbone. *J Phys Chem B* 113:15256–15262
  45. Yang K, Shao X, Ma YQ (2009) Shape deformation and fission route of the lipid domain in a multicomponent vesicle. *Phys Rev E* 79:051924
  46. Pivkin IV, Karniadakis GE (2008) Accurate coarse-grained modeling of red blood cells. *Phys Rev Lett* 101:118105
  47. Smith KA, Uspal WE (2007) Shear-driven release of a bud from a multicomponent vesicle. *J Chem Phys* 126:075102
  48. Avalos JB, Mackie AD (1997) Dissipative particle dynamics with energy conservation. *Europhys Lett* 40:141–146
  49. Español P (1997) Dissipative particle dynamics with energy conservation. *Europhys Lett* 40:631–636
  50. Mackie AD, Avalos JB, Navas V (1999) Dissipative particle dynamics with energy conservation: modeling of heat flow. *Phys Chem Chem Phys* 1:2039–2049
  51. Español P, Revenga M (2003) Smoothed dissipative particle dynamics. *Phys Rev E* 67:026705
  52. Pagonabarraga I, Frenkel D (2001) Dissipative particle dynamics for interacting systems. *J Chem Phys* 115:5015–5026
  53. Trofimov SY, Nies ELF, Michels MAJ (2002) Thermodynamics consistency in dissipative particle dynamics simulations of strongly non-ideal liquids and liquid mixtures. *J Chem Phys* 117:9383–9394
  54. Trofimov SY, Nies ELF, Michels MAJ (2005) Constant-pressure simulations with dissipative particle dynamics. *J Chem Phys* 123:144102
  55. Warren PB (2003) Vapor-liquid coexistence in many-body dissipative particle dynamics. *Phys Rev E* 68:066702

# Chapter 25

## Multiscale Molecular Dynamics Simulations of Membrane Proteins

Syma Khalid and Peter J. Bond

### Abstract

The time and length scales accessible by biomolecular simulations continue to increase. This is in part due to improvements in algorithms and computing performance, but is also the result of the emergence of coarse-grained (CG) potentials, which complement and extend the information obtainable from fully detailed models. CG methods have already proven successful for a range of applications that benefit from the ability to rapidly simulate spontaneous self-assembly within a lipid membrane environment, including the insertion and/or oligomerization of a range of “toy models,” transmembrane peptides, and single- and multi-domain proteins. While these simplified approaches sacrifice atomistic level detail, it is now straightforward to “reverse map” from CG to atomistic descriptions, providing a strategy to assemble membrane proteins within a lipid environment, prior to all-atom simulation. Moreover, recent developments have been made in “dual resolution” techniques, allowing different molecules in the system to be modeled with atomistic or CG resolution simultaneously.

**Key words:** Coarse grain, Multiscale, Molecular dynamics, Membrane protein

---

### 1. Introduction

Over the past 10 years or so, advances in the accuracy of molecular mechanics force fields coupled with the availability of ever-increasing computational power have enabled molecular dynamics (MD) simulation to develop into a well-established tool for the study of membrane protein conformational dynamics, structure refinement, and interaction with environments that mimic *in vitro* and *in vivo* conditions. In particular, atomistic (AT) simulations in which every atom is explicitly represented (all-atom), or atoms with the exception of some hydrogen atoms (united-atom models), have provided insights into many aspects of membrane protein biophysics (1–6); such simulations have been employed to explore, e.g., (1) the origins of selectivity and gating in ion channels (7–11);

(2) structural and mechanistic details of antimicrobial and related peptides (12–16); (3) structure–function relationship of bacterial outer membrane proteins (17–20); (4) water permeation through, e.g., aquaporins (6, 21–23); and (5) mechanistic aspects of transporters (24–26).

Coarse-grained (CG) force fields, in which the atomistic resolution of the system studied is sacrificed for a gain in simulation speed, allow the study of much larger systems than those possible by traditional atomistic methods. Whilst coarse-grained molecular dynamics (CG-MD) methods are a relatively new development in the field of biomolecular simulations, they have already successfully been employed to study membrane proteins and peptides. In particular, CG-MD simulations have extended the scope of molecular dynamics simulations such that protein–protein aggregation in a range of environments, lipid domain formation, and other self-assembly phenomena can now be routinely explored. However, the extended time and length scales afforded by CG-MD come at a price: atomistic level detail. While for large-scale changes, such as the formation of a lipid bilayer around a membrane protein, detail at the level of the individual atom-level can be safely sacrificed, for more subtle changes, e.g., in protein conformation in response to the local lipid environment, these details are desirable. Thus a “best of both worlds” approach is needed. Recently reports of multiscale, CG/AT simulation studies have appeared in the literature, in which a combination of atomistic and coarse-grained methods is employed (27). CG simulations are used to explore self-assembly or other longer timescale phenomena, while AT simulations are used to study the specific protein–protein or protein–lipid interactions of the same systems, in greater detail. Thus the atomistic level simulations are initiated from a self-assembled system; this eliminates the bias imposed upon the system by manual bilayer setup, and from a practical perspective, is less laborious, enabling a high throughput approach to simulations.

In the following chapter we review multiscale simulation studies of membrane protein and peptides. We particularly focus on combined atomistic and coarse-grained simulation studies, including a discussion of recent developments of “hybrid” or “dual resolution” methods in which two levels of resolution are combined in a single simulation. This is not intended to be an exhaustive review of all coarse-grained membrane simulation studies, and does not cover in detail the application of other computational methods, such as the treatment of the membrane as an implicit or elastic medium (28–30). Nevertheless, it should be noted that much progress has been made in this area: mean field models have provided insights into the structural, energetic (31, 32), dynamic (33), and mechanical/elastic (34–38) characteristics of interactions between membranes and embedded proteins, whilst the widespread use of implicit membrane representations based on generalized Born solvation has yielded insight into the

location (39), folding (40, 41), and assembly (42, 43) of membrane-associated peptides and proteins. Our intention here is to provide the reader with an overview of recent particle-based multiscale simulation studies reported in the literature, covering a range of membrane protein types, from small peptides to larger, multi-domain proteins.

---

## 2. Antimicrobial and Viral Peptides

Antimicrobial peptides (AMPs) are short amphipathic, cationic membrane-associated peptides secreted by a variety of organisms, which exhibit broad-spectrum antimicrobial activity and play a key role in host defense and innate immune response (44). They are thought to induce cell lysis by disrupting the lipid component of membranes, with their cationic nature aiding in the specific binding of anionic bacterial cell membranes. AMPs are generally disordered in solution, but adopt ordered conformations at the membrane interface, prior to cooperative membrane lysis and/or induced leakage. While the precise details of AMP-induced lysis are still unknown, a number of possible mechanisms have been proposed (45), which may be broadly divided into three models. In the “pore-forming” models, above a critical threshold concentration, a permeable channel is assembled to puncture the membrane; these pores are lined by either transmembrane (TM) peptides (the “barrel-stave” model, e.g., alamethicin) or lipid head groups (the “toroidal pore” model, e.g., magainins). Similar models are also thought to be responsible for the activity of many viral fusion peptides, which is crucial for the destabilization of target membranes prior to infection. In contrast, the “carpet” model describes the formation of a peptide monolayer on the membrane surface, leading to formation of defects, and detergent-like solubilization of the lipid bilayer. The three models may not be mutually exclusive, and likely depend upon the local membrane environment.

In the following section we provide a description of some of the successes to date of coarse-grained and multiscale MD simulations in providing insights into the dynamics and mechanism of action of antimicrobial and viral fusion peptides.

### 2.1. Carpet-Like AMP Models

Maculatin 1.1 (M1.1) is a broad-spectrum AMP (related to magainin; see above) from Australian tree frogs (46), whose  $\alpha$ -helical structure is characterized by a central kink at a Pro residue. CG-MD simulations of the M1.1 peptide in a phospholipid bilayer environment have been performed to reveal details of its membrane lytic mechanism (47). In planar membranes, cooperative insertion and aggregation of M1.1 peptides were observed, but no pore formation was evident, discounting this mechanism as a cause of membrane lysis. In contrast,

in the presence of whole lipid vesicles, insertion and assembly of M1.1 (but critically, not “generic” model WALP) peptides led to substantial changes in membrane curvature as a result of peptide-induced tension; over longer timescales, this may lead to detergent-like membrane solubilization.

Certain cyclic peptides have also been shown to exhibit antimicrobial properties, and have been studied via CG approaches (48, 49). These peptides usually contain an even number of L- and D-amino acids and adopt a characteristic flat conformation. Under appropriate conditions, they self-assemble into hollow nanotubes stabilized via a network of hydrogen bonds. The [RRKWLWLW] cyclic peptides are known to possess antimicrobial activity. As they form hollow nanotubes that lie parallel to the membrane interface at low concentrations, it has been suggested that they operate via a carpet-like mechanism. Khalfa and Tarek (50) performed coarse-grained MD simulations to investigate the interaction of the RRKWLWLW peptide with mixtures of POPG and POPE lipid bilayers. The simulations revealed mechanistic differences of these cyclic peptides compared to linear AMPs. Specifically, the first step is self-assembly of the peptides into amphipathic nanotubes at the bilayer/water interface. At high peptide-to-lipid concentrations the simulations revealed solubilization of the membrane without the formation of pores. Instead, as the membrane became coated with peptides, lipids were able to extrude into the now-exposed bilayer leaflet, ultimately leading to the release of micelle-like aggregates of phospholipids. The effect on the mechanical properties of the bilayer was highlighted through the drastically modified lateral pressure profile of the exposed leaflet when coated with cyclic peptides.

## 2.2. Pore-Forming AMPs

Magainins are AMPs found in the skin of the African frog *Xenopus laevis*. To explore the generally accepted hypothesis of cell lysis via toroidal pore formation by magainin (51), Marrink and coworkers have studied the long-time relaxation properties of the pores formed by magainin, by evolving simulation systems using a multi-scale approach. They had previously employed atomistic simulations to study magainin and related peptides, over timescales of up to a few hundred nanoseconds, but at high resolution (12, 52), revealing the spontaneous formation of disordered toroidal pores. Subsequently, CG-MD simulations of four magainin-H1 peptides in POPC lipids were performed for a total of over 20  $\mu$ s (53). Key peptide–lipid configurations extracted from the CG simulation trajectories were reverse-mapped to atomistic resolution using a recently described resolution–transformation protocol (54). Encouragingly, the pores formed in the CG-MD and atomistic simulations were comparable. Similarly, the pore retained its characteristic disorder in the CG simulations, and evidenced water permeation on the microsecond timescale. The long timescales

accessible to the CG-MD approach enabled the authors to systematically study the effect of peptide helicity on pore formation, revealing that partial helicity is likely to be necessary for pore formation, consistent with the disordered pores observed atomistically, and in contrast to the classical toroidal pore model. In summary, these studies highlight the potential for multiscale simulation approaches to yield insight into the lytic mechanism of pore-forming AMPs.

Alamethicin is a 20-amino acid AMP from the fungus *Trichoderma viride* (55). It is thought to form nonselective channels in cellular membranes, which result in leakage of cell contents into the extracellular milieu. Atomistic simulations of alamethicin have previously been reported (14), but most of these were restricted to relatively short timescales owing in part to the larger number of particles in atomistic simulation but also, in the case of the older simulations, to the smaller amounts of computational power available at the time (56, 57). To probe the details of the alamethicin mechanism of action on longer timescales, Tajkhorshid and coworkers performed a combined atomistic and CG-MD simulation study of the alamethicin channel in lipid bilayers (58). Microsecond timescale CG-MD simulations were employed to explore the preferred membrane orientation of alamethicin. These simulations predicted the peptide to adopt a transmembrane configuration with occasional transitions to the surface-bound state. Peptides within the membrane were reported to diffuse laterally, leading to self-assembly into clusters. The clusters slowly grew in size as additional peptides were incorporated. Thus the CG simulations revealed steps along the cluster-formation pathway and also enabled characterization of the architecture of the clusters. In addition to the study of alamethicin, the CG simulations of Tajkhorshid and coworkers were important in providing a comparison of two commonly used CG protein force fields (59, 60). Peptide-peptide packing and water penetration into peptide aggregates were explored via simulation of the systems reverse-mapped to atomistic detail. The combined use of all-atom and CG-MD simulations enabled the observation of penetration of water molecules into the core of the lipid bilayer, around the alamethicin peptides; indeed the water channels fully spanned the bilayer rendering it leaky in the presence of alamethicin. They also revealed some loss of helicity in the peptides, building up a picture of diverse, dynamic pores, consistent with NMR data.

### 2.3. Flu Peptides

Along with AMPs, recently reported microsecond-timescale CG-MD simulations have added to the understanding of the self-assembly and dynamics of small TM oligomers from influenza viruses such as M2, building on previous all-atom simulations, which focused on mechanisms of conduction and gating (61–65). The  $\alpha$ -helical M2 transmembrane domain of influenza A forms a homotetramer within host membranes, and acidification within

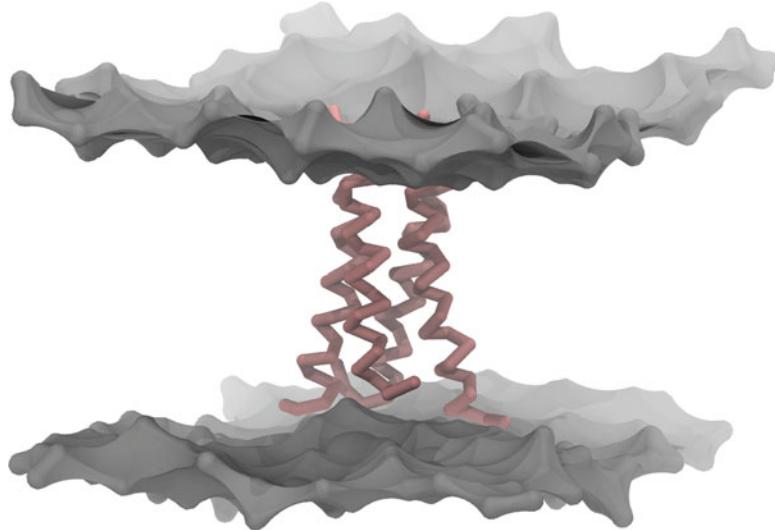


Fig. 1. 5  $\mu$ s CG-MD simulation of four M2 transmembrane peptides in a phospholipid bilayer (66). The snapshot shows the final state, in which the four peptides have spontaneously assembled into a left-handed tetrameric bundle in agreement with experimentally determined structures. Peptides are shown in red backbone format; the molecular surface of lipid head groups is colored grey. Lipid tails and solvent are omitted for clarity. Figure courtesy of Timothy Carpenter.

endosomal compartments leads to pH-gated opening of a proton channel within the M2 bundle, which triggers release of viral RNA into the cell. A similar role is played by the BM2 protein from influenza B. A combination of atomistic and CG simulations has been performed for both proteins to gain insight into their structure–function relationships (66, 67).

CG-MD simulations confirmed that both proteins self-assemble primarily in a transmembrane orientation, and subsequently, multiple microsecond-timescale simulations were used to observe the spontaneous oligomerization of left-handed tetramers (Fig. 1), in agreement with available experimental data. In the case of M2, evidence was found for two- rather than fourfold rotational symmetry within the bundle, consistent with recent X-ray studies suggesting that deviation from exact fourfold symmetry is possible. Subsequent reverse-mapping enabled the refinement of the CG bundles via atomic simulations; however, the ~microsecond-timescale conformational switching between related twofold symmetrical tetramers observed during CG simulations was not observed, highlighting the strength of a combined atomistic/CG approach (66).

Because these channels are known to be pH-gated, an atomistic model of the BM2 tetramer was reverse-mapped from the converged CG structure, and the ionization state of its pore-lining His-19 residues was varied prior to AT simulation. Upon transition

from two to three protonated His-19 side chains, a spontaneous conformational change occurred resulting in channel opening (67). This, along with the studies of M2, demonstrates the effectiveness of the combined atomistic/CG MD approach; CG-MD can yield unbiased, self-assembled conformations of oligomeric transmembrane bundles, whilst atomistic simulations enable one to ask more subtle mechanistic questions.

---

### 3. Other Peptides

CG-MD approaches have been used to study an extensive range of membrane peptides. Simplified “toy” models have provided significant insight into generic features of lipid–peptide interactions, in the context of phenomena such as AMP-like assembly and transmembrane pore formation (68–71), lipid deformation/clustering and hydrophobic mismatch (72–74), and membrane remodeling (75–77). Similar CG models have also been generated for more “realistic” membrane peptides, with applications ranging from translocation/permeation mechanisms of cell-penetrating peptides (13, 78–80) and synthetic aryl amide polymers (81); membrane partitioning and assembly of naturally occurring peptide or peptide fragments such as CFTR helices (82, 83), SNARE fusion peptides (84), and ATP synthase c-subunit (85); and structural and thermodynamic patterns of interaction of synthetic model peptides such as WALP-like (60, 82, 86–89) and LS3 (90, 91) peptides. Recently, the CG-MD approach to peptide/membrane assembly has been extended in detail, in the context of using both integrated multiscale models and more complex descriptions of membrane composition. The former is exemplified here by studies of CG-based dimerization of membrane peptides prior to AT-based structure refinement, and the latter by recent simulations of membranes containing multiple lipids representing those within cellular organelles.

#### 3.1. CG/AT Approaches to Peptide Dimerization

Glycophorin A (GpA) is a relatively simple erythrocyte membrane containing a short (~25 residues) hydrophobic TM domain. The TM domain of GpA has been shown to form a homodimer in detergent micelles and lipid bilayers and has been extensively studied via simulation. The tightly packed dimerization interface is formed by a GXXXG motif, which leads to a characteristic right-handed crossing. The thermodynamics of GpA oligomerization has been extensively studied experimentally, making this a useful “reference” system. Thus, a CG approach was applied to observe the spontaneous membrane adsorption, insertion, and dimerization of GpA transmembrane peptides, in support of the two-state model for membrane protein folding (92). The converged dimer

conformation was remarkably close to experimentally determined structures, packed in a right-handed fashion around the GXXXG motif. Extending this study to explore the thermodynamics of dimerization, GpA wild-type (WT) and mutant transmembrane helices were inserted within a preformed DPPC bilayer, and multiple  $>1\text{ }\mu\text{s}$  simulations were performed, during which the assembly equilibria were followed (93). Focussing on the WT GpA simulations, the general behavior of the helices was reported as random diffusion, followed by an encounter event after which the helices formed a long-lived dimer. CG simulations of mutants identified as “disruptive” or “nondisruptive” were also performed. Encouragingly it was found that the CG models were sensitive enough to predict less stable (short-lived association) dimers for the disruptive mutant sequences. Helix–helix association lifetimes were used to estimate a change in free energy of association of  $\sim 10\text{ kJ/mol}$ , and structurally, this resulted from a distorted packing of helices associated with variable crossing angles in comparison with WT GpA (93). Despite an apparent difference in helix–helix “stickiness,” a separate study of GpA mutants estimated a comparable disruptive change in free energy based on the calculation of the dimerization potential of mean force (PMF) (94). In an extension of (93), the final CG model of the WT GpA dimer was converted to atomistic resolution (95); a subsequent 30 ns simulation confirmed that the stabilizing packing interactions at the interface of the right-handed dimer arose largely from the glycine residues of the GXXXG motif along with inter-helix hydrogen bonds of Thr87, and led to a converged dimer conformation with a crossing angle of  $\sim 40^\circ$ , in agreement with previous experimental and computational approaches. Thus, a multiscale approach may be used to refine models of helix dimers (95).

Similar studies of dimerization were reported for transmembrane helices containing GXXXG or GXXXG-like motifs, whose association is thought to play a functional role in receptor signaling, including the syndecans and integrins (95). Integrins are heterodimeric transmembrane proteins composed of non-covalently associated  $\alpha$  and  $\beta$  subunits. Here, the MD studies were focused on the  $\alpha$ IIb/ $\beta$ 3 dimer; the  $\alpha$ IIb contains a GXXXG motif whereas the  $\beta$ 3 subunit contains a SXXXA motif. In contrast to the results of previous CG assembly simulations for GpA and syndecan, the helix crossing angle for the dimers exhibited a bimodal distribution with a bias (56% vs. 44%) towards right-handed dimers. To refine the CG dimers, an example of both a right- and left-handed dimer was reverse-mapped to atomistic resolution. The atomistic simulations revealed the left-handed dimer to be less stable than its right-handed counterpart, in agreement with recent NMR structural data. Recently, further details concerning integrin function have

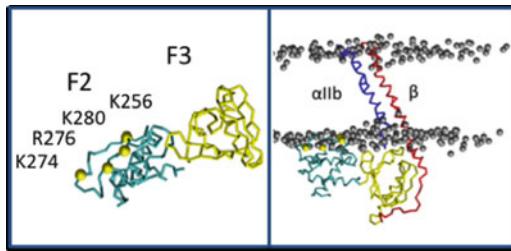


Fig. 2. Snapshot from a CG simulation of the talin/integrin complex (96). Interactions between a positively charged patch on the talin F2 subdomain and negatively charged phosphate moieties (red) on the membrane orient the talin F2–F3 domain for binding to the integrin  $\beta$  cytoplasm tail. The F2 and F3 domains are shown in cyan and yellow, and the integrin  $\alpha$  and  $\beta$  subunits are shown in blue and red, respectively. Lipid tails and water are omitted for clarity. Figure courtesy of Antreas Kalli.

been unraveled via a combined NMR and CG-MD simulation study of a talin/integrin complex (96), which may in future benefit from a multiscale refinement (Fig. 2).

### 3.2. A More Complex Membrane Model

Recently, multiscale simulations of interactions of the transmembrane domain of fukutin, a protein implicated in muscular dystrophy, with the Golgi Apparatus membrane were reported by Holdbrook et al. (97). The study focused on FKTMD1, a short 36-residue amino-terminal fragment of fukutin that is thought to insert into the membrane of the Golgi apparatus. In the absence of an X-ray or NMR structure, CD spectroscopic data was used as a guide to model FKTMD1 as a perfect  $\alpha$ -helix. First, atomistic simulations were performed to show that the secondary structure of a single peptide inserted within simple phospholipid bilayers is stable. Having established the structural integrity of the helical conformation at the atomistic level, the authors then performed coarse-grained simulations of the peptide in a range of more realistic mixtures of lipids reflecting the *in vivo* composition of the Golgi membrane (Fig. 3). Initially, the effect of the nature of the lipid tails and head groups on the orientation of the protein was investigated using bilayers composed of different lipids. Small tilt angles (with respect to the bilayer plane) were observed for longer lipid tail membranes, whereas for shorter lipids, substantial transmembrane tilting occurred to minimize hydrophobic mismatch. The protein was subsequently simulated in a mixture of lipids reflecting the Golgi membrane, including sphingomyelin, cholesterol, phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, and phosphatidylinositol bisphosphate (PIP<sub>2</sub>). A marked difference in the protein–lipid interactions at the amino- and carboxy-termini of FKTMD1 was observed. The amino-terminal domain exhibited a marked propensity to interact with PIP<sub>2</sub> lipids, such that local clustering of these lipids was observed. In contrast, the carboxy-terminal domain did not exhibit a preference for any particular lipid type.

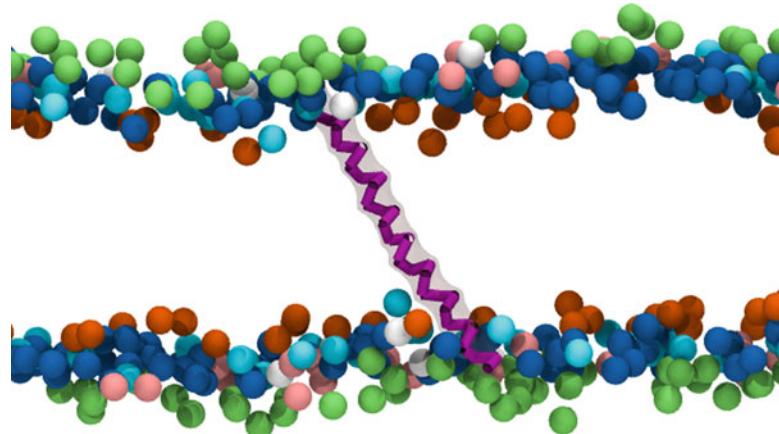


Fig. 3. Side view of the fukutin transmembrane domain (*mauve* and *white*) in a realistically composed Golgi membrane after 1  $\mu$ s of simulation (97). The head groups are shown for the membrane, which is composed of PIP<sub>2</sub> (*lime*), phosphatidylethanolamine (*cyan*), phosphatidylcholine (*blue*), phosphatidylserine (*white*), cholesterol (*orange*), and sphingomyelin (*pink*). Figure courtesy of Daniel Holdbrook.

These observations suggest that the interaction of the amino-terminal domain with the Golgi membrane may induce the formation of microdomains. In particular, this simulation study highlighted the importance of paying attention to the lipid composition of the membrane when studying mobility and orientation of membrane peptides and proteins.

#### 4. Membrane Proteins

The rate at which high-resolution structures of membrane proteins are solved continues to increase (98). Nevertheless, there is an ongoing dearth of experimental data concerning interactions between proteins and their membrane environment, which is unfortunate as both specific interactions and general bilayer properties, such as thickness or lipid composition, are likely to be critical for function. Computational approaches based on an implicit “slab” have been used to predict the likely location of proteins within membranes (39, 43, 99–102) but do not take into account the nonuniform complexity and dynamic nature of lipid bilayers. Recently, CG-MD approaches have provided a means to simulate the spontaneous assembly of proteins within realistic biological membrane models, and have proven successful for both  $\alpha$ -helical (82) and  $\beta$ -barrel (92) classes of membrane proteins, as well as for monotopic membrane-active enzymes (103, 104), and even aggregates of multiple proteins, such as G-protein-coupled receptors (105). In some cases, the approach has benefited from the

combination of a CG force field with an elastic network model in order to restrain the secondary and/or tertiary native protein structure (106, 107).

CG-MD and/or multiscale approaches have proven particularly useful when incorporation of the structural complexity of membranes is necessary, for example when mimicking conditions of hypo-osmotic shock to observe tension-driven gating in the mechanosensitive MscL channel (108), or when delineating the relationship between peptide composition and mechanisms of translocation/membrane insertion through the heterotrimeric Sec complex (109, 110). Moreover, CG-MD has provided a means for high-throughput, unbiased membrane assembly (111, 112), coordinates from which may be reverse-mapped to provide starting points for atomistic simulations (2, 104). The benefit of an explicit treatment of the complex, dynamic lipid environment via multiscale self-assembly simulations is highlighted by two particularly studied systems, namely, potassium-channel voltage sensor domains and lipoprotein particles; these are summarized in detail below.

#### **4.1. Potassium- Channels and Voltage Sensors**

The relationship between structure/dynamics and function of ion channels has been the focus of numerous MD simulation studies over two decades. Voltage-gated potassium ( $K_v$ ) channels are a particularly interesting class of ion channels that have been studied via atomistic and CG approaches. They play a central role in the physiology of electrically excitable cells, and are opened and closed in response to a change in transmembrane voltage, as a result of motion and/or conformational changes within the voltage sensor (VS) domain. CG-MD self-assembly simulations have been used to investigate the interaction of the VS domain from archaeabacterial KvAP with phospholipid bilayers (106). In the membrane-inserted state of the VS, considerable membrane deformation was evident in contrast with “canonical” membrane proteins, as a result of local interactions with the S4 helix. The S4 helix contains several positively charged amino acids which are thought to “sense” the transmembrane voltage; their side chains are accommodated within the hydrophobic membrane environment via lipid head groups and solvent molecules, in agreement with previous atomistic simulations. Treptow et al. used CG simulations to investigate how motions of the VS domain may induce channel gating (113). Spontaneous conformational changes were observed during the transition from open to closed forms of a  $K_v$  channel, which satisfied experimental constraints. Interestingly, only small motions of the S4 helices were observed.

Simulations have also been used to investigate the interactions between VS domains and gating-modifier toxins, which bind to and inhibit  $K_v$  channels. Such toxins can be used to probe the structure and orientation of the VS domain as it moves in response to changes in transmembrane voltage. For example, the

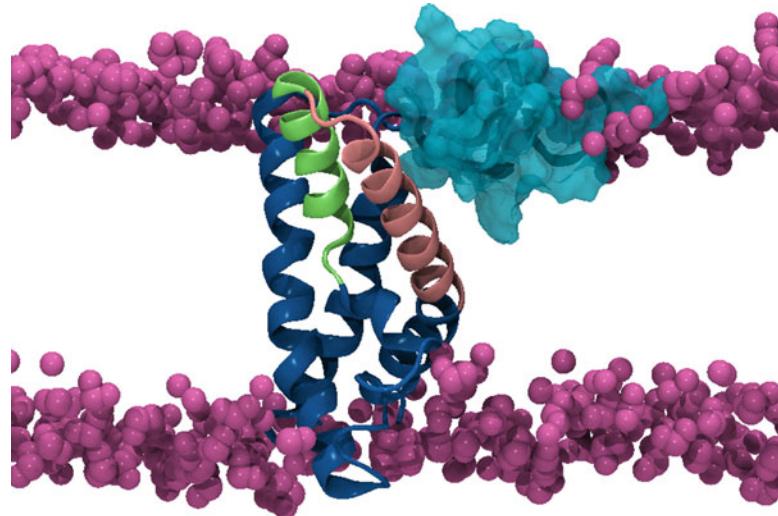


Fig. 4. Example of a simulation-refined (115) voltage sensor/toxin complex in a bilayer, showing the voltage sensor in blue (with the S3b and S4a helices of the paddle in green and pink, respectively), the toxin in cyan, and the phosphorus atoms of the lipid head groups in mauve. Figure courtesy of Mark Sansom.

amphipathic, 36-residue VSTx1 toxin is found in tarantula venom and binds to KvAP, stabilizing its open state. Wee et al. used serial multiscale MD simulations, employing a combination of atomistic and CG-MD, to study the interaction of toxins with membranes and the isolated voltage sensor domain of KvAP. Starting from an initial location in bulk solvent, partitioning of the SGTx1 toxin into the lipid head group/tail interface region of the bilayer was observed during CG simulations, equilibrating at locations consistent with those indicated from atomistic simulations (114). Similar observations were made during CG assembly simulations of VSTx1 (115), whilst microsecond-timescale umbrella sampling CG simulations were used to characterize the PMF associated with membrane interface interaction, highlighting the role of local bilayer deformation in partitioning (116). Following adsorption to the interface in CG assembly simulations, the VSTx1 toxin was subsequently observed to spontaneously bind to the voltage sensor domain (Fig. 4). System configurations from the CG-MD simulations were then reverse-mapped into atomistic detail to refine the toxin–VS complex (115). Specific protein–toxin interaction sites were identified that support a membrane-access mechanism of Kv channel inhibition by the VSTx1 toxin. Thus, mechanistic insights as well as structural refinement of the complex were achieved by employing a serial multiscale simulation approach (115). A similar study has also been reported for modeling the binding of regulatory PIP<sub>2</sub> molecules to inwardly rectifying potassium (Kir) channels (117), suggesting that such serial multiscale approaches may be

generally applicable to the identification and refinement of molecular binding sites on integral membrane protein channels and their associated domains.

#### **4.2. Lipoprotein Particles**

High-density lipoproteins (HDL) are protein–lipid particles which carry cholesterol, lipids, and fatty acids from arteries to the liver for degradation (59). They facilitate removal of excess cholesterol from the body, and hence are of some biomedical interest. Related “nanodiscs” are protein–lipid particles which are engineered to mimic nascent HDL particles, and have proven useful to embed and solubilize membrane proteins. HDL particles and nanodiscs contain two amphipathic lipid-binding chains, either apolipoprotein A-1 (A-1) or membrane scaffold proteins, respectively, which wrap around a lipid bilayer core. Because high-resolution structures for HDL particles and nanodiscs are unavailable, Shih, Schulten, and coworkers have reported a series of multiscale simulation studies which, combined with low-resolution structural data, provide a detailed picture for their likely structure and dynamics.

Various models have been proposed to describe the structure of HDL particles. The “picket fence” model implies that protein scaffold strands are oriented parallel to the disc normal, whereas in the “double belt” model, the strands are wrapped around the bilayer core in an antiparallel belt-like fashion. A preliminary CG simulation study revealed that preassembled double-belt nanodiscs were stable over long time periods, consistent with previous short atomistic simulations, and had temperature-dependent dimensions which correlated with experimental small-angle X-ray scattering (SAXS) data (118). On the other hand, microsecond-timescale CG self-assembly simulations, in which scaffold proteins were added to systems containing a lipid bilayer, micelle, or randomly placed lipids, led to discoidal particles approximating disordered versions of the picket fence model, rather than the expected belt-like conformation. A later study utilized an improved CG potential based on atomistic simulations, and resultant self-assembly simulations led to the spontaneous formation of more ordered, discoidal HDL conformations that resembled the double-belt model (119). Additional extended 10  $\mu$ s CG-MD simulations confirmed this model, and helped to provide a detailed description of HDL nanodisc assembly, consisting of an initial assembly phase driven by the hydrophobic effect, followed by a slower phase during which scaffold proteins adjusted their arrangement and optimized protein–protein interactions, resulting in a final double-belt structure (120). Because the assembly process of homogeneous nanodiscs is initiated by the removal of cholate from a mixture of cholate, scaffold protein, and lipids, CG simulations of preformed nanodiscs have also been performed in the presence of different concentrations of cholate (121). At high concentration, this led to the formation of spherical particles swelled by cholate molecules,

consistent with SAXS data; reverse-mapping followed by atomistic simulations revealed the presence of multiple pockets of water within the liposome, providing a molecular explanation for the role of cholate in nanodisc disassembly (121).

More recent studies of HDL particles have focused on the details of their shape (122), composition, and maturation (123). In particular, the enzymatic esterification of cholesterol plays a key physiological role in the maturation of discoidal HDL into spherical particles circulating in the blood, but the structural basis for this change is not well understood. Thus, a series of CG-MD simulations were performed during which cholesterol ester molecules were added in a stepwise manner to discoidal HDL particles (123). The esters spontaneously partitioned into the hydrophobic core over hundreds of nanoseconds, swelling an ester-containing core surrounded by phospholipids and A-1, and leading to a spherical, mature HDL particle. Similar spheroidal HDL particles were observed independently by Catte et al., when cholesterol ester molecules were manually added to the hydrophobic core (124). Shih et al. reverse-mapped their observed mature HDL particles, and subsequent atomistic simulations revealed reduced A-1 protein helicity, increased protein termini flexibility, and a loss of salt bridges between A-1 “belts.” Thus, mature HDL particles resemble spherical, disordered double-belt particles; the authors hypothesized that the less mobile A-1 regions may bind cholesterol-converting enzymes (123).

Collectively, these studies highlight the strength of combining simplified CG representations of protein–lipid complexes with resultant atomistically refined models and low-resolution experimental data in order to arrive at a consistent model for the structure and aggregation mechanism of physiologically or experimentally relevant assemblies. Moreover, even in the context of CG approaches, significant complexity may be captured within “simple” models, as highlighted by the recent CG-MD simulations of HDL particles using the full, realistic lipid composition, as found in human serum particles (125).

---

## 5. Dual Resolution Methods

In a further advancement of simulating membrane proteins at multiple levels of granularity, Voth and coworkers adopted an integrated multiscale approach to obtain CG force fields from atomistic simulations, via application of a force-matching procedure in which atomistic resolution forces are propagated in scale to the CG level to derive effective pair-wise parameters (126). The approach was used to simulate the antibacterial peptide Gramicidin A in phospholipid bilayers (127); a “hybrid” system was

constructed, in which the peptide was represented at atomistic resolution, whilst lipid and solvent were represented in CG detail. The system was stable in both the atomistic and CG regions over a timescale of 10 ns. Indeed, the peptide structural drift and peptide/lipid interaction distributions compared well with the corresponding pure atomistic simulations. The same kind of force-matching procedure has also been used to perform multiscale simulations of the interaction of BAR domains with membranes (128, 129). BAR domains play key roles in the sculpting of membranes *in vivo*, inducing bilayer curvature for, e.g., vesicle fission and endocytosis, but these processes take place over large time and length scales, necessitating multiscale approaches which build upon atomistic models (130, 131). This work may be contrasted with recent serial multiscale approaches which utilize atomistic, residue-based, shape-based, and continuum elastic resolutions to simulate protein-induced membrane remodeling (132, 133).

An alternative dual resolution approach has been adopted by Essex and coworkers (134). While this approach has not yet been applied to proteins, we have discussed it here as it represents a promising hybrid technique for modeling biological molecules. Initial work focused on the permeability of small organic molecules across lipid bilayers (Fig. 5). As with the Voth approach, the authors represent the phospholipid bilayer and solvent at the coarse-grained level of detail, while solutes of interest are represented in full atomistic detail. The lipid head groups are represented by two spherical Lennard-Jones particles, with the electrostatics treated by point charges for the head group particles (single positive charge for a choline group and a single negative charge for a phosphate group, for phosphatidylcholine). In contrast, the hydrocarbon and glycerol regions are modeled using the Gay–Berne potential, which may be considered as an extension of the isotropic Lennard-Jones potential, in which additional terms are included to allow the modeling of nonspherical particles. In the case of phosphatidylcholine lipids, the glycerol-ester region is described by two Gay–Berne ellipsoidal units, each embedded with a point-dipole to capture the charge distribution in this region, while hydrocarbon tails are modeled by chains of three neutral Gay–Berne ellipsoids. Incorporation of the point-dipole enables a more sophisticated representation of the electrostatics than many other CG models. Water is represented by the soft-sticky dipole (SSD) model, a single site model which includes an embedded point-dipole plus an additional term to capture hydrogen-bonding behavior. Standard mixing rules were used to combine the CG and atomistic parts of the models. In an extension to this study, the authors studied the permeability of two synthetic antimicrobial molecules, triclosan and triclocarban (135). Simulations of membranes incorporating increasing concentrations of antimicrobial molecules revealed substantial perturbation of the bilayer. In particular, bilayer stability was undermined at

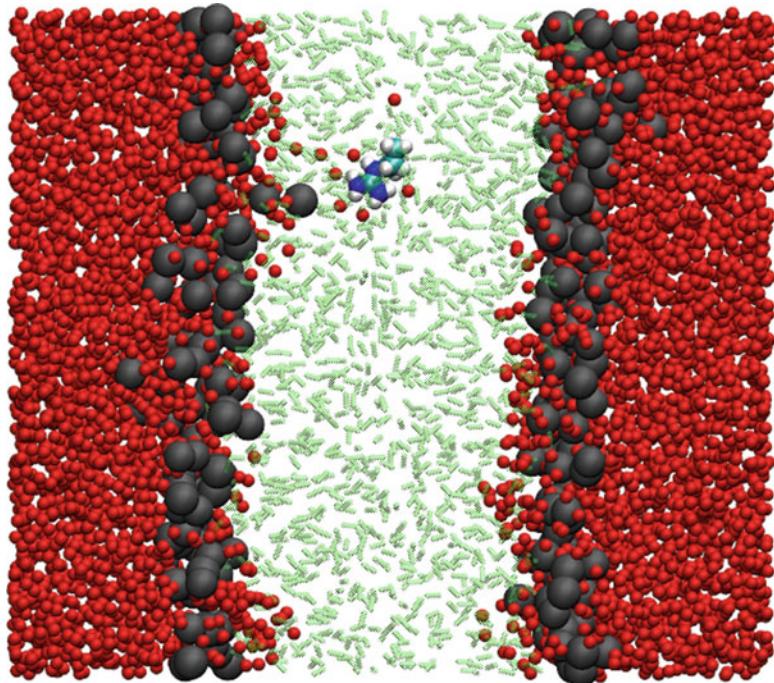


Fig. 5. Snapshot from a dual-resolution simulation (134). An atomistically modeled arginine molecule (shown in CPK colors) is embedded in a coarse-grained phospholipid bilayer. The water molecules are colored red; the lipid tails are green, while the head groups are grey. Figure courtesy of Mario Orsi.

concentrations corresponding to approximately one antimicrobial per two lipids. The authors noted that this antimicrobial molecule-to-lipid ratio is comparable to concentrations in real bacterial membrane systems in contact with antibacterial consumer formulations.

The integrated multiscale approaches (134, 136) clearly represent an exciting “best of both worlds” strategy in which the chemically sensitive part of the system is afforded atomistic level detail, whereas the surrounding environment is modeled at a lower resolution. Thus these approaches have the potential to reveal mechanistic insights into biomolecular function, without sacrificing important atomistic details.

## 6. Conclusions

Over the last two decades, molecular dynamics simulations have become a well-established tool for the study of conformational dynamics and membrane interactions of proteins and peptides. Studies that combine experimental techniques with simulations to

obtain complementary data have been particularly insightful (97, 137–139). The field of biomolecular simulation has seen some exciting advances in recent years, in particular the widespread use of CG models and more recently multiscale approaches. Such advancements have enabled a more thorough exploration of self-assembly of bilayers and micelles around membrane proteins and peptides, thus removing the possibility of any bias in the system introduced by the initial placement of the protein within the local environment (92, 112). Consequently, CG models are now routinely used to self-assemble suitable system configurations, which are then converted to atomistic detail for subsequent simulation. Furthermore, the smoother conformational landscapes of CG models enable simulations on the microsecond timescale; thus we can begin to move beyond conformational dynamics of single proteins to dynamics of biomolecular systems, e.g., to study the diffusion of membrane proteins (140). Simulations of much larger systems are key to making the link between the *in silico* and *in vivo* regimes. However, CG models on their own, while enabling longer and larger simulations, do also necessitate the loss of atomistic resolution. Two approaches have been developed to overcome this, the widely adopted method of performing CG simulations then converting key system configurations to atomistic detail for subsequent simulation and the more recently developed dual resolution techniques in which the molecule of key importance is represented at the atomistic level, while the solvent and other molecules are represented at the CG level. Both methods offer a route to achieving the “best of both worlds,” i.e., system evolution via long timescale simulations and atomistic detail when and where appropriate. In summary, recent advancements in models, methodologies, and computing technology have widened the scope of molecular dynamics simulations such that multiscale approaches, along with the microsecond- and even millisecond-timescale sampling now accessible (141), have rapidly reduced the gap between simulation, experimental, and biological systems.

---

## Acknowledgments

PJB and SK would like to acknowledge help from Jon Essex, Mario Orsi, Daniel Holdbrook, Mark Sansom, and Antreas Kalli.

## References

1. Ash WL, Zlomislic MR, Oloo EO, Tielemans DP (2004) Computer simulations of membrane proteins. *Biochim Biophys Acta* 1666:158–189
2. Biggin PC, Bond PJ (2008) Molecular dynamics simulations of membrane proteins. *Methods Mol Biol* 443:147–160
3. Lindahl E, Sansom MS (2008) Membrane proteins: molecular dynamics simulations. *Curr Opin Struct Biol* 18:425–431
4. Gumbart J, Wang Y, Asksimentiev A, Tajkhorshid E, Schulten K (2005) Molecular dynamics simulations of proteins in lipid bilayers. *Curr Opin Struct Biol* 15:423–431
5. Domene C, Bond PJ, Sansom MSP (2003) Membrane protein simulations: ion channels and bacterial outer membrane proteins. *Adv Protein Chem* 66:159–193
6. Dror RO, Jensen MO, Shaw DE (2009) Elucidating membrane protein function through long-timescale molecular dynamics simulation. *Conf Proc IEEE Eng Med Biol Soc* 2009:2340–2342
7. Beckstein O, Sansom MSP (2006) A hydrophobic gate in an ion channel: the closed state of the nicotinic acetylcholine receptor. *Phys Biol* 3:147–159
8. Sansom MSP, Shrivastava IH, Bright JN, Tate J, Capener CE, Biggin PC (2002) Potassium channels: structures, models, simulations. *Biochim Biophys Acta* 1565(2):294–307
9. Roux B, Allen T, Berneche S, Im W (2004) Theoretical and computational models of biological ion channels. *Q Rev Biophys* 37:15–103
10. Jensen MO, Borhani DW, Lindorff-Larsen K, Maragakis P, Jogini V, Eastwood MP, Dror RO, Shaw DE (2010) Principles of conduction and hydrophobic gating in K<sup>+</sup> channels. *Proc Natl Acad Sci USA* 107:5833–5838
11. Khalili-Araghi F, Tajkhorshid E, Schulten K (2006) Dynamics of K<sup>+</sup> ion conduction through Kv1.2. *Biophys J* 91:L72–L74
12. Leontiadou H, Mark AE, Marrink SJ (2006) Antimicrobial peptides in action. *J Am Chem Soc* 128:12156–12161
13. Herce HD, Garcia AE (2007) Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes. *Proc Natl Acad Sci USA* 104:20805–20810
14. Pan J, Tielemans DP, Nagle JF, Kucerka N, Tristram-Nagle S (2009) Alamethicin in lipid bilayers: combined use of X-ray scattering and MD simulations. *Biochim Biophys Acta* 1788:1387–1397
15. La Rocca P, Biggin PC, Tielemans DP, Sansom MSP (1999) Simulation studies of the interaction of antimicrobial peptides and lipid bilayers. *Biochim Biophys Acta* 1462:185–200
16. Matyus E, Kandt C, Tielemans DP (2007) Computer simulation of antimicrobial peptides. *Curr Med Chem* 14:2789–2798
17. Bond PJ, Sansom MSP (2004) The simulation approach to bacterial outer membrane proteins. *Mol Membr Biol* 21:151–161
18. Khalid S, Bond PJ, Carpenter T, Sansom MSP (2007) OmpA: gating and dynamics via molecular dynamics simulations. *Biochim Biophys Acta* 1778(9):1871–1880. doi:[10.1016/j.bbamem.2007.1005.1024](https://doi.org/10.1016/j.bbamem.2007.1005.1024)
19. Straatsma TP, Soares TA (2009) Characterization of the outer membrane protein OprF of *Pseudomonas aeruginosa* in a lipopolysaccharide membrane by computer simulation. *Proteins* 74:475–488
20. Tielemans DP, Berendsen HJC (1998) A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. *Biophys J* 74:2786–2801
21. Qiu H, Ma S, Shen R, Guo W (2010) Dynamic and energetic mechanisms for the distinct permeation rate in AQP1 and AQP0. *Biochim Biophys Acta* 1798:318–326
22. Hub JS, Grubmuller H, de Groot BL (2009) Dynamics and energetics of permeation through aquaporins. What do we learn from molecular dynamics simulations? *Handb Exp Pharmacol* 190:57–76
23. Wang Y, Schulten K, Tajkhorshid E (2005) What makes an aquaporin a glycerol channel? A comparative study of AqpZ and GlpF. *Structure* 13:1107–1118
24. Khalili-Araghi F, Gumbart J, Wen PC, Sotomayor M, Tajkhorshid E, Schulten K (2009) Molecular dynamics simulations of membrane channels and transporters. *Curr Opin Struct Biol* 19:128–137
25. DeChancie J, Shrivastava IH, Bahar I (2011) The mechanism of substrate release by the aspartate transporter GltPh: insights from simulations. *Mol Biosyst* 7:832–842
26. Shi L, Weinstein H (2010) Conformational rearrangements to the intracellular open states of the LeuT and ApcT transporters are modulated by common mechanisms. *Biophys J* 99: L103–L105

27. Hills RD Jr, Lu L, Voth GA (2010) Multiscale coarse-graining of the protein energy landscape. *PLoS Comput Biol* 6:e1000827
28. Chen J, Brooks CL 3rd, Khandogin J (2008) Recent advances in implicit solvent-based methods for biomolecular simulations. *Curr Opin Struct Biol* 18:140–148
29. Sperotto MM, May S, Baumgaertner A (2006) Modelling of proteins in membranes. *Chem Phys Lipids* 141:2–29
30. West B, Brown FL, Schmid F (2009) Membrane-protein interactions in a generic coarse-grained model for lipid bilayers. *Biophys J* 96:101–115
31. Zemel A, Ben-Shaul A, May S (2005) Perturbation of a lipid membrane by amphipathic peptides and its role in pore formation. *Eur Biophys J* 34:230–242
32. Zemel A, Ben-Shaul A, May S (2008) Modulation of the spontaneous curvature and bending rigidity of lipid membranes by interfacially adsorbed amphipathic peptides. *J Phys Chem B* 112:6988–6996
33. Khelashvili G, Weinstein H, Harries D (2008) Protein diffusion on charged membranes: a dynamic mean-field model describes time evolution and lipid reorganization. *Biophys J* 94:2580–2597
34. Chen X, Cui Q, Tang Y, Yoo J, Yethiraj A (2008) Gating mechanisms of mechanosensitive channels of large conductance, I: a continuum mechanics-based hierarchical framework. *Biophys J* 95:563–580
35. Brannigan G, Brown FL (2006) A consistent model for thermal fluctuations and protein-induced deformations in lipid bilayers. *Biophys J* 90:1501–1520
36. Choe S, Hecht KA, Grabe M (2008) A continuum method for determining membrane protein insertion energies and the problem of charged residues. *J Gen Physiol* 131:563–573
37. Tang Y, Yoo J, Yethiraj A, Cui Q, Chen X (2008) Mechanosensitive channels: insights from continuum-based simulations. *Cell Biochem Biophys* 52:1–18
38. Wiggins P, Phillips R (2005) Membrane-protein interactions in mechanosensitive channels. *Biophys J* 88:880–902
39. Ulmschneider MB, Ulmschneider JP, Sansom MS, Di Nola A (2007) A generalized born implicit-membrane representation compared to experimental insertion free energies. *Biophys J* 92:2338–2349
40. Im W, Brooks CL 3rd (2005) Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proc Natl Acad Sci USA* 102:6771–6776
41. Im W, Brooks CL 3rd (2004) De novo folding of membrane proteins: an exploration of the structure and NMR properties of the fd coat protein. *J Mol Biol* 337:513–519
42. Bu L, Im W, Brooks CL 3rd (2007) Membrane assembly of simple helix homooligomers studied via molecular dynamics simulations. *Biophys J* 92:854–863
43. Im W, Feig M, Brooks CL 3rd (2003) An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys J* 85:2900–2918
44. Brown KL, Hancock RE (2006) Cationic host defense (antimicrobial) peptides. *Curr Opin Immunol* 18:24–30
45. Shai Y (2002) Mode of action of membrane active antimicrobial peptides. *Biopolymers* 66:236–248
46. Boland MP, Separovic F (2006) Membrane interactions of antimicrobial peptides from Australian tree frogs. *Biochim Biophys Acta* 1758:1178–1183
47. Bond PJ, Parton DL, Clark JF, Sansom MS (2008) Coarse-grained simulations of the membrane-active antimicrobial peptide maculatin 1.1. *Biophys J* 95:3802–3815
48. Hwang H (2009) Coarse-grained molecular dynamics study of cyclic peptide nanotube insertion into a lipid bilayer. *J Phys Chem A* 113:4780–4787
49. Khalfa A, Treptow W, Maigret B, Tarek M (2009) Self assembly of peptides near or within membranes using coarse grained MD simulations. *Chem Phys* 358:161–170
50. Khalfa A, Tarek M (2010) On the antibacterial action of cyclic peptides: insights from coarse-grained MD simulations. *J Phys Chem B* 114:2676–2684
51. Ludtke SJ, Heller WT, Harroun TA, Yang L, Huang HW (1996) Membrane pores induced by magainin. *Biochemistry* 35:13723–13728
52. Sengupta D, Leontiadou H, Mark AE, Marrink SJ (2008) Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochim Biophys Acta* 1778:2308–2317
53. Rzepiela AJ, Sengupta D, Goga N, Marrink SJ (2010) Membrane poration by antimicrobial peptides combining atomistic and coarse-grained descriptions. *Faraday Discuss* 144:431–443, discussion 445–481
54. Rzepiela AJ, Schafer LV, Goga N, Risselada HJ, De Vries AH, Marrink SJ (2010) Reconstruction of atomistic details from coarse-

- grained structures. *J Comput Chem* 31:1333–1343
55. Qian S, Wang W, Yang L, Huang HW (2008) Structure of the alamethicin pore reconstructed by x-ray diffraction analysis. *Biophys J* 94:3512–3522
56. Sansom MSP, Bond P, Beckstein O, Biggin PC, Faraldo-Gómez J, Law RJ, Patargias G, Tieleman DP (2002) Water in ion channels and pores—simulation studies. *Novartis Found Symp* 245:66–78
57. He K, Ludtke SJ, Heller WT, Huang HW (1996) Mechanism of alamethicin insertion into lipid bilayers. *Biophys J* 71:2669–2679
58. Thøgersen L, Schiott B, Vossgaard T, Nielsen NC, Tajkhorshid E (2008) Peptide aggregation and pore formation in a lipid bilayer: a combined coarse-grained and all atom molecular dynamics study. *Biophys J* 95:4337–4347
59. Shih AY, Arkhipov A, Freddolino PL, Schulten K (2006) Coarse grained protein-lipid model with application to lipoprotein particles. *J Phys Chem B* 110:3674–3684
60. Monticelli L, Kandasamy SK, Periole X, Larson SB, Tieleman DP, Marrink SJ (2008) The MARTINI coarse-grained force field: extension to proteins. *J Chem Theory Comput* 4:819–834
61. Chen H, Wu Y, Voth GA (2007) Proton transport behavior through the influenza A M2 channel: insights from molecular simulation. *Biophys J* 93:3470–3479
62. Leonov H, Arkin IT (2009) pH-driven helix rotations in the influenza M2 H(+) channel: a potential gating mechanism. *Eur Biophys J* 39 (7):1043–1049
63. Yi M, Cross TA, Zhou HX (2009) Conformational heterogeneity of the M2 proton channel and a structural model for channel activation. *Proc Natl Acad Sci USA* 106:13311–13316
64. Mustafa M, Henderson DJ, Busath DD (2009) Free-energy profiles for ions in the influenza M2-TMD channel. *Proteins* 76:794–807
65. Khurana E, Dal Peraro M, DeVane R, Vemparala S, DeGrado WF, Klein ML (2009) Molecular dynamics calculations suggest a conduction mechanism for the M2 proton channel from influenza A virus. *Proc Natl Acad Sci USA* 106:1069–1074
66. Carpenter T, Bond PJ, Khalid S, Sansom MS (2008) Self-assembly of a simple membrane protein: coarse-grained molecular dynamics simulations of the influenza M2 channel. *Biophys J* 95:3790–3801
67. Rouse SL, Carpenter T, Stansfeld PJ, Sansom MS (2009) Simulations of the BM2 proton channel transmembrane domain from influenza virus B. *Biochemistry* 48:9949–9951
68. Illya G, Deserno M (2008) Coarse-grained simulation studies of peptide-induced pore formation. *Biophys J* 95:4163–4173
69. Lopez CF, Nielsen SO, Ensing B, Moore PB, Klein ML (2005) Structure and dynamics of model pore insertion into a membrane. *Biophys J* 88:3083–3094
70. Lopez CF, Nielsen SO, Moore PB, Klein ML (2004) Understanding nature's design for a nanosyringe. *Proc Natl Acad Sci USA* 101:4431–4434
71. Nielsen SO, Ensing B, Ortiz V, Moore PB, Klein ML (2005) Lipid bilayer perturbations around a transmembrane nanotube: a coarse grain molecular dynamics study. *Biophys J* 88:3822–3828
72. Klingelhoefer JW, Carpenter T, Sansom MS (2009) Peptide nanopores and lipid bilayers: interactions by coarse-grained molecular-dynamics simulations. *Biophys J* 96:3519–3528
73. Nielsen SO, Lopez CF, Ivanov I, Moore PB, Shelley JC, Klein ML (2004) Transmembrane peptide-induced lipid sorting and mechanism of Lalpha-to-inverted phase transition using coarse-grain molecular dynamics. *Biophys J* 87:2107–2115
74. Venturoli M, Smit B, Sperotto MM (2005) Simulation studies of protein-induced bilayer deformations, and lipid-induced protein tilting, on a mesoscopic model for lipid bilayers with embedded proteins. *Biophys J* 88:1778–1798
75. Smeijers AF, Pieterse K, Markvoort AJ, Hilbers PA (2006) Coarse-grained transmembrane proteins: hydrophobic matching, aggregation, and their effect on fusion. *J Phys Chem B* 110:13614–13623
76. de Meyer FJ, Venturoli M, Smit B (2008) Molecular simulations of lipid-mediated protein-protein interactions. *Biophys J* 95:1851–1865
77. Reynwar BJ, Illya G, Harmandaris VA, Muller MM, Kremer K, Deserno M (2007) Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature* 447:461–464
78. Herce HD, Garcia AE (2007) Cell penetrating peptides: how do they do it? *J Biol Phys* 33:345–356
79. Herce HD, Garcia AE, Litt J, Kane RS, Martin P, Enrique N, Rebollo A, Milesi V (2009) Arginine-rich peptides destabilize the plasma

- membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides. *Biophys J* 97:1917–1925
80. Yesylevskyy S, Marrink SJ, Mark AE (2009) Alternative mechanisms for the interaction of the cell-penetrating peptides penetratin and the TAT peptide with lipid bilayers. *Biophys J* 97:40–49
  81. Lopez CF, Nielsen SO, Srinivas G, Degrado WF, Klein ML (2006) Probing membrane insertion activity of antimicrobial polymers via coarse-grain molecular dynamics. *J Chem Theory Comput* 2:649–655
  82. Bond PJ, Holyoake J, Ivetac A, Khalid S, Sansom MS (2007) Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *J Struct Biol* 157:593–605
  83. Chetwynd A, Wee CL, Hall BA, Sansom MS (2010) The energetics of transmembrane helix insertion into a lipid bilayer. *Biophys J* 99:2534–2540
  84. Durrieu MP, Bond PJ, Sansom MS, Lavery R, Baaden M (2009) Coarse-grain simulations of the R-SNARE fusion protein in its membrane environment detect long-lived conformational sub-states. *Chemphyschem* 10:1548–1552
  85. Sengupta D, Rampioni A, Marrink SJ (2009) Simulations of the c-subunit of ATP-synthase reveal helix rearrangements. *Mol Membr Biol* 26:422–434
  86. Ramadurai S, Holt A, Schafer LV, Krasnikov VV, Rijkers DT, Marrink SJ, Killian JA, Poolman B (2010) Influence of hydrophobic mismatch and amino acid composition on the lateral diffusion of transmembrane peptides. *Biophys J* 99:1447–1454
  87. Bond PJ, Wee CL, Sansom MS (2008) Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. *Biochemistry* 47:11321–11331
  88. Vostrikov VV, Hall BA, Greathouse DV, Koeppe RE 2nd, Sansom MS (2010) Changes in transmembrane helix alignment by arginine residues revealed by solid-state NMR experiments and coarse-grained MD simulations. *J Am Chem Soc* 132:5803–5811
  89. Spijker P, van Hoof B, Debertrand M, Markvoort AJ, Vaidehi N, Hilbers PA (2010) Coarse grained molecular dynamics simulations of transmembrane protein-lipid systems. *Int J Mol Sci* 11:2393–2420
  90. Gkeka P, Sarkisov L (2009) Spontaneous formation of a barrel-stave pore in a coarse-grained model of the synthetic LS3 peptide and a DPPC lipid bilayer. *J Phys Chem B* 113:6–8
  91. Gkeka P, Sarkisov L (2010) Interactions of phospholipid bilayers with several classes of amphiphilic alpha-helical peptides: insights from coarse-grained molecular dynamics simulations. *J Phys Chem B* 114:826–839
  92. Bond PJ, Sansom MS (2006) Insertion and assembly of membrane proteins via simulation. *J Am Chem Soc* 128:2697–2704
  93. Psachoulia E, Fowler PW, Bond PJ, Sansom MS (2008) Helix-helix interactions in membrane proteins: coarse-grained simulations of glycophorin a helix dimerization. *Biochemistry* 47:10503–10512
  94. Sengupta D, Marrink SJ (2010) Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes. *Phys Chem Chem Phys* 12:12987–12996
  95. Psachoulia E, Marshall DP, Sansom MS (2010) Molecular dynamics simulations of the dimerization of transmembrane alpha-helices. *Acc Chem Res* 43:388–396
  96. Kalli AC, Wegener KL, Goult BT, Anthis NJ, Campbell ID, Sansom MS (2010) The structure of the talin/integrin complex at a lipid bilayer: an NMR and MD simulation study. *Structure* 18:1280–1288
  97. Holdbrook DA, Leung YM, Piggot TJ, Marius P, Williamson PT, Khalid S (2010) Stability and membrane orientation of the fukutin transmembrane domain: a combined multiscale molecular dynamics and circular dichroism study. *Biochemistry* 49:10796–10802
  98. White SH (2009) Biophysical dissection of membrane proteins. *Nature* 459:344–346
  99. Basyn F, Charlotteaux B, Thomas A, Brasseur R (2001) Prediction of membrane protein orientation in lipid bilayers: a theoretical approach. *J Mol Graph Model* 20:235–244
  100. Ducarme P, Rahman N, Brasseur R (1998) IMPALA: a simple restraint field to simulate the biological membrane in structure studies. *Proteins* 30:357–371
  101. Lomize AL, Pogozheva ID, Lomize MA, Mosberg HI (2006) Positioning of proteins in membranes: a computational approach. *Protein Sci* 15:1318–1333
  102. Ulmschneider MB, Sansom MS, Di Nola A (2006) Evaluating tilt angles of membrane-associated helices: comparison of computational and NMR techniques. *Biophys J* 90:1650–1660

103. Balali-Mood K, Bond PJ, Sansom MS (2009) Interaction of monotopic membrane enzymes with a lipid bilayer: a coarse-grained MD simulation study. *Biochemistry* 48:2135–2145
104. Wee CL, Balali-Mood K, Gavaghan D, Sansom MS (2008) The interaction of phospholipase A2 with a phospholipid bilayer: coarse-grained molecular dynamics simulations. *Biophys J* 95:1649–1657
105. Periole X, Huber T, Marrink SJ, Sakmar TP (2007) G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J Am Chem Soc* 129:10126–10132
106. Bond PJ, Sansom MS (2007) Bilayer deformation by the Kv channel voltage sensor domain revealed by self-assembly simulations. *Proc Natl Acad Sci USA* 104:2631–2636
107. Periole X, Cavalli M, Marrink SJ, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* 5:2531–2543
108. Yefimov S, van der Giessen E, Onck PR, Marrink SJ (2008) Mechanosensitive membrane channels in action. *Biophys J* 94:2994–3002
109. Lycklama ANJA, Bulacu M, Marrink SJ, Driessens AJ (2010) Immobilization of the plug domain inside the SecY channel allows unrestricted protein translocation. *J Biol Chem* 285:23747–23754
110. Zhang B, Miller TF 3rd (2010) Hydrophobically stabilized open state for the lateral gate of the Sec translocon. *Proc Natl Acad Sci USA* 107:5399–5404
111. Sansom MS, Scott KA, Bond PJ (2008) Coarse-grained simulation: a high-throughput computational approach to membrane proteins. *Biochem Soc Trans* 36:27–32
112. Scott KA, Bond PJ, Ivetac A, Chetwynd AP, Khalid S, Sansom MS (2008) Coarse-grained MD simulations of membrane protein-bilayer self-assembly. *Structure* 16:621–630
113. Treptow W, Marrink SJ, Tarek M (2008) Gating motions in voltage-gated potassium channels revealed by coarse-grained molecular dynamics simulations. *J Phys Chem B* 112:3277–3282
114. Wee CL, Bemporad D, Sands ZA, Gavaghan D, Sansom MS (2007) SGTx1, a Kv channel gating-modifier toxin, binds to the interfacial region of lipid bilayers. *Biophys J* 92: L07–L09
115. Wee CL, Gavaghan D, Sansom MS (2010) Interactions between a voltage sensor and a toxin via multiscale simulations. *Biophys J* 98:1558–1565
116. Wee CL, Gavaghan D, Sansom MS (2008) Lipid bilayer deformation and the free energy of interaction of a Kv channel gating-modifier toxin. *Biophys J* 95:3816–3826
117. Stansfeld PJ, Hopkinson R, Ashcroft FM, Sansom MS (2009) PIP(2)-binding site in Kir channels: definition by multiscale biomolecular simulations. *Biochemistry* 48:10926–10933
118. Shih AY, Denisov IG, Phillips JC, Sligar SG, Schulten K (2005) Molecular dynamics simulations of discoidal bilayers assembled from truncated human lipoproteins. *Biophys J* 88:548–556
119. Shih AY, Freddolino PL, Arkhipov A, Schulten K (2007) Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. *J Struct Biol* 157:579–592
120. Shih AY, Arkhipov A, Freddolino PL, Sligar SG, Schulten K (2007) Assembly of lipids and proteins into lipoprotein particles. *J Phys Chem B* 111:11095–11104
121. Shih AY, Freddolino PL, Sligar SG, Schulten K (2007) Disassembly of nanodiscs with cholate. *Nano Lett* 7:1692–1696
122. Shih AY, Sligar SG, Schulten K (2008) Molecular models need to be tested: the case of a solar flares discoidal HDL model. *Biophys J* 94:L87–L89
123. Shih AY, Sligar SG, Schulten K (2009) Maturation of high-density lipoproteins. *J R Soc Interface* 6:863–871
124. Catte A, Patterson JC, Bashtonyy D, Jones MK, Gu F, Li L, Rampioni A, Sengupta D, Vuorela T, Niemela P, Karttunen M, Marrink SJ, Vattulainen I, Segrest JP (2008) Structure of spheroidal HDL particles revealed by combined atomistic and coarse-grained simulations. *Biophys J* 94:2306–2319
125. Vuorela T, Catte A, Niemela PS, Hall A, Hyvonen MT, Marrink SJ, Karttunen M, Vattulainen I (2010) Role of lipids in spheroidal high density lipoproteins. *PLoS Comput Biol* 6:e1000964
126. Izvekov S, Voth GA (2005) A multiscale coarse-graining method for biomolecular systems. *J Phys Chem B* 109:2469–2473
127. Shi Q, Izvekov S, Voth GA (2006) Mixed atomistic and coarse-grained molecular dynamics: simulation of a membrane-bound ion channel. *J Phys Chem B* 110:15045–15048
128. Ayton GS, Lyman E, Voth GA (2010) Hierarchical coarse-graining strategy for protein-membrane systems to access mesoscopic

- scales. *Faraday Discuss.* 144:347–357, discussion 445–381
- 129. Ayton GS, Voth GA (2010) Multiscale simulation of protein mediated membrane remodeling. *Semin Cell Dev Biol* 21:357–362
  - 130. Blood PD, Voth GA (2006) Direct observation of Bin/amphiphysin/Rvs (BAR) domain-induced membrane curvature by means of molecular dynamics simulations. *Proc Natl Acad Sci USA* 103:15068–15072
  - 131. Cui H, Ayton GS, Voth GA (2009) Membrane binding by the endophilin N-BAR domain. *Biophys J* 97:2746–2753
  - 132. Arkhipov A, Yin Y, Schulten K (2008) Four-scale description of membrane sculpting by BAR domains. *Biophys J* 95:2806–2821
  - 133. Arkhipov A, Yin Y, Schulten K (2009) Membrane-bending mechanism of amphiphysin N-BAR domains. *Biophys J* 97:2727–2735
  - 134. Orsi M, Sanderson WE, Essex JW (2009) Permeability of small molecules through a lipid bilayer: a multiscale simulation study. *J Phys Chem B* 113:12019–12029
  - 135. Orsi M, Noro MG, Essex JW (2010) Dual-resolution molecular dynamics simulation of antimicrobials in biomembranes. *J R Soc Interface* 8(59):826–841
  - 136. Ayton GS, Voth GA (2007) Multiscale simulation of transmembrane proteins. *J Struct Biol* 157:570–578
  - 137. Shimamura T, Weyand S, Beckstein O, Rutherford NG, Hadden JM, Sharples D, Sansom MS, Iwata S, Henderson PJ, Cameron AD (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* 328:470–473
  - 138. Bond PJ, Derrick JP, Sansom MSP (2007) Membrane simulations of OpcA: gating in the loops? *Biophys J* 92:L23–L25
  - 139. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK (2011) Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* 469:236–240
  - 140. Niemela PS, Miettinen MS, Monticelli L, Hammaren H, Bjelkmar P, Murtola T, Lindahl E, Vattulainen I (2010) Membrane proteins diffuse as dynamic complexes with lipids. *J Am Chem Soc* 132:7574–7575
  - 141. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346

# Chapter 26

## Vesicles and Vesicle Fusion: Coarse-Grained Simulations

Julian C. Shillcock

### Abstract

Biological cells are highly dynamic, and continually move material around their own volume and between their interior and exterior. Much of this transport encapsulates the material inside phospholipid vesicles that shuttle to and from, fusing with, and budding from, other membranes. A feature of vesicles that is crucial for this transport is their ability to fuse to target membranes and release their contents to the distal side. In industry, some personal care products contain vesicles to help transport reagents across the skin, and research on drug formulation shows that packaging active compounds inside vesicles delays their clearance from the blood stream. In this chapter, we survey the biological role and physicochemical properties of phospholipids, and describe progress in coarse-grained simulations of vesicles and vesicle fusion. Because coarse-grained simulations retain only those molecular details that are thought to influence the large-scale processes of interest, they act as a model embodying our current understanding. Comparing the predictions of these models with experiments reveals the importance of the retained microscopic details and also the deficiencies that can suggest missing details, thereby furthering our understanding of the complex dynamic world of vesicles.

**Key words:** Vesicle, Vesicle fusion, Phospholipid, Bilayer, Membrane, Coarse-grained simulation, Dissipative particle dynamics

---

### 1. Introduction

The fundamental division of the molecules of life is usually presented as the triad of DNA, RNA, and proteins. Of these, proteins are the machines of nature, carrying out almost all of the mechanical functions of a living cell. Notwithstanding the recent discovery that RNA also plays an active role in cellular function, from the ribosomes that construct new proteins to pieces of RNA that regulate gene expression, proteins are recognized as the most dynamic molecules in a cell.

Largely overlooked until recently is the role of lipids. These are small compared to proteins, typically 1–2 nm in size, with a molecular weight around 700 Da. There are hundreds of different types of lipid, and they are the major constituent of all membranes in a

cell, from the Plasma Membrane itself, to those surrounding internal organelles such as the Endoplasmic Reticulum, Golgi apparatus, and nuclear membrane, to transport vesicles that shuttle around the cell carrying material from place to place (1). Phospholipids are the most abundant type of lipid in mammalian cells, and are amphiphiles, consisting of a hydrophilic (water-loving) headgroup to which are attached (via a glycerol moiety) one, two, or more hydrophobic (water-hating), fatty acid tails. When lipids are dispersed in water, their hydrophilic headgroups want to remain solvated, but their hydrophobic tails want to separate from the water. Because the two parts are tied together by covalent bonds, lipids are frustrated, and out of this frustration they spontaneously assemble into amazingly complex, thin yet robust, fluid structures (2). The shapes of these structures include small spheres (micelles) only a few nanometer in diameter; planar sheets that can be microns across but only are a few nanometer thick consisting of two apposed lipid monolayers (bilayers); and spherical sacs (vesicles) that range from a few tens of nanometer in diameter for transport vesicles to the 10  $\mu\text{m}$  diameter cellular Plasma Membrane that are essentially a planar bilayer that has closed on itself. Even more complicated structures exist that are less easily described, such as the tubular, smooth Endoplasmic Reticulum.

In this chapter, we describe how vesicles are formed from phospholipids, and how they are investigated experimentally and represented in theory and computer simulations. We use vesicle fusion (3, 4) to illustrate the increasingly important use of coarse-grained, particle-based simulations for modeling and visualizing complex dynamical cellular processes. Because biological membranes are composed mainly of phospholipid molecules, we look first at the physicochemical properties of lipid membranes.

Lipid bilayer vesicles are the most important lipid aggregate in biology as each cell is surrounded by its Plasma Membrane (PM), and contains many internal organelles each of which has its own bounding lipid bilayer. The PM represents only 2% of the total membrane area in a cell, the rest being part of the organelles of the cell (5). The average separation of these internal membranes is only about 50 nm, which is the same scale as that of large macromolecules. Cells are therefore dominated by membranes, and phospholipids have largely been viewed as an inert structural component. But lipids are increasingly being recognized as more than a passive chorus in a drama played by proteins, and their physicochemical properties are essential to many nonstructural functions in a cell (6) including signaling (7). Their importance does not end with their natural functions. Artificial vesicles composed of molecules formed from a short sequence of DNA bound to hydrocarbon chains (8) have been constructed that bind together in programmable ways (9). The increasing production of nanoparticles for industrial and pharmaceutical purposes has made understanding the interactions

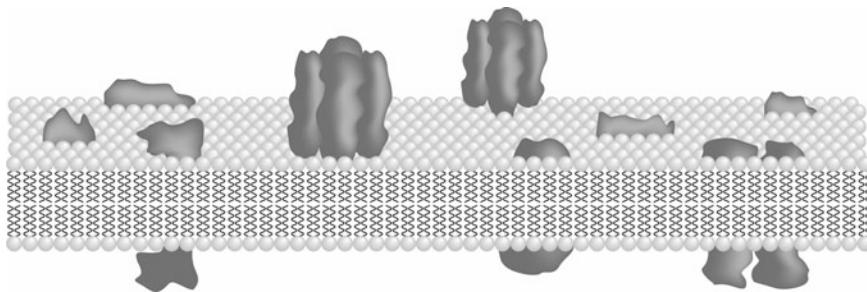


Fig. 1. Cartoon of the plasma membrane (PM) of a cell showing transmembrane proteins embedded in the fluid lipid bilayer. Proteins occur singly and oligomerized in groups. This is an updated view of the Singer–Nicolson “fluid-mosaic” model (17), which showed widely separated proteins floating in a sea of lipid. Increasing experimental evidence shows that a large fraction of the PM area is either occupied, or occluded, by proteins, and that many transmembrane proteins occur as oligomers. Additionally, although all lipids look alike in this cartoon, the PM contains hundreds of types of lipids that differ in their headgroup, charge, tail length, presence of double-bonds, etc. (picture courtesy of Thomas E. Rasmussen).

of such particles with living cells, and particularly the cellular PM, an urgent task (10). Potentially toxic effects arising from ingestion of nanoparticles by cells (11–13) must be quantified and monitored. Industrial sectors have also started to see the value of lipid vesicles in fields such as personal care products and pharmaceutical drug delivery. These applications require optimizing the size and shape of nanoparticles (14) so as to maximize their effect and minimize any undesired interactions with biological membranes (15). Evidence of the growing interest in lipids is that a recent issue of Soft Matter was devoted to the biophysics of membranes (16).

The Fluid Mosaic model of Singer and Nicholson (17) is the most common illustration of biological phospholipid membranes (such as the cellular Plasma Membrane) that appears in textbooks (1). In this model, the lipids constitute an oriented fluid bilayer matrix within which proteins float essentially independently. But recently it has become clear that this picture is oversimplified. An updated view of the Fluid-Mosaic model was given by Engelmann (18), and is illustrated by the cartoon in Fig. 1. The principal insight that the membrane is a fluid lipid bilayer is unaltered, but it is now envisaged as being crowded with proteins, some occurring singly but others oligomerized into functional groups, and others, although having only a thin transmembrane anchor, occluding a large area of membrane because of their cytoplasmic or extracellular domains. Similar to the plasma membrane, the surface of synaptic vesicles is largely obscured by transmembrane proteins (19). Whereas the earlier view of lipids assigned them only a barrier function, separating the cytoplasm from the exterior world and providing proteins with a platform on which to work, they are now seen as more active components of the cell. Membranes help to localize and organize other cellular components so that interactions that would be rare become more frequent. Proteins, for example, may diffuse through the cytoplasm of a cell until they

encounter, and bind to, another protein, thereby initiating a step in a reaction. But the same proteins diffusing in or near a membrane are more likely to encounter each other sooner as they effectively diffuse in two dimensions instead of three.

Membranes act as organizing agents in a cell, as well as barriers isolating distinct compartments. But the barrier function of membranes must be modified at times. Cells need to take in nutrients and expel waste material to survive. If the PM were an impenetrable wall, the cell would die. Conversely, if anything were allowed to pass through the PM, the cell would also die. Endo- and exocytosis are mechanisms whereby a cell ingests and expels material across its PM, and are accordingly highly regulated. Incoming material is packaged into lipid vesicles that pinch off from the PM and carry their payload into the cell. Waste material is expelled through pores created by vesicles fusing to the cytoplasmic side of the PM. Vesicle fusion also occurs in signal transmission across synapses in the brain, where neurotransmitter is released into the synaptic cleft between neurons, and hormones such as insulin enter the bloodstream via the fusion of insulin-containing granules to the PM of  $\beta$  cells in the pancreas. All of these processes require that a pore be produced in the normally intact lipid membrane, and vesicle fusion is a major focus of experiments, theory, and computer simulations.

Collective properties of lipids can be examined using monolayers in a Langmuir trough. A single monolayer of lipids is laid on the surface of a water reservoir that is exposed to air and bounded by three fixed walls and one movable barrier that separates the monolayer from a clean water surface. The net force on the movable barrier is in the direction of the clean water surface because the surface tension of the air–water interface always exceeds that of the monolayer–air interface. Another way of saying this is that the lipid monolayer would spread out to cover the whole water–air surface if the movable wall were removed. An equation of state of the monolayer can be expressed as a Pressure–Area isotherm, where the pressure is varied by moving the barrier and the subsequent area is measured. If the number of lipid molecules is known, the area can be expressed as an area per molecule.

Lipid monolayers are subject to various forces, both internal tensions due to oil–water and oil–air interfaces that try to minimize the monolayer area and entropic forces due to chain conformations and entropy of mixing of water molecules in the lipid headgroup region that try to expand the area. The net result of the opposing attractive and repulsive forces gives rise to the surface pressure of a monolayer that expands its area unless opposed by a mechanical constraint. If the pressure on the monolayer is low, around 1 mN/m, the lipids behave essentially as a gas at the air–water interface, and each lipid molecule has ample area to move around. As the pressure is increased, by reducing the monolayer area, the lipids start to notice each other and orient normal to the interfacial plane.

At higher pressures a phase transition occurs as the lipids lose translational freedom and are forced into a vertical orientation in which they can no longer translate but their hydrocarbon tails can still wiggle. At still higher pressure, typically around 45 mN/m, the monolayer buckles out of the plane and lipids enter the water phase. Measuring the Pressure-Area isotherms for different lipids has been an important tool for understanding the effects of lipid molecular structure on the collective material properties of lipid monolayers.

Although a lipid bilayer is not exactly equivalent to a monolayer, most obviously because there is no oil-air interface in a lipid bilayer, much has been learned about the physicochemical properties of lipids from Langmuir trough experiments that informs experiments on lipid bilayers and vesicles (20). Giant vesicles are frequently used in experiments on lipid bilayers because they are (relatively) simple to prepare and can be observed in a light microscope. Their typical size is 10–20  $\mu\text{m}$  in diameter, and their shape fluctuates continually due to thermal motion of the constituent lipids and surrounding solvent (20, 21). These fluctuations show that they are extremely soft objects whose properties on length scales much larger than the size of their constituent lipid molecules are mainly controlled by their membrane bending stiffness. Unlike a lipid monolayer, which expands if no mechanical constraint is imposed, the opposed forces of attraction arising from the hydrophobic effect that drives the lipid tails to aggregate and repulsion of the hydrophilic headgroups result in a preferred area per molecule in a lipid bilayer. It costs energy to increase and decrease the area of such a bilayer, so the surface tension of unstressed vesicles is typically close to zero.

The elastic properties of vesicles have been measured using micropipette aspiration experiments (22). In these experiments, a vesicle is partially sucked into a pipette, and the length of the extended neck measured as a function of the suction pressure. The result is a measurement of the vesicle membrane's area stretch modulus (23), which can be theoretically related to its bending modulus. It is found that the area stretch modulus of a single-component vesicle varies weakly with the lipid hydrocarbon chain length and their degree of unsaturation. The bending modulus, however, increases progressively with chain length for saturated lipids, and, for longer chains, drops sharply when two or more *cis* double bonds are present in a chain. Experiments such as these on lipid monolayers and vesicles have quantified the dependence of membrane properties on their constituent phospholipid molecular structure, and created a need for theoretical models to explain this dependence.

Theoretical treatments of bilayers and vesicles frequently start from the large-scale view that a membrane is a thin, structureless sheet, whose behavior is controlled by appropriate elastic moduli and a set of external constraints. Typical constraints are that the

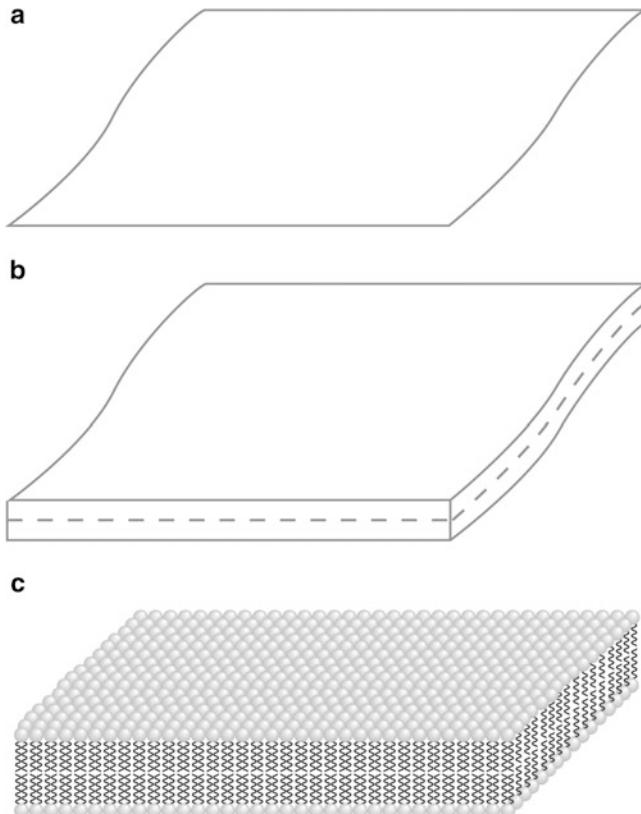
total number of lipids in the membrane is constant, and that the volume of a vesicle is constant. Figure 2 shows cartoons of three different representations of a lipid membrane. If the membrane thickness is regarded as infinitesimal (Fig. 2a), the membrane shape is defined in terms of its curvature at each point, and thermal fluctuations cause it to fluctuate continually. The simplest expression for the energy density arising from these shape fluctuations is the Canham–Helfrich Hamiltonian (24, 25):

$$f = 1/2\kappa(2H - C_0)^2 + \kappa_G K, \quad (1)$$

where the mean curvature  $H = 1/2(1/R_1 + 1/R_2)$  and gaussian curvature  $K = 1/(R_1 R_2)$  and  $R_1, R_2$  are the principal curvatures of the membrane at each point (26). The variable  $C_0$ , which can vary throughout the membrane, is the spontaneous curvature that quantifies the membrane's tendency to be flat or curved. The bending modulus,  $\kappa$ , sets the energy scale for fluctuations of the membrane that cause it to bend, while the gaussian curvature modulus  $\kappa_G$  is constant for membranes whose topology is constant, and can be ignored. This model defines the membrane's energy as a power series in its curvature, and is a good approximation if the membrane thickness is negligible compared to its length and its shape fluctuations are not too large.

Various alternative forms of the so-called curvature elastic models exist, for example the Area-Difference Elasticity model (21), in which the control parameters for the vesicle's energy are its reduced volume and the difference in the area of each monolayer from its preferred value. They have been used very successfully to generate and classify the shapes of vesicles as a function of the relevant state parameters (26). Curvature models of vesicles have been extended to include local fields embedded on the membrane surface (Fig. 2b). Each field adds extra terms to Eq. 1 with appropriate coupling constants. This approach has been used to explore the effects on membrane equilibrium shape of a local tilt of the lipid molecules. The molecular tilt induces a nematic order in the membrane that is coupled to the local membrane curvature, and can drive an instability in the flat phase of the membrane so that it becomes curved (27).

Treating membranes as fluctuating, thin sheets places them in field of fluctuating surfaces, both fluid and tethered, which has been reviewed in (28, 29). This approach is best suited to analyzing global properties of a vesicle, such as global shape transitions as a function of temperature and composition (30–32). It has recently been extended to allow simulations of a vesicle that has a nematic field embedded on its surface (33). It is less well-suited to situations where the membrane thickness is not negligible, such as membrane fusion, in which molecular conformations change over distances comparable to the membrane thickness.



**Fig. 2.** Representations of a fluid lipid membrane at different resolutions. **(a)** Infinitely thin elastic sheet described by the Canham–Helfrich Hamiltonian, Eq. 1. A lipid membrane in this model is a thin, flexible, uniform, elastic sheet that has no bilayer structure per se. The degrees of freedom of this membrane are the membrane curvature and spontaneous curvature that may vary from point to point. Its physical properties are completely described by its bending elastic constant, spontaneous curvature (if nonzero), and mechanical constraint on its area. This model is the first in an expansion of the membrane's elastic energy in powers of its curvature, so it is restricted to gently curved membranes only, and extra terms must be added to the Hamiltonian to describe energy changes that occur during, for example, pore formation. **(b)** Membrane with finite thickness and local field variables as used in recent continuum elastic models of membrane fusion (146). Compared to the infinitely thin membrane in **(a)**, this model possesses a bilayer structure and has a field defined at each point on the membrane that represents the tilt angle between the lipids at that point and the local normal vector to the membrane. Other fields can be added by including extra terms in the Hamiltonian as for the previous model. **(c)** Lipid bilayer in a coarse-grained, particle-based simulation such as DPD, in which a lipid molecule is represented by a set of particles, each of which represents several atoms or atomic groups. Note that for simplicity, we represent the tails of the lipid by thick lines, but in the simulation they consist of particles. The headgroup particles of the lipid are hydrophilic, whereas the tail particles are hydrophobic. These properties are captured in the force fields used in the simulation. Particles are bonded together into molecules using elastic springs. Each lipid has many degrees of freedom corresponding to the positions and velocities of its constituent particles, and these degrees of freedom allow membrane conformations that are not present in the lower resolution models **(a)** and **(b)**. For example, a lipid with two tails may splay so that one of its tails protrudes outside the hydrophobic core of its owning membrane and embeds itself in a closely apposed membrane (picture courtesy of Thomas E. Rasmussen).

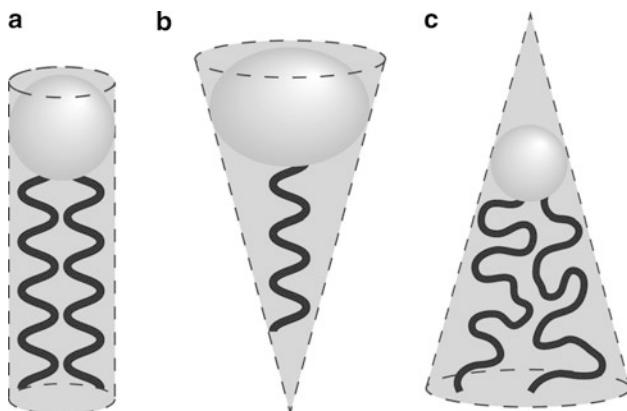


Fig. 3. Schematic representation of different shapes of lipid molecule that occur in the Plasma Membrane and vesicles of mammalian cells. (a) Bilayer-forming lipids, such as DMPC, are typically cylindrical so that their headgroups occupy an area of the membrane's surface that approximately matches the cross-sectional area of its two hydrocarbon tails. A symmetric bilayer composed of such lipids has zero spontaneous curvature and has a flat state as its lowest energy conformation. (b) Lipids with larger headgroups than tail area, such as lysophosphatidylcholine, impart a positive spontaneous curvature to a monolayer that contains them, causing it to curve outwards from tail region to headgroup region. (c) Lipids with smaller headgroups than tail area, such as ethanolamine-containing lipids, generate a negative spontaneous curvature that causes a monolayer containing them to curve inwards from the headgroup area to the tail region. If a lipid bilayer contains an asymmetric distribution of non-cylindrical lipids in its two monolayers, it will try to adopt a shape in which its curvature matches the net spontaneous curvature of its components at each point. If this is not possible at all points, the membrane will be in a frustrated state, and may be able to relieve the stress by adopting a non-bilayer state either locally or globally (picture courtesy of Thomas E. Rasmussen).

Particle-based computer simulations start from the small-scale view that the molecules composing the membrane cannot be ignored (Fig. 2c), and are explicitly represented with more or less accuracy. Whereas elastic curvature models of a membrane require extra terms in the Hamiltonian to represent the effects of additional local properties in the membrane, such properties arise naturally in a simulation from the degrees of freedom of the molecules. As an example, the spontaneous curvature of a fluid lipid membrane is a property of the packing of the lipid molecules in the lamellar structure of the membrane, and determines whether the membrane's equilibrium state is flat or curved. Figure 3 illustrates the effective shape that different types of lipid molecule can adopt in a bilayer leaflet (2). Those whose molecular shape is roughly that of a cylinder tend to pack into a flat monolayer (Fig. 3a). An example is dimyristoyl phosphatidylcholine (DMPC), whereas those whose headgroup is larger than the cross-sectional area of their tails (Fig. 3b) or smaller (Fig. 3c) will naturally form curved monolayers. If such molecules are placed asymmetrically in the two leaflets of a vesicle with excess

area, it will adopt a nonspherical shape spontaneously (but see ref. (34) for caveats regarding too literal an approach to the shape of lipid molecules).

Phospholipids dispersed in water spontaneously assemble into aggregates such as vesicles. But what does it mean to say that pieces of a system self-assemble? It does not mean that the pieces contain information about the larger structure of which they are parts, still less the function that this structure performs. It is rather the assertion that the material, shape, and interactions of the pieces cause them, when immersed in the appropriate environment, to respond to each other by attaching in just those ways that lead to the building up of the larger structure in a precise conformation (35). How does this happen? In random thermal fluctuations, nature has an unlimited supply of motive force to drive the construction of aggregates. At the macroscopic level energy must be expended to hold, manipulate, orient, and attach materials in the construction process; otherwise our building blocks would remain inert where they were last placed, gradually eroding over time or breaking up. But in the cellular world, thermal motion continually drives proteins, membranes, and other materials around, diffusing, bumping into each other, and jiggling in all manner of orientations until they find just that arrangement that allows them to interact in their preferred way, and so leads them to bind strongly together because no other arrangement lowers their free energy as much. The physics of random thermal motion and the chemistry of molecular interactions allow, or rather, drive, simpler objects to form more complex ones without there being an overall blueprint that defines their arrangement. Once a structure exists that performs a task, it is available for reuse in larger structures, and as time passes larger and more complex aggregates emerge.

The emergence of increasing complexity as amphiphiles aggregate into membranes, and membranes partition space into disconnected volumes, is vital to the viability of cellular life. Only in the last 20 years have computers become powerful enough to allow their application to the study of such processes. But progress in computational modeling of soft matter has been so rapid that the molecular dance that leads to vesicle formation out of randomly dispersed lipids in water, and their subsequent fusion to form larger and larger vesicles, is now visible in coarse-grained simulations after modest computational effort.

In the next section, we describe the features of several coarse-grained simulation techniques suitable for membranes and vesicles, and illustrate the results that have been obtained. We mention recent attempts to include nanoparticles and (highly simplified models of) proteins in the simulations so as to study important biological problems such as membrane shaping by proteins (36, 37). In the final section, we focus on vesicles as dynamic platforms for the simulation of biological processes and describe

recent results in simulating vesicle fusion. Membrane fusion is a vital process for cells, and is being actively studied experimentally and computationally. In cells, it occurs on a length scale of 50–100 nm, which puts it at the boundary of what can be captured in current computational models within a reasonable time frame. Most results published using coarse-grained simulations use a single level of coarse-graining technique carried out on (biologically) small systems, typically less than 50 nm in linear dimension. This is far less than the micron scale of a living cell. In order to have any hope of simulating biologically relevant length and timescales, a combination of coarse-graining techniques must be used (38). Advances are currently being made in this direction (39–42), but the field of multi-scale coarse-graining of biological materials (43) is still very much in its infancy.

---

## 2. Coarse-Grained Simulation Methodologies for Vesicles

Computer simulations of lipid membranes can be carried out at very different levels of detail. Atomistic Molecular Dynamics (MD) is the most accurate, requiring complex interatomic force fields, and is consequently limited to small systems containing perhaps a few hundred lipids plus surrounding solvent molecules (see Chapter 6). While this is ideal for certain problems, examples of which are given in the Applications chapters of part 2 of this volume, the desire to study larger systems has driven the development of various coarse-grained simulation techniques (44). A detailed introduction to coarse-graining is given in Chapter 19, so we only summarize the various techniques here.

In all coarse-grained simulation techniques some microscopic degrees of freedom of the molecules are averaged over and the atom–atom interaction potentials replaced by softer, usually short-ranged, effective potentials. It is important to realize that the choice of coarse-graining scheme largely determines the accuracy of the results that can be obtained. Broadly speaking, a greater degree of coarse-graining leads to a more generic model that is harder to relate to a specific molecule, but which reveals universal properties that are less affected by detailed chemistry and therefore more widely applicable. An excellent discussion of the choices that must be made in selecting a coarse-grained simulation technique for a specific system is given in the concluding remarks from a recent Faraday Discussion workshop (45). A recent review of computer simulations of lipid membranes also gives an overview of the current state of the art (46).

Coarse-grained MD has been used to study the self-assembly and material properties of generic lipid bilayers (47) and non-lipidic vesicles (48), while an entire force field has been developed

specifically for simulating lipids and sterols, including planar ring compounds (49), and is described in Chapter 20. Common force fields used in Molecular Dynamics simulations of proteins and lipids are presented in Chapters 9 and 10 of this volume. MD simulations have also proved valuable in following dynamic processes in vesicles such as fission, where daughter vesicles pinch off from a parent vesicle due to repulsion between two different lipid types (50).

Increasing the scale of coarse-graining leads to simulation types such as Brownian Dynamics BD (51), solvent-free techniques (52–55), and Dissipative Particle Dynamics DBD (56–60). The latter technique, which is described in detail in Chapter 24, has become popular for simulating processes that are too large for the more accurate techniques of atomistic and coarse-grained MD, but which require the presence of explicit solvent to ensure that the correct dynamics are being modeled.

An important aspect of all coarse-grained models is that each may be better suited than others for capturing certain properties of a modeled system. In fact, some properties of a physical system may be completely absent in a simulation. For example, solvent-free simulation techniques such as Brownian Dynamics do away with the solvent molecules entirely. The hydrophobic effect that drives amphiphiles to self-assemble into aggregates is manually inserted into the model by using complex force fields instead of relying on the simple repulsion of hydrophobic molecular fragments from water. It is important therefore to understand how each type of coarse-graining restricts the physicochemical phenomena that can be represented within a simulation type.

The principal property of lipids that makes them so important biologically, their amphiphilic nature, is remarkably robust in that different coarse-grained representations give quite similar predictions for their behavior in bilayers and vesicles (61, 62). A recent review was devoted to the study and comparison of various mesoscopic simulation techniques for lipid bilayers (63), and concluded that the various coarse-graining strategies produced surprisingly consistent results. This suggests that only a few microscopic lipid properties must be retained in order to describe a membrane accurately on the nanometer to micron scale. This in turn raises the hope that large systems, perhaps approaching biological scales, can be simulated using an appropriate choice of coarse-graining. But first we must answer the question of why one should not simply use the most accurate simulation technique available to simulate the system of interest.

Why coarse-grain? Even though we must clearly keep some molecular details, viz, their amphiphilicity, we expect that not all atomic-scale properties of lipids are relevant for long length and timescale phenomena like vesicle fusion (64). It is a waste of computational resources to retain details that are too small or short-lived to

affect the processes we are interested in studying. Coarse-grained models of lipid membranes have been developed over the last 20 years because they are far more efficient in their use of computational resources than traditional Molecular Dynamics, and so can approach biological systems in size if not yet in complexity.

Coarse-graining typically takes two forms: first, several atoms or atomic groups are lumped into a single particle or “bead,” thereby reducing the number of degrees of freedom (positions and momenta) that must be integrated in the system’s equations of motion; second, the forces between particles are generally softer than the Lennard-Jones potentials used in MD, and so allow a larger integration step-size to be used. As an example, Dissipative Particle Dynamics is more than six orders of magnitude faster than atomistic Molecular Dynamics and yet the largest lipid membrane simulated by DPD so far is only 100 nm in linear extent (61). The smallest vesicles involved in synaptic vesicle fusion are about 40 nm in diameter. Because the size of the simulation box containing a vesicle of diameter  $D$  grows as  $D^3$ , many groups have studied planar patches of lipid membrane representing a piece taken from a large vesicle. Very recently, attempts have been made to simulate vesicles with diameters of hundreds of nanometers using a hybrid coarse-grained approach (41), or by treating small patches of membrane as the fundamental unit of the simulation instead of lipid molecules (65). These are promising beginnings, but are not yet in widespread use.

Even most Molecular Dynamics simulations ignore some atomic degrees of freedom: nonpolar hydrogen atoms for example in the united atom model. The question is not whether but how to coarse-grain a system. The art of coarse-graining lies in identifying those degrees of freedom that are, in some sense, less important than others, and then subsuming them into the retained properties.

Viewed on length scales between nanometers and microns, the self-assembly and equilibrium properties of a lipid bilayer are essentially the same when simulated by coarse-grained MD, DPD, BD, and many types of solvent-free models (52–54). This suggests that coarse-grained models do indeed capture much of the physics of lipid bilayers in a technique-independent way. Because coarse-grained simulations are orders of magnitude faster than atomistic MD, they are frequently used to obtain a preliminary overview of the phase space of amphiphilic systems, to get a feel for their behavior. Of course, such coarse-grained models cannot reproduce the accuracy of atomistic MD, and can never replace it for certain problems. The aim is to apply them to phenomena on length and timescales beyond the reach of MD, for which the important molecular details can be identified. Slow processes, such as phase transitions, and long wavelength motions, such as membrane remodeling during fusion or exocytosis, are cellular processes that call out for the application of coarse-grained models.

Another reason for simplifying the representation of a system via coarse-graining is that new interactions appear on mesoscopic scales that are not present at smaller scales (66). Entropic forces, such as fluctuation-induced repulsion between membranes (67, 68) and the depletion effect (2), can make a large contribution to the stability and properties of vesicles and supramolecular aggregates. These effects should be visible in simulations on scales of 100 nm or more, and their independence from molecular details means that they can be represented within coarse-grained models.

An important consequence of coarse-graining that must be faced whenever degrees of freedom are integrated out in a simulated system is the loss of a direct link between the simulated and experimental observables. Coarse-graining weakens the link between the entities in the model (whether molecules, molecular groups, or membrane “particles”) and their physical counterparts. Instead of atoms interacting via highly complex but accurate force fields, we must deal with somewhat ill-defined particles or beads interacting via effective forces. Attempts are beginning to be made to relate accurate coarse-grained Molecular Dynamics simulations to high-level, more generic methods (41), but currently most coarse-grained simulations have units inserted “by hand.” This requires matching one length and timescale for known processes occurring in the physical and simulated systems and using these units to scale other quantities in the simulation. For amphiphilic membrane systems, the most common choice of units involves matching the equilibrium area per molecule of the amphiphiles in the membrane to set the length scale, and using the in-plane diffusion constant of the molecules to set the timescale. Other choices are, of course, possible.

A second consequence of coarse-graining applies particularly to DPD, and is a warning to ensure that the simulated system faithfully represents the physical system. As mentioned at the start of this section, integrating out the high-frequency, short-distance degrees of freedom in atomic force fields results in soft forces between the mesoscopic particles that can be integrated using a larger integration step-size. This has tempted some groups to use step sizes that may, actually, be too large so that the trajectories followed by the DPD particles deviate from those expected for a system at equilibrium in the canonical ensemble. Specifically, particles of different types (lipid headgroups and tail particles) may experience different temperatures in their interactions with each other. This problem has been highlighted by Jakobsen et al. (69) and Allen (70), who warn that it may not be obvious from the simulation results that the system is not in equilibrium because the DPD thermostat will damp out the effects. Allen (70) recommends checking for this type of error by comparing the configurational temperatures for each bead type in the simulation. If they are all the same (to within statistical errors), this indicates that the system is most likely in equilibrium, but if they

differ, the system has left equilibrium and the integration time-step may be too large. As a rule of thumb, integration step-sizes around 0.01 are considered by the author to be suitable for lipid membrane and vesicle simulations.

At still larger length scales, membranes are treated as continuous elastic sheets, and studied using Monte Carlo simulations (33), or triangulated surface models (29). Eventually, at macroscopic length scales, all molecular details are lost, and only the conservation laws of mass, momentum, and energy are retained in continuum fluid theories including the Navier–Stokes equations. But at these scales, it is difficult to incorporate topology-changing processes, which are crucial at the cellular scale, thereby precluding the study of vesicle fusion, membrane budding, viral and bacterial infection of a cell, etc.

In the remainder of this section, we show how mesoscopic simulation techniques have been used to understand some of the properties of vesicles. We describe their use in studying vesicle self-assembly and equilibrium structural properties, and the dynamic phenomena of domain formation in two-component vesicles and nanoparticles interacting with membranes.

Early simulation studies of the self-assembly of amphiphilic molecules used Monte Carlo simulations (71), or Molecular Dynamics simulations of highly simplified lipid models (72, 73). Simplified models were necessary because the available computing power was limited, and simulations that attempted to keep molecular details could not reach the length and timescales needed for self-assembly. During the 1990s computer power increased fast enough to allow more detailed coarse-grained models to be developed, and even atomistic MD simulations began to push up the number of molecules that could be represented. Around the turn of the millennium, many groups were performing simulations of lipid molecules self-assembling into micelles (74) or planar bilayers (47, 59, 60, 75), and exploring the phase behavior of lipid membranes containing cholesterol (76). Planar membranes are not common in biology, whereas vesicles are ubiquitous in the cell. But as the size of vesicles ranges from 40 nm for synaptic vesicles in neurons to hundreds of nanometers for endosomes and lysosomes, simulating a vesicle requires more computational resources. However, in defiance of the historical development, we first present results on the self-assembly of vesicles before turning to the equilibrium properties of planar lipid bilayers.

Vesicle self-assembly from a random distribution of amphiphiles is very slow. Lipid vesicles with diameters around 10–20 nm have been observed to form in coarse-grained MD (77) and atomistic MD (78). The numbers of lipids in these systems were 1,500–2,500 and 1,017, respectively, which are well below the number expected (~15,000) in synaptic vesicles of diameter 40 nm. Coarse-grained MD simulations of single vesicles up to 60 nm in diameter have been performed using novel boundary conditions that replace water

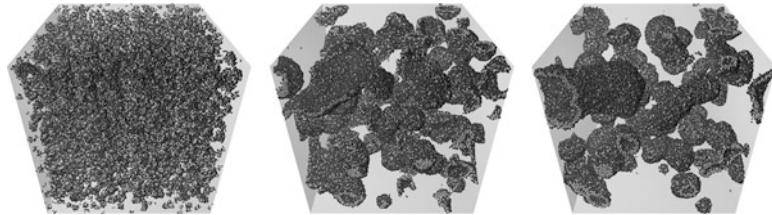


Fig. 4. Sequence of snapshots taken from a DPD simulation (*left to right*) showing the self-assembly of vesicles from an initially random distribution of lipid and solvent molecules. There are 120,000 lipid molecules and more than five million solvent particles in a volume (90 nm)<sup>3</sup>. The self-assembly process occurs in several stages illustrated here. The elapsed time corresponding to each snapshot is given in brackets. First, the amphiphilic lipids rapidly form small micelles (320 ns) so as to hide their hydrophobic tails from the solvent. These micelles aggregate and form undulating, quasi-planar lipid bilayer patches that amalgamate and grow larger, curling up as they do so (10  $\mu$ s). Eventually, they seal off forming closed quasi-spherical vesicles (30  $\mu$ s). The final state shown is probably not the equilibrium distribution, because the time required for the vesicles to encounter each other and fuse increases with their size (snapshots produced using PovRay [www.povray.org](http://www.povray.org)).

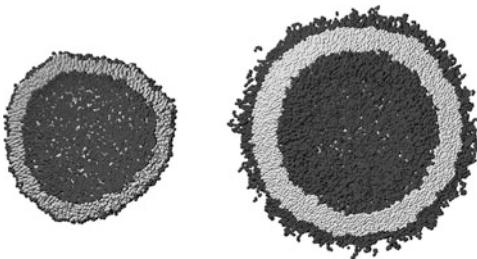
particles distant from the vesicle membrane (both inside and outside) by a mean field force boundary potential (79). While this method is very efficient for studying the equilibrium properties of a single vesicle, it is unlikely to be able to simulate systems hundreds of nanometers in size. Various coarse-grained simulation techniques have been used to go beyond these sizes. Monte Carlo simulations have been used to study the formation of a single vesicle (80, 81), as have Brownian Dynamics (51). Self-consistent field theory (82, 83) is able to follow the evolution of many vesicles simultaneously, and to extract the elastic properties of a model phospholipid membrane (84). We do not discuss these models further here because we are interested in simulations in which the lipid molecular motion follows Newton's laws as closely as possible. DPD can capture dynamics over larger length and timescales than MD, while still retaining near-molecular details. Early DPD simulations of vesicle formation followed the assembly of vesicles out of 1,008 single-tailed amphiphiles in a simulation box around 20 nm along each side (85). The increase in computing power since that time allows much larger systems to be simulated, as we now show.

The self-assembly of multiple vesicles on a length scale of 90 nm from a parallel DPD simulation of the author is illustrated in Fig. 4. The simulation begins with 120,000 amphiphiles randomly dispersed in more than five million water particles, corresponding to a concentration of 165 mM, and represents a real time of approximately 30  $\mu$ s. This simulation consumed 200 cpu-hours on a cluster of 64 processors using a parallel DPD code. The amphiphiles are constructed out of hydrophilic H particles bonded to hydrophobic T particles. Water is represented by a single bead W. Following Yamamoto et al. (85), the amphiphiles have the molecular shape, H<sub>2</sub>T<sub>4</sub>, that represents a generic single-tailed surfactant.

The process of vesicle formation proceeds through a series of stages, each having its own timescale. The stages reflect the distinct processes of diffusion and aggregation adopted by single amphiphiles, small micelles, larger membrane patches, and, finally, vesicles. First, the amphiphiles rapidly form many small micelles that diffuse more slowly than the individual molecules (Fig. 4a). This stage is short lasting only 500 ns. The micelles aggregate into larger ones that form quasi-planar bilayer patches in a few microseconds (Fig 4b), and these in turn slowly curl up and seal into vesicles in a few tens of microseconds (Fig 4c). This stage is not yet the equilibrium state of the system because the vesicles themselves can encounter and fuse with each other. But at this stage a new timescale enters: the mean time for two vesicles to fuse. The durations of the previous stages of the self-assembly process are governed by the time required for objects of radius  $R$  to diffuse over their mean separation. The Einstein relation for the diffusion constant of an object of radius  $R$  is  $D \sim 1/R$ . As the mean separation between the objects scales linearly with  $R$ , the time required for diffusing objects to encounter each other (according to  $\langle X^2 \rangle = 6.D.T$ ) scales as  $T \sim R^3$ . But the mean fusion time for vesicles can be much longer than this. Grafmüller et al. have found (86) that tense vesicles with diameters around 30 nm fuse over times ranging from hundreds of nanoseconds to many microseconds.

It is interesting to note at this point that self-assembly of vesicles composed of nonbiological amphiphiles has recently been simulated using coarse-grained MD (48), and the stages of vesicle formation are very similar in both types of simulation. This indicates that vesicle formation is a generic property of amphiphiles, and does not depend on their detailed chemistry.

The appearance of intermediate structures with their associated timescale during vesicle formation on large length scales reveals a cloud on the horizon of coarse-grained simulations. It is often assumed that as computing resources increase in power and drop in price, more problems will become amenable to the brute-force approach of Molecular Dynamics and other particle-based simulation techniques. But as longer length scales come within range of coarse-grained simulations, new physical phenomena appear that have their own, still longer, timescales. Capturing these new phenomena with simulation schemes that retain near-atomic detail is impossible. Instead, hybrid simulations are needed that combine distinct coarse-graining techniques that are coupled to each other (38). Steps in this direction are currently underway (41), and represent an aggressive attempt to span length and timescales of hundreds of nanometers and hundreds of nanoseconds. Future developments will need to couple several coarse-grained techniques together, and such hybrid schemes may finally be able to span scales from the molecular to the macroscopic. For now, we return to the properties of amphiphilic aggregates studied within a single coarse-grained model.



**Fig. 5.** (a) Simulation snapshot of a lipid vesicle with outer diameter 28 nm containing 6,594 lipid molecules representing DMPC. The vesicle surface is relatively smooth on a molecular scale as the lipids are constrained to the membrane plane by the high energy cost of exposing their hydrocarbon tails to the solvent. The membrane core is a relatively unstructured hydrocarbon block, and the lipids can diffuse fairly easily in the plane of the membrane. Note that water particles are present inside and outside the vesicle, but only those inside are visible here for clarity. (b) Simulation snapshot of a polymersome (or polymer vesicle) with outer diameter 40 nm containing 5,444 molecules of the diblock copolymer PEO40-PEE37 that has a mean molecular weight of 3,900 gm/mol. Unlike lipid vesicles, the surface of the polymersome is rough as the longer hydrophilic blocks (here PEO) protrude out into the aqueous solvent. The higher molecular weight of the hydrophobic PEE block (compared to lipid hydrocarbon tails) leads to greater entanglement of the polymers in the membrane and a lower in-plane diffusion constant than for a lipid vesicle. Note that water particles are present inside and outside the polymersome but are invisible for clarity, and that the two images are not to the same scale (snapshots produced using PovRay [www.povray.org](http://www.povray.org)).

The most common means of studying the material properties of a lipid bilayer or a vesicle in particle-based simulations is to construct it in the initial state and allow it to equilibrate before measuring observable properties. This is much faster than waiting for it to self-assemble, and the initial state of the aggregate can be more precisely specified. For example, creating a vesicle with zero spontaneous curvature in equilibrium requires a certain amount of care. The correct number of amphiphiles must be placed in the inner and outer leaflets or the vesicle will possess a spontaneous curvature. The area per molecule must be chosen appropriately for the type of lipid or the vesicle will be either deflated or under tension. Even when such states are desired, one should know which one obtains in the initial simulation state. Figure 5 shows an equilibrated lipid vesicle (Fig. 5a) and a polymersome (Fig. 5b) at different length scales in different simulations. Although both vesicles are superficially similar, they have quite different material properties.

Nonbiological vesicles can be constructed out of amphiphilic polymers. Block copolymers contain hydrophilic segments covalently linked to hydrophobic segments, and have been known for at least a decade (87, 88) to self-assemble into vesicles (often referred to as polymersomes to bring out their similarity to liposomes). Because the molecular weight of diblock copolymers can vary more widely than that of phospholipids, the material

properties of polymersomes can be modified more than those of phospholipid vesicles. This makes them promising vehicles for applications that require them to withstand stresses beyond those sustainable by lipid vesicles, such as drug delivery. DPD simulations of a planar patch of diblock copolymer and a complete polymersome have shown that their structural and elastic properties agree well with experimental results (89). Although not as accurate as MD simulations of similar systems (90), DPD is able to reach much larger length scales at a smaller computational cost, and provides a useful first pass at modeling novel complex materials such as polymersomes.

Simulations of bilayers are usually performed in the NVT ensemble. The bilayer is connected to itself across the fixed periodic boundaries of the simulation box, and the number of particles ( $N$ ), volume ( $V$ ), and temperature ( $T$ ) are constant. The constraint that the bilayer can only take up configurations that are periodic in the simulation box side lengths reduces its entropy below that of the same sized patch of bilayer in a larger membrane, and creates a lateral tension in the bilayer. The number of molecules in the bilayer can be adjusted so that this tension is close to zero, and this simulation setup is taken as representative of a much larger membrane. Simulations performed in the NPT ensemble, or more precisely the  $N\sigma T$  ensemble, where  $\sigma$  is the surface tension of the membrane, typically allow the simulation box area to fluctuate as the membrane undulates, while keeping the surface tension of the membrane constant. While the NVT ensemble is computationally easier, the  $N\sigma T$  ensemble is preferred for simulating tensionless (or, more generally, constant tension) membranes.

The molecular structure of lipids is known to modify the material properties of membranes and vesicles. Aspiration experiments measure the bending stiffness and area stretch modulus of fluid-phase lipid vesicles (22, 23). In these experiments, a vesicle is partially aspirated in a micropipette and the length of its tubular extension into the pipette measured. Owing to thermal fluctuations, there is more area in the vesicle's membrane than a measurement of its radius suggests, and as these fluctuations are removed by increased aspiration the stiffness of the membrane varies and can be measured. Once the fluctuations are removed, the direct expansion modulus of the membrane can be quantified. The bending stiffness of the membrane is found to increase with lipid hydrocarbon tail length, but decrease with the degree of unsaturation. It also increases with increasing concentration of cholesterol in the membrane (91). These results have been qualitatively reproduced in DPD simulations of planar lipid bilayers (92), particularly the weakening effect of unequal hydrocarbon tail lengths on the rigidity of a lipid bilayer. Several groups have studied the material properties of planar lipid bilayers using a variety of coarse-grained simulation techniques including MD (47), solvent-free MD (53–55), and DPD (59, 60, 69, 92),

and the results are quite consistent. Having calibrated the structural properties of lipid bilayers within coarse-grained simulations, more complicated physicochemical properties became of interest; among these is the distribution of pressure within the bilayer.

The pressure in the bulk phase of a single-component fluid is an isotropic quantity that can be represented by a single number. But at an interface between two fluids, or in an inhomogeneous fluid phase, the pressure is a symmetric, second-rank tensor whose components depend on the symmetry of the interface, the coordinate system used, and the interactions between the atoms or particles making up the fluids (93). The diagonal elements of the pressure tensor represent the compressive forces acting at each point in the fluid, while the off-diagonal elements represent shear forces, and are zero for fluids that cannot support a shear. The pressure tensor for a translationally invariant, planar interface between two isotropic fluids is diagonal and has three components, but only two of these are independent: the normal pressure,  $P_N$ , and the lateral pressure,  $P_L$ . The two components in the plane of the interface are equal because of translational symmetry. The normal pressure is constant across the interface, because if it were not the interface would move until it became constant. But the lateral pressure must vary across the interface; otherwise the interface would break up and disappear. The pressure profile across an interface is defined as the difference between the two independent components of the pressure tensor ( $P_N - P_L$ ), averaged over planar slabs parallel to the interface, and is a function of the coordinate through the interface. The integral of the pressure profile from a point far in the bulk phase on one side of the interface to another point far in the bulk phase on the opposite side yields the interfacial surface tension.

The calculation of the lateral pressure profile across a phospholipid bilayer is particularly interesting from a simulation point of view, because direct measurements of it are not currently possible in experiments without disturbing the profile (94). It has been suggested that proteins may be sensitive to the distribution of pressure in a membrane (95), and that this may be a key factor in general anesthesia (96). A direct calculation of the pressure profile across a bilayer or a vesicle would be an important tool for predicting the effects of membrane state on the properties of embedded proteins. Performing this calculation in a simulation is complicated (97), but mean field calculations have been performed and the effect of the profile on the aggregation of membrane-bound peptides predicted (98). Theoretically, the pressure profile across an interface is not uniquely defined as it depends on an arbitrarily chosen integration contour. Different contours lead to different values for some properties that depend on the profile, although the surface tension of a planar interface is independent of the chosen contour. A full discussion of the calculation of the pressure profile is given by

Schofield and Henderson (93). The calculation is straightforward to implement for planar lipid bilayers, and gives at least qualitatively similar results when calculated in MD simulations (47, 97, 99, 100) and DPD simulations (59, 60, 69).

The spherical symmetry of a vesicle makes it more difficult to calculate its stress profile than for a planar bilayer, but it has recently been calculated by Ollila et al. (101) using an extended form of the Schofield and Henderson technique (93). We have implemented this calculation in a DPD simulation code, and used it to measure the pressure profile for a lipid vesicle. In Fig. 6, we show the pressure profiles for a planar lipid bilayer and a vesicle together with a snapshot of each simulated system. Whereas the pressure profile for the planar bilayer is symmetric, reflecting the symmetry of the membrane, that for the vesicle is clearly not symmetric, because the vesicle membrane is curved.

Unlike planar membranes, which are translationally invariant, the interior of a vesicle membrane is not in a uniform state. The inner monolayer is compressed near the lipid headgroup region, while the lipid tails can expand slightly, whereas the outer monolayer is expanded in the headgroup region and the tails are compressed. The different chemical nature of the parts of lipid molecules, and their arrangement in the two monolayers, changes their state from the same lipids arranged in a flat bilayer patch. The nonuniform compression of the lipids in a vesicle has a particularly strong effect on its lateral pressure profile. The positive peaks in the pressure profiles of the planar membrane and vesicle reflect the repulsion of the hydrophobic lipid tails from the surrounding aqueous solvent. But the lipids in the inner leaflet of the vesicle are pushed slightly closer together and so shield their tails slightly more than lipids in the outer leaflet. Hence, the positive peak for the inner leaflet of the vesicle is lower than that for the outer leaflet. This asymmetry depends on the curvature, and therefore size, of the vesicle, and may be of biological importance (101). Because simulations of vesicles larger than 30 nm in diameter are only now becoming routine, further work is needed to see if the pressure profiles of vesicles of different size can be used to make quantitative predictions useful in experiments.

Biological membranes usually contain many different types of lipid, sometimes containing hundreds of species with distinct headgroups, tail lengths, number of double bonds, charge, etc. Simulating such huge multicomponent systems is almost impossible because accurate values are needed for all possible pairs of interacting particles. A simulation of just 10 lipid types, assuming that only their headgroups are distinct, requires the specification of 45 interaction parameters ( $^{10}C_2$  possible pairs). For simplicity, most membrane simulations use a single lipid species as, frequently, do experiments on model systems such as giant vesicles. Although Monte Carlo simulations of domain formation in planar lipid

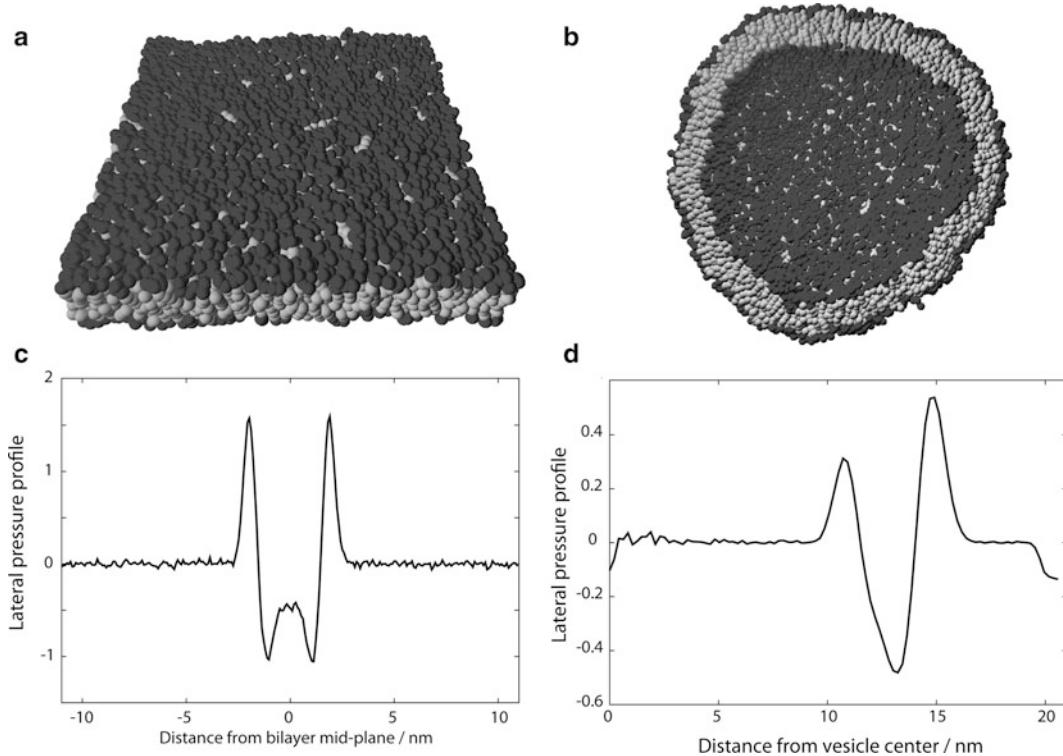


Fig. 6. (a) Oblique view of a  $22 \times 22$  nm planar lipid bilayer in a DPD simulation that contains 1,637 lipids. The lipid architecture is  $H_3(T_4)_2$  and represents a DMPC lipid. The remainder of the simulation box is filled with water particles that are here (and in b) invisible for clarity. (b) Cross section through a 28 nm diameter lipid vesicle in a DPD simulation that contains 6,594 lipids of the same type as in (a). (c) Lateral pressure profile averaged over thin, planar slices through the lipid bilayer shown in (a). The curve shows the difference between the normal and lateral components of the pressure tensor as a function of the distance from the bilayer mid-plane. In the bulk water phase, the pressure is isotropic, so the profile vanishes. The profile across the bilayer is symmetric reflecting its symmetry, but shows a strong variation with depth in the membrane that arises from the different bonded and nonbonded interactions between the beads in the lipid molecules at different depths in the membrane. Error bars are not shown because the pressure profile fluctuates greatly, but the structure of peaks and troughs seen in the figure is quite reproducible between runs. (d) Lateral pressure profile averaged over spherical shells as a function of the radial coordinate from the center of the vesicle shown in (b). In contrast to the planar bilayer profile in (c), a vesicle has an asymmetric pressure profile that results from the different packing of the lipids in the inner and outer leaflets. The pressure profile vanishes in the bulk water phases inside and outside the vesicle. Note that the small dips in the profile at the center of the vesicle and at the boundary of the simulation box are far from the vesicle membrane, and do not affect the calculation of the vesicle's pressure profile. They arise from the small number of particles in the center shell and from not counting bead-bead interactions that cross the periodic boundaries, respectively. As the vesicle is distant from the simulation box boundaries, we do not count these interactions to save processor time (snapshots produced using PovRay [www.povray.org](http://www.povray.org)).

membranes have been performed for some time (102), it is only recently that multicomponent vesicles have been simulated using particle-based methods.

Experiments have clearly shown that different phospholipids do not mix ideally if they have different chain-melting temperatures (103–105). This temperature depends on the lipids' hydrophobic tail lengths and unsaturation, and these properties in turn influence

their area per molecule and the packing of the lipid tails in the hydrophobic core of the bilayer. Depending on the temperature and composition, a mixture of lipids may undergo phase separation in a mixed membrane leading to domain formation. If the domains are large enough, and their bending stiffness is not too large, they may buckle out of the membrane plane and form buds that subsequently detach as daughter vesicles. Several groups have studied domain formation in two-component lipid vesicles using DPD and MD, measuring the change in membrane thickness and elastic properties (106) of the domains, and their budding and fission (107), or the breakup of a two-component vesicle into compositionally distinct daughter vesicles (50). A systematic DPD study of the dynamics of domain formation was carried out by Laradji and Kumar (108), who found that the hydrodynamic effects of the surrounding solvent, and the volume constraint of the vesicle, affected the coarsening dynamics at all timescales. Such processes are beyond the reach of atomistic MD simulations, so they are ideal candidates for study using explicit-solvent coarse-grained simulations.

Another field of study that is only now coming within the reach of particle-based simulations is the interaction of proteins or nanoparticles with lipid membranes. Transmembrane proteins have a membrane-spanning section of hydrophobic residues whose length usually matches that of the membrane. If there is a size mismatch between the protein's hydrophobic region and the membrane's hydrophobic core, the energy cost of exposing hydrophobic moieties to the aqueous solvent forces a rearrangement so as to reduce their exposure. Recent coarse-grained simulations of (highly simplified models of) proteins tilting in a lipid bilayer (109) or clustering (110) to reduce their hydrophobic exposure to the solvent have shown that such simulations can make measurable predictions that may guide future experiments.

The rapid increase in production and use of artificial nanoparticles has led to concerns about their potential toxicity (10). Particles whose sizes are around one nanometer can be simulated using atomistic MD, and this has been performed for fullerenes ( $C_{60}$  molecules) inside a small patch of lipid bilayer (111). But many types of nanoparticles have sizes in excess of tens of nanometers, and this is far beyond what can be simulated using atomistic Molecular Dynamics. Coarse-grained simulations are able to capture the lower end of this range, and are being used to explore the interactions between nanoparticles and model membranes. A recent review (112) summarizes both atomistic and coarse-grained simulations of carbon nanoparticles interacting with lipid membranes.

The first stage in this exploration is to obtain a force field that accurately represents the nanoparticle–nanoparticle interaction compared to nanoparticle–solvent and nanoparticle–lipid interactions. The first steps have only recently been taken in this direction, and even simple questions remain open. For example: Do fullerenes

aggregate inside the hydrophobic core of a lipid membrane or not? Early work using atomistic (111) and coarse-grained (113) MD simulations suggests that they do not aggregate even at high concentrations of one fullerene to 9 lipids. However, another group using a different force-field parametrization predicts that they do indeed aggregate when immersed in a hydrocarbon melt (114). The difference between these predictions may lie in the more ordered environment of the lipid bilayer core compared to the isotropic melt, or it may be due to differences in the force-field parametrization (114).

On a more generic level, the ability of nanoparticles to adsorb or translocate across a lipid membrane has been studied using DPD simulations for spherical (115) and rodlike (116) particles. The results show that the shape of a nanoparticle strongly influences its penetrative ability. This supports recent recognition (14) that the physical properties of nanoparticles (size, shape, and mechanical response) as well as their chemistry must be carefully incorporated into the rational design of nanoparticle-based drug delivery vehicles. Additionally, proteins are known to bend or shape membranes during processes such as endocytosis and vesicle fusion (37, 117), and attempts are being made to include model proteins in coarse-grained simulations of fusion (61, 118). Invagination of a lipid bilayer driven by adsorbed rigid particles has been simulated using a solvent-free MD scheme (119), and represents one of the largest particle-based simulations of a lipid bilayer so far published. A total of 36 curved cap-like particles were uniformly distributed on the surface of  $160 \times 160$  nm planar bilayer patch. Each cap curves the membrane locally to maximize its binding area, and these local deformations induce an attractive interaction between the caps. As the caps come closer together (without, however, aggregating into a single domain) the membrane deforms into a large funnel-shaped protrusion that looks remarkably like a budding vesicle. Such a membrane-mediated, generic interaction may be important for protein clustering in biological membranes (110).

In the next section, we describe simulation studies of vesicle fusion to illustrate the power and limitations of coarse-grained particle-based simulations.

---

### 3. Vesicle Fusion

Unperturbed vesicles are very resistant to forming pores. Fluid-phase lipids rapidly diffuse in the plane of the membrane and are driven to reseal a pore by the hydrophobic effect that leads the vesicle to form in the first place. If a pore appears in the membrane, the free energy cost of exposing the hydrocarbon tails of the lipids on the pore perimeter to the aqueous solvent creates a line tension that quickly

shrinks the pore until it reseals. In order to fuse, the lamellar structure of two apposed membranes must be destabilized so that the lipids can rearrange into a channel, or fusion pore, connecting them before the pore is resealed by the hydrophobic effect. Various ways of achieving this experimentally include the following: osmotically stressing vesicles (120); dehydrating two apposed membranes by adding PEG polymers (121, 122) or multivalent cations (123, 124) to the external solution; adding oppositely charged amphiphiles to a mixture of giant vesicles (125); applying electric fields to the membranes (126); or reconstituting soluble *N*-ethyl-maleimide-sensitive factor attachment protein receptor (SNARE) complexes in cells (127). The formation of multiple hourglass-shaped connections between lamellae in a multi-lamellar stack has also been observed by the relatively simple process of dehydrating the stack (128). All of these methods result in one or more pores connecting the distal regions of the membranes.

In spite of their impressive stability, biological lipid membranes must be able to fuse under specific circumstances because fusion is a crucial process for cells and viruses. For example, synaptic vesicles fuse with the presynaptic membrane in neurons and release their neurotransmitter contents when an action potential arrives; sperm fuse with eggs; and  $\beta$  cells in the pancreas release insulin by exocytosis. These processes are tightly regulated to ensure the cell's viability (129). Viral fusion is the subject of intense experimental interest because of its role in disease. Viruses are obliged to enter their host cell at some point in their life cycle, and they do this either by fusing directly with the cell's Plasma Membrane or by hijacking the cell's own endocytic pathways (130). Recent experiments (131) have shown that rigid amphiphilic peptides can inhibit viral fusion when they embed into fusing membranes because of their effect on the membrane curvature. Experiments such as these are important target systems for the simulated fusion protocols we describe here.

The common nanoscale features of vesicle fusion are that two initially separate phospholipid bilayers are brought to within a few nanometers of each other, and perturbed so as to cause them to merge at precise locations (132). Yet there are differences in how this is achieved. Influenza virus requires on average 9 hemagglutinin (HA) trimers for successful insertion of its DNA into a target cell whereas HIV appears to be able to infect a cell with just a single viral envelope glycoprotein (133). Synaptic vesicle fusion requires at least several SNARE complexes to drive full fusion (4, 134). In spite of great progress in identifying the proteins involved in SNARE-mediated fusion (135, 136), the precise sequence of protein-driven molecular movements that transforms two separated membranes into two merged membranes containing an expanded fusion pore is still unknown. Questions such as whether the initial fusion pore (137) is lined only with lipids or lipids and

proteins are actively debated (138–140), as is the precise role of the various proteins involved (141). Whereas SNARE complexes appear to be the minimal fusion machinery required in synaptic vesicle fusion (4), experiments on sea-urchin cortical exocytosis indicate that SNAREs are not sufficient but require additional proteins acting downstream of the SNARE complex to complete fusion (142).

The complexity of experiments involving protein-driven fusion has led experimentalists to explore protein-free fusion assays in order to understand the lipid dynamics that occurs. But the fastest video cameras currently available are unable to resolve the initial fusion pore formation as it appears to take place in less than 100  $\mu$ s. Also, synaptic vesicle fusion occurs for vesicles with diameters around 60–80 nm, whereas exocytosis, an example of which takes place in  $\beta$  cells of the pancreas, has vesicles (or granules) with diameters 245–280 nm. The speed of pore formation and small vesicle size means that direct experimental evidence of lipid rearrangements during fusion pore formation is almost impossible to obtain. Computer simulations can in principle make unbiased predictions of lipid molecular conformations during any stage of the fusion process, and reveal experimentally invisible motion. Particle-based computer simulations are now powerful enough to test hypotheses made by theoretical fusion models. They can be used to explore the influence of lipid (or peptide) molecular structure and membrane composition on the growth of a fusion pore. And because no assumptions are made about the pathway taken by the membranes during fusion, simulations can reveal dynamical details such as unexpected energy barriers or unusual molecular conformations that are not foreseen in theoretical models.

Vesicle fusion takes place through a series of three generic steps illustrated by the cartoons in Fig. 7. First, the membranes approach each other (Fig. 7a) and are held in a closely apposed state (Fig. 7b), until, by some means, a pore is formed in the merged membranes (Fig. 7c). The transition from the first stage to the second requires that water between the apposed membranes be removed, and is

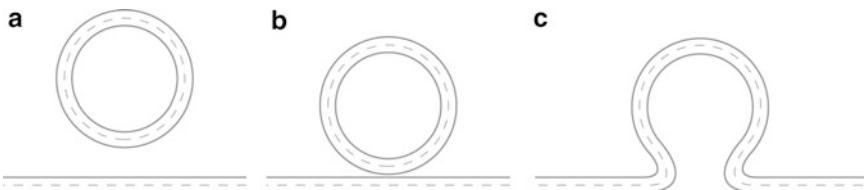


Fig. 7. Sequence of cartoons showing a vesicle approaching and fusing to a planar membrane patch. The complex process of fusion is here divided into three generic stages. (a) The vesicle is brought close to the target membrane. (b) The vesicle is subsequently held close to the proximal side of the target membrane in a “docked” state until it is signaled to fuse. (c) Once a fusion pore has formed, it expands to release the vesicle contents to the distal side of the target membrane. The final transition between stages (b and c) has been the target of many coarse-grained and atomistic simulations in the last decade (picture courtesy of Thomas E. Rasmussen).

opposed by strong hydration forces between the membranes. The transition to the third stage requires that a hole be produced in the merged membranes in spite of the tendency of lipids to rapidly reseal defects in the membrane. Finally, the pore expands to allow the distal spaces beyond the membranes to connect.

Once two membranes have been brought close together, and restrained from moving apart, what precisely happens as they fuse? This question has been addressed theoretically and in a variety of simulation protocols. The commonly accepted theoretical mechanism of protein-free lipid bilayer fusion (and viral fusion) is known as the “stalk hypothesis.” First postulated more than 20 years ago (143, 144), it has since undergone several reformulations (145–147). The stalk hypothesis proposes that two separated membranes pass through specific structures on their way to fuse. Once placed in close contact, their proximal monolayers merge at a point and form an hourglass shape, known as a stalk, while their distal monolayers are still discrete. This stage is referred to as hemifusion. Subsequently, the stalk enlarges in some fashion, and the distal monolayers touch and form a contact zone, or hemifusion diaphragm, consisting of a single bilayer within which, eventually, a pore forms and expands to create a direct connection between the spaces outside the two bilayers. The subsequent dynamics of an expanding fusion pore have been studied theoretically using continuum fluid mechanical methods (148).

The stalk model postulates that the major energy barrier that must be overcome by fusing membranes is the formation of the stalk conformation in the hemifused state, and it uses continuum elasticity theory to calculate the height of this barrier. The stalk is assumed to have a specific shape, for example, a semi-toroidal one (147), and the energy of the membranes in this state is calculated. Unfortunately, in the original model (143) this barrier was calculated to be hundreds of  $k_B T$  in size, far above the work that could reasonably be performed by proteins during biological fusion. Various modifications to the model were proposed to reduce this barrier. In particular, assuming that the lipid molecules tilt as the proximal leaflets merge reduces the barrier height to a few tens of  $k_B T$ , which is biologically feasible (146). An excellent summary of the fusion-through-hemifusion hypothesis has been given by Chernomordik and Kozlov (149).

The stalk hypothesis has also lately been augmented with the addition of a pre-stalk stage that addresses the question of how the stalk is brought into being given that strong hydration forces should prevent the two membranes from forming an extended, dehydrated contact zone. Efrat et al. (150) postulate that a point-like protrusion (PLP) forms between two closely apposed bilayers as a precursor to the stalk. The narrowness of the PLP lowers its energy below that of an extended contact zone because the strong repulsive hydration forces between the membranes are limited to a

region the size of a single lipid, or about  $1\text{ nm}^2$ . One caveat that should be kept in mind here is that membranes are quite rough on the length scale of single lipids due to lipid protrusions (151). Continuum elastic theory provides a means of calculating the energies of deformations that are imposed a priori on the membranes, but other mechanisms for forming membrane–membrane contacts may exist, as we discuss later when we describe recent DPD and MD simulations of protein-free lipid bilayer fusion.

Key features of the stalk hypothesis have experimental support. The hemifusion state in which the proximal leaflets have merged while the distal leaflets are still separate has been observed in the fusion of protein-free lipid bilayers (120). The authors note that the pore conductance in protein-free bilayer fusion is similar to that in exocytosis, suggesting that the initial fusion pore may be purely lipidic. If the protein machinery for fusion evolved to take advantage of an existing lipid mechanism of fusion pore formation, the role of proteins would be to regulate and create the pre-fusion state of the fusing bilayers. However, inferring the structure of the initial fusion pore from conductance measurements is complicated by the fact that even intact lipid bilayers can show ion channel-like conductance traces, as has been known for many years (152), and recently discussed in a review (153).

Experiments have shown that dehydrated phospholipid membranes exhibit a “stalk phase” (128), in which many hourglass-shaped connections between two membranes are stable. This shows that the assumptions of the stalk model are plausible, but it remains to be seen if its central assumption that the rate-limiting energy barrier is due to the hemifusion stalk is indeed correct.

Simulations of protein-free lipid bilayer fusion have tended to use the membrane tension or an external force to perturb the membranes as these are the simplest means of causing a fusion pore to appear in two nearby membranes. The simulations typically start by placing two vesicles near each other in the simulation box, with a small solvent-filled gap between them. After an equilibration period, an external force is applied to one or other of the vesicles to bring them into contact at the point of closest approach. The subsequent evolution of the merged membranes is then observed. Both Molecular Dynamics simulations (154–158), Brownian Dynamics (159) and DPD (86, 160, 161), have been used to explore the fusion of lipid bilayer membranes and vesicles. The computational demands of MD limit the method to vesicles with diameters around 15 nm. This is well below the diameter of synaptic vesicles. In order to go beyond this size, a more highly coarse-grained technique must be used.

The fusion of non-lipidic membranes, composed of amphiphilic copolymers, has also been explored using various simulation techniques. DPD has been used to study the fusion of vesicles composed of amphiphilic comb polymers (162) and triblock copolymers (163).

The fusion of diblock copolymer membranes has been studied by Monte Carlo simulations (164). This work revealed that a stalk in two apposed membranes modifies the nominal line tension around a pore that forms in its neighborhood (165). Because pore formation in a membrane must occur at some point along the fusion pathway, a perturbation that lowers the line tension can greatly increase the probability of pore formation. Self-consistent field theory has also been used to follow the spontaneous formation and fusion of amphiphilic polymeric vesicles (82).

We focus here on explicit-solvent, particle-based simulations of membrane fusion because they give information on the conformation and motion of molecules, and retain hydrodynamic effects of the solvent, such as viscosity, that may be important for cooperative effects that take place as a fusion pore is created. We do not discuss Monte Carlo or Brownian Dynamics simulations further, but refer the reader to several recent review articles (46, 61, 64).

In an effort to simulate the fusion of more realistic vesicle sizes than is possible in MD, DPD has been used to study the fusion of a 28 nm diameter vesicle to a 50 nm square planar membrane patch (160). This protocol represents the fusion of a vesicle to a much larger membrane such as the presynaptic membrane in neurons. Although two planar membranes might appear to be a simpler choice, the vesicle-planar membrane geometry has the advantage that the solvent can be displaced from between the membranes as they approach each other. Simulations typically use periodic boundary conditions to remove the influence of the finite simulation box, and two planar membranes in such a simulation are separated by a constant amount of solvent. If they are to fuse, there is nowhere for the solvent particles to go. This problem can be circumvented by dehydrating the membranes initially (157), and observing their subsequent evolution. But the initial state is then highly perturbed and represents a stage already well along the fusion pathway.

In the DPD fusion protocol of Shillcock and Lipowsky (160), both membranes are initially under tension (for example, osmotically swollen) and placed close to each other separated by a small solvent-filled gap. Thermal fluctuations of the membranes allow them to touch at the point of minimum separation, and the transient removal of the solvent is sufficient to allow a few lipids to cross from the less tense vesicle to the more tense planar membrane. Once this bridgehead is established, more lipids pass between the vesicle and the planar membrane, and the perturbation of this transport on the membranes creates a pore that rapidly removes the tension by expanding.

This systematic study of tension-induced vesicle fusion showed that it was a fast process, the pore growing rapidly in a few hundred nanoseconds after first contact of the membranes, but also rather unreliable. Only about one-half of the attempts succeeded in producing a fusion pore, while the other attempts resulted in a

burst vesicle or a planar membrane, adhered but not fused membranes, or the vesicle and planar membrane simply drifted apart. A subsequent study of the same geometry, but with a modified force field between the lipids, showed that the unreliability was a function of the interaction potential of the lipids (86). A new force field was obtained by improving the fit of the elastic properties of the simulated lipid bilayer to those of experimental membranes containing lipids such as DMPC. This work revealed that the simulated fusion process is sensitive to the interactions of the lipids, in particular the repulsive interaction between the lipid's hydrophilic headgroup and its hydrophobic tails determined the rate of flip-flop of lipids across the solvent-filled gap between the two membranes. Increasing this parameter from its value in the original work (160) stabilized the two adhering bilayers and resulted in longer fusion times. Obviously, chemical properties of lipids are essentially immutable in experiments, but can easily be modified in simulations to explore different hypotheses. The refined force field yielded fusion times ranging from hundreds of nanoseconds to tens of microseconds, and revealed that the pore lifetime depended on the membrane tension.

A key result from the work of Grafmüller et al. (86) is that fusion is initiated by a splayed lipid conformation in which one of the hydrocarbon tails of a single lipid remains in its original leaflet while the other moves across the solvent gap and embeds itself in the proximal leaflet of the apposed membrane. This splayed conformation was subsequently also found to initiate vesicle fusion in atomistic (156) and coarse-grained (158) Molecular Dynamics simulations. In the fluid lamellar phase, a phospholipid has both of its hydrocarbon tails embedded in a single leaflet of a bilayer, and the energy cost for moving one tail into the surrounding solvent is high. But under conditions that arise during fusion, a lipid may splay so that it has one tail in each of the apposed leaflets of the fusing bilayers. Such conformations were predicted almost 20 years ago (166), and observed in experiments later (167, 168). That an unusual lipid conformation occurs in both DPD and MD suggests that it does not depend sensitively on the details of the simulation force fields but only on generic lipid properties, viz, their amphiphilicity and double-tailed molecular shape.

This result illustrates the importance of identifying the “relevant” molecular details in a coarse-grained simulation, and supports the utilitarian goal of devoting computational resources to the simplest model that contains the correct physics. It is common practice in simulations to simplify a system as much as possible, but there is a limit. Whereas single-tailed lipids easily form bilayers in simulations, and are often used in studies of generic vesicle properties, a splayed lipid conformation is obviously only possible for lipids with (at least) two tails. As the molecular details relevant to a problem are not always obvious *a priori*, the choice of such details is still largely an art.

Molecular Dynamics simulations (157) of the fusion of two planar membranes subject to low hydration predict the appearance of splayed lipids that span the inter-bilayer gap, and suggest that the rate-limiting barrier to fusion is a pre-stalk transition state involving splayed lipids. But these simulations were performed for very small membrane separation, so they are already partially dehydrated. Predictions of continuum elastic theory then suggest that the stalk energy for such a state is indeed small. It is known experimentally (128) that the stalk phase forms spontaneously and remains stable for highly dehydrated lamellae. Hence, a crucial experimental question still to be answered is: What is the minimum separation of the fusing membranes, and their hydration state, just before a stalk appears? There is reason to hope that more refined simulations might be able to shed light on this question while it continues to elude experimental determination.

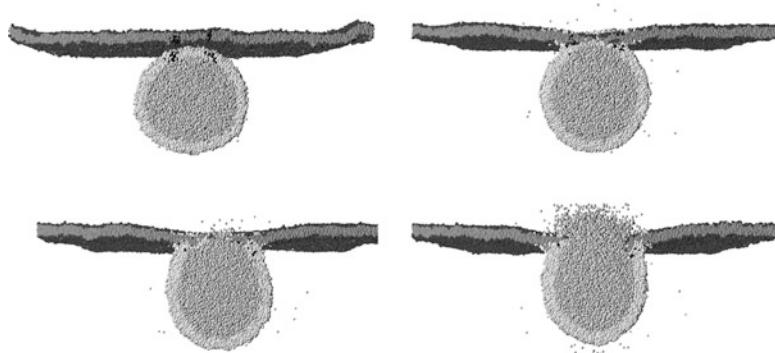
We end this section by describing the first attempts to include in simulations the proteins involved in synaptic vesicle fusion.

Synaptic vesicle fusion in neurons is driven by protein complexes called SNAREs (1). SNARE complexes assemble from proteins that are bound to one or other of the fusing membranes (169). The so-called v-SNARE (synaptobrevin 2 also known as VAMP2) is bound to the vesicle, and t-SNARES (syntaxin and SNAP-25) are bound to the neuronal presynaptic membrane. When a synaptic vesicle is brought close to the presynaptic membrane, these proteins form a SNARE complex that drives the fusion process. SNARE proteins reconstituted into lipid vesicles also promote fusion showing that they are the minimal machinery needed for successful fusion (170). Although there are a large number of SNARE proteins in different organisms, only a dozen or so combinations are fusogenic, indicating that cognate SNAREs impart selectivity to the fusion process. They are not however the only proteins involved in fusion. Others include Sec1/Munc18, calmodulin, synaptotagmin, etc. Some of these are involved in regulating fusion or in responding to  $\text{Ca}^{2+}$  signals (171), but a recent review of the field of SNARE-mediated fusion (141) discusses how Sec1/Munc18, originally thought only to regulate fusion, appears to play a direct role in stimulating fusion. The structures of some of the many types of SNARE proteins are known, but the precise mechanism by which they manipulate their associated membranes so as to drive them through the sequence of pre-stalk state, stalk formation, hemifusion diaphragm, and appearance of the fusion pore is still experimentally obscure because it takes place on length and timescales below the resolution of experimental apparatus. Unfortunately, it is also too large and slow to be fully captured in Molecular Dynamics simulations, although detailed MD simulations of parts of the SNARE complex have been performed (172). This has led some groups to represent only a subset of the SNARE proteins' structure in a

simulation and to impose external forces on these simplified proteins to see whether a series of forces localized in space and time can drive fusion.

At a coarse-grained level, SNARE protein complexes can be viewed as consisting of two parts: a rodlike extracellular part that represents the coiled-coil structure of the complex attached to a transmembrane anchor that binds the complex to its associated membrane. The extracellular, coiled-coil piece is known from experiments to zip up from its distal end to its proximal end, in the process drawing the two apposed membranes closer together. This zipping up process is exergonic, releasing about  $35\text{ kT}$  ( $20\text{ kcal/mol}$ ) per SNAREpin (4). The inverse process of taking apart the used SNARE complex requires energy in the form of ATP consumption, and because it is not part of fusion pore formation we do not discuss it further here. The transmembrane anchors of SNARE proteins require specific physical properties in order for fusion to occur. McNew et al. showed that replacing the transmembrane anchor of SNARE protein by a phospholipid rendered the SNARE complex unable to fuse (173). They concluded that one function of the SNARE complex in promoting fusion is to exert force on its anchor as the extracellular coiled-coil zips up. Attempts have been made to incorporate these facts into highly simplified coarse-grained simulation models (61, 118), and to ascertain what sequence of external forces localized in space and time could be applied by SNARE complexes to their attached membranes in order to promote fusion.

Shillcock and Lipowsky (61) explicitly represented only the transmembrane anchors of a SNARE complex in a DPD simulation of a vesicle fusing to a planar membrane patch, and applied a force protocol consisting of a sequence of bending and stretching forces to them. This protocol is illustrated in Fig. 8. Six transmembrane anchors, or barrels, were embedded in a circular fashion in the membranes of a vesicle and an adjacent planar membrane patch. The region of membrane lying within the circle defines the contact zone. The action of the SNARE proteins on the fusing membranes is simulated using two stages. First, the vesicle and planar membrane are bent towards each other by partially dehydrating the lipids in a circular region in both membranes. This has the effect of expelling the solvent from the gap between the membranes and causing their proximal monolayers to partially merge. This stage is a highly simplified representation of the effects of the external segments of the SNARE complexes “zipping up” from their distal ends, and their dehydrating effect caused by interactions with their associated membranes. When the proximal monolayers of the two membranes have merged, a transient, radial pulling force is applied to the six transmembrane “barrels” embedded in each membrane. This raises the local tension in the contact zone between the barrels and represents the action of SNARE proteins



**Fig. 8.** Sequence of snapshots at times 16, 160, 260, and 320 ns (row-wise from *top left*) from a DPD simulation of a 28 nm diameter vesicle fusing to a  $100 \times 100$  nm planar membrane patch. The vesicle and planar membrane contain 5,887 and 28,000 lipid molecules, respectively, and are under a small tension that is insufficient to cause them to fuse without further perturbation. Only water initially inside the vesicle is shown for clarity. They are initially adjacent with a 2 nm solvent-filled gap between their outer leaflets. Six transmembrane “barrels” are constructed in each bilayer to represent the membrane-bound anchors of the SNARE complexes (the *dark shapes* near the point of closest contact). A specific force protocol is employed to explore the transition between stages (**b** and **c**) of Fig. 7 by hypothesizing actions that SNARE proteins might take as described in the text (snapshots produced using PovRay [www.povray.org](http://www.povray.org)).

in locally perturbing the contacting membranes. The increased tension opens up a pore in the already perturbed contact zone, and the vesicle releases its contents to the exterior through the growing pore.

Note that even though the transmembrane barrels are effectively floating in the fluid membrane, there is a range of force, of the order of 100–200 pico-Newtons, which is less than the measured unbinding force of a SNARE complex (174, 175), for which the barrels transiently drag a patch of the surrounding membrane along with them. This range is limited by physical factors. Too large a force requires unrealistically large energy expenditure by the SNARE complexes, while too small a force is swamped by the thermal motion of the lipids.

This force protocol is only one hypothesis about the actions of the SNARE complexes. A different approach was taken by Wu and Guo (118) who explicitly represent the SNAREs’ extracellular coiled-coils as a rod with a hydrophobic segment embedded in a vesicle membrane and a hydrophilic end protruding into the solvent. Two vesicles containing several such rods each are placed adjacent to each other so that the distal ends of the rods are in close proximity in the solvent. On touching, the rods start to bind together along their length because of attractive forces between the particles composing them. This has the effect of pulling the attached vesicles together, forcing their membranes to merge and, subsequently, to fuse.

These models are of necessity highly simplified, and it is not clear that they contain the correct physics of SNARE proteins in fusing vesicles. Recent small-angle X-ray scattering experiments (19) have revealed that the surface of synaptic vesicles is almost entirely obscured by membrane-anchored proteins, many of which may play important roles in fusion. In the fusion simulations discussed here, no particles possess electrostatic charges, but membrane fusion is strongly influenced by charge, from the calcium ions that trigger the final stage of fusion to the complex interactions between charged proteins and the cellular Plasma Membrane. The assumption that all of this complexity can be reduced to simple forces applied to membrane-bound barrels or discrete extracellular rods is obviously too simplistic. But comparing the predictions of such simulations to experimental data, and observing the detailed molecular motion only visible in the simulations, can expose hidden assumptions in our mental picture of the dynamical world of cellular membranes, and so allow a better understanding to be developed.

---

## Acknowledgements

The author is grateful to Thomas E Rasmussen for drawing the pictures shown in Figures 1, 2, 3, and 7, and to MEMPHYS for financial support. MEMPHYS is supported by the Danish National Research Foundation.

## References

- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1989) Molecular biology of the cell, 2nd edn. Garland Publishing, New York
- Israelachvili J (1992) Intermolecular and surface forces, 2nd edn. Academic, press London
- Jahn R, Grubmuller H (2002) Membrane fusion. *Curr Op Cell Biol* 14:488–495
- Südhof TC, Rothman JE (2009) Membrane fusion: grappling with SNARE and SM proteins. *Science* 323:474–477
- Jones RAL (2002) Soft condensed matter. Oxford University Press, Oxford, UK
- Sprong H, van der Sluijs P, van Meer G (2001) How proteins move lipids and lipids move proteins. *Nat Mol Cell Biol* 2:504–513
- Simons K, Toomre D (2000) Lipid rafts and signal transduction. *Nat Mol Cell Biol* 1:31–39
- Thompson MP, Chien M-P, Ku T-H, Rush AM, Gianneschi NC (2010) Smart lipids for programmable nanomaterials. *Nanoletters* 10:2690–2693
- Hadorn M, Hotz PE (2010) DNA-mediated self-assembly of artificial vesicles. *PLoS One* 5:e9886
- Nel A, Xia T, Mädler L, Li N (2006) Toxic potential of materials at the nanolevel. *Science* 311:622–629
- Jiang W, Kim BYS, Rutka JT, Chan WCW (2008) Nanoparticle-mediated cellular response is size-dependent. *Nat Nanotechnol* 3:145–150
- Zhang S, Li J, Lykotrafitis G, Bao G, Suresh S (2009) Size-dependent endocytosis of nanoparticles. *Adv Mater* 21:419–424

13. Mailänder V, Landfester K (2009) Interaction of nanoparticles with cells. *Biomacromolecules* 10:2379–2400
14. Mitragotri S, Lahann J (2009) Physical approaches to biomaterial design. *Nat Mater* 8:15–23
15. Nel AE, Mädler L, Velegol D, Xia T, Hoek EMV, Somasundaran P, Klaessig F, Castranova V, Thompson M (2009) Understanding biophysicochemical interactions at the nano-bio interface. *Nat Mater* 8:543–557
16. Heimborg T (2009) Membrane biophysics. *Soft Matter* 5:3129–3364
17. Singer SJ, Nicolson GL (1972) The fluid mosaic model of the structure of cell membranes. *Science* 175:720–731
18. Engelman DM (2005) Membranes are more mosaic than fluid. *Nature* 438:578–580
19. Castorph S, Riedel D, Arleth L, Sztucki M, Jahn R, Holt M, Salditt T (2010) Structure parameters of synaptic vesicles quantified by small-angle X-ray scattering. *Biophys J* 98:1200–1208
20. Rosoff M (1996) Vesicles. Surfactant science series. Marcel Dekker, New York
21. Döbereiner H-G (2000) Fluctuating vesicle shapes. In: Luisi PL, Walde P (eds) Giant vesicles. John Wiley and Sons Ltd, New York
22. Evans E, Rawicz W (1990) Entropy-driven tension and bending elasticity in condensed-fluid membranes. *Phys Rev Lett* 64:2094–2097
23. Rawicz W, Olbrich KC, McIntosh T, Needham D, Evans E (2000) Effect of chain length and unsaturation on elasticity of lipid bilayers. *Biophys J* 79:328–339
24. Canham PB (1970) The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *J Theor Biol* 26:61–81
25. Helfrich W (1973) Elastic properties of lipid bilayers: theory and possible experiments. *Z Naturforschung c* 28:693–703
26. Seifert U (1997) Configurations of fluid membranes and vesicles. *Adv Phys* 46:13–137
27. Seifert U, Shillcock J, Nelson P (1996) Role of bilayer tilt difference in equilibrium membrane shapes. *Phys Rev Lett* 77:5237–5240
28. Gompper G, Kroll DM (1997) Network models of fluid, hexatic and polymerized membranes. *J Phys Condens Matter* 9:8795–8834
29. Noguchi H (2009) Membrane simulation models from nanometer to micrometer scale. *J Phys Soc Japan* 78:041007-1–041007-9
30. Jülicher F, Lipowsky R (1993) Domain-induced budding of vesicles. *Phys Rev Lett* 70:2964–2967
31. Seifert U (1993) Curvature-induced lateral phase segregation in two-component vesicles. *Phys Rev Lett* 70:1335–1338
32. Taniguchi T (1996) Shape deformation and phase separation dynamics of two-component vesicles. *Phys Rev Lett* 76:4444–4447
33. Ramakrishnan N, Sunil Kumar PB, Ipsen JH (2010) Monte Carlo simulations of fluid vesicles with in plane orientational ordering. *Phys Rev E* 81:041922-1–041922-9
34. Gruner SM (1989) Stability of lyotropic phases with curved interfaces. *J Phys Chem* 93:7562–7570
35. Jones RAL (2004) Soft Machines Nanotechnology and Life. Oxford University Press, Oxford, UK
36. Arkhipov A, Yin Y, Schulten K (2008) Four-scale description of membrane sculpting by BAR domains. *Biophys J* 95:2806–2821
37. Campelo F, Fabrikant G, McMahon HT, Kozlov MM (2009) Modeling membrane shaping by proteins: focus on EHD2 and N-BAR domains. *FEBS Lett* 584:1830–1839
38. Shillcock JC (2008) Insight or illusion? seeing inside the cell with mesoscopic simulations. *HFSP J* 2:1–6
39. Ayton G, Smolyrev AM, Bardenhagen SG, McMurtry P, Voth GA (2002) Interfacing molecular dynamics and macro-scale simulations for lipid bilayer vesicles. *Biophys J* 83:1026–1038
40. Delgado-Buscalioni R, Kremer K, Prapotnik M (2008) Concurrent triple-scale simulation of molecular liquids. *J Chem Phys* 128:114110-1–114110-9
41. Ayton GS, Voth GA (2009) Hybrid coarse-graining approach for lipid bilayers at large length and time scales. *J Phys Chem B* 113:4413–4424
42. Peter C, Kremer K (2009) Multiscale simulation of soft matter systems—from the atomistic to the coarse-grained level and back. *Soft Matter* 5:4357–4366
43. Murtola T, Bunker A, Vattulainen IV, Deserno M, Karttunen M (2009) Multiscale modeling of emergent materials: biological and soft matter. *Phys Chem Chem Phys* 11:1869–1892
44. Bennun SV, Hoopes MI, Xing C, Faller R (2009) Coarse-grained modeling of lipids. *Chem Phys Lipids* 159:59–66
45. Berendsen HJC (2009) Concluding remarks. *Faraday Discuss* 144:467–481
46. Marrink SJ, de Vries AH, Tieleman DP (2009) Lipids on the move: simulations of

- membrane pores, domains, stalks, and curves. *Biochim Biophys Acta* 1788:149–168
47. Goetz R, Lipowsky R (1998) Computer simulations of bilayer membranes: self-assembly and interfacial tension. *J Chem Phys* 108:7397–7409
  48. Wu R, Deng M, Kong B, Yang X (2009) Coarse-grained molecular dynamics simulation of ammonium surfactant self-assemblies: micelles and vesicles. *J Phys Chem B* 113:15010–15016
  49. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse-grained model for biomolecular simulations. *J Phys Chem B* 111:7812–7824
  50. Markvoort AJ, Smeijers AF, Pieterse K, van Santen RA, Hilbers PAJ (2007) Lipid-based mechanisms for vesicle fission. *J Phys Chem B* 111:5719–5725
  51. Noguchi H, Takasu M (2001) Self-assembly of amphiphiles into vesicles: a Brownian dynamics simulation. *Phys Rev E* 64:041913-1–041913-7
  52. Farago O (2003) “Water-free” computer model for fluid bilayer membranes. *J Chem Phys* 119:596–605
  53. Brannigan G, Brown FLH (2004) Solvent-free simulations of fluid membrane bilayers. *J Chem Phys* 120:1059–1071
  54. Cooke IR, Kremer K, Deserno M (2005) Tunable generic model for fluid bilayer membranes. *Phys Rev E* 72:011506-1–011506-4
  55. Wang Z-J, Deserno M (2010) A systematically coarse-grained solvent-free model for quantitative phospholipid bilayer simulations. *J Phys Chem B* 114:11207–11220
  56. Hoogerbrugge PJ, Koelman JMVA (1992) Simulating microscopic hydrodynamic phenomena with Dissipative Particle Dynamics. *Europhys Lett* 19:155–160
  57. Espagñol P, Warren PB (1995) Statistical properties of Dissipative Particle Dynamics. *Europhys Lett* 30:191–196
  58. Groot RD, Warren PB (1997) Dissipative Particle Dynamics: bridging the gap between atomistic and mesoscopic simulation. *J Chem Phys* 107:4423–4435
  59. Venturoli M, Smit B (1999) Simulating the self-assembly of model membranes. *Phys Chem Commun* 10:1–10
  60. Shillcock JC, Lipowsky R (2002) Equilibrium structure and lateral stress distribution of amphiphilic bilayers from Dissipative Particle Dynamics simulations. *J Chem Phys* 117:5048–5061
  61. Shillcock JC, Lipowsky R (2006) The computational route from bilayer membranes to vesicle fusion. *J Phys Condens Matter* 18: S1191–S1219
  62. Schmid F (2009) Toy amphiphiles on the computer: what can we learn from generic models? *Macromol Rapid Commun* 30:741–751
  63. Venturoli M, Sperotto MM, Kranenburg M, Smit B (2006) Mesoscopic models of biological membranes. *Physics Reports* 437:1–54
  64. Müller M, Katsov K, Schick M (2006) Biological and synthetic membranes: what can be learned from a coarse-grained description? *Physics Reports* 434:113–176
  65. Füchsli RM, Maeke T, McCaskill JS (2009) Spatially resolved simulations of membrane reactions and dynamics: multipolar reaction DPD. *Eur Phys J E* 29:431–448
  66. Laughlin RB, Pines D, Schmalian J, Stojkovic BP, Wolynes P (2000) The middle way. *PNAS* 97:32–37
  67. Lipowsky R, Leibler S (1986) Unbinding transitions of interacting membranes. *Phys Rev Lett* 56:2541–2544
  68. Safinya CR, Roux D, Smith GS, Sinha SK, Dimon P, Clark NA, Bellocq AM (1986) Steric interactions in a model multimembrane system: a synchrotron X-ray study. *Phys Rev Lett* 57:2718–2721
  69. Jakobsen AF, Mouritsen OG, Besold G (2005) Artifacts in dynamical simulations of coarse-grained lipid bilayers. *J Chem Phys* 122:204901-1–204901-11
  70. Allen MP (2006) Configurational temperature in membrane simulations using dissipative particle Dynamics. *J Phys Chem B* 110:3823–3830
  71. Larson RG (1989) Self-assembly of surfactant liquid crystalline phases by Monte Carlo simulation *J Chem Phys* 91:2479–2488
  72. Drouffe J-M, Maggs AC, Leibler S (1991) Computer simulations of self-assembled membranes. *Science* 254:1353–1356
  73. Smit B, Hilbers PAJ, Esselink K, Rupert LAM, van Os NM, Schlijper AG (1990) Computer simulations of a water/oil interface in the presences of micelles. *Nature* 348:624–625
  74. Tieleman DP, van der Spoel D, Berendsen HJC (2000) Molecular dynamics simulations of dodecylphosphocholine micelles at three different aggregate sizes: micellar structure and chain relaxation. *J Phys Chem B* 104:6380–6388

75. Marrink SJ, Lindahl E, Edholm O, Mark AE (2001) Simulation of the spontaneous aggregation of phospholipids into bilayers. *JACS* 123:8638–8639
76. Nielsen M, Miao L, Ipsen JH, Zuckermann MJ, Mouritsen OG (1999) Off-lattice model for the phase behavior of lipid-cholesterol bilayers. *Phys Rev E* 59:5790–5803
77. Marrink SJ, Mark AE (2003) Molecular dynamics simulation of the formation, structure and dynamics of small phospholipid vesicles. *JACS* 125:15233–15242
78. de Vries AH, Mark AM, Marrink SJ (2004) Molecular dynamics simulation of the spontaneous formation of a small DPPC vesicle in water in atomistic detail. *JACS* 126:4488–4489
79. Risselada HJ, Mark AE, Marrink SJ (2008) Application of mean field boundary potentials in simulations of lipid vesicles. *J Phys Chem B* 112:7438–7447
80. Bernardes AT (1996) Computer simulations of spontaneous vesicle formation. *Langmuir* 12:5763–5767
81. Huang J, Wang Y, Qian C (2009) Simulation study on the formation of vesicle and influence of solvent. *J Chem Phys* 131:234902-1–234902-5
82. Sevink GJA, Zvelindovsky AV (2005) Self-assembly of complex vesicles. *Macromolecules* 38:7502–7513
83. He X, Schmid F (2006) Dynamics of spontaneous vesicle formation in dilute solutions of amphiphilic diblock copolymers. *Macromolecules* 39:2654–2662
84. Pogodin S, Baulin VA (2010) Coarse-grained models of biological membranes within the single chain mean field theory. *Soft Matter* 6:1–12
85. Yamamoto S, Maruyama Y, Hyodo S (2002) Dissipative Particle Dynamics study of spontaneous vesicle formation of amphiphilic molecules. *J Chem Phys* 116:5842–5849
86. Graßmüller A, Shillcock J, Lipowsky R (2009) The fusion of membranes and vesicles: pathway and energy barriers from Dissipative Particle Dynamics. *Biophys J* 96:2658–2675
87. Discher BM, Won YY, Ege DS, Lee JCM, Bates FS, Discher DE, Hammer DA (1999) Polymersomes: tough vesicles made from diblock copolymers. *Science* 284:1143–1146
88. Yu S, Azzam T, Rouiller I, Eisenberg A (2009) Breathing vesicles. *JACS* 131:10557–10566
89. Ortiz V, Nielsen SO, Discher DE, Klein ML, Lipowsky R, Shillcock J (2005) Dissipative Particle Dynamics simulations of polymersomes. *J Phys Chem B* 109:17708–17714
90. Srinivas G, Discher DE, Klein ML (2004) Self-assembly and properties of diblock copolymers by coarse-grained molecular dynamics. *Nat Mater* 3:638–644
91. Meleard P, Gerbeaud C, Pott T, Fernandez-Puente L, Bivas I, Mitov MD, Dufourcq J, Bothore P (1997) Bending elasticities of model membranes: influences of temperature and sterol content. *Biophys J* 72:2616–2629
92. Illya G, Lipowsky R, Shillcock JC (2005) Effect of chain length and asymmetry on material properties of bilayer membranes. *J Chem Phys* 122:244901-1–224901-6
93. Schofield P, Henderson JR (1982) Statistical mechanics of inhomogeneous fluids. *Proc R Soc London A* 379:231–246
94. Templer RH, Castle SJ, Curran AR, Rumbles G, Klug DR (1998) Seeing isothermal changes in the lateral pressure in model membranes using di-pyrenyl phosphatidylcholine. *Faraday Disc* 111:41–53
95. Cantor RS (1999) Lipid composition and the lateral pressure profile in bilayers. *Biophys J* 76:2625–2639
96. Cantor RS (2001) Breaking the Meyer-Overton rule: predicted effects of varying stiffness and interfacial activity on the intrinsic potency of anesthetics. *Biophys J* 80:2284–2297
97. Sonne J, Hansen FY, Peters GH (2005) Methodological problems in pressure profile calculations for lipid bilayers. *J Chem Phys* 122:124903-1–124903-9
98. Cantor RS (2002) Size distribution of barrel-stave aggregates of membrane peptides: influence of the bilayer lateral pressure profile. *Biophys J* 82:2520–2525
99. Lindahl E, Edholm O (2000) Spatial and energetic-entropic decomposition of surface tension in lipid bilayers from Molecular Dynamics simulations. *J Chem Phys* 113:3882–3893
100. Gullingsrud J, Schulten K (2004) Lipid bilayer pressure profiles and mechanosensitive channel gating. *Biophys J* 86:3496–3509
101. Ollila OHS, Risselada HJ, Louhivuori M, Lindahl E, Vattulainen I, Marrink SJ (2009) 3D pressure field in lipid membranes and membrane-protein complexes. *Phys Rev Lett* 102:078101-1–078101-4
102. Jørgensen K, Mouritsen OG (1995) Phase separation dynamics and lateral organization of two-component lipid membranes. *Biophys J* 69:942–954
103. Wu SH, McConnell HM (1975) Phase separations in phospholipid membranes. *Biochemistry* 14:847–854

104. Wilkinson DA, Nagle JF (1979) Dilatometric study of binary mixtures of phosphatidylcholines. *Biochemistry* 18:4244–4249
105. Bagatolli LA, Gratton E (2000) A correlation between lipid domain shape and binary phospholipid mixture composition in free standing bilayers: a two-photon fluorescence microscopy study. *Biophys J* 79:434–447
106. Illya G, Lipowsky R, Shillcock JC (2006) Two-component membrane material properties and domain formation from Dissipative Particle Dynamics. *J Chem Phys* 125:114710-1–114710-9
107. Yamamoto S, Hyodo S (2003) Budding and fission dynamics of two-component vesicles. *J Chem Phys* 118:7937–7943
108. Laradji M, Sunil Kumar PB (2005) Domain growth, budding, and fission in phase-separating self-assembled fluid bilayers. *J Chem Phys* 123:224902-1–224902-10
109. Venturoli M, Smit B, Sperotto MM (2005) Simulation studies of protein-induced bilayer deformations, and lipid-induced protein tilting, on a mesoscopic model for lipid bilayers with embedded proteins. *Biophys J* 88:1778–1798
110. Schmidt U, Guigas G, Weiss M (2008) Cluster formation of transmembrane proteins due to hydrophobic mismatching. *Phys Rev Lett* 101:128104-1–128104-4
111. Li L, Davande H, Bedrov D, Smith GD (2007) A molecular dynamics simulation study of C<sub>60</sub> fullerenes inside a dimyristoyl-phosphatidylcholine lipid bilayer. *J Phys Chem B* 111:4067–4072
112. Monticelli L, Salonen E, Ke PC, Vattulainen I (2009) Effects of carbon nanoparticles on lipid membranes: a molecular simulation perspective. *Soft Matter* 5:4433–4445
113. Wong-Ekkabut J, Baoukina S, Triampo W, Tang I-M, Tielemans DP, Monticelli L (2008) Computer simulation study of fullerene translocation through lipid membranes. *Nat Nanotechnol* 3:363–368
114. Chiu CC, deVane R, Klein ML, Shinoda W, Moore PB, Nielsen SO (2010) Coarse-grained potential models for phenyl-based molecules: II. application to fullerenes. *J Phys Chem B* 114:6394–6400
115. Smith KA, Jasnow D, Balazs AC (2007) Designing synthetic vesicles that engulf nanoscopic particles. *J Chem Phys* 127:084703-1–084703-10
116. Yang K, Ma Y-Q (2010) Computer simulation of the translocation of nanoparticles with different shapes across a lipid bilayer. *Nature Nanotechnology* 5:579–583
117. Graham TR, Kozlov MM (2010) Interplay of proteins and lipids in generating membrane curvature. *Curr Op Cell Biology* 22:430–436
118. Wu S, Guo H (2009) Simulation study of protein-mediated vesicle fusion. *J Phys Chem B* 113:589–591
119. Reynwar BJ, Illya G, Harmandaris VA, Müller MM, Kremer K, Deserno M (2007) Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature* 447:461–464
120. Chanturiya A, Chernomordik LV, Zimmerberg J (1997) Flickering fusion pores comparable with initial exocytic pores occur in protein-free phospholipid bilayers. *PNAS* 94:14423–14428
121. Lentz BR, Lee J (1999) Poly(ethylene glycol) (PEG)-mediated fusion between pure lipid bilayers: a mechanism in common with viral fusion and secretory vesicle release? *Mol Membrane Biol* 16:279–296
122. Safran SA, Kuhl TL, Israelachvili JN (2001) Polymer-induced membrane contraction, phase separation, and fusion via Marangoni flow. *Biophys J* 81:659–666
123. Bentz J, Alford D, Cohen J, Düzung N (1988) La<sup>3+</sup> induced fusion of phosphatidylserine liposomes. *Biophys J* 53:593–607
124. Haluska CK, Riske KA, Marchi-Artzner V, Lehn J-M, Lipowsky R, Dimova R (2006) Time scales of membrane fusion revealed by direct imaging of vesicle fusion with high temporal resolution. *PNAS* 103:15841–15846
125. Sunami T, Caschera F, Morita Y, Toyota T, Nishimura K, Matsuura T, Suzuki H, Hanczyc MM, Yomo T (2010) Detection of association and fusion of giant vesicles using a fluorescence-activated cell sorter. *Langmuir* 26:15098–15103
126. Dimova R, Bezlyepkina N, Jordö MD, Knorr RL, Riske KA, Staykova M, Vlahovska PM, Yamamoto T, Yang P, Lipowsky R (2009) Vesicles in electric fields: some novel aspects of membrane behavior. *Soft Matter* 5:3201–3212
127. Hu C, Ahmed M, Melia TJ, Söllner TH, Mayer T, Rothman JE (2003) Fusion of cells by flipped SNAREs. *Science* 300:1745–1748
128. Yang L, Huang HW (2002) Observation of a membrane fusion intermediate structure. *Science* 297:1877–1879
129. McNew JA (2008) Regulation of SNARE-mediated membrane fusion during exocytosis. *Chem Rev* 108:1669–1686
130. Thorley JA, McKeating JA, Rappoport JZ (2010) Mechanisms of viral entry: sneaking in the front door. *Protoplasma* 244:15–24

131. St. Vincent MR, Colpitts CC, Ustinov AV, Muqadas M, Joyce MA, Barsby NL, Epand RF, Epand RM, Khramyshev SA, Valueva OA, Korshun VA, Tyrrell DLJ, Schang LM (2010) Rigid amphipathic fusion inhibitors, small molecule antiviral compounds against enveloped viruses. *PNAS* 107:17339–17344
132. Jahn R, Lang T, Südhof TC (2003) Membrane fusion. *Cell* 112:519–533
133. Yang Y, Kurteva S, Ren X, Lee S, Sodroski J (2005) Stoichiometry of envelope glycoprotein trimers in the entry of human immunodeficiency virus type 1. *J of Virology* 79:12132–12147
134. Mohrmann R, de Wit H, Verhage M, Neher E, Sørensen JB (2010) Fast vesicle fusion in living cells requires at least three SNARE complexes. *Science* 330:502–505
135. Tamm LK, Crane J, Kiessling V (2003) Membrane fusion: a structural perspective on the interplay of lipids and proteins. *Curr Op Struct Biol* 13:453–466
136. Chernomordik LV, Zimmerberg J, Kozlov MM (2006) Membranes of the world unite. *J Cell Biology* 175:201–207
137. Lindau M, de Toledo GA (2003) The fusion pore. *Biochim Biophys Acta* 1641:167–173
138. Han X, Wang C-T, Bai J, Chapman ER, Jackson MB (2004) Transmembrane segments of syntaxin line the fusion pore of  $\text{Ca}^{2+}$ -triggered exocytosis. *Science* 304:289–292
139. Szule JA, Coorssen JR (2004) Comment on Transmembrane segments of syntaxin line the fusion pore of  $\text{Ca}^{2+}$ -triggered exocytosis. *Science* 306:813b
140. Han X, Jackson MB (2004) Response to comment on Transmembrane segments of syntaxin line the fusion pore of  $\text{Ca}^{2+}$ -triggered exocytosis. *Science* 306:813c
141. Carr CM, Rizo J (2010) At the junction of SNARE and SM protein function. *Curr Op Cell Biol* 22:488–495
142. Coorssen JR, Blank PS, Albertorio F, Bezrukov L, Kolosova I, Chen X, Backlund PS Jr, Zimmerberg J (2003) Regulated secretion: SNARE density, vesicle fusion and calcium dependence. *J Cell Science* 116:2087–2096
143. Kozlov MM, Markin VS (1983) Possible mechanism of membrane fusion. *Biofizika* 28:243–247
144. Chernomordik LV, Melikyan GB, Chizmadzhev YA (1987) Biomembrane fusion: a new concept derived from model studies using two interacting planar lipid bilayers. *Biochim Biophys Acta* 906:309–352
145. Siegel DP (1993) Energetics of intermediates in membrane fusion: comparison of stalk and inverted micellar intermediate mechanisms. *Biophys J* 65:2124–2140
146. Kozlovsky Y, Kozlov MM (2002) Stalk model of membrane fusion: solution of energy crisis. *Biophys J* 82:882–895
147. Markin VS, Albanesi JP (2002) Membrane fusion: stalk model revisited. *Biophys J* 82:693–712
148. Tajparast M, Glavinović MI (2009) Forces and stresses acting on fusion pore membrane during secretion. *Biochim Biophys Acta* 1788:1009–1023
149. Chernomordik LV, Kozlov MM (2005) Membrane hemifusion: crossing a chasm in two leaps. *Cell* 123:375–382
150. Efrat A, Chernomordik L, Kozlov MM (2007) Point-like protrusion as a pre-stalk intermediate in membrane fusion pathway. *Biophys J* 92:L61–L63
151. Lipowsky R, Grotehans S (1993) Hydration vs protrusion forces between lipid bilayers. *Europhys Lett* 23:599–604
152. Woodbury DJ (1989) Pure lipid vesicles can induce channel-like conductances in planar bilayers. *J Membrane Biol* 109:145–150
153. Heimburg T (2010) Lipid ion channels. *Bioophys Chem* 150:2–22
154. Marrink SJ, Mark AE (2003) The mechanism of vesicle fusion as revealed by molecular dynamics simulations. *JACS* 125:11144–11145
155. Stevens MJ, Hoh JH, Woolf TB (2003) Insights into the molecular mechanism of membrane fusion from simulation: evidence for the association of splayed tails. *Phys Rev Lett* 91:188102-1–188102-4
156. Kasson PM, Lindahl E, Pande VS (2010) Atomic-resolution simulations predict a transition state for vesicle fusion defined by contact of a few lipid tails. *PLoS Comput Biol* 6: e1000829
157. Smirnova YG, Marrink SJ, Lipowsky R, Knecht V (2009) Solvent-exposed tails as pre-stalk transition states for membrane fusion at low hydration. *JACS* 132:6710–6718
158. Mirjanian D, Dickey AN, Hoh JH, Woolf TB, Stevens MJ (2010) Splaying of aliphatic tails plays a central role in barrier crossing during liposome fusion. *J Phys Chem B* 114:11061–11068
159. Noguchi H, Takasu M (2001) Fusion pathways of vesicles: a Brownian dynamics simulation. *J Chem Phys* 115:9547–9551
160. Shillcock JC, Lipowsky R (2005) Tension-induced fusion of bilayer membranes and vesicles. *Nat Mater* 4:225–228

161. Gao L, Lipowsky R, Shillcock JC (2008) Tension-induced vesicle fusion: pathways and pore dynamics. *Soft Matter* 4:1208–1214
162. Liu Y-T, Zhao Y, Liu H, Liu Y-H, Lu Z-Y (2009) Spontaneous fusion between the vesicles formed by A<sub>2n</sub>(B<sub>2</sub>)<sub>n</sub> type comb-like block copolymers with a semiflexible hydrophobic backbone. *J Phys Chem B* 113:15256–15262
163. Li X, Liu Y, Wang L, Deng M, Liang H (2009) Fusion and fission pathways of vesicles from amphiphilic triblock copolymers: a dissipative particle dynamics simulation study. *Phys Chem Chem Phys* 11:4051–4059
164. Müller M, Katsov K, Schick M (2003) A new mechanism of model membrane fusion determined from Monte Carlo simulation. *Biophys J* 85:1611–1623
165. Schick M, Katsov K, Müller M (2005) The central role of line tension in the fusion of biological membranes. *Mol Phys* 103:3055–3059
166. Kinnunen PKJ (1992) Fusion of lipid bilayers: a model involving mechanistic connection to H<sub>II</sub> phase forming lipids. *Chem Phys Lipids* 63:251–258
167. Holopainen JM, Lehtonen JYA, Kinnunen PKJ (1999) Evidence for the extended phospholipid conformation in membrane fusion and hemifusion. *Biophys J* 76:2111–2120
168. Kinnunen PKJ, Holopainen JM (2000) Mechanisms of initiation of membrane fusion: role of lipids. *Biosci Rep* 20:465–482
169. Stein A, Weber G, Wahl MC, Jahn R (2009) Helical extension of the neuronal SNARE complex into the membrane. *Nature* 460:525–528
170. Weber T, Zemelman BV, McNew JA, Westermann B, Gmachi M, Parlati F, Söllner TH, Rothman JE (1998) SNAREpins: minimal machinery for membrane fusion. *Cell* 92:759–772
171. Pang ZP, Südhof TC (2010) Cell biology of Ca<sup>2+</sup>-triggered exocytosis. *Curr Op Cell Biology* 22:496–505
172. Knecht V, Grubmüller H (2003) Mechanical coupling via the membrane fusion SNARE protein Syntaxin 1A: a molecular dynamics study. *Biophys J* 84:1527–1547
173. McNew JA, Weber T, Parlati F, Johnston RJ, Melia TJ, Söllner TH, Rothman JE (2000) Close is not enough: SNARE-dependent membrane fusion requires an active mechanism that transduces force to membrane anchors. *J Cell Biol* 150:105–117
174. Yersin A, Hirling H, Steiner P, Magnin S, Regazzi R, Hüni B, Huguenot P, de Los RP, Dietler G, Catsicas S, Kasas S (2003) Interactions between synaptic vesicle fusion proteins explored by atomic force microscopy. *PNAS* 100:8736–8741
175. Abdulreda MH, Bhalla A, Rico F, Berggren P-O, Chapman ER, Moy VT (2009) Pulling force generated by interacting SNAREs facilitates membrane hemifusion. *Integr Biol* 1:301–310

# INDEX

## A

- ABC transporter ..... 366–371  
Ab initio methods ..... 3–9, 17, 18, 37, 57, 72, 101, 354  
Ab initio molecular dynamics ..... 29–41  
Adaptive resolution simulation (AdResS) ..... 568–581  
Air-water interface ..... 235, 432, 433, 435, 662  
Alchemical methods ..... 275–277, 279, 281, 286, 287, 296, 299, 300  
AMBER ..... 71, 100, 112, 135, 148, 149, 199, 200, 219, 234, 236, 251, 258–260, 282, 300, 307, 316, 320, 334, 342, 343, 436, 446, 448, 449, 451, 453, 457, 462–464, 471, 472, 474, 476–482, 591  
Anisotropic network model  
    (ANM) ..... 381, 382, 607–610  
*Apo*-state ..... 373–376  
ATP hydrolysis ..... 366–369  
Autodock ..... 472–474, 477, 478, 482

## B

- BAR. *See* Bennett acceptance ratio (BAR)  
Basis set ..... 4, 6, 7, 14, 18–22, 32, 34, 35, 39, 50, 92, 94, 114, 260, 263, 511–513, 516, 518, 519, 522–523, 605, 612  
Bennett acceptance ratio (BAR) ..... 275, 278–281, 292, 294, 296–297, 302, 303, 305, 319, 329  
B-factors ..... 610–611, 614  
Binding free energy ..... 107, 108, 283, 307, 342, 344  
Binding site ..... 106–108, 274, 288, 305, 306, 340, 341, 347, 351, 352, 356, 366, 368, 369, 372–380, 383–386, 388–391, 394–398, 477, 534, 557, 647  
Biomacromolecules ..... 630, 631  
Biomolecular simulation,  
Biomolecular system ..... 3, 5–7, 9, 44, 148, 153–155, 187, 188, 218, 236, 259, 313–334, 351, 362, 398, 448, 534, 617  
Boltzmann inversion ..... 500–503, 515, 574, 629

## C

- Carbohydrate ..... 201, 209, 320, 469–482, 509, 546–547, 559  
Car-Parrinello molecular dynamics ..... 31–32

- Catalysis ..... 67–70, 97  
CHARMM ..... 49, 71, 75, 76, 79, 112, 135, 148, 185, 199, 200, 219, 226–228, 233–236, 251, 259, 280, 282, 307, 365, 436, 446–449, 451, 452, 457–459, 471, 472  
Classical mechanics ..... 127, 128, 198, 211, 579  
Coarse-grained (CG) ..... 128, 148, 199, 211, 215, 254, 320, 381, 435, 438, 439, 487–524, 533–561, 568, 569, 572–576, 578, 580, 581, 585–597, 601, 602, 613, 614, 618, 622, 624, 625, 627–630, 636–638, 643, 649, 650, 659–691  
models ..... 128, 199, 490, 550, 559, 618, 625, 629, 630, 674  
Computer simulation ..... 197, 433, 617, 618, 660, 662, 666, 668, 683  
Conformational change ..... 68–70, 73, 81, 104, 106, 216, 217, 233, 362–364, 366, 367, 370, 372–374, 377–380, 382–386, 390, 392, 393, 395, 397, 398, 470, 585, 586, 612, 614, 641, 645  
Conformational coupling ..... 370  
Coupled-cluster linear response methods ..... 8, 16  
Coupled-cluster methods ..... 5, 6
- D**
- Density functional theory (DFT) ..... 3–22, 30, 33, 38, 53, 72–75, 78, 79, 92–94, 97, 99, 100, 103, 107, 110–112, 115, 224, 257  
Dipalmitoyl-phosphatidylcholine (DPPC) ..... 207, 235, 408–415, 418–420, 422–425, 427, 428, 436, 437, 536, 543, 550, 553, 554, 642  
Dipole moment ..... 13, 57, 105, 110, 112, 204, 217, 219–221, 224, 226, 227, 231, 233, 235, 245, 346, 374, 375  
Dissipative particle dynamics (DPD) ..... 128, 136, 139, 254, 572, 576–578, 617–631, 665, 669–671, 673, 676, 678–681, 685–687, 689, 690  
DNA ..... 93, 206, 231, 236, 243, 445–464, 534, 543, 549, 550, 559, 587, 612, 659, 660, 682  
    dynamics ..... 445–464  
    structure ..... 446, 447, 449–451, 453, 456–461  
DPD. *See* Dissipative particle dynamics (DPD)  
DPPC. *See* Dipalmitoyl-phosphatidylcholine (DPPC)

Dynamics ..... 30, 44, 68, 91, 127, 153, 197, 217, 254, 299, 314, 354, 362, 414, 438, 445, 469–482, 489, 534, 575, 586, 602, 617–631, 635, 659

**E**

Electronic excited states ..... 21  
Electronic polarization ..... 71, 74, 106, 215–219, 224, 232–237  
Electrostatics ..... 44, 71, 95, 135, 185, 198, 216, 243–263, 290, 322, 340, 365, 414, 433, 451, 475, 513, 536, 601, 649, 691  
Enzyme ..... 45, 67–83, 128, 471, 473, 611, 617, 644, 648  
Ewald summation ..... 200, 244–250, 258, 451

**F**

Fast multipole method (FMM) ..... 253–256  
First-principles approach ..... 15, 197, 204, 205, 446, 457  
FMM. *See* Fast multipole method (FMM)  
Force field ..... 14, 33, 44, 70, 91, 128, 185, 197–211, 215–237, 244, 321, 341, 365, 409, 433, 446, 471, 487, 533–560, 573, 586, 605, 627, 635, 665  
Free energy ..... 18, 34, 54–56, 69, 97, 156, 207, 218, 271–307, 313–334, 431, 453, 478, 504, 534, 569, 585, 603, 626, 642, 667

**G**

Gaussian network model (GNM) ..... 607–610  
Generalized ensemble algorithm ..... 154, 156–182, 187, 188  
Generalized-Yvon-Born-Green theory ..... 517, 524  
GlpT. *See* Glycerol-3-phosphate transporter (GlpT)  
Glutamate transporter (GLT) ..... 363, 364, 371–376  
GLYCAM ..... 321, 471–478, 481  
Glycerol-3-phosphate transporter (GlpT) ..... 363, 364, 376–383  
GNM. *See* Gaussian network model (GNM)  
Groningen Machine for Chemical Simulations (GROMACS) ..... 148, 149, 228, 251, 280, 282, 286, 307, 342, 408–411, 413–417, 419, 420, 423, 425–428, 480, 537, 539, 553

**H**

Harmonic potential ..... 47, 82, 135, 201, 319, 395, 396

**I**

IMC. *See* Inverse Monte Carlo (IMC)  
Information theory ..... 497  
Inverse Monte Carlo (IMC) ..... 503–505, 508, 509, 514–516, 520, 521

**K**

Kinetics ..... 9, 12, 13, 49, 136–140, 156, 157, 160, 166, 172, 178, 273, 299, 303, 331, 452, 522, 571, 572, 586, 587, 589, 590, 597, 604, 622–624

**L**

Langmuir monolayer ..... 433, 662  
Leap-frog integrator ..... 130, 134, 623  
Lennard-Jones ..... 47, 50, 71, 76, 102, 105, 206, 208–211, 229, 249, 273, 274, 285, 286, 289, 302, 305, 343, 352, 472, 476, 501, 589, 593, 601, 649  
Lennard-Jones potential ..... 44, 47, 48, 91, 144, 204, 206, 227, 229, 286, 296, 438, 489, 500, 501, 535, 588, 619, 649, 670  
Lipid bilayer ..... 128, 206, 207, 211, 217, 218, 235, 363–365, 378, 387, 407–428, 433, 503, 509, 547–549, 553, 628, 629, 636–639, 641, 644, 647, 649, 660, 661, 663–666, 668–670, 672, 673, 675–681, 684, 685, 687

Lipid membrane ..... 243, 365, 553, 554, 556, 557, 617, 618, 628, 660, 662, 664–666, 668, 670, 672, 680–682

Lipid monolayer ..... 235, 431–441, 540, 660, 662, 663

Low-frequency modes ..... 382, 453, 612, 613  
Lung surfactant ..... 436

**M**

Major facilitator superfamily (MFS) ..... 376, 377  
Maltose transporter ..... 367, 370, 371  
Martini force field ..... 436, 533, 534, 537, 540, 542, 543, 546, 547, 549, 550, 556, 560  
MBAR. *See* Multistate Bennett acceptance ratio (MBAR)  
Membrane protein ..... 106, 107, 113, 210, 218, 349, 372, 388, 552, 553, 556, 628, 635–651  
Membrane simulation ..... 207, 210, 413, 554, 636, 678  
MFS. *See* Major facilitator superfamily (MFS)  
Misfolding ..... 104  
Molecular docking ..... 74, 339–356, 469–482, 560

- Molecular dynamics (MD) ..... 29–41, 44, 69, 96, 127–149, 153, 197–211, 221, 252, 285, 315–317, 341, 362, 407–428, 445–464, 471, 569, 586, 606, 618, 635–651, 668
- Molecular mechanics (MM) ..... 14, 34, 43–63, 67, 91–115, 215, 273, 344, 560
- Monolayer collapse ..... 437, 439–441
- Monte Carlo (MC) ..... 69, 81, 83, 128, 153–155, 157–159, 161, 165–168, 170, 173–175, 178, 179–184, 188, 282, 300, 345, 487, 502–504, 514, 591–593, 595, 624, 672, 673, 678, 686
- Multicanonical algorithm (MUCA) ..... 154–156, 167–172, 174–176, 181, 187, 188
- Multigrid ..... 244, 251–257, 259
- Multiscale coarse-graining (MS-CG) ..... 509–520, 524, 668
- Multiscale methods ..... 94–95, 559, 567
- Multistate Bennett acceptance ratio (MBAR) ..... 275, 280–281, 292, 296, 297, 299, 302, 305
- N**
- $\text{Na}^+$ /Betaine symporter (BetP) ..... 364, 393–397
- $\text{Na}^+$ -coupled galactose transporter ..... 388–393
- Nanoscale molecular dynamics (NAMD) ..... 71, 148, 149, 186, 251, 252, 307, 365, 448, 472, 479, 480, 537
- NBDs. *See* Nucleotide binding domains (NBDs)
- Neighbor list ..... 144, 145, 542
- Normal mode analysis (NMA) ..... 57, 110, 112, 128, 233, 234, 344, 381–383, 603–607, 609–614
- Nucleic acid ..... 71, 201, 211, 218, 236, 362, 448, 451, 559
- Nucleotide binding domains (NBDs) ..... 366–371
- O**
- Oligosaccharide ..... 470, 546, 547
- P**
- Parameterization ..... 74, 78, 107, 199, 200, 208–210, 284, 285, 413–414, 451, 537–540, 542–544, 546, 547, 559, 591, 625–631
- Particle mesh Ewald (PME) ..... 185, 206–208, 245, 250–252, 255, 258–260, 262, 291, 322, 325, 365, 415, 451, 452, 463, 464, 478
- Particle-particle particle-mesh ..... 245, 250
- Peptide ..... 7, 93, 182, 203, 233–235, 243, 314, 353, 509, 534, 592, 636, 677
- Phospholipid ..... 201, 231, 423, 436, 539, 627, 637, 638, 640, 643, 645, 648–650, 660, 661, 663, 667, 673, 675–677, 679, 682, 685, 687, 689
- P3M ..... 250, 251, 255
- Polarizable force field ..... 51, 104, 204, 215–237, 258
- Primary transporter ..... 362
- Proline ..... 68, 320, 323–333, 549, 595
- Protein ..... 8, 38, 44, 68, 91–115, 127, 154, 199, 218, 243, 283, 324, 339, 362, 407, 436, 445, 470, 488, 534, 585, 601, 617, 635, 659
- aggregation ..... 636
- conformations ..... 68, 69, 81, 342, 356, 362, 364, 365, 380, 382–384, 398, 587, 595, 613, 635, 636
- dynamics ..... 69–70, 101, 381, 383, 586, 606, 613
- folding ..... 93, 97, 127, 155, 200, 210, 489, 585–597, 602, 641
- structure ..... 70, 101–106, 235, 339, 350, 356, 365, 382, 473, 476–478, 546, 590, 606, 609, 645, 688
- Proton transfer ..... 19, 36, 38, 63, 106, 109
- Q**
- QM/MM. *See* Quantum mechanics/molecular mechanics (QM/MM)
- Quantum chemistry ..... 5, 16, 33, 36, 45, 46, 49, 56, 92–94, 108, 155, 474, 586
- Quantum mechanics/molecular mechanics (QM/MM) ..... 17, 18, 34, 37, 38, 43–63, 67, 73–82, 91–115, 231, 232, 273, 560
- R**
- Replica exchange method (REM) ..... 155–165, 169–173, 179–185, 187, 188
- RESP charge derivation ..... 471, 481
- Restrained electrostatic potential (RESP) ..... 105, 300, 474–476, 481
- Rocker-switch model ..... 378, 382
- S**
- Scoring function ..... 104, 107, 108, 340, 342–343, 346–348, 352
- Secondary transporter ..... 362, 373, 386, 388–390, 393
- Second-order perturbation theory ..... 5
- Semi-empirical methods ..... 18–20, 53, 57, 93, 100
- Sequence dependent DNA flexibility ..... 446, 447
- Simulated tempering ..... 156, 165–166, 169–170
- Simulations of liquid water ..... 572

Solvation free energy ..... 182–184, 284, 539, 552  
Solvent-accessible surface ..... 182, 344, 345, 593  
Statistical mechanics ..... 136, 272, 488, 490, 574  
Structure-based drug design ..... 340  
Structure refinement ..... 102, 610, 635, 641  
Sugar ..... 201, 314, 320, 334, 386, 391, 407, 446, 447, 454–457, 461, 470, 471, 473, 474, 476–480, 534, 543–547

**T**

Theoretical spectroscopy ..... 108  
Thermal motion ..... 29, 425, 663, 667, 690  
Thermodynamic integration (TI) ..... 36, 55, 56, 275, 277–279, 292–297, 299, 302, 305, 306, 356, 538, 539  
Thermodynamics ..... 318, 342, 346, 488, 491, 521, 522, 549, 586–588, 594, 596, 641, 642  
Time-dependent density functional theory ..... 224  
Transmembrane helices ..... 372, 373, 377, 386, 389, 554, 555, 642

**U**

Umbrella sampling ..... 54–55, 82, 97, 154, 172, 288, 299, 314, 316, 327, 479, 646

**V**

Validation of simulations ..... 210, 234, 417–419, 445–464

Verlet algorithm ..... 130, 132–134, 622–623

Vesicle ..... 555, 556, 659–691  
fusion ..... 552, 617, 659–691

**W**

Wave-function methods ..... 4, 9, 21

Weak coupling algorithm ..... 137

Weighted histogram analysis method  
(WHAM) ..... 82, 154, 176, 184, 186, 279–280, 292, 297, 314, 317, 588