

**REVIEWS IN COMPUTATIONAL
CHEMISTRY, VOLUME 32**

REVIEWS IN COMPUTATIONAL CHEMISTRY, VOLUME 32

Edited by

ABBY L. PARRILL

College of Arts and Sciences,
The University of Memphis,
Memphis, TN, USA

KENNY B. LIPKOWITZ

Office of Naval Research,
Arlington, VA, USA

WILEY

This edition first published 2022
© 2022 John Wiley & Sons, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Abby L. Parrill and Kenny B. Lipkowitz to be identified as the editors of the editorial material in this work has been asserted in accordance with law.

Registered Office

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty

In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data Applied for:

ISBN: 9781119625896

Cover design by Wiley

Cover image: © Abby L. Parrill

Set in 10/12pt TimesLTStd by Straive, Chennai, India

10 9 8 7 6 5 4 3 2 1

CONTENTS

| | |
|--|-----------|
| List of Contributors | ix |
| Preface | xi |
| Contributors to Previous Volumes | xv |
| | |
| 1 Non-Deterministic Global Structure Optimization: An Introductory Tutorial | 1 |
| <i>Bernd Hartke</i> | |
| List of abbreviations | 1 |
| Introduction | 2 |
| The Need for Structural Optimization | 2 |
| Search Space is Vast | 3 |
| Deterministic vs Non-Deterministic Search | 5 |
| Fundamental Take-Home Lessons | 8 |
| A Closer Look at Some NDGO Background Details | 8 |
| Too Inspired by Nature | 8 |
| No Free Lunch | 11 |
| NDGO Algorithm Comparisons | 14 |

| | |
|---|-----------|
| Barrier Crossing | 15 |
| Old vs New Machine Learning | 19 |
| Take-Home Lessons for NDGO Background Details | 20 |
| General Guidelines for NDGO Applications | 21 |
| Brief Summary of Some Fundamental NDGO Algorithm | |
| Ideas | 21 |
| NDGO Method Design Choices | 22 |
| NDGO Tips for Absolute Beginners | 28 |
| Things to Do, and Pitfalls to Avoid | 31 |
| Recent Highlights | 32 |
| References | 34 |
| | |
| 2 Density Functional Tight Binding Calculations for Probing Electronic-Excited States of Large Systems | 45 |
| | |
| <i>Sharma S.R.K.C. Yamijala, Ma. Belén Oviedo, and Bryan M. Wong</i> | |
| Introduction | 45 |
| Real-Time Time-Dependent DFTB (RT-TDDFTB) | 46 |
| Theory and Methodology | 46 |
| Tutorial on RT-TDDFTB Electron Dynamics for a Naphthalene Molecule | 49 |
| Absorption Spectrum for Naphthalene | 49 |
| Electron Dynamics of Naphthalene with a Laser-Type Perturbation | 51 |
| RT-TDDFTB Electron Dynamics of a Realistic Large Systems | 51 |
| DFTB-Based Nonadiabatic Electron Dynamics | 59 |
| Adiabatic vs Nonadiabatic Dynamics | 59 |
| Equations Governing Nonadiabatic Electron Dynamics | 61 |
| The Classical Path Approximation | 62 |
| Surface Hopping and Fewest Switches Criterion | 63 |
| Implementation Details of CPA-FSSH-DFTB | 65 |
| Post-processing Tools | 67 |
| Computational Details | 67 |
| An Example on Charge Transfer Dynamics in Organic Photovoltaics | 68 |
| Conclusion and Outlook | 72 |
| Acknowledgments | 72 |
| References | 73 |

| | |
|---|------------|
| 3 Advances in the Molecular Simulation of Microphase Formers | 81 |
| <i>Patrick Charbonneau and Kai Zhang</i> | |
| Introduction | 81 |
| Block Copolymers | 83 |
| Surfactants and Microemulsions | 84 |
| Lattice Spin Systems | 87 |
| Colloidal Suspensions | 87 |
| Other Examples | 90 |
| Field Theory of Microphase Formation | 90 |
| Molecular Simulations and Challenges | 91 |
| Simulating Periodic Microphases | 93 |
| Expanded Thermodynamics | 94 |
| Thermodynamic Integration for Microphases | 95 |
| Ghost Particle/Cluster Switching Method | 100 |
| Cluster Volume Moves | 103 |
| Determining Phase Transitions | 105 |
| Simulations of Disordered Microphases | 106 |
| Wolff-Like Cluster Algorithms | 106 |
| Virtual Cluster Moves | 107 |
| Aggregation Volume Biased (AVB) Moves | 109 |
| Morphological Crossovers in the Disordered Regime | 110 |
| Microphase Formers Solved by Molecular Simulations | 112 |
| One-Dimensional Models | 112 |
| Lattice Spin Models | 113 |
| Colloidal Models | 117 |
| Conclusion | 118 |
| Free Energy of an Ideal Gas in a Field | 119 |
| Constant pressure Simulations of Particles in A Field | 120 |
| Virial Coefficients of Particles in a Field | 120 |
| Acknowledgments | 122 |
| References | 122 |
| 4 Molecular Simulations of Deep Eutectic Solvents: A Perspective on Structure, Dynamics, and Physical Properties | 135 |
| <i>Shalini J. Rukmani, Brian W. Doherty, Orlando Acevedo, and Coray M. Colina</i> | |
| Introduction | 135 |
| Deep Eutectic Solvents | 137 |

| | |
|--|-----|
| Definition of Deep Eutectic Solvents | 137 |
| DES as Ionic Liquid Analogues | 137 |
| Molecular Structure of DESs and Type of Interactions | 140 |
| Types of DES | 142 |
| Molecular Simulation Methods | 143 |
| An Overview of <i>Ab Initio</i> Methods | 145 |
| Classical Molecular Dynamics at the Atomic Level | 149 |
| Nonpolarizable Force Fields used for DES Simulations | 153 |
| Physical Properties | 159 |
| Liquid Density | 159 |
| Volume Expansivity | 162 |
| Surface Tension | 162 |
| Thermodynamic Properties | 164 |
| Heat Capacity | 164 |
| Heats of Vaporization | 168 |
| Isothermal Compressibility | 169 |
| Transport Properties | 170 |
| Viscosity | 170 |
| Diffusion Coefficients | 178 |
| Deep Eutectic Solvent Structure | 183 |
| Radial Distribution Functions | 183 |
| Hydrogen Bond Analysis | 189 |
| Spatial Distribution Functions | 196 |
| Application of DES Through Simulation | 196 |
| Gas Sorption Studies on DES | 196 |
| DES Interactions at Metal Surfaces | 198 |
| Proteins in DES | 199 |
| Summary | 200 |
| Acknowledgments | 201 |
| References | 201 |

LIST OF CONTRIBUTORS

Orlando Acevedo, Department of Chemistry, University of Miami, Coral Gables, USA (Electronic mail: orlando.acevedo@miami.edu).

Patrick Charbonneau, Department of Chemistry and Department of Physics, Duke University, Durham, NC, USA (Electronic mail: Patrick.charbonneau@duke.edu).

Coray M. Colina, Department of Materials Science and Engineering, George and Josephine Butler Polymer Research Laboratory, Department of Chemistry, University of Florida, Gainesville, USA (Electronic mail: colina@chem.ufl.edu).

Brian W. Doherty, Department of Chemistry, University of Miami, Coral Gables, USA (Electronic Mail: brian.doherty@nyu.edu).

Bernd Hartke, Theoretical Chemistry, Institute for Physical Chemistry, Christian-Albrechts-University, Olshausenstr. 40, 24098 Kiel, Germany (Electronic mail: hartke@pctc.uni-kiel.de).

Ma. Belén Oviedo, Departamento de Química Teórica y Computacional, Facultad de Ciencias Químicas, Instituto de Investigaciones en Físicoquímica de Córdoba (INFIQC), UNC-CONICET, Universidad Nacional de Córdoba, Córdoba, Argentina (Electronic mail: belovedo@gmail.com).

Shalini J. Rukmani, Department of Materials Science and Engineering, George and Josephine Butler Polymer Research Laboratory, Department of Chemistry, University of Florida, Gainesville, Florida, USA (Electronic mail: jayaramanrus@ornl.gov).

Bryan M. Wong, Department of Chemical & Environmental Engineering, Materials Science & Engineering Program, Department of Chemistry, and Department of Physics & Astronomy, University of California-Riverside, Riverside, CA, USA (Electronic mail: bmwong@enr.ucr.edu).

Sharma S. R. K. C. Yamijala, Department of Chemical & Environmental Engineering, Materials Science & Engineering Program, Department of Chemistry, and Department of Physics & Astronomy, University of California-Riverside, Riverside, CA, USA (Electronic mail: sharmay@ucr.edu).

Kai Zhang, Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, Jiangsu, China (Electronic mail: kai.zhang.statmech@gmail.com).

PREFACE

This book series seeks to aid researchers in selecting and applying new computational chemistry methods to their own research problems. This aim is achieved through tutorial-style chapters that provide solid starting points and advice for novices as well as critical literature reviews highlighting advanced applications to illustrate current state of the art. Volume 32 continues this longstanding tradition. While each chapter has a unique focus, two themes connect the chapters in this volume. The first theme centers on methods that can be broadly applied to a variety of systems in Chapters 1 and 2, and the second theme emphasizes special considerations required when modeling very specific system types in Chapters 3 and 4.

Chapter 1 highlights the vast space of local minimum energy structures of complex molecules to illustrate the importance of non-deterministic global optimization (NDGO) approaches. Such approaches avoid visiting every region of search space, thus inevitably allowing for the possibility that the global minimum has not been found (otherwise the method would be deterministic!). Because NDGO methods lack a true convergence criterion, it is essential to use them properly to ensure meaningful results are obtained. Thus, the subsection titled “NDGO Tips for Absolute Beginners” should be bookmarked and reviewed regularly by new computational chemists applying NDGO methods to any problem. Bernd Hartke illustrates the humorous practice of naming NDGO algorithms after natural processes that have no parallels to the NDGO problem, but then breaks down the fundamental algorithm ideas into simple ingredients to identify similarities and differences between such algorithms. Important guidance on the use of

NDGO methods include use of a two-level strategy with combining exhaustive global search using inexpensive methods to calculate energies followed by local post-optimization of selected results at high levels of theory as well as avoiding single runs when using nondeterministic algorithms. This latter advice is one that would be well-taken in many areas of computational chemistry and a piece of advice one of the book editors personally gives her students regularly! Chapter 1 closes with a set of recent highlights such as the applicability of NDGO to a variety of optimization targets ranging from force-field parameters to reaction networks.

Chapter 2 focuses on the excited-state dynamics calculations required to calculate electronic absorption spectra or to investigate electron dynamics of chemical systems irradiated by laser light. In particular, real-time time-dependent (RT-TD) and non-adiabatic dynamics calculations using the density functional tight binding (DFTB) formalism are explored. Stepwise tutorials on the molecule naphthalene, are given to provide researchers with practice applying these techniques to probe and understand the chemical dynamics exhibited in a simple system to prepare them for work on larger systems. Silver nanoparticles and nanoparticle chains illustrate applications of the method to large systems. After thoroughly exploring the electron dynamics of adiabatic systems in external electric fields using a single potential energy surface (PES), the theory and methods used to allow nonadiabatic evolution of nuclear position on different PESs are considered. Nonadiabatic evolution is essential for accurate modeling of photochemical and photovoltaic processes that involve transitions between PESs. Computational efficiency of DFTB has advanced to the point that applications in emerging areas of excited-state chemical dynamics in large, complex systems are now within reach.

Chapter 3 transitions from a focus on methodology to applications of methodologies to investigate a specialized type of chemical system, namely chemical systems that form microphases with periodic morphologies such as lamellae and cylinders. Microphases can form in diverse systems, ranging from aggregated colloidal particles to diblock copolymers. Charbonneau and Khang focus on the phenomenological field-theory description of the order-disorder transition between ordered and disordered mesoscopic domains. Key challenges identified in the simulation of microphase-forming systems are largely addressed using very coarse-grained systems. Such models are capable of describing phase transitions with mesoscale patterns that occur over long time scales, a feat that is currently not computationally tractable using atomic or molecular systems. The theoretical and computational foundations of periodic microphase simulations are presented systematically, culminating in a lesson on determining phase transitions from such simulations. A series of classical Monte Carlo algorithms used to enhance sampling efficiency required to simulate disordered microphases

and distinguish morphological regimes are discussed. Applications of free energy-based methods to obtain equilibrium phase information in a range of systems illustrate the scope of applications currently within reach.

Chapter 4 provides a comprehensive overview of deep eutectic solvents (DESS) and simulation methods to studying such systems. DESSs are mixtures with far lower melting points than the individual components comprising the mixture, often formed by mixing a strong hydrogen bond acceptor with a strong hydrogen bond donor. The critical contributions of polarization effects to the behavior of such systems, and the long timescale of simulations needed to investigate many properties of interest, define the key trade-off: While *ab initio* methods provide explicit treatment of polarization with limitations on simulation timescales and system size, nonpolarizable force fields can be applied to large systems and long timescales but lack explicit treatment of polarization. Following a comprehensive consideration of obtaining physical, thermodynamic, transport, and structural properties of DESSs from simulations, Shalini Rukmani, Brian Doherty, Orlando Acevedo, and Coray Colina summarize non-polarizable force fields as performing well for bulk properties, but poorly for reproduction of self-diffusion coefficients. Charge scaling provided significant improvements while raising concerns about treatment of additives that may alter charge transfer magnitudes. Specific charge models from DFT cluster or *ab initio* MD simulations have also been implemented successfully, yet limit generalizability and transferability of the nonpolarizable force field they are implement into. Methods to study DESSs are under ongoing development to address these issues.

The value of *Reviews in Computational Chemistry* stems from the pedagogically driven reviews that have made this ongoing book series so popular. We are grateful to the authors featured in this volume for continuing the tradition of providing not only comprehensive reviews, but also highlighting best practices and factors to consider in performing similar modeling studies.

Volumes of *Reviews in Computational Chemistry* are available in an online form through Wiley InterScience. Please consult the Web (<http://www.interscience.wiley.com/onlinebooks>) or contact reference@wiley.com for the latest information.

We thank the authors of this and previous volumes for their excellent chapters.

Abby L. Parrill
Memphis, TN

Kenny B. Lipkowitz
Washington, DC
October 2020

CONTRIBUTORS TO PREVIOUS VOLUMES

Volume 1 (1990)

David Feller and **Ernest R. Davidson**, Basis Sets for Ab Initio Molecular Orbital Calculations and Intermolecular Interactions.

James J. P. Stewart, Semiempirical Molecular Orbital Methods.

Clifford E. Dykstra, **Joseph D. Augspurger**, **Bernard Kirtman**, and **David J. Malik**, Properties of Molecules by Direct Calculation.

Ernest L. Plummer, The Application of Quantitative Design Strategies in Pesticide Design.

Peter C. Jurs, Chemometrics and Multivariate Analysis in Analytical Chemistry.

Yvonne C. Martin, **Mark G. Bures**, and **Peter Willett**, Searching Databases of Three-Dimensional Structures.

Paul G. Mezey, Molecular Surfaces.

Terry P. Lybrand, Computer Simulation of Biomolecular Systems Using Molecular Dynamics and Free Energy Perturbation Methods.

Donald B. Boyd, Aspects of Molecular Modeling.

Donald B. Boyd, Successes of Computer-Assisted Molecular Design.

Ernest R. Davidson, Perspectives on Ab Initio Calculations.

Volume 2 (1991)

Andrew R. Leach, A Survey of Methods for Searching the Conformational Space of Small and Medium-Sized Molecules.

John M. Troyer and **Fred E. Cohen**, Simplified Models for Understanding and Predicting Protein Structure.

J. Phillip Bowen and **Norman L. Allinger**, Molecular Mechanics: The Art and Science of Parameterization.

Uri Dinur and **Arnold T. Hagler**, New Approaches to Empirical Force Fields.

Steve Scheiner, Calculating the Properties of Hydrogen Bonds by Ab Initio Methods.

Donald E. Williams, Net Atomic Charge and Multipole Models for the Ab Initio Molecular Electric Potential.

Peter Politzer and **Jane S. Murray**, Molecular Electrostatic Potentials and Chemical Reactivity.

Michael C. Zerner, Semiempirical Molecular Orbital Methods.

Lowell H. Hall and **Lemont B. Kier**, The Molecular Connectivity Chi Indexes and Kappa Shape Indexes in Structure-Property Modeling.

I. B. Bersuker and **A. S. Dimoglo**, The Electron-Topological Approach to the QSAR Problem.

Donald B. Boyd, The Computational Chemistry Literature.

Volume 3 (1992)

Tamar Schlick, Optimization Methods in Computational Chemistry.

Harold A. Scheraga, Predicting Three-Dimensional Structures of Oligopeptides.

Andrew E. Torda and **Wilfred F. van Gunsteren**, Molecular Modeling Using NMR Data.

David F. V. Lewis, Computer-Assisted Methods in the Evaluation of Chemical Toxicity.

Volume 4 (1993)

Jerzy Cioslowski, Ab Initio Calculations on Large Molecules: Methodology and Applications.

Michael L. McKee and **Michael Page**, Computing Reaction Pathways on Molecular Potential Energy Surfaces.

Robert M. Whitnell and **Kent R. Wilson**, Computational Molecular Dynamics of Chemical Reactions in Solution.

Roger L. DeKock, **Jeffry D. Madura**, **Frank Rioux**, and **Joseph Casanova**, Computational Chemistry in the Undergraduate Curriculum.

Volume 5 (1994)

John D. Bolcer and **Robert B. Hermann**, The Development of Computational Chemistry in the United States.

Rodney J. Bartlett and **John F. Stanton**, Applications of Post-Hartree-Fock Methods: A Tutorial.

Steven M. Bachrach, Population Analysis and Electron Densities from Quantum Mechanics.

Jeffry D. Madura, **Malcolm E. Davis**, **Michael K. Gilson**, **Rebecca C. Wade**, **Brock A. Luty**, and **J. Andrew McCammon**, Biological Applications of Electrostatic Calculations and Brownian Dynamics Simulations.

K. V. Damodaran and **Kenneth M. Merz Jr.**, Computer Simulation of Lipid Systems.

Jeffrey M. Blaney and **J. Scott Dixon**, Distance Geometry in Molecular Modeling.

Lisa M. Balbes, **S. Wayne Mascarella**, and **Donald B. Boyd**, A Perspective of Modern Methods in Computer-Aided Drug Design.

Volume 6 (1995)

Christopher J. Cramer and **Donald G. Truhlar**, Continuum Solvation Models: Classical and Quantum Mechanical Implementations.

Clark R. Landis, **Daniel M. Root**, and **Thomas Cleveland**, Molecular Mechanics Force Fields for Modeling Inorganic and Organometallic Compounds.

Vassilios Galiatsatos, Computational Methods for Modeling Polymers: An Introduction.

Rick A. Kendall, Robert J. Harrison, Rik J. Littlefield, and Martyn F. Guest, High Performance Computing in Computational Chemistry: Methods and Machines.

Donald B. Boyd, Molecular Modeling Software in Use: Publication Trends.

Eiji Ōsawa and Kenny B. Lipkowitz, Appendix: Published Force Field Parameters.

Volume 7 (1996)

Geoffrey M. Downs and Peter Willett, Similarity Searching in Databases of Chemical Structures.

Andrew C. Good and Jonathan S. Mason, Three-Dimensional Structure Database Searches.

Jiali Gao, Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials.

Libero J. Bartolotti and Ken Flurchick, An Introduction to Density Functional Theory.

Alain St-Amant, Density Functional Methods in Biomolecular Modeling.

Danya Yang and Arvi Rauk, The A Priori Calculation of Vibrational Circular Dichroism Intensities.

Donald B. Boyd, Appendix: Compendium of Software for Molecular Modeling.

Volume 8 (1996)

Zdenek Slanina, Shyi-Long Lee, and Chin-hui Yu, Computations in Treating Fullerenes and Carbon Aggregates.

Gernot Frenking, Iris Antes, Marlis Böhme, Stefan Dapprich, Andreas W. Ehlers, Volker Jonas, Arndt Neuhaus, Michael Otto, Ralf Stegmann, Achim Veldkamp, and Sergei F. Vyboishchikov, Pseudopotential Calculations of Transition Metal Compounds: Scope and Limitations.

Thomas R. Cundari, Michael T. Benson, M. Leigh Lutz, and Shaun O. Sommerer, Effective Core Potential Approaches to the Chemistry of the Heavier Elements.

Jan Almlöf and **Odd Gropen**, Relativistic Effects in Chemistry.

Donald B. Chesnut, The Ab Initio Computation of Nuclear Magnetic Resonance Chemical Shielding.

Volume 9 (1996)

James R. Damewood, Jr., Peptide Mimetic Design with the Aid of Computational Chemistry.

T. P. Straatsma, Free Energy by Molecular Simulation.

Robert J. Woods, The Application of Molecular Modeling Techniques to the Determination of Oligosaccharide Solution Conformations.

Ingrid Pettersson and **Tommy Liljefors**, Molecular Mechanics Calculated Conformational Energies of Organic Molecules: A Comparison of Force Fields.

Gustavo A. Arteca, Molecular Shape Descriptors.

Volume 10 (1997)

Richard Judson, Genetic Algorithms and Their Use in Chemistry.

Eric C. Martin, **David C. Spellmeyer**, **Roger E. Critchlow Jr.**, and **Jeffrey M. Blaney**, Does Combinatorial Chemistry Obviate Computer-Aided Drug Design?

Robert Q. Topper, Visualizing Molecular Phase Space: Nonstatistical Effects in Reaction Dynamics.

Raima Larter and **Kenneth Showalter**, Computational Studies in Nonlinear Dynamics.

Stephen J. Smith and **Brian T. Sutcliffe**, The Development of Computational Chemistry in the United Kingdom.

Volume 11 (1997)

Mark A. Murcko, Recent Advances in Ligand Design Methods.

David E. Clark, **Christopher W. Murray**, and **Jin Li**, Current Issues in De Novo Molecular Design.

Tudor I. Oprea and **Chris L. Waller**, Theoretical and Practical Aspects of Three-Dimensional Quantitative Structure–Activity Relationships.

Giovanni Greco, Ettore Novellino, and Yvonne Connolly Martin, Approaches to Three-Dimensional Quantitative Structure–Activity Relationships.

Pierre-Alain Carrupt, Bernard Testa, and Patrick Gaillard, Computational Approaches to Lipophilicity: Methods and Applications.

Ganesan Ravishanker, Pascal Auffinger, David R. Langley, Bhyravahotla Jayaram, Matthew A. Young, and David L. Beveridge, Treatment of Counterions in Computer Simulations of DNA.

Donald B. Boyd, Appendix: Compendium of Software and Internet Tools for Computational Chemistry.

Volume 12 (1998)

Hagai Meirovitch, Calculation of the Free Energy and the Entropy of Macromolecular Systems by Computer Simulation.

Ramzi Kuttah and T. P. Straatsma, Molecular Dynamics with General Holonomic Constraints and Application to Internal Coordinate Constraints.

John C. Shelley and Daniel R. Bérard, Computer Simulation of Water Physisorption at Metal–Water Interfaces.

Donald W. Brenner, Olga A. Shenderova, and Denis A. Areshkin, Quantum-Based Analytic Interatomic Forces and Materials Simulation.

Henry A. Kurtz and Douglas S. Dudis, Quantum Mechanical Methods for Predicting Nonlinear Optical Properties.

Chung F. Wong, Tom Thacher, and Herschel Rabitz, Sensitivity Analysis in Biomolecular Simulation.

Paul Verwer and Frank J. J. Leusen, Computer Simulation to Predict Possible Crystal Polymorphs.

Jean-Louis Rivail and Bernard Maigret, Computational Chemistry in France: A Historical Survey.

Volume 13 (1999)

Thomas Bally and Weston Thatcher Borden, Calculations on Open-Shell Molecules: A Beginner's Guide.

Neil R. Kestner and Jaime E. Combariza, Basis Set Superposition Errors: Theory and Practice.

James B. Anderson, Quantum Monte Carlo: Atoms, Molecules, Clusters, Liquids, and Solids.

Anders Wallqvist and **Raymond D. Mountain**, Molecular Models of Water: Derivation and Description.

James M. Briggs and **Jan Antosiewicz**, Simulation of pH-dependent Properties of Proteins Using Mesoscopic Models.

Harold E. Helson, Structure Diagram Generation.

Volume 14 (2000)

Michelle Miller Franci and **Lisa Emily Chirlian**, The Pluses and Minuses of Mapping Atomic Charges to Electrostatic Potentials.

T. Daniel Crawford and **Henry F. Schaefer III**, An Introduction to Coupled Cluster Theory for Computational Chemists.

Bastiaan van de Graaf, **Swie Lan Njo**, and **Konstantin S. Smirnov**, Introduction to Zeolite Modeling.

Sarah L. Price, Toward More Accurate Model Intermolecular Potentials For Organic Molecules.

Christopher J. Mundy, **Sundaram Balasubramanian**, **Ken Bagchi**, **Mark E. Tuckerman**, **Glenn J. Martyna**, and **Michael L. Klein**, Nonequilibrium Molecular Dynamics.

Donald B. Boyd and **Kenny B. Lipkowitz**, History of the Gordon Research Conferences on Computational Chemistry.

Mehran Jalaie and **Kenny B. Lipkowitz**, Appendix: Published Force Field Parameters for Molecular Mechanics, Molecular Dynamics, and Monte Carlo Simulations.

Volume 15 (2000)

F. Matthias Bickelhaupt and **Evert Jan Baerends**, Kohn-Sham Density Functional Theory: Predicting and Understanding Chemistry.

Michael A. Robb, **Marco Garavelli**, **Massimo Olivucci**, and **Fernando Bernardi**, A Computational Strategy for Organic Photochemistry.

Larry A. Curtiss, **Paul C. Redfern**, and **David J. Frurip**, Theoretical Methods for Computing Enthalpies of Formation of Gaseous Compounds.

Russell J. Boyd, The Development of Computational Chemistry in Canada.

Volume 16 (2000)

Richard A. Lewis, Stephen D. Pickett, and David E. Clark, Computer-Aided Molecular Diversity Analysis and Combinatorial Library Design.

Keith L. Peterson, Artificial Neural Networks and Their Use in Chemistry.

Jörg-Rüdiger Hill, Clive M. Freeman, and Lalitha Subramanian, Use of Force Fields in Materials Modeling.

M. Rami Reddy, Mark D. Erion, and Atul Agarwal, Free Energy Calculations: Use and Limitations in Predicting Ligand Binding Affinities.

Volume 17 (2001)

Ingo Muegge and Matthias Rarey, Small Molecule Docking and Scoring.

Lutz P. Ehrlich and Rebecca C. Wade, Protein-Protein Docking.

Christel M. Marian, Spin-Orbit Coupling in Molecules.

Lemont B. Kier, Chao-Kun Cheng, and Paul G. Seybold, Cellular Automata Models of Aqueous Solution Systems.

Kenny B. Lipkowitz and Donald B. Boyd, Appendix: Books Published on the Topics of Computational Chemistry.

Volume 18 (2002)

Geoff M. Downs and John M. Barnard, Clustering Methods and Their Uses in Computational Chemistry.

Hans-Joachim Böhm and Martin Stahl, The Use of Scoring Functions in Drug Discovery Applications.

Steven W. Rick and Steven J. Stuart, Potentials and Algorithms for Incorporating Polarizability in Computer Simulations.

Dmitry V. Matyushov and Gregory A. Voth, New Developments in the Theoretical Description of Charge-Transfer Reactions in Condensed Phases.

George R. Famini and Leland Y. Wilson, Linear Free Energy Relationships Using Quantum Mechanical Descriptors.

Sigrid D. Peyerimhoff, The Development of Computational Chemistry in Germany.

Donald B. Boyd and **Kenny B. Lipkowitz**, Appendix: Examination of the Employment Environment for Computational Chemistry.

Volume 19 (2003)

Robert Q. Topper, **David, L. Freeman**, **Denise Bergin** and **Keirnan R. LaMarche**, Computational Techniques and Strategies for Monte Carlo Thermodynamic Calculations, with Applications to Nanoclusters.

David E. Smith and **Anthony D. J. Haymet**, Computing Hydrophobicity.

Lipeng Sun and **William L. Hase**, Born-Oppenheimer Direct Dynamics Classical Trajectory Simulations.

Gene Lamm, The Poisson-Boltzmann Equation.

Volume 20 (2004)

Sason Shaik and **Philippe C. Hiberty**, Valence Bond Theory: Its History, Fundamentals and Applications. A Primer.

Nikita Matsunaga and **Shiro Koseki**, Modeling of Spin Forbidden Reactions.

Stefan Grimme, Calculation of the Electronic Spectra of Large Molecules.

Raymond Kapral, Simulating Chemical Waves and Patterns.

Costel Sârbu and **Horia Pop**, Fuzzy Soft-Computing Methods and Their Applications in Chemistry.

Sean Ekins and **Peter Swaan**, Development of Computational Models for Enzymes, Transporters, Channels and Receptors Relevant to ADME/Tox.

Volume 21 (2005)

Roberto Dovesi, **Bartolomeo Civalleri**, **Roberto Orlando**, **Carla Roetti** and **Victor R. Saunders**, Ab Initio Quantum Simulation in Solid State Chemistry.

Patrick Bultinck, **Xavier Gironés** and **Ramon Carbó-Dorca**, Molecular Quantum Similarity: Theory and Applications.

Jean-Loup Faulon, **Donald P. Visco, Jr.** and **Diana Roe**, Enumerating Molecules.

David J. Livingstone and **David W. Salt**, Variable Selection- Spoilt for Choice.

Nathan A. Baker, Biomolecular Applications of Poisson-Boltzmann Methods.

Baltazar Aguda, Georghie Craciun and Rengul Cetin-Atalay, Data Sources and Computational Approaches for Generating Models of Gene Regulatory Networks.

Volume 22 (2006)

Patrice Koehl, Protein Structure Classification.

Emilio Esposito, Dror Tobi and Jeffry Madura, Comparative Protein Modeling.

Joan-Emma Shea, Miriam Friedel, and Andrij Baumketner, Simulations of Protein Folding.

Marco Saraniti, Shela Aboud, and Robert Eisenberg, The Simulation of Ionic Charge Transport in Biological Ion Channels: An Introduction to Numerical Methods.

C. Matthew Sundling, Nagamani Sukumar, Hongmei Zhang, Curt Breneman, and Mark Embrechts, Wavelets in Chemistry and Chemoinformatics.

Volume 23 (2007)

Christian Ochsenfeld, Jörg Kussmann, and Daniel Lambrecht, Linear Scaling in Quantum Chemistry.

Spiridoula Matsika, Conical Intersections in Molecular Systems.

Antonio Fernandez-Ramos, Benjamin Ellingson, Bruce Garrett, and Donald Truhlar, Variational Transition State Theory with Multidimensional Tunneling.

Roland Faller, Coarse Grain Modeling of Polymers.

Jeffrey Godden and Jürgen Bajorath, Analysis of Chemical Information Content using Shannon Entropy.

Ovidiu Ivanciuc, Applications of Support Vector Machines in Chemistry.

Donald Boyd, How Computational Chemistry Became Important in the Pharmaceutical Industry.

Volume 24 (2007)

Martin Schoen, and **Sabine H. L. Klapp**, Nanoconfined Fluids. Soft Matter Between Two and Three Dimensions.

Volume 25 (2007)

Wolfgang Paul, Determining the Glass Transition in Polymer Melts.

Nicholas J. Mosey and **Martin H. Müser**, Atomistic Modeling of Friction.

Jeetain Mittal, **William P. Krekelberg**, **Jeffrey R. Errington**, and **Thomas M. Truskett**, Computing Free Volume, Structured Order, and Entropy of Liquids and Glasses.

Laurence E. Fried, The Reactivity of Energetic Materials at Extreme Conditions.

Julio A. Alonso, Magnetic Properties of Atomic Clusters of the Transition Elements.

Laura Gagliardi, Transition Metal- and Actinide-Containing Systems Studied with Multiconfigurational Quantum Chemical Methods.

Hua Guo, Recursive Solutions to Large Eigenproblems in Molecular Spectroscopy and Reaction Dynamics.

Hugh Cartwright, Development and Uses of Artificial Intelligence in Chemistry.

Volume 26 (2009)

C. David Sherrill, Computations of Noncovalent π Interactions.

Gregory S. Tschumper, Reliable Electronic Structure Computations for Weak Noncovalent Interactions in Clusters.

Peter Elliott, **Filip Furche** and **Kieron Burke**, Excited States from Time-Dependent Density Functional Theory.

Thomas Vojta, Computing Quantum Phase Transitions.

Thomas L. Beck, Real-Space Multigrid Methods in Computational Chemistry.

Francesca Tavazza, **Lyle E. Levine** and **Anne M. Chaka**, Hybrid Methods for Atomic-Level Simulations Spanning Multi-Length Scales in the Solid State.

Alfredo E. Cárdenas and Eric Bath, Extending the Time Scale in Atomically Detailed Simulations.

Edward J. Maginn, Atomistic Simulation of Ionic Liquids.

Volume 27 (2011)

Stefano Giordano, Alessandro Mattoni, Luciano Colombo, Brittle Fracture: From Elasticity Theory to Atomistic Simulations.

Igor V. Pivkin, Bruce Caswell, George Em Karniadakis, Dissipative Particle Dynamics.

Peter G. Bolhuis and Christoph Dellago, Trajectory-Based Rare Event Simulation.

Douglas L. Irving, Understanding Metal/Metal Electrical Contact Conductance from the Atomic to Continuum Scales.

Max L. Berkowitz and James Kindt, Molecular Detailed Simulations of Lipid Bilayers.

Sophya Garaschuk, Vitaly Rassolov, Oleg Prezhdo, Semiclassical Bohmian Dynamics.

Donald B. Boyd, Employment Opportunities in Computational Chemistry.

Kenny B. Lipkowitz, Appendix: List of Computational Molecular Scientists

Volume 28 (2015)

Giovanni Bussi and Davide Branduardi, Free-energy Calculations with Metadynamics: Theory and Practice.

Yue Shi, Pengyu Ren, Michael Schnieders and Jean-Philip Piquemal, Polarizable Force Fields for Biomolecular Modeling.

Clare-Louise Towse and Valerie Daggett, Modeling Protein Folding Pathways.

Joël Janin, Shoshana J. Wodak, Marc F. Lensink and Sameer Velankar, Assessing Structural Predictions of Protein-Protein Recognition: The CAPRI Experiment.

C. Heath Turner, Zhongtao Zhang, Lev D. Gelb and Brett I. Dunlap, Kinetic Monte Carlo Simulation of Electrochemical Systems.

Ilan Benjamin, Reactivity and Dynamics at Liquid Interfaces.

John S. Tse, Computational Techniques in the Study of the Properties of Clathrate Hydrates.

John M. Herbert, The Quantum Chemistry of Loosely Bound Electrons.

Volume 29 (2016)

Gino A. DiLabio and **Alberto Otero-de-la-Roza**, Noncovalent Interactions in Density Functional Theory

Akbar Salam, Long-Range Interparticle Interactions: Insights from Molecular Quantum Electrodynamics (QED) Theory

Joshua Pottel and **Nicolas Moitessier**, Efficient Transition State Modeling Using Molecular Mechanics Force Fields for the Everyday Chemistry

Tim Mueller, **Aaron Gilad Kusne**, and **Rampi Ramprasad**, Machine Learning in Materials Science: Recent Progress and Emerging Applications

Eva Zurek, Discovering New Materials via *A Priori* Crystal Structure Prediction

Alberto Ambrosetti and **Pier Luigi Silvestrelli**, Introduction to Maximally Localized Wannier Functions

Zhanyong Guo and **Dieter Cremer**, Methods for Rapid and Automated Description of Proteins: Protein Structure, Protein Similarity, and Protein Folding

Volume 30 (2017)

Andreas Hermann, Chemical Bonding at High Pressure

Mitchell A. Wood, **Mathew J. Chrukara**, **Edwin Antillon**, and **Alejandro Strachan**, Molecular Dynamics Simulations of Shock Loading of Materials: A Review and Tutorial

Balazs Nagy and **Frank Jensen**, Basis Sets in Quantum Chemistry

Anna Krylov, The Quantum Chemistry of Open-Shell Species

Raghunathan Ramakrishnan and **O. Anatole von Lilienfeld**, Machine Learning, Quantum Chemistry, and Chemical Space

Dmitri Makarov, The Master Equation Approach to Problems in Chemical and Biological Physics

Pere Alemany, David Casanova, Santiago Alvarez, Chaim Dryzun and David Avnir, Continuous Symmetry Measures: A New Tool in Quantum Chemistry

Volume 31 (2018)

Ulf D. Schiller and Olga Kuksenok, Lattice-Boltzmann Modeling of Multicomponent Systems: An Introduction

David M Leitner and Takahisa Yamato, Mapping Energy Transport Networks in Proteins

Paul N. Patrone and Andrew Dienstfrey, Uncertainty Quantification for Molecular Dynamics

Horia Metiu, Vishal Agarwal, and Henrik H. Kristoffersen, The Role of Computations in Catalysis

Richard Dawes and Ernesto Quintas Sánchez, The Construction of Ab Initio Based Potential Energy Surfaces

Heather J. Kulik, Modeling Mechanochemistry from First Principles

1

NON-DETERMINISTIC GLOBAL STRUCTURE OPTIMIZATION: AN INTRODUCTORY TUTORIAL

BERND HARTKE

*Theoretical Chemistry, Institute for Physical Chemistry, Christian-
Albrechts-University, Olshausenstr. 40, 24098 Kiel, Germany*

LIST OF ABBREVIATIONS

| | |
|--------|---|
| ACO: | Ant-Colony Optimization |
| BH: | Basin Hopping (Monte Carlo with Minimization) |
| CD: | Collision Detection |
| CD/DD: | combined Collision and Dissociation Detection |
| CSP: | Crystal Structure Prediction |
| DD: | Dissociation Detection |
| DFT: | Density Functional Theory |
| DFTB: | Density Functional Tight-Binding |
| DOF: | Degree of Freedom |
| DGO: | Deterministic Global Optimization |
| FF: | Force Field (empirical potential) |
| GA: | Genetic Algorithm |
| GO: | Global Optimization |
| HPC: | High-Performance Computing |

Reviews in Computational Chemistry, Volume 32, First Edition.

Edited by Abby L. Parrill and Kenny B. Lipkowitz.

© 2022 John Wiley & Sons, Inc. Published 2022 by John Wiley & Sons, Inc.

| | |
|-------------------|---|
| IR: | infrared |
| LJ _n : | Lennard-Jones cluster with n atoms |
| MC: | Monte Carlo |
| MCM: | Monte Carlo with Minimization (Basin Hopping) |
| MD: | Molecular Dynamics |
| ML: | Machine Learning |
| MP2: | Møller-Plesset perturbation theory, second order |
| MPI: | message-passing interface (parallelization) |
| NDGO: | Non-Deterministic Global Optimization (algorithms) |
| NFL(T): | No Free Lunch (Theorem) for search and optimization |
| PES: | Potential Energy Surface |
| PSO: | Particle Swarm Optimization |
| SA: | Simulated Annealing |
| wNFL(T): | Weak No Free Lunch (Theorem) |

INTRODUCTION

The Need for Structural Optimization

For any calculation of static or dynamic properties of molecular systems, the (starting) structures of these molecules need to be known, at the level of theory to be used. Since full configuration interaction in a complete basis set extrapolation (and possibly with explicitly treated solvent molecules, at finite temperature, etc.) is not practical, every level of theory is approximate, in the sense that its results differ from experimental data. Because properties may change appreciably even with small structure distortions, using experimental molecular structures is not a good idea. Instead, they always need to be optimized at the given level of theory, before any further calculations can commence.

Several decades ago, computational chemistry could only deal with small, isolated molecules. For these, local optimizations from guessed starting structures were sufficient. These starting structures could be taken from chemical intuition or from experiment. With gradients and frequencies (1st and 2nd derivatives of the electronic energy with respect to the nuclear coordinates) at the given level of theory, a local minimum-energy* structure can then be found, using efficient standard procedures that are general and system-independent. This is taken for granted now, but significant

*Maximization of $-f(x)$ corresponds to minimization of $f(x)$, and for molecules usually minima are more interesting than maxima. Therefore, “optimization” and “minimization” (and sometimes also the designation “search”) are used interchangeably here and in the literature.

development efforts were required,¹ as well as an acknowledgment of the need for all this in the computational chemistry community.

For three decades, we have experienced complex systems in theoretical calculations, giving rise to very many local minimum-energy structures. Frequently, experimental information is insufficient to disentangle the signatures from several different species, and human chemical intuition may also struggle, at least outside the area of standard organic chemistry. In such situations, approaches are needed that can find many or all local minima and locate the best one of these, i.e., global optimization (GO).

It is frequently assumed that it may suffice to substitute true global optimization by simply performing series of local minimizations, from different guessed starting structures. However, with increasing complexity, this naive approach quickly loses any reliability. This was so well-known in the GO community already in the 1990s that it took a surprisingly long time until Avaltroni and Corminboeuf demonstrated this explicitly for a real-life test case² in the published literature. Until today, however, this finding frequently remains underappreciated. Astonishingly, even today, papers can be published³ in which not even the need for global structure optimization is recognized.

Search Space is Vast

This failure of series of local minimizations arises because the search space of molecular structures is huge: It scales exponentially with the number of degrees of freedom (DOF). The essence of the reason for this scaling is depicted in Figure 1.

Unless *a priori* information allows for general restrictions, we have to combine every possible coordinate value in one DOF with every other value in all other DOFs, in direct-product style. This obviously leads to exponential scaling of the search space size to be covered, with the number of coordinates or particles. Thus, e^n scaling is a basic feature of n -dimensional space, combined with the need to cover all of it. This need arises from our wish to be sure to find the true global minimum and from our lack of global information: Without reliable *a priori* information on how the function $f(\vec{x})$ to be optimized behaves at a new point \vec{x} , we have to visit this new point (i.e., we need to evaluate $f(\vec{x})$ at \vec{x})—but as soon as we do this for all new points \vec{x} , we are stuck in the e^n trap.

Figure 1 illustrates how exponential scaling arises, but fails to illustrate how bad exponential scaling really is. Because even seasoned experts sometimes underappreciate this, it is provided by Table 1, which shows linear, cubic, and exponential scaling with the number of DOFs, in terms

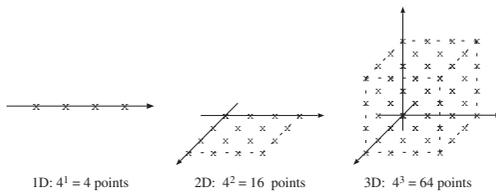


FIGURE 1 If 4 coordinate values per DOF have to be checked, a two-dimensional or three-dimensional search has to cover not 8 and 12 points (linear scaling) but 16 and 64, respectively (exponential scaling).

TABLE 1 Computation Wallclock Times (Rounded) to be Expected for Calculations With Varying Numbers of Degrees of Freedom, for Fictitious Algorithms That Scale Linearly, Cubically or Exponentially in the Number of Degrees of Freedom.

| # Degrees of freedom | Scaling | | |
|----------------------|---------|-----------|------------------|
| | Linear | Cubic | Exponential |
| 5 | 1 s | 1 s | 1 s |
| 10 | 2 s | 8 s | 2.5 min |
| 20 | 4 s | 1 min | 1 month |
| 40 | 8 s | 8.5 min | 50 million years |
| 80 | 16 s | 1 h 8 min | — |
| 160 | 32 s | 9 h | — |
| 320 | 1 min | 3 d | — |

of computational wallclock times. It starts with the assumption that a calculation for 5 DOFs needs one second, which is a completely arbitrary setting (but somewhat realistic, e.g., for quantum-chemistry calculations). The ensuing scaling with the number of DOFs (which also is chemically realistic, for electrons or nuclei), however, is very real and shows drastically that all exponentially scaling algorithms can be used only for small DOF numbers. Obviously, this situation cannot be changed fundamentally, even if the prefactor of the exponentially scaling algorithm can be made much smaller (i.e., if its wallclock time for 5 DOFs were much shorter than for the linear or cubic scalings). Note that things are even worse if we examine scaling in terms of particle numbers, with 3 DOFs per particle.

Deterministic vs Non-Deterministic Search

Since the above explanation how e^n scaling arises is so fundamental and abstract, it applies in most cases, regardless of the technical details and of the system under study. For the same reason, it is hard to find ways around exponential scaling, unless equally fundamental assumptions are dropped:

- A1: get rid of the idea to find the global minimum with certainty, and/or
- A2: deliberately avoid visiting every region of search space.

These two requirements are obviously interconnected. The only other ways out are

A3: to reduce problem dimensionality n , or

A4: to have the above-mentioned reliable *a priori* information on $f(\vec{x}_i)$, without ever calculating any of these function values $f(\vec{x}_i)$.

Problems with **A3** and **A4** are manifold: To counter exponential growth in any useful way, dimensionality reduction has to be big and the *a priori* information has to be certain and has to pertain to huge (and exponentially growing) parts of search space. Again, both items appear to be linked: In order to achieve any dimensionality reduction, one has to know that certain parts of the system under study are not relevant. Additionally, practical experience tells us that both strategies, if viable at all, will be system-specific^{4,5} and hence not transferable.

As a side note, the above seemingly convincing arguments are challenged by repeated findings that (some?) GO problems appear to become easier if embedded temporarily into a higher(!)-dimensional framework; for a recent example see Ref.⁶. The standard argument to rationalize such findings is that higher-dimensional spaces offer more low-barrier connections between local minima than lower-dimensional ones, making the higher-dimensional search space easier to navigate. Obviously, the above argument of Figure 1 still holds, i.e., the higher-dimensional space is (exponentially) larger. Hence, for this strategy to work, the advantage of a simpler search space has to overcompensate the size disadvantage. To the knowledge of the present author, it has not been shown yet that this works in a generalizable sense, beyond the particular cases studied so far.

In contrast, strategies **A1** and **A2** appear to be more transferable, but the extent to which they are useful may vary from system to system. Clearly, however, they have proven themselves to be very useful indeed, for a wide variety of concrete problems in many areas.

For the present text, sticking to finding the global minimum with certainty (and hence having to explore all of search space) corresponds to “deterministic global optimization” (DGO), while following strategies **A1** and **A2** is “Non-deterministic Global Optimization” (NDGO).*

There has been, and there continues to be, substantial development in DGO Algorithms.⁷⁻¹¹ However, from the above discussion, it follows that DGO has to scale exponentially, which confines its applicability to small problem instances. This is strikingly illustrated by the lack of substantial progress in problem size when DGO methods are applied to global structure optimization of atomic or molecular clusters, between earlier work¹²

*These definitions are widely but not universally used in this way.

on 7 particles and similar work 13 years later¹³ on 11 particles, while for the same problem, NDGO methods progressed from 4 particles¹⁴ to well over 150 particles¹⁵ in about half of this time. Because chemists are typically interested in particle numbers between 50 and 200 rather than 5 and 10, in real-life chemical applications GO is essentially always done by NDGO methods, and this will likely remain so for a few decades. Quantum computers may offer more practical prospects of treating exponentially scaling problems of nontrivial size in a deterministic fashion and they are making significant progress,¹⁶ but sadly they are still far away from routine, real-life applications. Therefore, the remainder of this chapter deals exclusively with NDGO approaches.

In contrast to DGO methods, NDGO approaches *have to* forgo any attempts to cover all of search space and to find the global optimum with certainty. If this is not done rigorously, NDGO mutates into DGO, but with less efficiency: NDGO methods are built on heuristic inference while DGO methods are designed to achieve complete search space coverage in the most efficient way. Hence, it is highly unlikely that DGO methods can be beaten on their home turf by NDGO methods that were designed for very different aims. And again, as shown in Figure 1 and Table 1, when facing the e^n cliff you desperately need all the efficiency you can possibly get. Nevertheless, many NDGO papers proclaim to strive toward complete search space coverage; as described, this is a fundamental misconception.

Another important corollary of this situation is that NDGO methods have no convergence criterion—at least not in the usual sense: some (approximate or rigorous) method to measure the difference between the currently best candidate solution and the true global optimum, or the rate of “approach to” the latter (quadratic, cubic, etc.) in terms of iterative step numbers. Both are available for standard local optimization algorithms,¹⁷ where they can be derived from the typical underlying quadratic approximation that also provides the iterative improvement recipes for those methods. Since this quadratic approximation simply is a (multidimensional) Taylor series, local optimization inherits the typical Taylor series generality. In global optimization, however, such very short, truncated Taylor series approximations cannot be employed in a general fashion, due to their strict locality. Nevertheless, in DGO methods, the degree of search space coverage (directly, or indirectly via upper and lower bounds to the function to be optimized) and/or these bounds themselves do provide at least an indirect measure for convergence. As argued above, even the idea of search space coverage as a proxy to estimate convergence is absent in NDGO methods (or should not be used). What can, and should, be done instead is discussed below. Sadly, again in many NDGO papers, this situation does not seem to be fully appreciated.

Of course, there are reasons for the continuing misuse of NDGO methods, as stated in the previous two paragraphs. Many researchers are used to getting a final answer from only one program run or from a short sequence of just a few program runs with well-defined differences between them (e.g., solving an integral or a differential equation with decreasing step sizes, until convergence, or executing a standard sequence of quantum-chemistry runs with increasing method sophistication until a proper balance between accuracy and computer time is found). This is not possible in the NDGO area, and the above-mentioned deviations from proper NDGO usage probably are futile attempts to get back to more familiar situations.

Fundamental Take-Home Lessons

Despite the complete lack of technical details in this first section, several insights could already be obtained that are of great importance for using NDGO methods in practice:

- DGO methods
 - deliver global optima with certainty
 - and come with convergence criteria (in incomplete runs);
 - but they have to cover all search space, hence they cannot avoid exponential scaling with problem size. This makes them impractical for most GO problems of interest in chemistry.
- NDGO methods
 - are not designed for full search space coverage and should not be used in this way,
 - hence their results are never certain
 - and they also lack a convergence criterion;
 - thus using them properly is somewhat more difficult than it may seem at first,
 - but they are the only possibility, since series of local optimizations will fail for all but the smallest problems.

A CLOSER LOOK AT SOME NDGO BACKGROUND DETAILS

Too Inspired by Nature

Despite good introductory literature on heuristics,¹⁸ finding good heuristics for a given problem may seem to be a daunting task, particularly if new

problems require new heuristics (cf. the “no free lunch” (NFL) theorem in section “No Free Lunch”). Maybe for this reason, problem-external “inspiration” has been a part of NDGO development from the start. For example, simulated annealing (SA)¹⁹ was likened to the annealing process in metallurgy and genetic algorithms (GA)²⁰ to natural evolution. While the first of these two analogies is physically largely correct, the second one may seem natural at first but becomes somewhat dubious upon closer inspection.²¹ Most biochemists or molecular biologists would not see sufficient similarities between GAs and what really happens with DNA, RNA, and proteins upon cell division. And at a far more coarse-grained level, natural evolution of species is not a GO at all; to survive and procreate you do not need to be optimal in any sense, it clearly suffices to be better than your “neighbor” (in a geographic sense, or in the competition over resources, or in evading predators). So why should something that provides winning (and losing) outcomes in repeated neighbor contests be a good blueprint for *global* optimization heuristics? Hence, it would be quite valid to ask if “genetic algorithm” is a good name at all.

Things become much weirder with an arbitrary, quick look at the NDGO literature, which abounds with strange new algorithm names. Even Wikipedia²² provides an astonishing list of them, containing glowworm swarm optimization (not to be confused with particle swarm and cat swarm optimization), shuffled frog leaping algorithm, imperialist competitive algorithm, intelligent water drops algorithm (not to be confused with rain water algorithm), flower pollination algorithm, duelist algorithm, etc. On Scholarpedia, Glover and Sörensen²³ state:

The list of natural or man-made processes that has been used as the basis for a metaheuristic framework now includes such diverse processes as bacterial foraging, river formation, biogeography, musicians playing together, electromagnetism, gravity, colonization by an empire, mine blasts, league championships, clouds, and so forth. An important subcategory is found in metaheuristics based on animal behavior. Ants, bees, bats, wolves, cats, fireflies, eagles, vultures, dolphins, frogs, salmon, vultures, termites, flies, and many others, have all been used to inspire a “novel” metaheuristic.

This is the product of uninhibited proliferation across 20–30 years. More recently, however, this practice was heavily criticized^{23–25} as being actually harmful to the field of NDGO research. The arguments are not only that it is hard to sell something with funny names as serious science but also that hiding algorithm design ingredients behind “nature-inspired” nomenclature makes it unnecessarily hard to compare different (or perhaps not-so-different) algorithm designs. It was even shown²³ that, at least in

one case, a seemingly “new” algorithm (coupled to a new “inspiration”) only was an old algorithm dressed up in new language. Additionally, the argument given in the previous paragraph also still applies: It is by no means clear *a priori* why and how colonization by an empire or the behavior of termites should have anything to do with GO, neither in an absolute sense nor as an improvement over inspiration from river formation or swarms of fireflies. However, instead of discussing such questions, most papers of this kind only present their “inspired” algorithm and some tests of it. Therefore, such “inspirations” may also serve as an excuse for not discussing why algorithm design choices were made in just this way.

Nevertheless, maybe a few important insights can be distilled from this otherwise embarrassingly wild mushrooming of “inspired” NDGO algorithms: If we assume (without actual tests) that at least some of these algorithms are at least somewhat different and that at least some of these performed quite well in at least some of their benchmarks or applications, then a possible conclusion could be that certain algorithm design choices are not very critical. For example, all swarm-based algorithms and most “animal-inspired” ones operate by iteratively improving not just one candidate solution but a whole set of simultaneous candidate solutions, with some sort of “information exchange” between them. Differences between those algorithms usually pertain to small details of how this information exchange is done and how it influences the next moves of the candidate solutions. The mere existence of many different ways of arranging these details could simply indicate that these details are not very important, compared to other algorithm design decisions (like using one vs many candidate solutions). Alternatively, they could merely reflect the need to micro-adapt one basic algorithm framework to the varying needs of different GO problem realizations (cf. the next section “No Free Lunch”). In both cases, however, clarity would be better served by not giving different names to such algorithm variations.

Actually, several frequently used algorithm design ideas can be readily justified without recourse to “inspiration from nature”: In the GA-concept of “crossover”, two candidate solutions are each partitioned into two or more pieces, and then, new candidate solutions are formed by exchanging some of these pieces between the old candidate solutions. Additionally, it can be verified easily that a GA with crossover normally performs better than one without and that both versions normally perform better than a simplistic random search^{2, 26}; in other words, crossover really enhances GO efficiency. Obviously, this can be fairly consistently true only if such a mixed reassembly of candidate solution pieces has a significantly nonzero chance of producing a new candidate solution that is at least as good or even

better than the old ones that provided these pieces. In other words, a piece of a good candidate solution is a good piece, and assembling two or more good pieces may result in an even better complete candidate solution. This is possible only if the problem is approximately (or exactly) separable. Hence, crossover exploits (approximate) separability.

In a similar vein, “mutation” in GAs (small parts of a candidate solution are changed, to produce a new one) and the basic move in Monte Carlo with minimization (MCM) or basin hopping (BH) are designed to get out of the present local minimum (basin of attraction) and into another one nearby. This exploits the frequently observed search space structure in which local minima occur in clusters, resulting in rough funnels. The roughness prevents local optimizations from reaching the bottom of the funnel, while this is less of a challenge for GO-algorithms incorporating such mutation-like moves.

As a final example, both BH and SA use the standard Metropolis criterion to allow not only downhill moves but also uphill moves, with a probability that decreases exponentially for moves that would lead more and more uphill. This assigns a nonzero possibility to climbing out even from the bottom of one funnel into the next one, which may then turn out to be even deeper. The strong imbalance between downhill moves (always allowed) and uphill moves (exponentially improbable) ensures that there still is a strong tendency to find the deepest sub-minimum in each funnel (after all, we want to find minima, not maxima) and that the overall algorithm is better than a random walk (which we would get if both uphill and downhill moves would be always accepted). My personal suspicion is that this particular uphill weighting (Metropolis) is used only because it is familiar from Monte Carlo (MC) simulations and generates a Boltzmann distribution (canonical ensemble) which seems to be reasonable for chemists and physicists. To my knowledge, it has not been tested if this really is the best choice for GO (instead of generating a canonical ensemble). I would not be surprised if it turned out that the uphill probability weighting should also vary with search space structure, upon addressing new problems.

Clearly, some of these explanations are (partially) speculative but compared to “inspirations by nature” they have the advantage of being testable and of providing a much more reasonable foundation for rational NDGO algorithm design.

No Free Lunch

As alluded to above, in typical applications in the natural sciences, all local optimizations of well-defined functions are created equal since they can be modeled with sufficient accuracy by short, truncated Taylor series. Hence,

they all have the same fundamental characteristics, can be treated with the same set of standard algorithms, and these algorithms themselves¹⁷ can be ranked according to their performance, in a largely problem-independent way. A user can then simply pick the best-performing algorithm from this list, possibly also taking into account secondary criteria (availability of 1st and/or 2nd derivatives, memory demands, etc.), and rely on getting the promised performance—possibly diminished somewhat only when starting points are chosen in which the truncated Taylor series is a much worse approximation than elsewhere, but even such problems often tend to disappear after several iterations. Note that the core achievement of such a truncated Taylor series model is to get approximate but sufficiently good information on what the function to be optimized does at points some distance away from a point \vec{x}_i where we are now, based on information obtained solely at \vec{x}_i (and possibly at \vec{x}_{i-1} etc.), in a problem/function-independent fashion. This works because these distances are small, and because the truncated series is able to cover these small distances well enough (and in a problem-independent way), providing sufficient quality assurance for the next iterative step.

For GO, the situation is fundamentally different—for obvious reasons. However, even seasoned GO practitioners tend to forget these differences and their consequences sometimes, hence we emphasize them here very strongly. When we stick to this picture of truncated Taylor series approximations to the function to be optimized, it is indeed obvious that the approximation quality of a series truncated at any finite order will degrade by construction upon venturing farther and farther away from the point about which the Taylor series was constructed, even if the radius of convergence of the series is technically infinite. In other words, information on the function to be optimized (values of the function itself and its derivatives) that is confined to a finite, local region cannot tell us something definite about the behavior of this function (far) outside of this local region, in the absence of strong, additional information on this function (e.g., that it is a sine function times an exponential function). The example of a Taylor series is an extreme one since the local information pertains to one point only, but the core argument remains the same if we relax this to a region of actually finite size. Therefore, for GO there is no general, problem-independent way of extracting information about the behavior of our function at some distance away from our local region—ultimately because these distances need to be much longer and because attempting to build a general iterative optimization scheme using a Taylor series truncated at an order far beyond the quadratic one is highly impractical (or even impossible, if the true Taylor series at the given point has a small, finite convergence radius).

Therefore, instead of the problem-independent universality of general algorithm performances in local optimization, we have the opposite thing in GO: We should expect dependence of algorithm performance on the problem at hand or on problem classes. With some GO experience for a given problem class, we may indeed be able to extrapolate far-away function behavior from local information, but chances are that this knowledge will become partly or completely useless upon transition to a totally different problem class. Given that search landscape characterization has come up with vastly different landscape types (single funnels that are smooth or rugged, multiple funnels, flat “golf courses” with widely spaced, small holes, and several more^{4, 27, 28}), it is intuitively clear that this is what we should have expected.

Additionally, this vague expectation has been cast into a full-blown theorem: the “no free lunch” theorem (NFLT)²⁹ for search and optimization. It essentially states that for the collection of all possible GO problems, no GO algorithm is better on average than any other one. In other words, if algorithm A beats algorithm B on average for a subset of all possible problems, algorithm B will beat algorithm A by the same margin on average over all other problems.

Note that “algorithm B” could well be blind, random search. Then, the NFLT appears to be quite depressing since it *seems* to tell us that even if we are very clever NDGO algorithm designers we will never come up with a powerful “algorithm A” that consistently beats stupid, blind, random search.

However, glaring loopholes are offered by the words “on average” and “all possible problems”. Upon second thought, the latter obviously contains problems of little or no practical interest to us, for example finding the global minimum in a 42-dimensional hypervolume of random noise or finding the leftmost point on a finite line segment, but also these problems have to be part of the problem performance averaging in NFLT. Thus, it can be sufficient for us to have an efficient algorithm A that consistently beats algorithm B (and/or random search) for the few problem classes of our interest while we quietly ignore all other problem classes. This would not violate the NFLT.

Such arguments have led some people to believe (or to hope) that while NFLT is nice, it has no practical value. Based on the arguments at the beginning of this section and on the empirical fact that problems of interest have been shown to have vastly different search space features, I tend to take a compromise position, called “weak no free lunch theorem” (wNFLT): Since there certainly are whole problem classes we can safely ignore, NFLT does not tell us that it is futile to attempt any NDGO algorithm development. However, NFLT does tell us that some of the fundamental expectations we might have (and which have been used frequently in many NDGO

publications) may not be justified. A few examples are mentioned in the remainder of this section.

If NDGO algorithm A performs better on problem X than algorithm B, it is not *a priori* clear that A will also beat B on problem Y—very definitely for full NFLT, but with a high probability also for wNFLT. Nevertheless, many published NDGO papers (for example, Ref.³⁰) make such bold claims implicitly or explicitly, or even the bolder one that if A beats B on problem X, it will beat B on every other problem, too. This would imply that the notion of a “universally best” NDGO algorithm is a valid one, but wNFLT tells us otherwise.

Likewise, striving for “unbiased” search seems to be a naturally valid aim in the NDGO area, and I am guilty of having advocated this in some of my key papers (e.g., in Refs.^{31–35}). However, both from wNFLT and from the discussions in sections “Search Space is Vast” and “Fundamental Take-Home Lessons” (i.e., we do *not* want to cover all search space with NDGO methods) it follows that on the contrary we *have to* be biased to avoid the e^n scaling cliff and to perform better than blind, random search (which is futile²¹) on the selected problem classes of our choice. Putting it bluntly, a search heuristic is a bias. Hence, we should not avoid bias, but we should avoid wrong biases for the problem at hand. How do we determine this? We may arrive at higher-level insights that help us here, but at the end of the day, we even should check them directly: Does our favorite NDGO algorithm equipped with our cleverly designed search-speedup bias A really perform better than without it? (remembering wNFL, you do not (only) want to do this for benchmark problems, but in particular (also) for several representative instances of your target problem). Or we can pit A against opposite-of-A, to lift the desired performance difference signal farther above the inevitable noise.

NDGO Algorithm Comparisons

Comparisons between NDGO algorithms have been rare,^{30, 36} and fair comparisons have been essentially nonexistent. Sadly, this will likely remain so, for several reasons. As discussed in the previous section “No Free Lunch”, even with wNFLT it likely makes little sense to ask what “the best” NDGO algorithm is; presumably, there only is a best algorithm for each and every rather narrow problem class.

As further illustrated later (in section “Brief Summary of some Fundamental NDGO Algorithm Ideas”), many NDGO algorithm ingredients can be combined at will in a “plug and play” fashion. Additionally, NDGO algorithms come with several parameters (some have more than others) to

“tweak” their action in a more fine-grained fashion. Hence, striving to compare “pure NDGO algorithm prototypes” is not a well-defined aim. Is algorithm X still the same algorithm after several changes and tweaks? Would anyone ever use the unchanged/untweaked version of algorithm X in practice? How well does the performance of the “pure” algorithm predict its performance after changes and tweaks? Instead you want to modify and tweak your NDGO algorithm so that it works best for your current problem. In this sense, ease of use is not a very relevant criterion; an algorithm with more tweak options should be better than one with less, which is the opposite of what was concluded in, e.g., Ref.³⁰.*

Algorithm comparisons are also hard to make because it is not clear on what to test. There is a big collection of multimodal, high-dimensional functions out there that is frequently taken as test set, but as we have shown³⁷ almost all of these functions are *not* challenging enough, compared to real-life GO application cases of interest (in chemistry and elsewhere). A reason for this observation is that these functions are defined via fairly simple, closed-form expressions. This automatically imposes a high degree of regularity on the (many) minima, which in turn is easily exploitable by any decent NDGO algorithm. Also, for the same reason and by (w)NFLT, good performance on these abstract test functions does not necessarily translate into good performance for a given application area. Instead, it appears to be more meaningful to do method development and method comparisons for benchmark cases close to the respective application area—for example, by doing benchmarks on Lennard-Jones³⁸ or water clusters if the application area will be atomic/molecular clusters.

Very rarely, our finding³⁷ that those “standard” global optimization benchmark functions simply are not challenging enough is explicitly acknowledged in the literature and better benchmark functions are proposed,³⁹ often similar to those we and others proposed.³⁷ Much more frequently, however, our benchmark paper³⁷ is cited for strangely wrong reasons.^{36,40} Sometimes it is even cited as a source for those benchmark functions^{41,42} which are then used for this purpose, although we did not invent these functions and pointed out that exactly this way of using them is not a good idea.

Barrier Crossing

An important subtask in global minimization is to get *out* of a current minimum quickly (because it very likely is not the global minimum) and

*Sorry, Stefan, for picking on your nice paper here.

into a new one. By definition, there is an energy barrier between these minima. Hence, when visualizing this process as a physical one at the atomistic/molecular level, crossing this barrier becomes exponentially harder with increasing barrier height. Hence, whole NDGO algorithm classes (MCM,⁴³ BH^{44, 45} and several further variations of this idea) focus on this particular subtask. In this context, it has been argued^{44, 45} that combining (“hybridizing”) such an NDGO strategy with local minimization at each iterative step effectively transforms the original potential energy surface (PES) into a different one, in which the usually smooth hills and valleys have been replaced by a rectangular “staircase” form and (much more importantly) some of the energy barriers have disappeared. This argument concludes with the statement that this elimination of (some, not all!) barriers obviously makes it easier to transition from one minimum to the next (as seen from the original PES) and that this explains the success of MCM/BH-like NDGO strategies.

This “staircase transformation” can work (in the sense of having this simplifying effect) only if the Metropolis criterion used for the MC steps is applied *after* the local minimization, not before it. This in turn means that each and every attempted MC step has to be followed by a local minimization (otherwise, steps rejected by the Metropolis criterion would not need to be followed by a local minimization). However, local minimization (by whichever algorithm) is computationally very costly compared to an MC step since it is an iterative procedure involving many evaluations of the target function, its gradient (and possibly 2nd derivatives). As a rule of thumb, in such global-local hybrid algorithms, 95–99% of the overall computer time is spent in local minimizations, even for cheap force fields (FF). Therefore, there is a significant prize to be paid for this “staircase transformation”. Frequently, however, this turns out to be a good investment.

Interestingly, NDGO methods share this aim of “getting across barriers more quickly than normally” with other areas of method development, mostly pertaining to calculations of free energies, e.g., accelerated molecular dynamics (MD), metadynamics, hyperdynamics, etc.^{46–48} Unsurprisingly, also several other important concepts of these areas overlap or are the same, for example “collective variables” or “order parameters”, “niches”, “broken ergodicity” and “deceptive landscapes”. Because these two areas (NDGO and free-energy calculations) are largely out of touch from each other, this potentially common vocabulary actually is used in a quite disparate way. Collective variables and order parameters (as low-dimensional description of structural changes upon going from one state (or PES minimum/funnel) to another) firmly belong in the free-energy camp. Deceptive landscapes and broken ergodicity describe the same thing

(a PES with a narrow funnel containing the global minimum, separated by a sizeable barrier from another, wide funnel that dominates the distribution at least at higher temperatures) with different names, for NDGO and for free energies, respectively. Niches are a NDGO (or more accurately, GA/evolutionary algorithms (EA)) concept to fight against the domination of wide funnels, somewhat akin to what metadynamics does for free energies. On some rare occasions, at least this sharable nomenclature is acknowledged and transferred.^{38,49} However, quite obviously a more intensive exchange between these areas would benefit both of them.

For a chemist, quantum-dynamical tunneling “through” a barrier is an everyday concept. Therefore, it is not surprising that quantum tunneling was frequently (re-)proposed (e.g., in Refs.^{50, 51} but also in many others) as a clever trick to enhance NDGO performance by making it easier to pass from one minimum to the next. Typically (as in the papers just cited), some degree of success for these barrier-tunneling algorithms was reported, in comparison with NDGO algorithms that have to “surmount” barriers. Characteristically, Gregor and Car⁵¹ found good performance in locating global minima of Lennard-Jones clusters with n atoms (LJ_n), but could not solve the case $n = 75$, for which they reported “ergodicity problems”. Hence, this performance is qualitatively on par with many other NDGO studies that had similar problems with this (and other) notoriously deceptive two-funnel cases (several are cited in the introduction of Ref.³⁸)—although the premise of these quantum-annealing algorithms was that such cases should become significantly easier to solve, by tunneling from one funnel into the other. Note that several other publications^{15, 38} (including earlier ones) managed to solve the $n = 75$ case readily, using niching instead of tunneling. More recent quantum-annealing overviews actually conceded that “there are drawbacks to quantum annealing” and that “it is an important problem to reveal a difference in performance between quantum annealing and classical SA”.⁵² The current expectation is that quantum annealing performed on a quantum computer may be the winning combination,^{52, 53} once sufficiently many qubits become available in real-life quantum-computer hardware. Once this happens, however, then also NDGO algorithms without tunneling may receive a significant performance boost, or even DGO approaches may become practical for problem sizes of practical interest.

Independent of quantum computers, I personally think that while tunneling through barriers may appear useful, a more fundamental premise of any “quantum”-based algorithm largely destroys its prospects for any GO task: In some way, either directly or approximately (e.g., via path integrals or ring-polymer dynamics), a Schrödinger equation has to be solved, with the PES-function to be globally optimized in the potential energy term,

resulting in a ground-state wavefunction which should have a significant amplitude maximum at the global PES minimum. However, this wavefunction is defined over all space, so we have to calculate its value everywhere, also at points at which it will then turn out to be essentially zero because the potential value is very large there—which we cannot know in advance, though. Therefore, according to section “Search Space is Vast”, such a calculation has to scale exponentially. Even worse, already checking/calculating values of the potential itself on a dense grid of points across all space scales exponentially. And after we have done that, why should we go on calculating, to construct a wavefunction and then check its maxima? We could simply relax the lowest potential value we encountered to its proper off-grid position and be done, without having obtained a single wavefunction value and hence without any quantum-based tunneling whatsoever.

At first, things seem to be different if we do not aim at obtaining a stationary ground-state wavefunction but rather start with a very compact, initial-guess wavefunction (in one minimum basin or across several of them) that is strictly zero elsewhere, and then follow its explicit time development. Such a wavefunction could then really tunnel (time-dependently) through surrounding barriers, successively finding new minima farther away from its starting region. To avoid having to represent this wavefunction again on an exponentially scaling grid covering all space, established methods could be used that spread a compact support of grid points or local basis functions only to regions where the evolving wavepacket currently is (or will be) nonzero,^{54–56} even in tunneling situations. However, in contrast to misleading textbook cartoons, in such a case (and almost generally) tunneling does not proceed by the wavepacket gradually disappearing in front of the barrier and reappearing behind it; instead, the wavepacket moves gradually through the barrier,⁵⁴ hence we also have to calculate potential energy values in the barrier region, in contrast to a more naive expectation. Additionally, as it is the nature of essentially all wavepacket movement, the initially compact wavepacket will want to spread out in all directions (with and without tunneling), and this alone will guarantee exponential scaling. Finally, as above, also for this initially compact wavepacket, potential energy values have to be pre-calculated at points to which the wavepacket may spread next, so that the ensuing calculation of the new wavepacket values themselves feels like an unnecessary add-on. Putting it succinctly, even for a time-dependent quantum-mechanical wavefunction, there is no way to “pre-conceive” far-away regions of low potential energy values, without actually calculating these potential energy values (and all in-between, and everywhere else).

By the way, note that it is *not* clear that non-tunneling algorithms like MCM/BH really have to go “over” all barriers in the intuitively assumed sense: They proceed by MC-style “jumps” that only care about their initial and final points, not about anything in-between. And a jump can find the next minimum only if it has left the catchment region of the first minimum, i.e., if it landed at least somewhat “beyond” the barrier. Hence, even if the jump could see the full barrier (i.e., in absence of the staircase transformation), it would always appear lower than the barrier top really is. For long jumps and/or narrow barriers, these barriers may even disappear almost completely, independent of their heights. This may contribute to the repeated findings that algorithms that incorporate “quantum tunneling through barriers” apparently have a smaller performance advantage than expected over those employing “classical” barrier crossing.

Old vs New Machine Learning

Currently, machine learning (ML), “deep learning”, neural networks and similar methods are used everywhere for everything, frequently hailed as the final answer to all problems and surrounded by a flood of mediocre literature.⁵⁷ Previously, in the 1980s–1990s, there was a similar but smaller hype around EA methods. In fact, back then, EA methods were also called “machine learning” occasionally. However, this should not fool us into assuming that ML merely is a reincarnation of EA or that the two are closely related, as sometimes stated.⁵⁸ In fact, these two directions can be construed as partially orthogonal: In essence, ML combines suitable representations (or possibly several layers of representations) with brute-force big-data interpolation, employing interpolation functions that are intentionally very flexible. Crucially, before any ML can start, these big data have to be available already and will be used in their entirety. Hence, by construction, ML has a prescribed field of applicability, spanned by these input data, to which it should stay confined. Turning an interpolation into an extrapolation (beyond the realm of given data) has always been highly dangerous, and this is still true if the interpolation method bears fancy, modern names (and even more true if the interpolation function is highly flexible). In contrast, EA-like methods (or more generally, also other NDGO algorithms) are all about inferential and extrapolating leaps into the (yet) unknown, purposefully starting from small data and typically avoiding to turn them into truly big data.

After having stated these conceptual differences, it is fair to point out that there also are commonalities. Choosing a suitable representation surely also is helpful for NDGOs. And of course, the aims of NDGOs and ML can also

be seen as closely related: In both cases, we try to infer “new information” from what we already have. Finally, in both cases and elsewhere, extrapolation reliability improves substantially if *a priori* information is available on what to expect in as yet uncharted regions (which implies that it is of course possible to design ML versions with extrapolation in mind, necessarily by restricting functional flexibility).

While there surely are many impressive ML successes, even in hype phases we should remember that true complexity cannot be tackled by one single approach, across all conceivable application areas. In this sense, I expect considerable benefits from combining modern ML techniques with “old-fashioned” nondeterministic search. In my opinion, this is one possible answer to the lament⁵⁸ that “the hundreds of thousands (or more) data points required for deep learning will be unavailable in many drug discovery projects. Alternative methods for equally robust feature extraction and hypothesis generation from small data sets need to be identified.” (similar difficulties exist in many other ML-application areas). In fact, a few steps in this direction have already been made.^{59–64}

Take-Home Lessons for NDGO Background Details

- NDGO algorithms continue to be advertised as “inspired by nature”, but this should stop since it has become detrimental to the whole field (e.g., it introduces apparently huge differences between closely related algorithm variations)
- NFLT tells us that all NDGO algorithms perform the same when averaged over all possible application cases. This may not be strictly applicable if we are interested only in a subset of (related) application cases, but it reminds us that the seemingly clear notion that one NDGO algorithm is “always” better than another one may be ill-founded and that adapting a standard NDGO algorithm to the specific case at hand may reap significant performance benefits.
- “Unbiased” seems to be a nice-to-have feature, but NDGO algorithms have to embrace bias to avoid being DGO algorithms. Test your bias ideas to find the good ones for your problem.
- For the above NFL reasons, comparing NDGO algorithms is not easy and the results of such comparisons probably are less transferable than we think. Malleability of an algorithm may be more important than benchmark performance (given some degree of the latter).
- There are interesting relations between global search and free-energy calculations, calling for fruitful cross-fertilization.

- Quantum tunneling through barriers only seems to be an attractive idea for NDGO.
- NDGO and ML have some similarities but also differences; therefore, combinations between the two will be fruitful.

GENERAL GUIDELINES FOR NDGO APPLICATIONS

Brief Summary of Some Fundamental NDGO Algorithm Ideas

Given the dazzling array of “inspired” NDGO methods, any attempt to really survey them is too painful. Instead, as reference for the following text, here is a brief reminder of a few key ingredients and their relations to some of the main methods:

Randomness: Obviously, fully random, blind search is not a good idea²¹ if our problem has *some* (albeit yet unknown) structure. And zero randomness is rarely realized in NDGOs; even the otherwise fully deterministic MD-SA usually starts from randomly selected points. Most of the successful NDGO paradigms are somewhere on the broad spectrum in-between.

Move class: How to change one trial solution into a modified one at each iterative step; usually with some randomness thrown in.

Population size: NDGOs can operate by iteratively improving (or at least changing) a single trial solution (population size 1), or several of them simultaneously (population size > 1). Obviously, the latter differs from several independent runs of the former only if there is some “on the fly” information exchange between the trial solutions within the population.

Kind of information exchange: Typical choices are (i) cut trial solutions into smaller pieces and exchange some of these pieces between different trial solutions (“crossover” in the GA/EA language), (ii) interpolate between trial solutions. Of course, this overlaps with the move class concept, to some extent.

Hybridization with local search: This can be done always (at each NDGO step), sometimes or never, and in principle local optimization also can or cannot change the current location of the trial solution in search space (although the former seems to be the typical choice now).

Memory: So-far-best function values of iteratively changing trial solutions or of the whole population can or cannot be stored and used internally. Of course, it would be self-defeating *not* to store the best-so-far function value externally (even if it is not used internally).

Diversity: In NDGOs with population size >1 , design attempts to deliberately prevent several or all trial solutions from meeting in a single minimum basin *not* escaping again.

For example, if the population size is 1 and moves are random (MC) jumps with a Metropolis criterion, local search is done at each step (and changes the trial solutions' position) and there is no memory, you have MCM/BH. If the population size is >1 , moves are random MC-style jumps (typically without a Metropolis criterion) of single trial solutions ("mutation") and exchange of pieces between (usually two) trial solutions ("crossover"), you have a basic GA or EA (to which local search, memory ("elitism") and diversity ("niches") can be tacked on). In contrast, if random "mutations" are replaced by a more "kinetic movement" and "crossover" by influencing this movement by an individual and collective memory, you land somewhere in the broad array of swarm-based NDGOs (particle-swarm optimization (PSO), ant-colony optimization (ACO), etc.).

NDGO Method Design Choices

As to be expected from sections "No Free Lunch" and "NDGO Algorithm Comparisons", fair comparisons between NDGO algorithms are rare and hard to make, and their outcome likely is not general but problem-specific. Nevertheless, general superiority of algorithm X over algorithm Y has been claimed on many occasions, typically by developers of algorithm X—which is psychologically understandable but tends to confuse potential users that are not NDGO experts.

To take some heat out of the debate and to provide some user guidance, it is a good start to realize that some apparently controversial topics may be less important (or even nonissues) and/or can be decided via other considerations.

For example, in the early decades of GA/EA developments and applications, it was the accepted standard that trial solutions ("individuals") had to be "represented" in some abstract fashion (closer to the presumed biological "genetic" inspiration of DNA strings). Long serious debates discussed the optimal nature of this representation, concluding that, e.g., strings of binary

numbers were better than strings of decimal numbers,* because this would provide more leverage to “crossover” and “mutation” operators. However, this discussion gradually disappeared at the end of the millennium, and in cluster structure optimization, it became standard to operate on real-number strings or even directly on cluster (sub)structures in their 3D-space (“no representation”, also known as “phenotype representation”^{15, 65–67}). With hindsight, it is clear that it should be possible in principle to reproduce exactly what operator A does in representation X, upon a switch to representation Y, if the operator is changed accordingly. Or, viewed differently, certain operations are easier to realize in one representation than in another one. However, an additional advantage of phenotype operators acting directly on clusters in their 3D-space is that they retain the 3D-neighborhood relations already present there—which is quite hard to realize in a 1D string representation: To achieve this in a robustly transferable fashion, advanced concepts like “linkage learning”^{68, 69} have to be invoked, which come with their own bag of additional problems.

It is frequently stated^{70, 71} that NDGO algorithms need to maintain a proper balance between exploration (venturing into new regions of search space) and exploitation (finding even lower points in the present basin or funnel). While this is obviously true, it is frequently difficult to pin down which algorithm features are responsible for which of the two and to which extent, when other algorithm ingredients than those of the move class are examined.⁷² Also, qualitative search space features may change from one problem instance to another, and this should require a concomitant adaption of the exploration–exploitation balance—but this can hardly be known in advance, before attempting to solve the problem instance. Therefore, exploration vs exploitation is a useful qualitative description, but hardly a quantitative algorithm design prescription. In extreme cases, too much exploitation can result in the “premature convergence” already mentioned above. A simple expedient against that is to check not only the single best end result of a long run, but to keep watching how the “population” (best, medium, worst trial solutions) changes during the run, at least when applying NDGO to a new problem instance.

An important part of “exploitation” is to find the bottom of a single (smooth) basin of attraction. The repeated finding is that NDGO algorithms are not good for that at all—unless combined (“hybridized”) with local optimization (which MCM/BH always is, by construction). As pointed out above, however, then >95% of the overall computational effort goes into this local search—which is likely wasteful in the initial stages of the

*Both representations hardly have any relation to how proteins are represented on DNA strings.

global search, where all local minima found are much worse than the global minimum. Hence, if possible, switching off local search in the initial stage tends to be a good idea that can save substantial computer time without deteriorating the final results too much. As discussed in section “Barrier Crossing”, skipping local search forgoes the potential advantages of the “staircase transformation”, which may be tolerable if search jumps are large enough to cross barriers easily. In the “end game”, doing local search ceases to be optional, and even fairly tight convergence criteria may become necessary: For larger atomic and molecular clusters, surprisingly many, structurally diverse minima exist within small energy intervals.

MC-style “mutation” moves (but also “crossover”-style information exchange moves) in global structure optimization can easily move several atoms too close to each other, resulting in strongly repulsive forces. Hence, an ensuing local optimization (in a global-local hybrid algorithm) should be able to lead the structure back to reasonable interatomic distances. However, it can also happen that the local optimization then “overshoots” into distances that are too long, corresponding to dissociation. If the attractive gradients there are too small, a continued local optimization will not be able to recover from that. Even if a recovery is possible, such excursions are an unproductive waste of valuable computer time (which is spent almost exclusively on local optimizations anyway). Hence, it is far more economical to pre- (and post-)screen every local optimization with a collision and dissociation detection (CD/DD)³⁴ that simply discards collided or dissociated structures and costs far less than a typical local optimization, even at the FF level. Of course, it is then paramount to properly parametrize these CD and DD operations, to avoid discarding good structures or structures that easily revert to good ones after short local optimizations, but this effort is well invested compared to the gains in computer time.

With NDGO on FF, it frequently happens that FF regions are explored that are never visited in MD (at reasonable temperatures). This is not necessarily a bad thing in itself but rather a sign of the exploratory power of the NDGO. However, it can have disruptive effects, for example, with strongly negative (infinite) energy contributions from a single FF term, leading to fusion of two atoms. This fatally distracts the search into these unphysical regions but has the advantage of being easily detectable. Possible remedies are to “patch up” the FF so that these unphysical FF branches are eliminated (while not affecting the physically correct FF regions), and/or to employ a CD with well-adjusted lower atom distance limits.

From section “Search Space is Vast” it is clear that GO scales exponentially, for fundamental reasons. For the same reasons, calculating

interatomic forces by exact ab-initio quantum chemistry (full configuration interaction in a sufficiently large one-particle basis) also scales exponentially, and still scales badly (formally at least n^3 to n^5) if approximate treatments like density functional theory (DFT) or Møller-Plesset perturbation theory of second order (MP2) are employed that cannot work well for systems with strong electron correlation. Hence, combining these two by performing NDGO directly at the low-end ab-initio or DFT level⁷³⁻⁷⁵ may be brave and well-intentioned but hardly is a wise choice. Only very small system sizes are accessible in this way,⁷³⁻⁷⁵ near the crossover region between legitimate use of multistart local search and true superiority of NDGO global search.² Simply put, the huge expense of ab-initio or DFT methods prevents the global search from achieving enough search space exploration. Within the NDGO paradigm, the only way to check this is doing more and longer runs, which is not possible at this expense level. A presumably more efficient and certainly better testable approach is a two-level strategy^{76, 77} where the global search is performed with an approximate method, followed by local post-optimization of the best results at the DFT or ab-initio level. There surely is a considerable danger that the lower-level global search is misled by quantitative or even qualitative distortions of the PES at this approximate level. However, this can be mitigated in several ways: All NDGOs deliver not just one final global minimum candidate but also large collections of local minima, making it easy to check the energy ordering and energy differences between them at higher levels of theory. From our experience, such checks are mandatory for a 2-level strategy, and these then offer a good way to judge the degree of (dis)similarity between theory levels. If the differences are too great, it is possible to adapt the lower-level description to this particular system under study, which essentially always generates a far better and more reliable description than the usual compromise approximations that merely try to achieve mediocre descriptions across large areas of chemistry. In any case, the possible gains in performance and/or system size are huge and well worth the additional human effort. For example, for self-assembly at surfaces, system sizes and agreement with experiment achievable with simple FFs on small-scale local computers⁷⁸ surely will not be addressable with NDGO at the DFT level on national supercomputers⁷³⁻⁷⁵ within the next 10–20 years.

Sometimes^{73, 75} performing NDGO at the DFT (or another semiempirical) level is justified by claims that a FF description cannot include the quantum-mechanical nature of electrons that clearly is necessary for describing chemical bonding and/or the making and breaking of bonds. However, in a FF-description electrons have been coarse-grained away completely,

only their final effects (providing attractive and repulsive forces between nuclei) survive and are captured by suitably parametrized functions. From this perspective, it even is misguided to call FFs “classical-mechanical”; if anything, they are a mixture of classical and quantum-mechanical notions, as exemplified by typical harmonic oscillator or Coulomb terms, which may justifiably be called “classical”, and Lennard-Jones terms, which have been derived from a purely quantum-mechanical model treatment.⁷⁹ Since the Schrödinger equation and the Hamiltonian operator are linear and quite well-behaved, there is no intuitive reason to assume that electronic eigenvalues as functions of nuclear coordinates cannot be described by suitably chosen model functions. This intuition is rigorously supported by theorems that show that every quantum-mechanical distribution can always be modeled with “classical” FFs, cf. Ref.⁸⁰ and references therein. The particularly weird notion that FFs cannot describe the making and breaking of bonds is refuted by the mere existence of reactive FFs,⁸¹ which can be fitted not just to DFT data of equilibrium structures^{35, 82} but also to high-end multireference ab-initio data⁸³ for bond-breaking situations.

Only a few NDGO algorithms, namely SA and its variants, are eminently serial; all others are not just parallel but in most cases even embarrassingly parallel, in particular all the function evaluations in the PSO and EA/GA paradigms. For many years, however, the latter were dominated by a prescription based on “generations”⁷⁰: A generation consisted of many applications of crossover and mutation operators, acting in parallel on many trial solution pairs (“parents”), producing an even larger number of “children”. These were then compared, based on their function values (“fitness”), and the best ones were selected to constitute the next generation. As noted in Ref.³¹, these comparison and selection phases produce serial bottlenecks in the otherwise embarrassingly parallel calculation. However, these can be eliminated by replacing the generational GA/EA model with a “pool model”,³¹ in which crossover/mutation, fitness calculation and selection are grouped into a small subtask involving only a few trial solutions. These subtasks can then again be parallelized efficiently, since fitness comparisons of the new children to the other trial solutions present in the pool and possible insertion of the children into the pool (both of which disrupt parallelizability) only take a negligible amount of time, compared to function evaluations (even on a cheap FF level). Note that such redesigned EA/GA models were not invented in Ref.³¹ or somewhat earlier works⁸⁴ (and for sure not in even later works^{73, 75}) but were rather well-known in the earliest GA history, under the name “steady-state GA”. Because back then computations were essentially always serial, the different parallelization efficiencies of steady-state and generational GAs were noticed only much

later.^{31, 84} In fact, due to its random nature and to many steps that are to a large extent independent of each other, GA/EA methods can (and probably should) be implemented not with old-fashioned, rigid message-passing interface (MPI) paradigms but with advanced parallelization techniques that allow for heterogeneous compute nodes on independent machines, for hardware fail-safe operation and for incredible (and incredibly simple) malleability/adaptability in parallel resource consumption.^{85, 86} Your high-performance computing (HPC) administrator may not be amused at the low amount of internode network traffic that your highly parallel NDGO run generates, since this hardly justifies the money invested for this network, but from your perspective, this is yet another big benefit (that most NDGO algorithms share).

The fundamental idea to break a big problem into smaller subproblems that are then solved iteratively also is an old one in the NDGO literature⁸⁷ and has been resurrected several times, cf. Ref.⁸⁸ and references therein. Quite obviously, this can work well if the pieces-breakdown corresponds to near-separability, and chances for failure are high if this condition is not met. Since such strategies have not had sweeping success, one may suspect that systems amenable to such an easy and obvious partitioning are rare, at least for small to medium-sized cases (for really large ones, NDGO applicability is questionable, and success of such a divide-and-conquer technique will be hard to check for lack of true benchmark cases).

Similarly, it has been rediscovered and reused repeatedly that (unsurprisingly) *huge* performance benefits accrue upon drastically cutting down on search space size. An extreme case of doing that is to discretize space, i.e., transitioning from continuous coordinates to (predefined) grids. This has been done in the earlier NDGO years, e.g., for LJ and Morse clusters^{89, 90} and for protein folding⁹¹ but also more recently for aggregation on surfaces,⁹²⁻⁹⁴ and recently again for LJ clusters⁹⁵ where it made large clusters up to $n = 1000$ accessible to NDGO.

A similarly frequently rediscovered and reused idea is to build starting structures for series of local optimizations for cluster size n by attaching another atom/molecule to the best cluster structures found previously for cluster size $n - 1$. For the LJ _{n} benchmark case, this is a great idea for very many sizes, as discovered a long time ago,⁹⁶ but necessarily fails for most of the interesting (hard) cases, for which adding one atom would have to trigger a restructuring of the whole cluster, crossing lots of high-energy barriers. Finding such transitions efficiently would require differently specialized global search methods.⁹⁷

In global cluster structure optimization, it was clear from the outset that low-energy structures are not the only possible optimization target

and that also targets combining several (vastly) different things can be set, but this simply was not done very frequently, compared to pure search for low-energy structures. Therefore, observable-targeting NDGOs have been rediscovered several times in different communities^{26, 98–104} but still are comparatively rarely used. Note, however, that the additional benefits are substantial: In global cluster structure optimization, if we have found a global minimum-energy structure with NDGO methods and it agrees with the one determined experimentally (directly or indirectly), we can conclude additionally that the experiment was at equilibrium conditions, i.e., the experimental cluster preparation did not influence the cluster structures at the detection stage. If there is disagreement, there typically is little we can conclude, besides diagnosing the nonequilibrium situation (unless the clusters are very small)—simply because the number of non-global minima is overwhelming, and their “density” per energy unit increases dramatically with increasing energy, while simultaneously the amount of experimental data is likely insufficient to discern between all the possible ways calculated properties can be averaged over local minimum-energy structures (with varying weights) to optimally match the experimental data. Of course, this can be turned into yet another (global) optimization problem, as illustrated in Ref.⁹⁸ for a real-life test case, with an infrared (IR) spectrum as target property (which is a huge amount of data, compared to a single number like, e.g., a cluster dipole moment). In this example, observable-targeting NDGO could locate a single (high-energy) cluster structure that provides a better fit to the experimental spectrum than a synthetic spectrum averaged over the 500 lowest-energy local minimum structures obtained from a standard lowest-energy-targeting NDGO, with weights optimized to fit this experimental spectrum. Of course, this still does not prove that this single structure is correct, and not a different mix of yet several other structures, but it does provide a first-order guess toward a property–structure relation in a nonequilibrium situation.

In NDGO algorithm development, useful paradigm-combinations^{49, 104, 105} should be more welcome than sticking religiously to ad hoc algorithm subclass definitions. We have had the latter for a long time, so that the basic algorithm classes (MCM/BH, GA/EA, etc.) probably cannot be improved further without leaving their originally set, arbitrary definition domains.

NDGO Tips for Absolute Beginners

NDGOs still are not broadly available in standard chemistry program packages; or if they are, then typically only in fairly simple standard forms.

As argued here frequently, this is suboptimal, but it may suffice to get started. A better approach for real production work would be to combine a standard stand-alone NDGO with the energy/force engine of your choice and then tune the former for your given GO problem (where “tuning” does not only mean tweaking parameter values that happen to be available in this particular NDGO method, but also freely borrowing ingredients from other NDGO methods).

Since NDGO algorithms and their handling are fuzzy by design, it is a good idea to start with test cases for which global minima are known. For cluster structure optimization, the best test case available still is LJ_n . Putative global minima for $n=2-1000$ and their energies are available online.¹⁰⁶ Great features of this test case are that the potential energy function is extremely simple, so coding it yourself really is a realistic option, and that by simply choosing different cluster sizes you obtain both very hard instances³⁸ with deceptive landscapes ($n=75,76,77,98,102,103,104$) and fairly easy ones despite nontrivial sizes (e.g., $n=55$). Hence, this is an ideal training ground to become acquainted with the behavior of NDGOs in a realistic setting—and this is not easy if you have only dealt with deterministic computing so far.

You will then experience quickly that the NDGO-typical absence of a true convergence criterion is challenging. Without knowing what to expect, you will have to choose in advance how many NDGO steps you allow (overall, or in the end while the best minimum found so far does not change anymore). And once you have done several runs, you will realize that one single NDGO run does not tell you anything. In cases where you do not know the answer in advance (i.e., not for LJ_n with $n < 150$ with near certainty), to get something like an (uncertain) estimate of the chance you have found a good candidate for the global minimum, you will need to do several runs (5, 10, 20, ...), varying your initial structures and/or the random number seed, and establish the percentage of runs in which the same, so far lowest-energy structure was found. If this is 100%, chances are pretty good that you have obtained an excellent global minimum candidate—independent of problem size and number of steps: In particular, for a huge search space and/or far too few steps, finding the same lowest-energy structure in each and every run is almost impossible. Far more likely in this case is that 0% of your runs find the same lowest-energy structure, indicating that you are very far away from “global convergence” and that the lowest-energy structure you have found so far presumably is *not* a good global minimum candidate. In fact, these seemingly simplistic percentages are a good indicator for you and others how far to trust your results. Hence, you should report them in your publications, even if you

see many NDGO papers out there that omit doing these simple tests and reporting these percentages.

Luckily, there is more orientation for you than a few final numbers: While still struggling with a new GO case (here energy minimization of structures), in every single run always follow how your NDGO algorithm behaves during its iterations, looking at least at two things: function values (energies) and structures. Record at least the best-so-far function value vs iteration number and examine it periodically while the algorithm is running. You should see an exponential-like fall-off. This is not a feature of any particular NDGO algorithm but simply what to expect in general: In the beginning, finding lower-energy structures is simple and happens frequently; this inevitably changes when the algorithm proceeds. While you still are in the steep fall-off region, the run should continue. If the decrease is leveling off and becomes a flat plateau, it is time to stop the run and start a new one. This is typically a better allocation of your computing resources than continuing the run even further. In theory, your move class should be good enough to break out of such a stagnation, but in practice possibly not, last but not least because getting further improvements has to become harder and harder.

In population-based NDGOs, it is also a good idea to examine not only the best-so-far function value but other population-based values, e.g., the mean or median value across the population and the worst one. Additionally, not just the function values but also the actual structures themselves should be examined, again across the iterations. If diversity measures (niching) are present, do examine if they really work as intended. All this provides very valuable information on the NDGO exploration and exploitation and how it can be improved for the problem at hand.

Finally, if you have designed your own NDGO (or have tweaked the NDGO code written by someone else), be again fully aware of the fact that NDGO methods are fundamentally different from deterministic computing in yet another respect: In the latter case, you presumably are used to even small bugs advertising themselves by end results (or intermediate results) that quite obviously cannot be right. This can be very different for NDGO methods: They are designed to find ways around obstacles, even in challengingly weird, high-dimensional search spaces. Hence, for them a program bug (unless it leads to self-advertising runtime exceptions like an array index out of bounds) is just another obstacle that could as well be an intended, legitimate feature of the search space. In other words, your NDGO program may be seriously buggy without you even noticing. Yes, there may be some bug-induced performance degradation, but without a direct comparison (due to wNFL: for *this* particular problem instance, performed with a correct implementation of *this* particular NDGO method

variant) you simply have no idea what the optimal performance could be. Hence, my hard-earned advice is to test all program ingredients in a non-NDGO setting *before* incorporating them into the full-blown global search NDGO program frame. (Note that this is totally independent of the random numbers that likely are involved in several locations in the code. Of course, for testing and debugging you want a fixed random number seed anyway, not the “quasi-random” one usually used in production runs, derived from the current date-and-time or from an internal machine counter.)

Things to Do, and Pitfalls to Avoid

- Avoid being overwhelmed by the bewildering, broad array of “inspired” NDGO methods; use the keywords in section “Brief Summary of Some Fundamental NDGO Algorithm Ideas” to filter out the differences and similarities that matter.
- If possible, freely choose your favorite ingredients across all NDGO methods, satisfying both exploration and exploitation. Religiously sticking to one “best” method may make little sense when tackling a new problem class (cf. section “A Closer Look at Some NDGO Background”).
- Local search beats global search within a given basin of attraction, by a wide margin. And the staircase-transformed PES is easier. Therefore, if you can afford to combine local search with global search, do so. If the local search is too expensive, try to sneak it in at least in the later stages of global search.
- Do not rely on local search to domesticate even the most wildly weird outcomes of your global search moves; this is too expensive. Instead, weed out the really bad ones directly (CD/DD).
- For global cluster structure optimization, do not be afraid of using a 2-level strategy, combining exhaustive global search on a cheap level of theory with local post-optimization of the best results at a high level of theory. For example, FFs have a 5–6 orders of magnitude speed advantage compared to quantum chemistry (including DFT)—which you badly need to explore the vast search space and/or to address systems that are not trivially small.
- NDGO algorithms are embarrassingly parallel. This is ideal for the typical many-core HPC hardware of today.
- If you have prior information on your problem (predefined positions, adsorption poses, etc.), use it—but remember that this is

high-risk / high-gain: On the e^n cliff, even seemingly small steps downhill bring huge performance benefits; but you are forbidding access to huge search-space regions, so you should better be sure that the global minimum is not hiding there.

- To match nonequilibrium experiments, you may want to try observable-targeting NDGO methods.
- If this is your first venture into NDGO territory, you first need to gain some familiarity with applying them, for a test case closely related to your ultimate target problem. For cluster structures, it is obligatory to try both simple and hard cases of LJ_n and/or $(H_2O)_n$.
- Never judge single runs. Instead, do batches of runs for the exact same case and examine (and report) their results statistics.
- Unless you are very experienced, do not only look at the final results of each run but also examine how characteristic items evolved iteratively during the run: best, average, worst energies; best (and some of the not so good) structures; etc.
- Be aware that your NDGO program may have bugs although it seems to perform quite well. Searching for bugs is much easier if code pieces are examined in a suitable non-NDGO mode.

RECENT HIGHLIGHTS

In this final section, a few notable publications and insights from recent years are pointed out that have not been mentioned in the above text. Complete reviews already are available.^{73, 107–109} Additionally, there are very comprehensive treatises¹¹⁰ on several NDGO classes, good guides to NDGO algorithm design,^{111, 112} and successful attempts to combine both.¹¹³

In many areas of chemistry, NDGOs are well-established or even indispensable standard tools. One of them is their traditional application field of global cluster structure search in vacuo, which is still done quite extensively, e.g., for $H^+(NH_3)_n$, $n = 18–30$, with IR spectra calculation¹¹⁴ or for Si_n with density functional tight-binding (DFTB) and DFT, up to $n = 80$.¹¹⁵ Another such field is crystal structure prediction (CSP), where huge blind tests¹¹⁶ and refined NDGO techniques¹¹⁷ are commonplace. Although CSP still is challenging,^{118, 119} it is considered to currently make the transition from an unsolved basic science endeavor to a standard applied science method.¹²⁰

For more than 10 years, finding low-energy surface reconstructions and low-energy structures of clusters on surfaces also has been in the

realm of NDGO applications.^{121, 122} Those pioneering works were followed by several similar algorithms and application studies,^{109, 123–125} ranging from simple benchmark systems¹²⁶ to surface assembly of complicated tether molecules,⁷⁸ and from EA methods^{78, 109, 124–126} to MCM/BH.¹²⁷

Besides low-energy structures, there are many more possible optimization objects and targets, including, for example, force-field parameters,^{35, 82, 128} external electric fields for abstract catalyst design,^{129, 130} exploration of reaction networks,^{131, 132} and many more.

It would be very helpful to associate different energy landscapes with different system properties. To some limited extent, this is already possible. For example, single/double/multifunnel landscapes have been associated with structure-seeking and multifunctional proteins.^{133, 134} It would also be very useful to be able to “pre-view” the search landscape before attempting GO on it, so that an appropriate GO algorithm can be selected beforehand. With disconnectivity graphs,^{5, 133, 135} a simplified view of search space structure is available for structure optimization. Obtaining this, however, requires an exhaustive search. Interestingly, it is possible to generate the next step (i.e., transition paths and activation energies between these minima) not only by brute-force^{136, 137} at high computational expense but also to precalculate this approximately^{97, 138} with much less effort, i.e., to find out if the more accurate but very costly full calculation will be useful or not.

In a blog post, Weise¹³⁹ has provided a nice collection of EA pro and contra arguments, which essentially also apply to most other NDGO algorithms and echo several of the general statements in the present chapter. They provide a nice background to the frequent observation that there appears to be a big gap between claims in academic research on what can be done with NDGOs and the extent to which they are really used in the real (business) world.¹⁴⁰ As argued above, part of the reasons for this discrepancy may be that in contrast to local optimization algorithms NDGOs either are efficient but system-specific or general-black-box but fairly inefficient. Hence, NDGO novices frequently tend to try an NDGO only once, in standard form, and then simply conclude that “algorithm X does not work” after this one attempt, without any further tests or modifications. However, from all the above text, it should be obvious that usage of a standard NDGOs for a given, specific problem may easily lead to an efficiency that is suboptimal by a few orders of magnitude. Therefore, I hope that the present text inspires NDGO newcomers to improve upon their first-shot NDGO attempts for their given problem by problem-specific adaptations and a substantial amount of experimentation—the potential benefits of this can be great.

REFERENCES

1. P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.*, **101**, 2550–2560 (1979). Systematic Ab-Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives.
2. F. Avaltroni and C. Corminboeuf, *J. Comput. Chem.*, **32**, 1869–1875 (2011). Efficiency of Random Search Procedures along the Silicon Cluster Series.
3. S. Anila and C. H. Suresh, *Phys. Chem. Chem. Phys.*, **21**, 23143–23153 (2019). Formation of Large Clusters of CO₂ Around Anions: DFT Study Reveals Cooperative CO₂ Adsorption.
4. D. J. Wales, *Energy Landscapes, with Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, Cambridge, 2003. ISBN 0521814154.
5. D. J. Wales, M. A. Miller, and T. R. Walsh, *Nature*, **394**, 758 (1998). Archetypal Energy Landscapes.
6. C. J. Pickard, *Phys. Rev. B*, **99**, 054102 (2019). Hyperspatial Optimization of Structures.
7. A. Neumaier, *Acta Numerica*, A. Iserles, Ed., pp. 271–369. Cambridge University Press, 2004. Complete Search in Continuous Global Optimization and Constraint Satisfaction.
8. Y. Lin and M. A. Stadtherr, *AIChE J.*, **53**, 866–875 (2007). Deterministic Global Optimization of Nonlinear Dynamic Systems.
9. A. Miro, C. Pozo, G. Guillen-Gosalbez, J. A. Egea, and L. Jimenez, *BMC Bioinform.*, **13**, 90 (2012). Deterministic Global Optimization Algorithm Based on Outer Approximation for the Parameter Estimation of Nonlinear Dynamic Biological Systems.
10. P. P. Janes and A. P. Rendell, *J. Global Optim.*, **56**, 537–558 (2013). Deterministic Global Optimization in Ab-Initio Quantum Chemistry.
11. D. E. Kvasov and Y. D. Sergeyev, *Adv. Eng. Softw.*, **80**, 58–66 (2015). Deterministic Approaches for Solving Practical Black-Box Global Optimization Problems.
12. C. D. Maranas and C. A. Floudas, *J. Chem. Phys.*, **97**, 7667–7678 (1992). A Global Optimization Approach for Lennard-Jones Microclusters.
13. Y. Lin and M. A. Stadtherr, *J. Comput. Chem.*, **26**, 1412–1420 (2005). Deterministic Global Optimization of Molecular Structures Using Interval Analysis.
14. B. Hartke, *J. Phys. Chem.*, **97**, 9973 (1993). Global Geometry Optimization of Clusters Using Genetic Algorithms.
15. B. Hartke, *J. Comput. Chem.*, **20**, 1752 (1999). Global Cluster Geometry Optimization by a Phenotype Algorithm with Niches: Location of Elusive Minima, and Low-Order Scaling with Cluster Size.

16. E. Gibney, *Nature*, **574**, 461–462 (2019). Hello Quantum World! Google Publishes Landmark Quantum Supremacy Claim.
17. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, University Press, Cambridge, 2003. ISBN 052143064X.
18. Z. Michalewicz and D. B. Fogel, *How to Solve it: Modern Heuristics*, Springer, Berlin, 2000. ISBN 3540660615.
19. S. Kirkpatrick, J. C. D. Gellat, and M. P. Vecchi, *Science*, **220**, 671–680 (1983). Optimization by Simulated Annealing.
20. J. H. Holland, *Adaption in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI, 1975. ISBN 0262082136.
21. J. C. Culberson, *Evol. Comput. J.*, **6**, 109 (1998). On the Futility of Blind Search.
22. List of Metaphor-Based Metaheuristics, https://en.wikipedia.org/wiki/List_of_metaphor-based_metaheuristics, accessed: 2019-10-11.
23. F. Glover and K. Sörensen, Metaheuristics, <http://www.scholarpedia.org/article/Metaheuristics>, accessed: 2019-01-21.
24. T. Weise, Why Research in Computational Intelligence should be Less Nature-Inspired, <http://iao.hfuu.edu.cn/blogs/science-blog/21-why-research-in-computational-intelligence-should-be-less-nature-inspired>, accessed: 2019-01-21.
25. K. Sörensen, *Int. Trans. Oper. Res.*, **22**, 3–18 (2015). Metaheuristics – the Metaphor Exposed.
26. A. Jain, I. Castelli, G. Hautier, D. Bailey, and K. Jacobsen, *J. Mater. Sci.*, **48**, 6519 (2013). Performance of Genetic Algorithms in Search for Water Splitting Perovskites.
27. D. J. Wales, *Adv. Rev. Phys. Chem.*, **69**, 401–425 (2018). Exploring Energy Landscapes.
28. Y. Cai and L. Cheng, *J. Chem. Phys.*, **149**, 084102 (2018). Single-Root Networks for Describing the Potential Energy Surface of Lennard-Jones Clusters.
29. D. H. Wolpert and W. G. Macready, *IEEE Trans. Evol. Comput.*, **1** (1997). No Free Lunch Theorems for Optimization.
30. S. E. Schönborn, S. Goedecker, S. Roy, and A. R. Oganov, *J. Chem. Phys.*, **130**, 144108 (2009). The Performance of Minima Hopping and Evolutionary Algorithms for Cluster Structure Prediction.
31. B. Bandow and B. Hartke, *J. Phys. Chem. A*, **110**, 5809 (2006). Larger Water Clusters with Edges and Corners on their Way to Ice: Structural Trends Elucidated with an Improved Parallel Evolutionary Algorithm.
32. F. Schulz and B. Hartke, *Theor. Chem. Acc.*, **114**, 357–379 (2005). A New Proposal for the Reason of Magic Numbers in Alkali Cation Microhydration Clusters.

33. B. Hartke, *J. Chem. Phys.*, **130**, 024905 (2009). Methane-Water Clusters Under Pressure: Are Clathrate Cages Optimal Clusters?
34. J. M. Dieterich and B. Hartke, *Mol. Phys.*, **108**, 279–291 (2010). OGOLEM: Global Cluster Structure Optimization for Arbitrary Mixtures of Flexible Molecules. A Multiscaling, Object-Oriented Approach.
35. H. R. Larsson, A. C. T. van Duin, and B. Hartke, *J. Comput. Chem.*, **34**, 2178–2189 (2013). Global Optimization of Parameters in the Reactive Force Field ReaxFF for SiOH.
36. A. E. Ezugwu, O. J. Adeleke, A. A. Akinyelu, and S. Viriri, *Neural Comput. Appl.*, **31**, 1–45 (2019). A Conceptual Comparison of Several Metaheuristics Algorithms on Continuous Optimization Problems.
37. J. M. Dieterich and B. Hartke, *Appl. Math.*, **3**, 1552–1564 (2012). Empirical Review of Standard Benchmark Functions using Evolutionary Global Optimization.
38. M. Dittner and B. Hartke, *Comput. Theor. Chem.*, **1107**, 7–13 (2017). Conquering the Hard Cases of Lennard-Jones Clusters with Simple Recipes.
39. R. Sala, N. Baldanzini, and M. Pierini, *Optim. Lett.*, **11**, 699–713 (2017). Global Optimization Test Problems Based on Random Field Composition.
40. I. Ibrahim, Z. Ibrahim, H. Ahmad, M. F. M. Jusof, S. W. Nawawi, and M. Mubin, *Int. J. Adv. Manuf. Technol.*, **79**, 1363–1376 (2015). An Assembly Sequence Planning Approach with a Rule-Based Multi-State Gravitational Search Algorithm.
41. Y. Huang and K. Jiang, *Adv. Appl. Math. Mech.*, **9**, 307–323 (2017). Hill-Climbing Algorithm with a Stick for Unconstrained Optimization Problems.
42. M. Pluhacek, R. Senkerik, A. Viktorin, and T. Kadavy, *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2018, pp. 770–776. Why Simple Population Restart Does Not Work in PSO.
43. Z. Li and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 6611–6615 (1987). Monte Carlo Minimization Approach to the Multiple-Minima Problem in Protein Folding.
44. J. P. K. Doye and D. J. Wales, *J. Phys. Chem. A*, **101**, 5111–5116 (1997). Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms.
45. D. J. Wales and H. A. Scheraga, *Science* **285**, 1368–1372 (1999). Global Optimization of Clusters, Crystals, and Biomolecules.
46. O. Valsson, P. Tiwary, and M. Parrinello, *Annual Review of Physical Chemistry*, vol. **67**, M. A. Johnson and T. J. Martinez, Eds., pp. 159–184, 2016. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. ISBN 978-0-8243-1067-7.
47. A. Barducci, M. Bonomi, and M. Parrinello, *WIREs Comput. Mol. Sci.*, **1**, 826–843 (2011). Metadynamics.

48. A. Laio and F. L. Gervasio, *Rep. Prog. Phys.*, **71**, 126601 (2008). Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science.
49. G. Rossi and R. Ferrando, *Chem. Phys. Lett.*, **423**, 17–22 (2006). Global Optimization by Excitable Walkers.
50. A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, *Chem. Phys. Lett.*, **219**, 343–348 (1994). Quantum Annealing: A New Method for Minimizing Multidimensional Functions.
51. T. Gregor and R. Car, *Chem. Phys. Lett.*, **412**, 125–130 (2005). Minimization of the Potential Energy Surface of Lennard-Jones Clusters by Quantum Optimization.
52. S. Suzuki, *Eur. Phys. J. Spec. Top.*, **224**, 51–61 (2015). Performance of Quantum Annealing in Solving Optimization Problems: A Review.
53. E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, *Science*, **292**, 472–476 (2001). A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem.
54. B. Hartke, *Phys. Chem. Chem. Phys.*, **8**, 3627–3635 (2006). Propagation with Distributed Gaussians as a Sparse, Adaptive Basis for Higher-Dimensional Quantum Dynamics.
55. H. R. Larsson and B. Hartke, and D. J. Tannor, *J. Chem. Phys.*, **145**, 204108 (2016). Efficient Molecular Quantum Dynamics in Coordinate and Phase Space Using Pruned Bases.
56. H. R. Larsson and D. J. Tannor, *J. Chem. Phys.*, **147**, 044103 (2017). Dynamical Pruning of the Multiconfiguration Time-Dependent Hartree Method (DP-MCTDH): An Efficient Approach for Multidimensional Quantum Dynamics.
57. F. Minhas, A. Asif, and A. Ben-Hur, Ten Ways to Fool the Masses with Machine Learning (2019), arXiv:1901.01686.
58. G. Schneider, *Nature Reviews*, **17**, 97–113 (2018). Automating Drug Discovery.
59. E. Uteva, R. S. Graham, R. D. Wilkinson, and R. J. Wheatley, *J. Chem. Phys.*, **149**, 174114 (2018). Active Learning in Gaussian Process Interpolation of Potential Energy Surfaces.
60. K. Gubaev, E. V. Podryabinkin, and A. V. Shapeev, *J. Chem. Phys.*, **148**, 241727 (2018). Machine Learning of Molecular Properties: Locality and Active Learning.
61. J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitberg, *J. Chem. Phys.*, **148**, 241733 (2018). Less is More: Sampling Chemical Space with Active Learning.
62. N. J. Browning, R. Ramakrishnan, O. A. von Lilienfeld, and U. Roethlisberger, *J. Phys. Chem. Lett.*, **8**, 1351–1359 (2017). Genetic Optimization of Training Sets for Improved Machine Learning Models of Molecular Properties.

63. A. Nigam, P. Friederich, M. Krenn, and A. Aspuru-Guzik, Augmenting Genetic Algorithms with Deep Neural Networks for Exploring the Chemical Space, <https://arxiv.org/abs/1909.11655> (2019), accessed: 2019-11-11.
64. T. L. Jacobsen, M. S. Jørgensen, and B. Hammer, *Phys. Rev. Lett.*, **120**, 026102 (2018). On-the-fly Machine Learning of Atomic Potential in Density Functional Theory Structure Optimization.
65. D. M. Deaven and K. M. Ho, *Phys. Rev. Lett.*, **75**, 288 (1995). Molecular Geometry Optimization with a Genetic Algorithm.
66. D. M. Deaven, N. Tit, J. R. Morris, and K. M. Ho, *Chem. Phys. Lett.*, **256**, 195–200 (1996). Structural Optimization of Lennard-Jones Clusters by a Genetic Algorithm.
67. W. J. Pullan, *Comput. Phys. Commun.*, **107**, 137–148 (1997). Genetic Operators for the Atomic Cluster Problem.
68. G. Harik and D. Goldberg, *Foundations of Genetic Algorithms*, vol. **4**, R. K. Belew and M. D. Vose, Eds., pp. 247–262, Morgan Kaufmann, San Francisco, 1997. Learning Linkage.
69. J. P. Martins, C. B. Neto, M. K. Crocomo, K. Vittori, and A. C. B. Delbem, *2013 IEEE Congress on Evolutionary Computation* (2013), pp. 502–509. A Comparison of Linkage-learning-based Genetic Algorithms in Multidimensional Knapsack Problems.
70. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989. ISBN 0-201-15767-5.
71. A. E. Eiben and C. A. Schippers, *Fundam. Inform.*, **35**, 35–50 (1998). On Evolutionary Exploration and Exploitation.
72. A. Hussain and Y. S. Muhammad, *Complex Intell. Syst.*, **6**, 1–14 (2019). Trade-off Between Exploration and Exploitation with Genetic Algorithm Using a Novel Selection Operator.
73. M. Jäger, R. Schäfer, and R. L. Johnston, *Adv. Phys.: X*, **3**, 1077–1107 (2018). First-principles Global Optimization of Metal Clusters and Nanoalloys.
74. H. A. Hussein, I. Demiroglu, and R. L. Johnston, *Eur. Phys. J. B*, **91**, 34 (2018). Application of a Parallel Genetic Algorithm to the Global Optimization of Medium-sized Au-Pd Sub-nanometre Clusters.
75. M. Jäger, R. Schäfer, and R. L. Johnston, *Nanoscale*, **11**, 9042–9052 (2019). GIGA: A Versatile Genetic Algorithm for Free and Supported Clusters and Nanoparticles in the Presence of Ligands.
76. B. Hartke, *Theor. Chem. Acc.*, **99**, 241–247 (1998). Global Geometry Optimization of Small Silicon Clusters at the Level of Density Functional Theory.
77. B. Hartke, *Chem. Phys. Lett.*, **258**, 144–148 (1996). Global Geometry Optimization of Clusters Guided by N-dependent Model Potentials.
78. A. Freibert, J. M. Dieterich, and B. Hartke, *J. Comput. Chem.*, **40**, 1978–1989 (2019). Exploring Self-organization of Molecular Tether Molecules on a Gold Surface by Global Structure Optimization.

79. A. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, Oxford, 2nd ed. (2013). ISBN 9780199672394.
80. R. Babbush, J. Parkhill, and A. Aspuru-Guzik, *Front. Chem.*, **1**, 26 (2013). Force-field Functor Theory: Classical Force-fields which Reproduce Equilibrium Quantum Distributions.
81. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, *J. Phys. Chem. A*, **105**, 9396 (2001). ReaxFF: A Reactive Force Field for Hydrocarbons.
82. M. Dittner, J. Müller, H. M. Aktulga, and B. Hartke, *J. Comput. Chem.*, **36**, 1550–1561 (2015). Efficient Global Optimization of ReaxFF Parameters.
83. J. Müller and B. Hartke, *J. Chem. Theory Comput.*, **12**, 3913–3925 (2016). A ReaxFF Reactive Force Field for Disulfide Mechanochemistry, Fitted to Multireference Ab-initio Data.
84. Y. Ge and J. D. Head, *Chem. Phys. Lett.*, **398**, 107 (2004). Global Optimization of H-Passivated Si Clusters with a Genetic Algorithm.
85. J. M. Dieterich and B. Hartke, *J. Chem. Theory Comput.*, **12**, 5226 (2016). An Error-safe, Portable, and Efficient Evolutionary Algorithms Implementation with High Scalability.
86. F. Spenke, K. Balzer, S. Frick, B. Hartke, and J. M. Dieterich, *Comput. Theor. Chem.*, **1151**, 72–77 (2019). Malleable Parallelism with Minimal Effort for Maximal Throughput and Maximal Hardware Load.
87. K. Michaelian, *Chem. Phys. Lett.*, **293**, 202 (1998). A Symbiotic Algorithm for Finding the Lowest Energy Isomers of Large Clusters and Molecules.
88. P. Yang, K. Tang, and X. Yao, A Parallel Divide-and-Conquer based Evolutionary Algorithm for Large-scale Optimization (2018).
89. X. G. Shao, L. J. Cheng, and W. S. Cai, *J. Comput. Chem.*, **25**, 1693–1698 (2004). A Dynamic Lattice Searching Method for Fast Optimization of Lennard-Jones Clusters.
90. L. Cheng and J. Yang, *J. Phys. Chem. A*, **111**, 2336–2342 (2007). Novel Lattice-searching Method for Modeling the Optimal Strain-free Close-packed Isomers of Clusters.
91. J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, *Annu. Rev. Phys. Chem.*, **48**, 545 (1997). Theory of Protein Folding: The Energy Landscape Perspective.
92. V. Obersteiner, M. Scherbela, L. Hörmann, D. Wegner, and O. T. Hofmann, *Nano Lett.*, **17**, 4453–4460 (2017). Structure Prediction for Surface-Induced Phases of Organic Monolayers: Overcoming the Combinatorial Bottleneck.
93. S. Fortuna, D. L. Cheung, and K. Johnston, *J. Chem. Phys.*, **144**, 134707 (2016). Phase Behavior of Self-assembled Monolayers Controlled by Tuning Physisorbed and Chemisorbed States: A Lattice-model View.
94. D. M. Packwood and T. Hitosugi, *Appl. Phys. Express*, **10**, 065502 (2017). Rapid Prediction of Molecular Arrangements on Metal Surfaces via Bayesian Optimization.

95. K. Yu, X. Wang, L. Chen, and L. Wang, *J. Chem. Phys.*, **151**, 214105 (2019). Unbiased Fuzzy Global Optimization of Lennard-Jones Clusters.
96. J. A. Northby, *J. Chem. Phys.*, **87**, 6166 (1987). Structure and Binding of Lennard-Jones Clusters.
97. D. S. De, M. Krummenacher, B. Schaefer, and S. Goedecker, *Phys. Rev. Lett.*, **123**, 206102 (2019). Finding Reaction Pathways with Optimal Atomic Index Mappings.
98. J. M. Dieterich and B. Hartke, *Phys. Chem. Chem. Phys.*, **17**, 11958–11961 (2015). Observable-targeting Global Cluster Structure Optimization.
99. S. Bhattacharya, B. H. Sonin, C. J. Jumonville, L. M. Ghiringhelli, and N. Marom, *Phys. Rev. B*, **91**, 241115 (2015). Computational Design of Nanoclusters by Property-based Genetic Algorithms: Tuning the Electronic Properties of $(\text{TiO}_2)_n$ Clusters.
100. N. O. Carstensen, J. M. Dieterich, and B. Hartke, *Phys. Chem. Chem. Phys.*, **13**, 2903 (2011). Design of Optimally Switchable Molecules by Genetic Algorithms.
101. N. M. O’Boyle, C. M. Campbell, and G. R. Hutchison, *J. Phys. Chem. C*, **115**, 16200 (2011). Computational Design and Selection of Optimal Organic Photovoltaic Materials.
102. M. d’Avezac, J.-W. Luo, A. Saraiva, B. Koiller, and A. Zunger, *Phys. Rev. Lett.*, **108**, 027401 (2012). Genetic-Algorithm Discovery of a Direct-Gap and Optically Allowed Superstructure from Indirect-Gap Si and Ge Semiconductors.
103. A. L. S. Chua, N. A. Benedek, L. Chen, M. W. Finnis, and A. P. Sutton, *Nat. Mater.*, **9**, 418 (2010). A Genetic Algorithm for Predicting the Structures of Interfaces in Multicomponent Systems.
104. K. Röder and D. J. Wales, *J. Chem. Phys. Lett.*, **9**, 6169–6173 (2018). Mutational Basin-Hopping: Combined Structure and Sequence Optimization for Biomolecules.
105. J. M. Dieterich and B. Hartke, *Inorganics*, **5**, 64 (2017). Improved Cluster Structure Optimization: Hybridizing Evolutionary Algorithms with Local Heat Pulses.
106. D. J. Wales, J. P. K. Doye, A. Dullweber, M. P. Hodges, F. Y. Naumkin, F. Calvo, J. Hernández-Rojas, and T. F. Middleton, The Cambridge Cluster Database, <http://www-wales.ch.cam.ac.uk/CCD.html>, accessed: 2019-11-10.
107. B. Hartke, *WIREs Comput. Mol. Sci.*, **1**, 879–887 (2011). Global Optimization.
108. R. Ferrando, *J. Nanopart. Res.*, **20**, 179 (2018). Determining the Equilibrium Structures of Nanoalloys by Computational Methods.
109. M. Sierka, *Progr. Surf. Sci.*, **85**, 398–434 (2010). Synergy Between Theory and Experiment in Structure Resolution of Low-Dimensional Oxides.
110. T. Weise, Global Optimization Algorithms, www.it-weise.de/projects/book.pdf, accessed: 2019-01-21.

111. T. Weise, M. Zapf, R. Chiong, and A. J. Nebro, *Nature-Inspired Algorithms for Optimization*, R. Chiong, Ed., pp. 1–50, Springer, Berlin, 2009. Why is Optimization Difficult?
112. T. Weise, R. Chiong, and K. Tang, *J. Comput. Sci. Technol.*, **27**, 907–936 (2012). Evolutionary Optimization: Pitfalls and Booby Traps.
113. T. Weise, *An Introduction to Optimization Algorithms*, Institute of Applied Optimization (IAO), Faculty of Computer Science and Technology, Hefei University, Hefei, Anhui, China, 2019-11-16 ed. (2018–2019), <http://thomasweise.github.io/aitoa/>.
114. A. Malloum, J. J. Fifen, and J. Conradie, *J. Chem. Phys.*, **149**, 244301 (2018). Structures and Infrared Spectroscopy of Large Sized Protonated Ammonia Clusters.
115. S. Heydariyan, M. R. Nouri, M. Alaei, Z. Allahyari, and T. A. Niehaus, *J. Chem. Phys.*, **149**, 074313 (2018). New Candidates for the Global Minimum of Medium-sized Silicon Clusters: A Hybrid DFTB/DFT Genetic Algorithm Applied to Si_n , $n = 8–80$.
116. A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. D. A. Dzyabchenko Jr., B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhuqq, and C. R. Groom, *Acta Cryst.*, **872**, 439–459 (2016). Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods.
117. F. Curtis, T. Rose, and N. Marom, *Faraday Discuss.*, **211**, 61–77 (2018). Evolutionary Niching in the GATOR Genetic Algorithm for Molecular Crystal Structure Prediction.
118. S. L. Price, *Faraday Discuss.*, **211**, 9–30 (2018). Is Zeroth Order Crystal Structure Prediction Coming to Maturity?
119. A. R. Oganov, *Faraday Discuss.*, **211**, 643–660 (2018). Crystal Structure Prediction: Reflections on Present Status and Challenges.

120. J. Nyman and S. M. Reutzel-Edens, *Faraday Discuss.*, **211**, 459–476 (2018). Crystal Structure Prediction is Changing from Basic Science to Applied Technology.
121. F. C. Chuang, C. V. Ciobanu, V. B. Shenoy, C. Z. Wang, and K. M. Ho, *Surf. Sci.*, **573**, L375–L381 (2004). Finding the Reconstructions of Semiconductor Surfaces via a Genetic Algorithm.
122. F.-C. Chuang, B. Liu, C.-Z. Wang, T.-L. Chan, and K.-M. Ho, *Surf. Sci.*, **598**, L339–L346 (2005). Global Structure Optimization of Si Magic Clusters on the Si(111)7x7 Surface.
123. H. Shi, S. M. Auerbach, and A. Ramasubramaniam, *J. Phys. Chem. C*, **120**, 11899–11909 (2016). First-principles Predictions of Structure-function Relationships of Graphene-supported Platinum Nanoclusters.
124. R. Włodarczyk, M. Sierka, K. Kwapien, and J. Sauer, *J. Phys. Chem. C*, **115**, 6764–6774 (2011). Structures of the Ordered Water Monolayer on MgO(001).
125. X. Zhao, X. Shao, Y. Fujimori, S. Bhattacharya, L. M. Ghiringhelli, H.-J. Freund, M. Sterrer, N. Nilius, and S. V. Levchenko, *J. Phys. Chem. Lett.*, **6**, 1204–1208 (2015). Formation of Water Chains on CaO(001): What Drives the 1D Growth?
126. C. Witt, J. M. Dieterich, and B. Hartke, *Phys. Chem. Chem. Phys.*, **20**, 15661 (2018). Cluster Structures Influenced by Interaction with a Surface.
127. X. Chen, Y.-F. Zhao, Y.-Y. Zhang, and J. Li, *J. Comput. Chem.*, **40**, 1105–1112 (2019). TGMIn: an Efficient Global Minimum Searching Program for Free and Surface-supported Clusters.
128. G. Shchygol, A. Yakovlev, T. Trnka, A. C. T. van Duin, and T. Verstraelen, *J. Chem. Theory Comput.*, **15**, 6799–6812 (2019). ReaxFF Parameter Optimization with Monte Carlo and Evolutionary Algorithms: Guidelines and Insights.
129. M. Dittner and B. Hartke, *J. Chem. Theory Comput.*, **14**, 3547–3564 (2018). Globally Optimal Catalytic Fields: Inverse Design of Abstract Embeddings for Maximum Reaction Rate Acceleration.
130. M. Dittner and B. Hartke, *J. Chem. Phys.*, **152**, 119106 (2020). Optimal Catalytic Fields for a Diels-Alder Reaction.
131. X.-J. Zhang and Z.-P. Liu, *Phys. Chem. Chem. Phys.*, **17**, 2757–2769 (2015). Reaction Sampling and Reactivity Prediction Using the Stochastic Surface Walking Method.
132. P.-L. Kang, C. Shang, and Z.-P. Liu, *J. Am. Chem. Soc.*, **141**, 20525–20536 (2019). Glucose to 5-Hydroxymethylfurfural: Origin of Site-selectivity Resolved by Machine Learning Based Reaction Sampling.
133. O. M. Becker and M. Karplus, *J. Chem. Phys.*, **106**, 1495 (1997). The Topology of Multidimensional Potential Energy Surfaces: Theory and Application to Peptide Structure and Kinetics.

134. K. Röder, J. A. Joseph, B. E. Husic, and D. J. Wales, *Adv. Theory Simul.*, **2**, 1800175 (2019). Energy Landscapes for Proteins: From Single Funnel to Multifunctional Systems.
135. L. C. Smeeton, M. T. Oakley, and R. L. Johnston, *J. Comput. Chem.*, **35**, 1481–1490 (2014). Visualizing Energy Landscapes with Metric Disconnectivity Graphs.
136. J. M. Carr, S. A. Trygubenko, and D. J. Wales, *J. Chem. Phys.*, **122**, 234903 (2005). Finding Pathways Between Distant Local Minima.
137. J. P. K. Doye, M. A. Miller, and D. J. Wales, *J. Chem. Phys.*, **111**, 8417 (1999). Evolution of the Potential Energy Surface with Size for Lennard-Jones Clusters.
138. B. Schaefer and S. Goedecker, *J. Phys. Chem.*, **145**, 034101 (2016). Computationally Efficient Characterization of Potential Energy Surfaces Based on Fingerprint Distances.
139. T. Weise, Why You Should Use Evolutionary Algorithms to Solve Your Optimization Problems (and why not), <http://iao.hfuu.edu.cn/blogs/science-blog/48-why-use-evolutionary-algorithms>, accessed: 2019-11-03.
140. Z. Michalewicz, *ACM Ubiquity*, **2012**, 1–13 (2012). The Emperor is Naked: Evolutionary Algorithms for Real-world Applications.

2

DENSITY FUNCTIONAL TIGHT BINDING CALCULATIONS FOR PROBING ELECTRONIC-EXCITED STATES OF LARGE SYSTEMS

SHARMA S.R.K.C. YAMIJALA^{1,2,3,5,6}, MA. BELÉN OVIEDO⁴ and BRYAN M. WONG^{1,2,3}

¹*Department of Chemical & Environmental Engineering, Materials Science & Engineering Program, University of California-Riverside, Riverside, CA, USA*

²*Department of Chemistry, University of California-Riverside, Riverside, CA, USA*

³*Department of Physics & Astronomy, University of California-Riverside, Riverside, CA, USA*

⁴*Departamento de Química Teórica y Computacional, Facultad de Ciencias Químicas, Instituto de Investigaciones en Fisicoquímica de Córdoba (INFIQC), UNC-CONICET, Universidad Nacional de Córdoba, Córdoba, Argentina*

⁵*Department of Chemistry, Indian Institute of Technology Madras, Chennai 600036, India*

⁶*Center for Atomistic Modelling and Materials Design, Indian Institute of Technology Madras, Chennai 600036, India*

INTRODUCTION

Over the past decade, researchers in computational chemistry have witnessed a resurgence in the development and application of semi-empirical methods for treating large chemical/material systems. In particular, the

Reviews in Computational Chemistry, Volume 32, First Edition.

Edited by Abby L. Parrill and Kenny B. Lipkowitz.

© 2022 John Wiley & Sons, Inc. Published 2022 by John Wiley & Sons, Inc.

density functional tight binding (DFTB) formalism¹⁻⁴ has garnered immense popularity for probing the electronic properties of biomolecules,⁵ molecules/clusters with numerous conformations,^{6,7} and immense nanostructures.⁸ While classical molecular dynamics can handle hundreds of thousands of atoms, it cannot provide a first-principles-based description of large systems at a quantum mechanical (i.e., electronic) level of detail. At the other extreme, conventional Kohn-Sham density functional theory (DFT) methods can access the true quantum mechanical nature of matter, but they cannot tackle the large sizes and complex chemical environments relevant to many of the large systems mentioned previously. To bridge these immense size scales, the DFTB formalism was developed to probe these chemical/materials systems with a viable approach that is both computationally efficient and quantum mechanical in nature.

While there have been a number of good reviews on DFTB,⁹⁻¹² in this chapter, we focus on the application of DFTB to electronic-excited states, which has attracted significant attention for extending this computationally efficient approach to the time domain. Although the field of chemical dynamics is incredibly vast, we concentrate our attention on the application of DFTB to real-time electron dynamics and nonadiabatic dynamics calculations. We first give a brief discussion of the underlying theory in each of these areas, followed by a didactic tutorial on simple molecules or model systems to show how these computational approaches and techniques are carried out in practice. These simple tutorials are written in a stepwise, instructive fashion to provide practicing researchers a detailed “look under the hood” to obtain a deeper understanding of how these techniques can be used to understand, probe, and even control the chemical dynamics of large systems. We then conclude each of these respective sections with a “real-life” application of these DFTB-based approaches on large dynamical systems to demonstrate their usefulness in contemporary areas of chemistry and materials science.

REAL-TIME TIME-DEPENDENT DFTB (RT-TDDFTB)

Theory and Methodology

Over the past few years, the use of real-time time-dependent DFTB (RT-TDDFTB) has attracted significant attention as a promising approach for extending the DFTB formalism to the nonequilibrium electron dynamics of extremely large chemical systems. As specific examples, this method has been used to calculate photo-injection dynamics in dye-sensitized TiO₂ solar cells,¹³⁻¹⁵ optical properties of photosynthetic pigments,^{16,17}

molecular aggregates,¹⁸ graphene nanoflakes,¹⁹ DNA intercalation complexes,²⁰ many-body interactions in solvated nanodroplets,²¹ and excitation energy transfer (EET) dynamics in plasmonic arrays.^{22, 23} To carry out an RT-TDDFTB dynamics calculation, one must first compute the ground-state Hamiltonian, overlap matrix elements, and the initial single-electron density matrix within the self-consistent DFTB approach. The open-source DFTB+ code²⁴ provides a practical way to obtain these quantities, and a more detailed description of ground-state DFTB can be found in previous publications.^{4, 25}

Once the ground-state Hamiltonian, overlap matrix elements, and the initial single-electron density matrix are calculated, they can be used as initial input conditions in subsequent real-time quantum dynamics calculations. These RT-TDDFTB quantum dynamics calculations are carried out in practice by applying a time-dependent electric field to the initial ground state (GS) density matrix, resulting in the Hamiltonian

$$\hat{H}(t) = \hat{H}^0 - E_0(t) \cdot \hat{\mu}(t) \quad [1]$$

where $E_0(t)$ is the applied electric field, and $\hat{\mu}$ is the dipole moment operator. Because the Hamiltonian in Eq. [1] is explicitly time dependent, the density matrix, $\hat{\rho}$, evolves according to the Liouville–von Neumann equation of motion which, in the nonorthogonal-DFTB basis, is given by²⁶

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i\hbar} (S^{-1} \cdot \hat{H}[\hat{\rho}] \cdot \hat{\rho} - \hat{\rho} \cdot \hat{H}[\hat{\rho}] \cdot S^{-1}) \quad [2]$$

where \hat{H} is the Hamiltonian matrix (which implicitly depends on the density matrix), S^{-1} is the inverse of the overlap matrix, and \hbar is Planck’s constant.

Furthermore, because the quantum system is directly propagated in the time domain, one can choose $E_0(t)$ to have any time-dependent form. For example, if $E_0(t)$ is chosen to have the form of a Dirac delta function given by

$$E_{\text{delta}}(t) = E_0 \delta(t - t_0) \quad [3]$$

this “kick” perturbation instantaneously changes the velocity field of the electrons and causes a time-varying dipole moment. As a result, the state of the system is no longer an eigenfunction of the Hamiltonian. In practice, E_{delta} is implemented as a linearly polarized Gaussian-type perturbation or as a phase in the initial electron density²⁷ given by:

$$E_{\text{delta}}(t) = E_0 \exp \left[-\frac{(t - t_0)^2 w^2}{2} \right] \hat{n} \quad [4]$$

where \hat{n} is the polarization vector. The temporal evolution of the density matrix is carried out iteratively where the atomic charge on atom α is first calculated using the Mulliken approximation:

$$q_\alpha = \sum_{\mu \in \alpha} [\rho(t)S]_{\mu\mu} \quad [5]$$

Afterward, the self-consistent charge Hamiltonian matrix is constructed as⁴

$$H_{\mu\nu} = \langle \phi_\mu | \hat{H}^0 | \phi_\nu \rangle + \frac{1}{2} S \sum_{\xi} (\gamma_{\alpha\xi} + \gamma_{\beta\xi}) \Delta q_\xi \quad [6]$$

where ϕ_μ are DFTB Slater-type orbital basis functions centered on the atomic sites, and $\gamma_{\alpha\xi}$ is a function of the interatomic separation and the Hubbard parameter U .²⁸ Finally, the density matrix is updated using a three-point integration algorithm:

$$\hat{\rho}(t_0 + \Delta t) = \hat{\rho}(t_0 - \Delta t) - 2 \frac{i}{\hbar} \hat{\rho} \Delta t + O[(\Delta t)^2] \quad [7]$$

which is accurate to order $(\Delta t)^2$.

However, if one chooses $E_0(t)$ to take the form of a sinusoidal perturbation given by

$$E_{\text{laser}} = E_0 \sin(\omega t) \quad [8]$$

it represents a continuous interaction of the system with monochromatic light of frequency ω (i.e., a laser) in the time domain, and this perturbation is added to the Hamiltonian in the iterative procedure:

$$\hat{H}(t) = \hat{H}_0 + \frac{1}{2} [S \hat{V}(t) + \hat{V}(t) S] \quad [9]$$

where $\hat{V}(t)$ is set to $-E_0 \sin(\omega t) \cdot \hat{\mu}(t)$ or the expression for the delta perturbation (Eq. [4]). It is worth mentioning that both of these different choices for the electric field give different but complementary viewpoints of quantum dynamics, as will be explained in both the tutorial and example sections that follow. To more easily understand this entire procedure, the following pseudocode flowchart shows the sequential steps for carrying out an RT-TDDFTB electron dynamics calculation.²⁹

In summary, the RT-TDDFTB algorithm commences by reading the following files from a ground-state DFTB+ calculation: the non-self-consistent Hamiltonian matrix (H_0), the self-consistent Hamiltonian matrix (H_{SCC}), the

Algorithm 1 Pseudocode for RT-TDDFTB dynamics

Read H_0 , H_{SCC} , ρ_{MO} , and S from DFTB+ output filesCompute S^{-1} Solve equation: $H_{\text{SCC}}C = SC E$ Calculate $\rho_0 = C\rho_{\text{MO}}C^\dagger$ **do** $i = 0$ to N $t = i\Delta t$ Compute $q(t)$ (Eq. [5])Compute $H(t)$ (Eq. [6])Add electric field to $H(t)$ Update ρ (Eqs. [2] and [7])

Compute electronic properties of the system.

end do

single-electron density matrix in the molecular orbital representation (ρ_{MO}), and the overlap matrix (S). Next, the inverse of the overlap matrix (S^{-1}) is computed, and H_{SCC} is diagonalized. The diagonalization of this matrix gives the one-electron reduced density matrix of the GS in the atomic orbital (AO) basis.

Tutorial on RT-TDDFTB Electron Dynamics for a Naphthalene Molecule

With the basic theoretical concepts of RT-TDDFTB outlined in the previous section, we now give two tutorials showing how to (i) compute an absorption spectrum of a simple naphthalene molecule and (ii) probe the time-dependent dynamics of naphthalene in the presence of monochromatic light (i.e., a laser perturbation) using the RT-TDDFTB approach.

Absorption Spectrum for Naphthalene

As mentioned in the RT-TDDFTB theory and methodology section, a ground-state DFTB+ calculation for naphthalene must be carried out first before performing an RT-TDDFTB calculation. With the ground-state DFTB+ calculation for naphthalene properly converged, one obtains the non-self-consistent Hamiltonian matrix, the self-consistent Hamiltonian matrix, the initial single-electron density matrix, and the overlap matrix as output files. Next, to compute an absorption spectrum of naphthalene, three

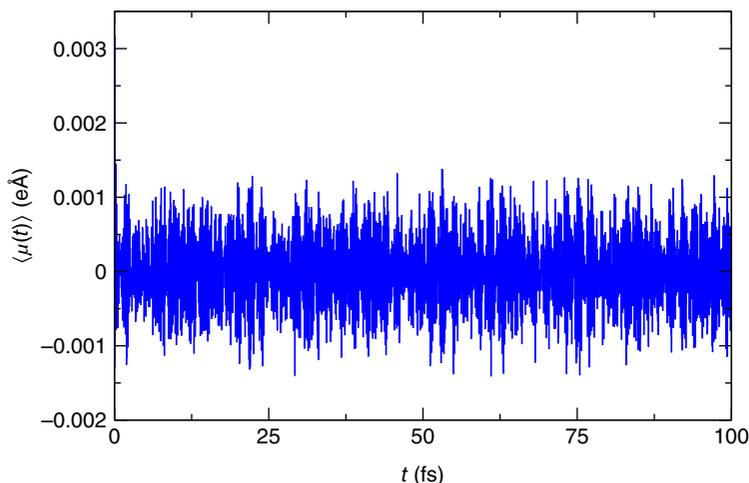


FIGURE 1 Time-dependent dipole moment for a naphthalene molecule obtained by applying a Dirac delta electric field pulse to the system.

independent simulations must be carried out in which the system is excited with a very short electric pulse (i.e., a Dirac delta pulse) that is applied in three mutually orthogonal directions to compute the polarizability tensor. Once the RT-TDDFTB electron dynamics calculation finishes, the time-dependent dipole moment is obtained, as shown in Figure 1.

In the limit of very weak perturbations, the system is said to be in the linear response regime, and the induced dipole moment of the system is given by

$$\mu(t) = \int_{-\infty}^{\infty} \alpha(t - \tau) E(\tau) d\tau \quad [10]$$

where τ is the time difference between the electric field and induced dipole moment, E is the electric field used to induce a rearrangement of charges inside the system, and α is the polarizability tensor. The quantity most easily accessible experimentally is the photo-absorption cross section given by:

$$\sigma(\omega) = \frac{4\pi\omega}{c} \text{Im}(\alpha) \quad [11]$$

where c is the speed of light, and $\text{Im}(\alpha)$ is the imaginary part of the average polarizability. $\text{Im}(\alpha)$ is obtained by the application of the convolution theorem, and Eq. [10] can be expressed in the frequency domain as

$$\mu(\omega) = \alpha(\omega) E(\omega) \quad [12]$$

To mimic the experimental absorption spectrum within this formalism, a damping term is typically included in Eq. [12]. This effectively incorporates a finite lifetime of the excited state in the response function, which produces a broadening in the absorption spectrum peaks.^{30, 31} The damping factor used in this tutorial on naphthalene is 0.01 fs^{-1} (see Figure 2a). This approach gives the polarizability along the direction of the initially applied field, and the full polarizability tensor is obtained from three independent time propagations where the only quantity that is altered is the direction of the external electric field.³² The average of the polarizability along the three Cartesian axes is taken as the absorption spectrum of naphthalene, as shown in Figure 2b.

Electron Dynamics of Naphthalene with a Laser-Type Perturbation

With the absorption spectrum of naphthalene properly computed, one can apply a laser-type perturbation tuned to the lowest excitation energy (i.e., 5.63 eV) of the system. As in the case of the Dirac delta pulse, one can also set the polarization of the laser field to any orientation with respect to the molecule; however, for this example on naphthalene, we have oriented the laser field in the direction of maximum polarizability to produce the maximum variation of the dipole moment as a function of time, as shown in Figure 3.

It can be noted from Figure 3 that the naphthalene molecule exhibits a linearly increasing dipole moment, as expected from a quantized system in the linear response regime that is continuously excited and in the absence of any dissipative mechanisms.³³ As such, this can be used to check that the laser is in tune with the electronic excitation energy and that the simulation is indeed in the linear response regime.

RT-TDDFTB Electron Dynamics of a Realistic Large Systems

In this final section on RT-TDDFTB dynamics, we give an example of the techniques used in the previous tutorial section on a “real-life” application to a large system. In particular, we focus on the real-time electron dynamics of EET in a large plasmonic nanoantenna system using the RT-TDDFTB formalism.^{22, 23} Understanding and achieving a controlled transfer of energy in these novel systems has been a continual area of interest in various technologies including nanophotonic circuits,^{34–37} waveguide materials,^{38–40} and other natural light-harvesting antenna systems.^{41, 42}

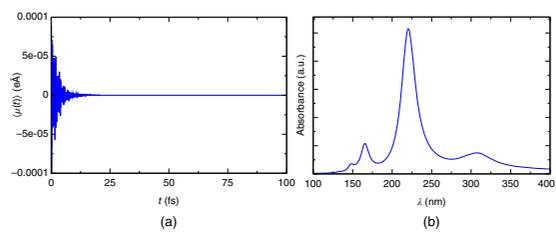


FIGURE 2 (a) Damped time-dependent dipole moment for a naphthalene molecule, and (b) absorption spectrum for naphthalene obtained from the Fourier transform of the damped time-dependent dipole moment.

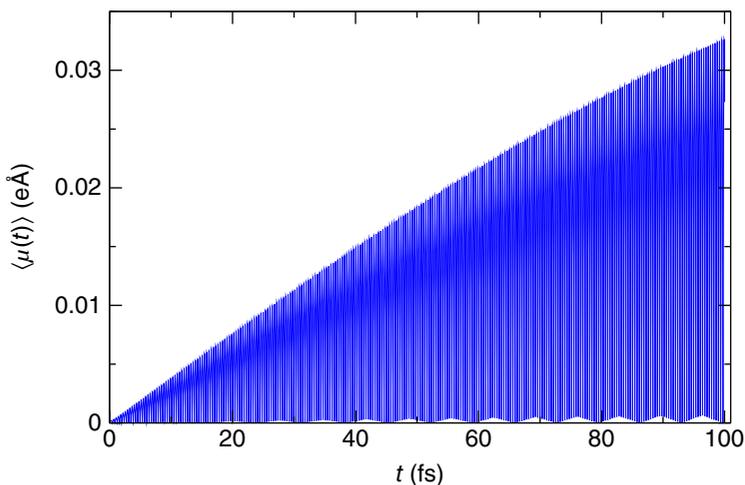


FIGURE 3 Time-dependent dipole moment for a naphthalene molecule obtained by applying a laser field tuned to the lowest excitation energy (5.63 eV) of the system.

Absorption Spectrum of a Single Plasmonic Nanoparticle

We begin our analysis of EET by first characterizing the plasmon resonance energy of a single nanoparticle (NP) containing 55 silver atoms and having an icosahedral shape. The geometry of this NP was optimized with the DFTB+ package using the hyb-0-2 set of DFTB parameters (available at dftb.org), and its absorption spectrum was obtained in the same manner as described in the tutorial section on naphthalene. As can be seen in Figure 4, a prominent peak, corresponding to the plasmon resonance is observed around 3.23 eV. This result is in good agreement with a time-dependent DFT calculation of 3.6 eV⁴³ and a recent experimental result of 3.8 eV⁴⁴ for similar-sized Ag NPs.

Exploring Excitation Energy Transfer in Ag Nanoparticle Chains

With the energy of a single Ag NP characterized, we can now proceed to an analysis of EET in plasmonic NP assemblies. Accordingly, we can use the single Ag NP to construct a variety of NP antenna configurations, each with interparticle distances (d) varying from 5 to 0.5 Å. We define the interparticle distance as the edge-to-edge distance between the NPs, and two of the model NP waveguides are shown in Figure 5. We can also construct a NP chain where the NPs “touch” each other ($d = 0$ Å) in which

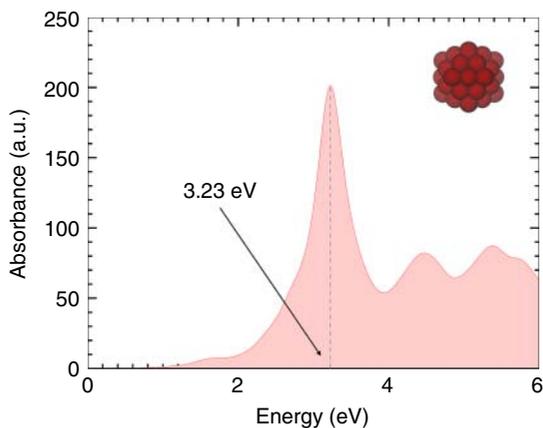


FIGURE 4 Absorption spectrum of a 55 atom icosahedral silver nanoparticle. A prominent plasmon resonance peak is observed around 3.23 eV. *Source:* From Ref. 23 with permission from The Royal Society of Chemistry.

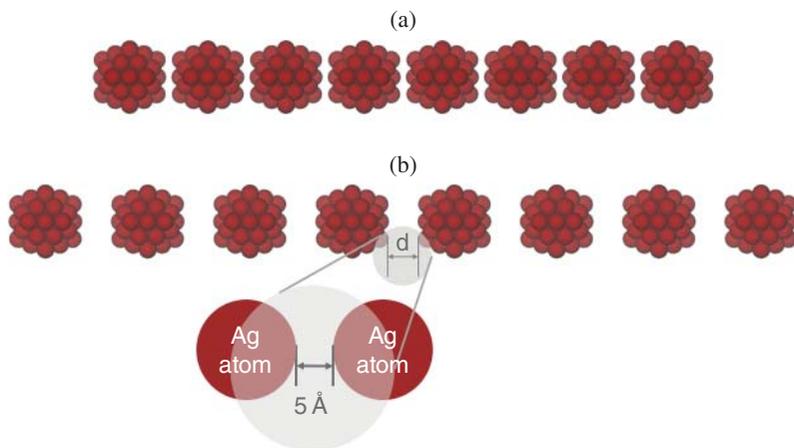


FIGURE 5 Pictorial representation of two finite chains with 8 Ag NPs of radius ≈ 1.23 nm and interparticle (edge-to-edge) distances equal to (a) 1 Å and (b) 5 Å. *Source:* From Ref. 23 with permission from The Royal Society of Chemistry.

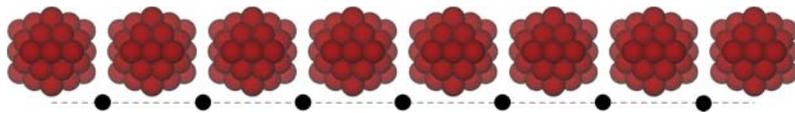


FIGURE 6 The field intensity values are taken at identical positions in each nanoparticle chain as shown by the black dots. The points lie exactly between two nanoparticles and on a line approximately 1 Å below the lowest atom in the NP. *Source:* From Ref. 23 with permission from The Royal Society of Chemistry.

the center-to-center distance between two atoms from adjacent NPs is less than the Ag–Ag bond-forming distance (the Ag–Ag atom bond length is 3.00 Å). It should be mentioned that each of these chains are extremely large systems containing a total of 440 atoms and, therefore, would be computationally prohibitive to calculate with conventional real-time TDDFT approaches.

To simulate EET along the NP chains, we excite only the first Ag NP in the chain using a monochromatic laser (using a similar approach discussed in the tutorial section) with an energy equal to the plasmonic resonance energy of a single Ag NP (3.23 eV). With this chosen initial condition, the entire system is allowed to evolve in time according to Eq. [2]. To quantify the EET efficiency along the chain, we can compute the electric field intensities, $I = \sqrt{\epsilon_0/\mu_0} \times |\mathbf{E}|^2$, at identical points between each of the NPs along the axial direction shown in Figure 6. \mathbf{E} is the total electric field, and ϵ_0 and μ_0 are the permittivity and permeability of free space, respectively.

Figure 7 shows the intensity trends of the NP chains with interparticle distances ranging from 0 to 5 Å. From the intensity trends in Figure 7, we observe a monotonic increase in the EET efficiency (i.e., the slope of the intensity lines decreases) as the interparticle distance is reduced from 5 to about 2 Å. This result is in qualitative agreement with previous studies on similar systems using classical electrodynamic methods.^{45, 46} This increase in EET efficiency can be attributed to an increase in capacitive coupling between the Ag NPs as the interparticle distance between them is reduced. In other words, this phenomenon is analogous to a charged capacitor,⁴⁷ where the capacitance of a capacitor increases as the charged plates are brought closer together. However, as the interparticle distance is further reduced below 2 Å, we observe an opposite trend of the EET efficiency. In particular, we see a sudden drop in EET efficiency for interparticle distances below 2 Å (i.e., the slope of the intensity line increases). This result is qualitatively opposite to what has been predicted by previous computational studies that have observed a decrease in EET when the NPs directly touch each other.^{45, 46}

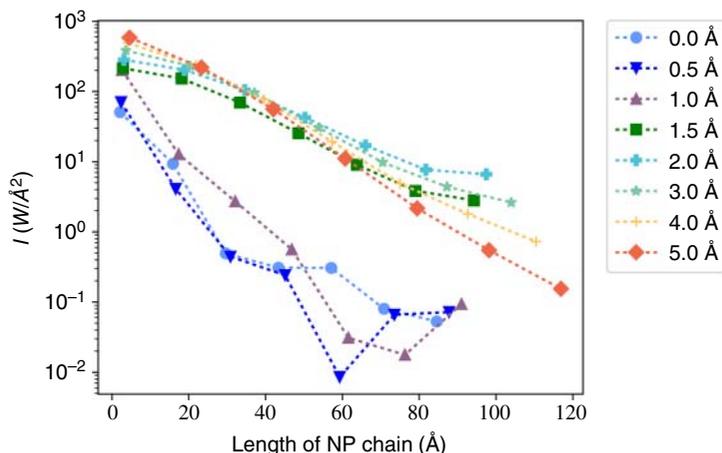


FIGURE 7 Field intensities along silver NP chains with varying interparticle distances. The first nanoparticle in each of the chains is excited at the plasmon resonance energy, and the intensity values are computed at the interparticle gaps of the NPs as shown in Figure 6. The excitation energy used in the simulation is equal to the plasmon resonance energy of the single Ag nanoparticle. A drastic drop in the field intensity is seen for Ag chains with interparticle spacings less than 2 Å. *Source:* From Ref. 23 with permission from The Royal Society of Chemistry.

Analyzing the Electronic Couplings in NP Chains

To understand these interesting dynamical effects, we can use RT-TDDFTB to analyze in detail the electronic couplings between the NPs in the plasmonic chain. To this end, we plot the RT-TDDFTB absorption spectra of Ag NP dimers with varying interparticle distances in Figure 8. On careful observation of Figure 8, we note that a single prominent peak, close to the value of the single NP plasmonic energy, can be observed for all interparticle distances. However, for interparticle spacings less than 2 Å, an additional peak (marked with arrows in Figure 8) forms in the absorption spectrum. The prominent peak normally arises due to interactions (hybridizations) between the basic plasmon resonances of the elementary nanostructures (in this case, the single Ag NP). This excitation is the bonding (symmetric) mode, normally known as the bonding dipole plasmon, or BDP, and is characterized by charge oscillations of the NPs in phase with each other.⁴⁸ The other peak appearing at lower energies and smaller interparticle distances is normally observed when an optical-frequency conductive pathway is established between two NPs, enabling the transfer of charge between them. This

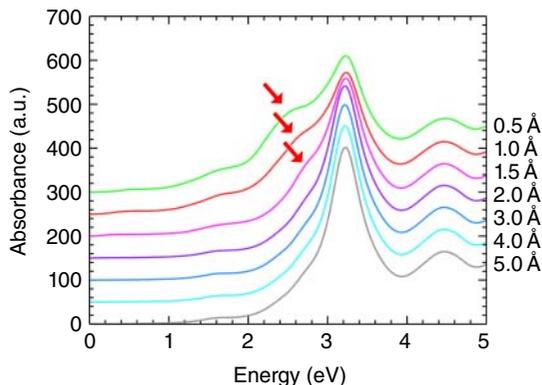


FIGURE 8 Absorption spectrum for Ag NP dimers with varying interparticle separations. An additional lower-energy peak (corresponding to a charge transfer plasmon excitation) emerges in the absorption spectrum for dimers having an interparticle spacing less than 2 Å, denoted by arrows. *Source:* From Ref. 23 with permission from The Royal Society of Chemistry.

conductive pathway can be physical, due to a physical bridge or due to quantum tunneling. This is known as a Charge Transfer Plasmon, or CTP.⁴⁹ Unlike the BDP, the CTP is characterized by a total charge moving between the two NPs of the dimer, which we observe as the lower-energy peak in our absorption spectrum. In our case of non-touching NPs, the CTP excitation can be attributed completely to quantum tunneling that establishes a conductive pathway between the two NPs of the dimer. While charge transfer plasmons have been previously observed theoretically in DFT and quantum-corrected classical models,^{47, 49, 50} this is the first example of predicting CTPs using RT-TDDFTB calculations.

Investigating the Nature of Plasmonic Excitations

To understand these drastic drops in EET efficiency for smaller interparticle spacings, we need an intuitive way to analyze these excitations. Figure 9 shows the charge distributions and the time-dependent changes in Mulliken charges for the NP dimer with an interparticle spacing of 1 Å. In the panels of this figure, we compare the time-dependent dynamics when the NP dimer is excited at either the BDP or CTP energy peak. When the NP dimer is excited at the CTP peak, one of the NPs shows a predominantly positive charge, while the other one shows a negative charge (Figure 9a). The time-dependent changes in Mulliken charges confirm this observation in

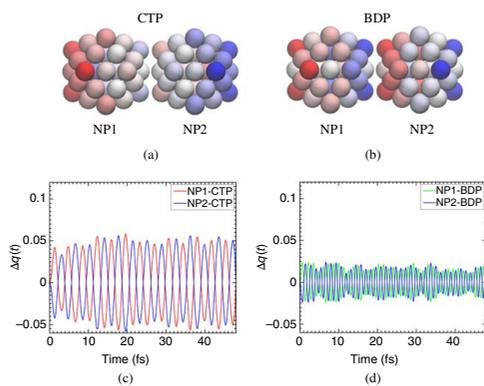


FIGURE 9 Snapshot of charge distributions at one instance in time for a Ag NP dimer with an interparticle distance equal to 1 Å excited at (a) the CTP peak and (b) the BDP peak. The CTP peak distributions show a total charge separation between the two NPs, while the BDP peak distributions show dipolar charge distributions within each of the NPs. The time-dependent changes in Mulliken charges are shown for the (c) CTP and (d) BDP peak for the same Ag NP dimer. For both the CTP and the BDP excitations, a net charge fluctuation is seen between the NPs, which indicates a hybridized nature of the BDP peak at subnanometer spacings. *Source:* From Ref. 23 with permission from The Royal Society of Chemistry.

Figure 9c. This behavior is characteristic of a CTP excitation, where an oscillating current occurs between the two NPs of the dimer. Also note in Figure 9a that we observe a slight dipolar nature of charge distributions near the particle edges. This can be attributed to the atomistic treatment of the NPs, whereby the charge transfer plasmon induced on the NP dimers also establishes a small opposing dipole on the inner edges of the same NPs due to interatomic electrodynamic interactions.

However, when the NP is excited at the BDP peak, we observe some charge transfer from one NP to the other, which is uncharacteristic of a BDP excitation.²² In particular, we observe that at subnanometer interparticle spacings, the pure BDP excitation forms a hybridized excitation that has some CTP character. As such, the decrease in the EET efficiency in smaller interparticle spacing chains can be attributed to the formation of this hybridized BDP. Because the hybridized BDP allows for a small charge transfer between the NPs, it reduces the capacitive coupling between the NPs. Going back to the capacitor analogy used previously, this can be thought of as a leaking capacitor. This, in turn, is ultimately responsible for the reduction in capacitive coupling between the NPs and hence the drop in EET efficiency, as revealed by these RT-TDDFTB calculations.

DFTB-BASED NONADIABATIC ELECTRON DYNAMICS

Adiabatic vs Nonadiabatic Dynamics

Until now, we have highlighted the use of RT-TDDFTB to probe the electron dynamics of large systems in external electric fields where the nuclei are held fixed. However, in this section, we discuss and give examples where this constraint is relaxed and the nuclei are allowed to evolve nonadiabatically on different potential surfaces. We first give a general overview of nonadiabatic dynamics and present a specific example of how DFTB can be further extended to give mechanistic insight in these excited-state processes.

In conventional electronic structure methods (such as DFT or even DFTB), one can solve the time-independent Schrödinger equation (TISE), $\hat{H}\Phi = \epsilon\Phi$, for a given set of nuclear coordinates (\mathbf{R}). Specifically, the electronic Hamiltonian is diagonalized to obtain a set of eigenvectors and eigenvalues that depend on the nuclear coordinates and, therefore, are known as the *adiabatic* eigenvectors and eigenvalues. When the TISE is solved for several sets of nuclear coordinates, we obtain the adiabatic potential energy surfaces (PESs).

In many chemical reactions, the wavefunction of the entire system can be expressed as a single adiabatic PES (instead of a linear combination of

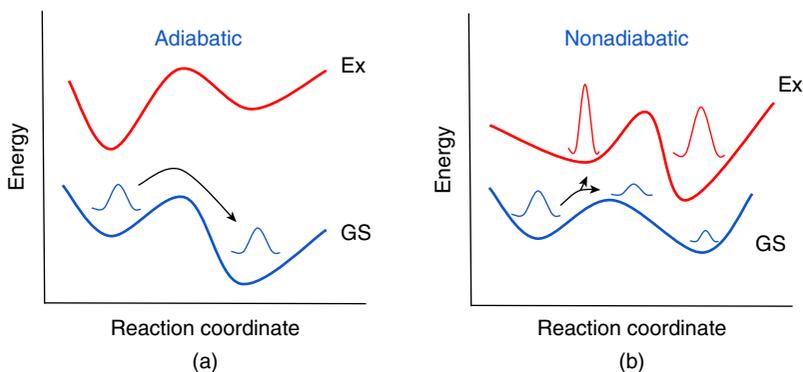


FIGURE 10 In an adiabatic processes (a), the wavefunction of the entire system is confined to a single PES, typically the ground-state (GS) surface. In a nonadiabatic process (b), the electronic amplitudes of the wavefunction will evolve over many PESs (i.e., the wavefunction is expressed as a linear combination of several adiabatic states).

several adiabatic surfaces). Often, this adiabatic PES corresponds to the ground-state of the species involved in the chemical reaction. For example, when we draw a reaction path depicting the transformation of the reactants ($\text{HCl} + \text{C}_2\text{H}_5\text{Br}$) into the products ($\text{HBr} + \text{C}_2\text{H}_5\text{Cl}$) through a transition state ($\text{C}_2\text{H}_5\text{BrCl}$), we inherently assume that these reactions occur along the ground-state PES (through a specific reaction coordinate). These chemical processes, where the wavefunction is confined to a single PES, are termed as adiabatic processes (see left panel of Figure 10a) and can be accurately described by the TISE (i.e., when the nuclear dynamics are governed by a single adiabatic PES, it is within the realm of the adiabatic dynamics). Born-Oppenheimer molecular dynamics and Car–Parrinello molecular dynamics are examples of such *ab initio* adiabatic dynamics, and many electronic structure packages such as CP2K,⁵¹ Quantum Espresso,⁵² VASP,⁵³ SIESTA,⁵⁴ NWChem,⁵⁵ and DFTB+²⁴ are routinely used by many researchers to calculate on-the-fly adiabatic dynamics.

In contrast to adiabatic dynamics, in most photochemical processes, the wavefunction of the entire system spans several PESs. A well-known example of this is fluorescence (i.e., the decay of a system from an excited (Ex) to the GS by emitting radiation), and the wavefunction of the entire system *cannot* be solely described as either the ground or excited state, but can only be expressed as a superposition of these states. These processes, where the electronic amplitudes of the wavefunction evolve over many

PESs, are known as nonadiabatic processes (Figure 10b). Furthermore, the nuclear motion that is described by several adiabatic PESs is known as nonadiabatic molecular dynamics (NAMD). Compared to adiabatic dynamics, only a few software packages such as Q-Chem⁵⁶ and Octopus³² have implemented direct on-the-fly NAMD (i.e., the entire NAMD simulation can be performed within the same package). Although direct on-the-fly implementations are scarce, many software packages such as PYXAID,⁵⁷ NEWTON-X,⁵⁸ SHARC,⁵⁹ QMFlows,⁶⁰ and LIBRA⁶¹ have implemented various NAMD methods. Also, most of these packages have an interface with traditional electronic structure packages such as Quantum Espresso and others, enabling us to perform NAMD calculations on both molecular and periodic systems. In this chapter, we present a tutorial and example of the fewest switches (FS) surface hopping (SH) method (a NAMD method), which we implemented in the DFTB+ software package.

To accurately capture nonadiabatic effects, a full quantum mechanical treatment of both the electronic and nuclear degrees of freedom is necessary. However, such a treatment is computationally prohibitive even for systems with moderate sizes (~ 50 -100 atoms). To circumvent this computational burden, only the electrons are treated quantum mechanically, and the nuclear motion is described classically. Among these so-called “mixed quantum-classical” (MQC) approaches, the Fewest-Switches Surface-Hopping (FSSH) method is one of the most popular methods, and in this chapter, we use it to study the nonadiabatic electron dynamics in organic systems.

Equations Governing Nonadiabatic Electron Dynamics

In any MQC method, the electron dynamics is captured by a time-dependent total wavefunction, $\Psi(\mathbf{r}, t)$, which satisfies the time-dependent electronic Schrödinger equation (TDSE):

$$i\hbar \frac{\partial}{\partial t} |\Psi(\mathbf{r}, t)\rangle = \hat{H}_{\text{el}}(\mathbf{r}, \mathbf{R}(t)) |\Psi(\mathbf{r}, t)\rangle \quad [13]$$

Here, $\hat{H}_{\text{el}}(\mathbf{r}, \mathbf{R}(t))$ is the electronic Hamiltonian operator for a set of nuclear coordinates, \mathbf{R} , at time t . Using an electronic structure method such as DFT, the electronic Hamiltonian can be solved to obtain the adiabatic eigenvectors $\Phi_i(\mathbf{r}, \mathbf{R}(t))$ and eigenvalues $\epsilon_i(\mathbf{R}(t))$. These adiabatic eigenvectors can correspond to molecular orbitals, Slater determinants, etc., depending on the employed electronic structure method. Accordingly, the corresponding eigenvalues can represent the molecular orbital energies, ground or excited-state energies, etc.

By assuming that these adiabatic eigenvectors form a complete orthonormal basis, we can express the total wavefunction as their linear combination, i.e.,

$$|\Psi(\mathbf{r}, t)\rangle = \sum_i a_i(t) |\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle \quad [14]$$

Inserting the above ansatz (Eq. [14]) into the TDSE (Eq. [13]), and multiplying the resulting expression with $\langle \Phi_j(\mathbf{r}, \mathbf{R}(t)) |$ from the left gives us the following differential equation:

$$\dot{a}_j(t) = -\frac{i}{\hbar} a_j(t) \epsilon_j(\mathbf{R}(t)) - \sum_i a_i(t) \langle \Phi_j(t) | \frac{\partial}{\partial t} \Phi_i(t) \rangle \quad [15]$$

In this expression, we have used the shorthand notation $|\Phi_i(t)\rangle \equiv |\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle$, and the nonadiabatic coupling elements, $C_{ji} \equiv \langle \Phi_j(t) | \frac{\partial}{\partial t} \Phi_i(t) \rangle$ are related to the derivative coupling vectors, \mathbf{d}_{ji} , through

$$C_{ji} = \dot{\mathbf{R}} \langle \Phi_j(\mathbf{r}, \mathbf{R}(t)) | \nabla_{\mathbf{R}} | \Phi_i(\mathbf{r}, \mathbf{R}(t)) \rangle \equiv \dot{\mathbf{R}} \mathbf{d}_{ji} \quad [16]$$

It is important to note that the coupling elements, C_{ji} , are responsible for the nonadiabatic electronic transitions between any two adiabatic states j and i . Integrating Eq. [15] gives us the wavefunction coefficients, $a_j(t)$, at each time step, allowing us to completely specify the total wavefunction, $|\Psi(\mathbf{r}, t)\rangle$ (we already know the adiabatic eigenfunctions at each nuclear step by solving the TISE). As a reminder, the set of $\{|\Phi_p(\mathbf{r}, \mathbf{R}(t))\rangle\}$ is known as an adiabatic basis set because it explicitly depends on the nuclear positions, \mathbf{R} .

The Classical Path Approximation

As mentioned earlier, in MQC methods, the nuclei are propagated according to classical mechanics, and the forces acting on the nuclei at each nuclear time step can be obtained using the Hellmann–Feynman theorem:

$$F_I = -\langle \Psi(\mathbf{r}, t) | \nabla_I \hat{H}_{\text{el}}(\mathbf{r}, \mathbf{R}(t)) | \Psi(\mathbf{r}, t) \rangle \quad [17]$$

Thus, the changes in the total wavefunction with time will have a direct influence on the nuclear motion. Often, to decrease computational costs, the electronic “back reaction” on the nuclear motion is neglected, and this approximation is known as the classical path approximation (CPA). As such, within the CPA, the nuclear motion is not affected by the electron dynamics. Nonetheless, the electron dynamics still depends on the nuclear

coordinates (due to the parametric dependence of the adiabatic eigenvectors on the nuclear coordinates at each time step), and the CPA is generally valid when the ground and excited-state PESs differ slightly.

Surface Hopping and Fewest Switches Criterion

SH is a general MQC nonadiabatic dynamics methodology with many variants, such as FSSH,^{62, 63} Decoherence Induced Surface Hopping (DISH),⁶⁴ Independent Electron Surface Hopping (IESH),⁶⁵ and others.⁶⁶ The similarities in most of these variants include the following:

1. The nuclei are propagated according to classical mechanics, and the forces on the nuclei, at any given instant of time, arise from a single adiabatic PES.
2. The nonadiabatic electron dynamics are captured by evolving the wavefunction using a stochastic algorithm for each trajectory and by averaging the results over a swarm of trajectories.

For each SH trajectory, we start our simulation from a single adiabatic PES (known as the active state). Next, we compute the probabilities for an electron to hop from this active state to all the other states (i.e., the adiabatic PESs). These computed probabilities are then compared with a uniformly generated random number. If the probability to jump from the active state to any other state is greater than the generated random number, the hop is accepted; otherwise, it will be rejected. This process is repeated for a swarm of trajectories, and the results are averaged over them to obtain the electron dynamics (as further explained in the next section).

Among the variants of SH approaches, the FSSH method is the most successful.^{57, 66, 69} In this method, the number of hops (switches) between the states is minimized, as described in the following example. Let us assume that our system has only two states or PESs (namely, ground and excited states), and we are running 10 SH trajectories. Let us also consider that at a particular instant of time (say, $t = 0$), 5 of these 10 trajectories are in the excited state, while the other 5 are in the GS; i.e., $[N_{\text{GS}}, N_{\text{Ex}}] = [5, 5]$ (see Figure 11a). Finally, let us assume that in the next time step ($t = 1$), there are six trajectories in the GS ($N_{\text{GS}} = 6$), and four in the excited state ($N_{\text{Ex}} = 4$). We can obtain this configuration, $[N_{\text{GS}}, N_{\text{Ex}}] = [6, 4]$, in several ways including: (i) switching three trajectories from the excited state to the GS, and only two trajectories from the ground to the excited state (Figure 11c) or (ii) allowing two trajectories to hop from the excited state to the GS, and only one trajectory to hop from the ground to the excited state, and so on. There

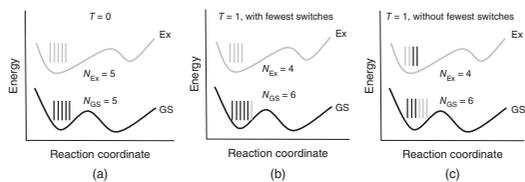


FIGURE 11 Number of trajectories in the excited (N_{Ex}) and ground state (N_{GS}) at times (a) $t = 0$, (b) $t = 1$, satisfying the fewest switches criterion, and (c) $t = 1$, without satisfying the fewest switches criterion. Trajectories that are present in the excited (ground) state at time $t = 0$ are shown in gray (black) to emphasize the fewest switches criterion at time $t = 1$.

will always be a way to achieve the desired configuration *with the least number of switches/hops* between the PESs. To achieve the desired configuration with minimum switches in the present example, only one trajectory in the excited state should switch to the GS, and zero trajectories should switch from the GS to the excited state (Figure 11b). This criterion is known as the FS criterion, and using this constraint, Tully proposed the FSSH mechanism to capture nonadiabatic electron dynamics phenomena. We use the CPA version of FSSH in conjunction with DFTB (as the underlying electronic structure method), as explained below.

Implementation Details of CPA-FSSH-DFTB

To understand charge transfer dynamics in large organic photovoltaic systems, we use the CPA-FSSH method in the DFTB+ software package (version 17.1). In particular, we use the DFTB3 variant⁷⁰ as our electronic structure method to obtain the adiabatic eigenvectors and eigenvalues at each nuclear time step. For this implementation, we assume that the photoinduced excited-state dynamics can be well described with a single-electron wavefunction. Hence, we adapt the single particle version of the time-dependent Kohn-Sham (TDKS) approximation,⁷¹ which assumes the time-dependent excited-state wavefunction $|\Psi(\mathbf{r}, t)\rangle$ can be represented as a linear combination of the GS KS orbitals $|\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle$, as given in Eq. [14]. This definition of $|\Psi(\mathbf{r}, t)\rangle$ has been shown to provide reasonable photoinduced charge transfer dynamics.^{72–76} Using the above definition, Eq. [15] can be numerically integrated using the fourth-order Runge–Kutta (RK4) method to obtain the expansion coefficients, $a_i(t)$, at each time step. During the integration, to compute the nonadiabatic coupling elements (Eq. [16]), we use the following well-established approximation⁶²

$$\left\langle \Phi_j(t) \left| \frac{\partial}{\partial t} \Phi_i(t) \right. \right\rangle = \frac{1}{2\tau} [\langle \Phi_j(t) | \Phi_i(t + \tau) \rangle - \langle \Phi_j(t + \tau) | \Phi_i(t) \rangle] \quad [18]$$

This quantity needs to be carefully calculated by following the random phases generated during the electronic structure calculations to obtain the adiabatic eigenvectors, $|\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle$.^{77–79} In DFTB+, these eigenvectors (MOs) are expanded as a linear combination of the AOs; i.e.,

$$|\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle = \sum_{\mu} C_{\mu i}(\mathbf{r}, \mathbf{R}(t)) |\varphi_{\mu}(\mathbf{r}, \mathbf{R}(t))\rangle \quad [19]$$

and in Eq. [18], the overlap between the adiabatic basis at two different time steps is computed as

$$\Phi_j(\mathbf{r}, \mathbf{R}(t + \tau)) | \Phi_i(\mathbf{r}, \mathbf{R}(t)) = \sum_{\mu\nu} C_{\mu j}(\mathbf{r}, \mathbf{R}(t + \tau)) C_{\nu i}(\mathbf{r}, \mathbf{R}(t)) S_{\mu\nu}(t + \tau, t) \quad [20]$$

Here, $S_{\mu\nu}(t + \tau, t)$ is the overlap between two AOs at two different time steps

$$S_{\mu\nu}(t + \tau, t) = \varphi_{\mu}(\mathbf{r}, \mathbf{R}(t + \tau)) | \varphi_{\nu}(\mathbf{r}, \mathbf{R}(t)). \quad [21]$$

These AO overlap integrals are explicitly evaluated with our in-house version of the DFTB+ code.

After obtaining the eigenvectors, eigenvalues, wavefunction expansion coefficients, and the coupling elements at each time step, the SH simulation is carried out using the CPA-FSSH scheme proposed by Akimov and Prezhdo,⁵⁷ which is an adaptation of the original FSSH scheme proposed by Tully and Hammes-Schiffer (with some modifications).⁶² In CPA-FSSH, an instantaneous active state is assigned and the probability of switching from the current state, $|\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle$, to any other state, $|\Phi_j(\mathbf{r}, \mathbf{R}(t))\rangle$, during a small time interval, $t \in [t, t + \delta]$ is calculated as

$$\tilde{g}_{ij}(t) = -\frac{2\text{Re} \left[\rho_{ij}(t) \langle \Phi_j(t) | \frac{\partial}{\partial t} \Phi_i(t) \rangle \right]}{\rho_{ii}(t)} \delta t \quad [22]$$

where $\rho_{ij}(t) = a_i^*(t) a_j(t)$ are the adiabatic electronic density matrix elements. Because of the CPA, one can ignore the back reaction of the electronic-nonadiabatic transition on the nuclear degrees of freedom. Hence, we do not rescale the velocity, which is one of the key ingredients in the original FSSH algorithm. Instead, following earlier work,^{57, 80} the transition probabilities are rescaled to preserve the energy conservation with the following expression:

$$g_{ij}(t) = \max[\tilde{g}_{ij}(t) * b_{ij}(t), 0] \quad [23]$$

with $b_{ij}(t) = e^{-(\epsilon_j - \epsilon_i)/k_B T}$ for $\epsilon_j > \epsilon_i$, and $b_{ij}(t) = 1$ for $\epsilon_j \leq \epsilon_i$, where k_B is Boltzmann's constant, and T is the temperature of the system, which is assumed to be constant during the nonadiabatic dynamics. Eq. [23] considers only the positive probability flux. Finally, a switch from the state $|\Phi_i(\mathbf{r}, \mathbf{R}(t))\rangle$ to any other state, $|\Phi_j(\mathbf{r}, \mathbf{R}(t))\rangle$, is accepted only when

$$\sum_{k=1}^{k=j-1} g_{ik} < \xi \leq \sum_{k=1}^{k=j} g_{ik} \quad [24]$$

where ξ is a uniform random number between 0 and 1.

An ensemble of CPA-FSSH trajectories are generated by propagating the nuclei using Eq. [17], by computing the electronic amplitudes using Eq. [15], and by determining the active state using Eq. [24].

Post-processing Tools

With the generated ensemble of CPA-FSSH trajectories, we compute the adiabatic reduced density matrix as

$$\bar{\rho}_{ij}(t) = \langle \rho_{ij}(t) \rangle \quad [25]$$

where $\langle \dots \rangle$ represents the ensemble average. The estimator, $\rho_{ij}(t)$, is expressed as

$$\begin{aligned} \rho_{ii}(t) &= \langle \Phi_i(\mathbf{r}, \mathbf{R}(t)) | \Phi_\alpha(\mathbf{r}, \mathbf{R}(t)) \rangle = \delta_{i\alpha} \quad [26] \\ \rho_{ij}(t) &= a_i^*(t) a_j(t) \quad (\text{for } i \neq j) \end{aligned}$$

Here, the diagonal elements of $\hat{\rho}$ (i.e., $\rho_{ii}(t)$) are chosen based on the active state $|\Phi_\alpha(\mathbf{r}, \mathbf{R}(t))\rangle$. Along a specific trajectory, $\mathbf{R}(t)$, the instantaneous population at time t is considered as 1 for the active state $|\Phi_\alpha(\mathbf{r}, \mathbf{R}(t))\rangle$, and 0 for all other states. The off-diagonal elements are computed based on the wavefunction expansion coefficients, $a_j(t)$.

Apart from the adiabatic populations mentioned above, we also need the time-dependent diabatic/charge populations on each moiety of the entire system to characterize the photo-induced charge transfer dynamics. The charge population on a specific fragment is obtained by projecting the adiabatic reduced density matrix onto the AO basis associated with that molecular fragment, N , as:

$$P_N(t) = \text{Re} \left[\sum_{\mu \in N} \sum_{ij} \rho_{ij}(t) C_{\mu i}(t) S_{\mu\nu}(t) C_{\nu j}(t) \right] \quad [27]$$

where $S_{\mu\nu}(t) = \varphi_\mu(\mathbf{r}, \mathbf{R}(t)) | \varphi_\nu(\mathbf{r}, \mathbf{R}(t))$ is the AO overlap matrix at time t . The expectation value of the charge population is thus

$$\bar{P}_N(t) = \langle P_N(t) \rangle \quad [28]$$

where $\langle \dots \rangle$ represents the ensemble average over the CPA-FSSH trajectories.

COMPUTATIONAL DETAILS

To describe nonadiabatic electron dynamics in large organic systems, we use the MQC CPA-FSSH method in conjunction with DFTB, which has

been shown to give accurate electronic structures for model organic photovoltaic systems.^{81, 82, 90} The electronic structure calculations are performed at the DFTB3 level of theory⁷⁰ with the 3ob-3-1 Slater–Koster parameter set⁸³ as implemented in the DFTB+ package.²⁴ Dispersion interactions between the atoms are incorporated using the Lennard-Jones potential with UFF parameters.⁸⁴

The initialized wavefunction, $|\Psi(\mathbf{r}, 0)\rangle$, is the LUMO of the donor moiety, $|\Phi_{\text{LUMO}}^{\text{D}}\rangle$, which is a widely used approximation for simulating photoinduced charge transfer dynamics.^{72, 73} Here, the LUMO of the donor moiety is obtained from a separate DFTB calculation performed for the isolated donor. From the set of eigenvectors (MOs) of the entire system at the zeroth time step, $\{|\Phi_i(\mathbf{r}; \mathbf{R}(\mathbf{0}))\rangle\}$, an MO maximizing the overlap $\langle \Phi_{\text{LUMO}}^{\text{D}} | \Phi_i(\mathbf{r}; \mathbf{R}) \rangle$ is selected.

As explained above, the wavefunction $|\Psi(\mathbf{r}, 0)\rangle$ at the initial time step is represented as one of the MOs of the entire system. This choice provides a reasonable single-electron picture of the localized photo-excitation in the system. At all other time steps, to reduce computational cost, the size of the MO basis is truncated to a smaller set containing the LUMO to LUMO+9 orbitals, which are low-lying orbitals that participate directly in the photo-induced charge transfer process.

Several nuclear configurations are then generated with the following procedure. First, the system is equilibrated in an NVT (constant number of particles, volume, and temperature) ensemble for 50 ps with a 1 fs nuclear time step using the Nosé–Hoover chain thermostat as implemented in the DFTB+ package. From this NVT trajectory, 30 different nuclear conditions (coordinates and velocities at every 1 ps interval) are collected for the subsequent 4 ps-long quantum dynamics propagation. For each of these 30 nuclear trajectories, 10^4 FSSH trajectories should be carried out to achieve convergence. To compute the charge transfer population, an ensemble average over both the FSSH and nuclear trajectories is considered.

AN EXAMPLE ON CHARGE TRANSFER DYNAMICS IN ORGANIC PHOTOVOLTAICS

Above, we provided an overview of our implementation of MQC CPA-FSSH DFTB for treating CT dynamics. Here, we illustrate the CT dynamics of Phenyl-C61-butyrac acid methyl ester/polythiophene (PCBM-PT), which is a model system for understanding photo-induced charge transfer dynamics in organic photovoltaics.^{85–88}

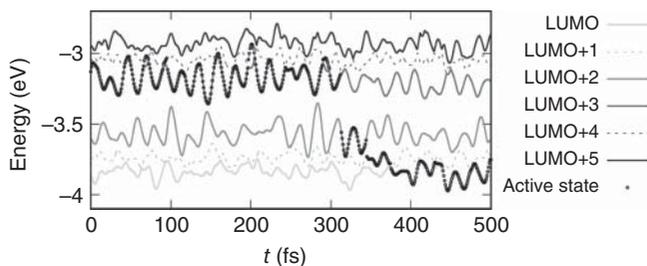


FIGURE 12 Variations in the energies of the LUMO to LUMO+5 orbitals (adiabatic states) of Phenyl-C61-butrylic acid methyl ester/Polythiophene (PCBM-PT) system are shown for a specific nuclear trajectory. Changes in the energy of the active state for a specific FSSH trajectory are also shown. For this FSSH trajectory, the active state switches between the LUMO+3 and LUMO+4 during the first 300 fs of the simulation. Later, it switches to the LUMO+2, the LUMO+1, and finally to the LUMO. It should be mentioned that 10 000 FSSH trajectories are calculated for each nuclear trajectory.

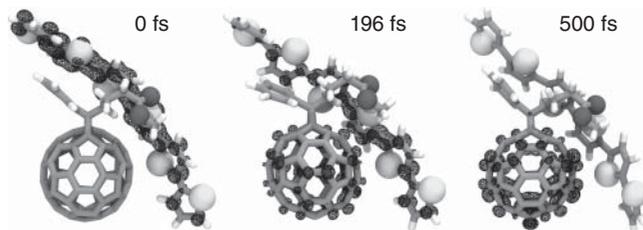


FIGURE 13 Charge density of the active state in Figure 12 at various time steps. Carbon and hydrogen (oxygen and sulfur) atoms are represented by gray and white sticks (balls), respectively. The charge density is shown as a black mesh. An iso-value of $0.001 \text{ e}/(\text{Bohr})^3$ is used.

The simulation begins by initializing the entire system's wavefunction, $|\Psi(\mathbf{r}, 0)\rangle$, by populating the LUMO of the PT molecule, $|\Phi_{\text{LUMO}}^{\text{PT}}\rangle$, and allowing it to evolve for a few picoseconds according to the CPA-FSSH-DFTB methodology. In Figure 12 the time-dependent MO energies of the LUMO to LUMO+5 are displayed for a single CPA-FSSH nuclear trajectory during the first 500 fs. Variations in the active state as a function of time for a single FSSH trajectory are also provided in the same panel. At 0 fs, the active state corresponds to the LUMO+3 orbital of the entire system; i.e., $|\Psi(\mathbf{r}, 0)\rangle \equiv |\Phi_{\text{LUMO}+3}^{\text{PCBM-PT}}\rangle$, which is primarily localized on the PT molecule ($|\Phi_{\text{LUMO}+3}^{\text{PCBM-PT}}\rangle \sim |\Phi_{\text{LUMO}}^{\text{PT}}\rangle$) as shown at the left of Figure 13.

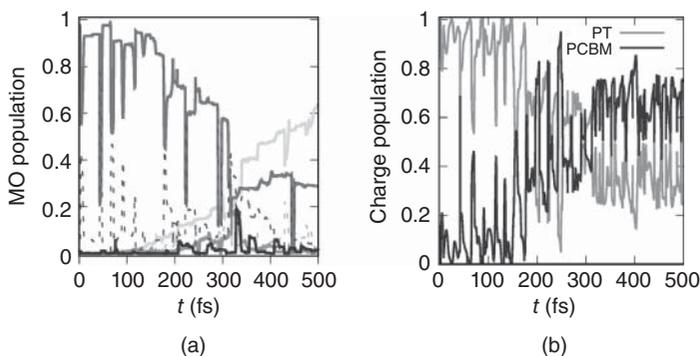


FIGURE 14 (a) Adiabatic MO and (b) charge population of the PCBM-PT system for a single nuclear trajectory. The same nuclear trajectory as in Figure 12 is used, but the results are averaged over 10 000 FSSH trajectories. MO colors in (a) are the same as in Figure 12. Due to the nonadiabatic transitions, the MO populations change with time. For this nuclear trajectory, an oscillation in the MO population between the LUMO+3 (localized on PT) and LUMO+4 orbitals (localized on PCBM) is observed during the first 100 fs. The same oscillation is also reflected in the charge population plot (panel b), where the gray and black curves represent the charge populations of the PT and PCBM molecules, respectively.

In contrast, the LUMO+2 and LUMO+4 orbitals are localized on the PCBM molecule. As the simulation proceeds, the LUMO+3 continues to be the active state until ~ 50 fs. At this stage, the active state hops from the LUMO+3 to the LUMO+4, but quickly returns to the LUMO+3. At the avoided crossings of the LUMO+3 and LUMO+4 (~ 50 or 190 fs), we find that both of these orbitals have a mixed character, as shown at the center of Figure 13. Hops between the LUMO+3 and LUMO+4 are also observed until 300 fs. Thereafter, the active state switches to the low-lying LUMOs (which have a strong PCBM character) and retains its acceptor character until the end of the simulation (as shown at the right of Figure 13). The hops between the LUMOs correspond to the nonadiabatic transitions.

In Figure 14a and b, we present the adiabatic MO and charge populations calculated for the same nuclear trajectory in Figure 12, but averaged over 10 000 FSSH trajectories. Due to the nonadiabatic transitions, the MO populations change with time. Also, due to the stochastic nature of the method, these changes are not smooth if we consider only a single CPA-FSSH nuclear trajectory. For example, as shown in Figure 14a, at 0 fs, the MO population is entirely on LUMO+3. However, as shown in Figure 12, until ~ 300 fs, an active state can continuously hop between the

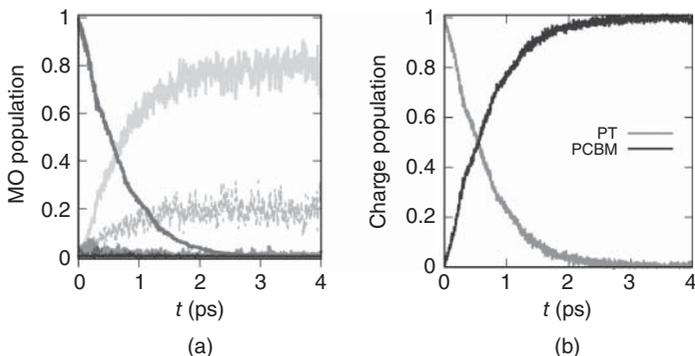


FIGURE 15 (a) Adiabatic MO and (b) charge population of the PCBM-PT system, averaged over several nuclear and FSSH trajectories. The colors in panel (a) are the same as those used in Figure 12. The LUMO+3 orbital (localized on PT) loses its population to the LUMO and LUMO+1 orbitals (localized on PCBM), suggesting a charge transfer from the donor to the acceptor. The gray and black curves in panel (b) denote the populations of the PT and PCBM molecules, respectively.

LUMO+3 (localized on PT) and LUMO+4 orbitals (localized on PCBM) due to the presence of several avoided crossings between these MOs for this nuclear trajectory. Due to these continuous hops, the MO population oscillates, and the same oscillation is also reflected in the charge population (Figure 14b). It should be mentioned that although the same nuclear trajectory is used in obtaining both Figures 12 and 14a, a one-to-one comparison cannot be made between them. This situation arises because the results in Figure 14a are averaged over 10 000 FSSH trajectories, whereas the results in Figure 12 are presented for a single FSSH trajectory. Since the MO energy fluctuations are the same for all 10 000 FSSH trajectories, understanding the nature of a single FSSH trajectory could be useful in understanding the averaged behavior of 10 000 FSSH trajectories. Finally, it is important to note that one cannot accurately assign the charge transfer time scales using the oscillating charge populations; to obtain any meaningful results, one needs to run at least a few tens of nuclear trajectories.

The MO and charge populations obtained after averaging over 30 nuclear trajectories and 10 000 FSSH trajectories for each nuclear trajectory are depicted in Figure 15. In Figure 15a, the decaying MO population corresponds to the PT-donor molecular orbital (the LUMO+3) while the increasing MO populations correspond to the PCBM-acceptor orbitals (the LUMO and LUMO+1). Due to the presence of more than one acceptor orbital, the

increasing MO populations continued to exhibit minor oscillations. However, as both of these MOs correspond to the PCBM moiety, such oscillations are not present in the charge populations (recall that the charge populations are obtained by projecting the adiabatic MO populations onto the diabatic AO basis). From the charge populations, we find a complete charge transfer within ~ 2 ps, but most of the population has been transferred by the first picosecond of the simulation. These simulated charge transfer time scales are in good agreement with earlier *ab initio* theoretical calculations on similar PCBM-PT models.^{85–88}

CONCLUSION AND OUTLOOK

In this chapter, we have presented an overview of DFTB-based excited-state dynamics with applications to both real-time time-dependent DFTB (RT-TDDFTB) and nonadiabatic dynamics. In both of these formalisms, a series of didactic tutorials and examples demonstrate how each approach is used in practice to reveal dynamical effects in chemical/material systems of contemporary interest. Within the sections on RT-TDDFTB, we demonstrated how this computational approach can be used to either calculate an electronic absorption spectrum or probe the electron dynamics of a chemical system in the presence of monochromatic light. Both of these choices give a different but complementary view of electron dynamics in large systems, as demonstrated by our example on plasmonic NPs. Within the sections on nonadiabatic dynamics, we explained our implementation of CPA-FSSH-DFTB, which was applied to study charge transfer dynamics in an organic photovoltaic system. By computing the charge transfer time scales in a model PCBM-PT system, which are in good agreement with earlier *ab initio* results, this approach shows immense promise for probing charge-transfer dynamics in even larger mesoscopic systems. Both of these examples in RT-TDDFTB and nonadiabatic dynamics extend the computational efficiency of DFTB to emerging areas of excited-state chemical dynamics (some of which have already been implemented in new computational hardware⁸⁹), creating an exciting opportunity for understanding these dynamical effects in large, complex systems.

ACKNOWLEDGMENTS

The RT-TDDFTB research and applications described in this work were supported by the U.S. Department of Energy, Office of Science, Early Career Research Program under Award No. DE-SC0016269. S.S.R.K.C.Y.

acknowledges Prof. Pengfei Huo at the University of Rochester for several helpful discussions related to the CPA-FSSH-DFTB part of the chapter and for providing computing resources through the Center for Integrated Research Computing (CIRC) at the University of Rochester. M.B.O. acknowledges financial support by the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT-FONCyT PICT-2017-0795). The authors also acknowledge the Office of Naval Research (grant N00014-18-1-2740) for financial support associated with the extensive research and preparation of this book chapter.

REFERENCES

1. T. Frauenheim, F. Weich, T. Köhler, S. Uhlmann, D. Porezag, and G. Seifert, *Phys. Rev. B*, **52**, 11492 (1995). Density-Functional-Based Construction of Transferable Nonorthogonal Tight-Binding Potentials for Si and SiH.
2. D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner, *Phys. Rev. B*, **51**, 12947 (1995). Construction of Tight-Binding-Like Potentials on the Basis of Density-Functional Theory: Application to Carbon.
3. J. Widany, T. Frauenheim, T. Köhler, M. Sternberg, D. Porezag, G. Jungnickel, and G. Seifert, *Phys. Rev. B*, **53**, 4443 (1996). Density-Functional-Based Construction of Transferable Nonorthogonal Tight-Binding Potentials for B, N, BN, BH, and NH.
4. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Phys. Rev. B*, **58**, 7260 (1998). Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties.
5. S. I. Allec, Y. Sun, J. Sun, C.-E. A. Chang, and B. M. Wong, *J. Chem. Theory Comput.*, **15**, 2807 (2019). Heterogeneous CPU+ GPU-Enabled Simulations for DFTB Molecular Dynamics of Large Chemical and Biological Systems.
6. S. M. Islam and P.-N. Roy, *J. Chem. Theory Comput.*, **8**, 2412 (2012). Performance of the SCC-DFTB Model for Description of Five-Membered Ring Carbohydrate Conformations: Comparison to Force Fields, High-Level Electronic Structure Methods, and Experiment.
7. K. Leong, M. E. Foster, B. M. Wong, E. D. Spoecke, D. Van Gough, J. C. Deaton, and M. D. Allendorf, *J. Mater. Chem. A*, **2**, 3389 (2014). Energy and Charge Transfer by Donor-Acceptor Pairs Confined in a Metal-Organic Framework: A Spectroscopic and Computational Investigation.
8. Q. Cui and M. Elstner, *Phys. Chem. Chem. Phys.*, **16**, 14368 (2014). Density Functional Tight Binding: Values of Semi-Empirical Methods in an Ab Initio Era.

9. G. Seifert and J.-O. Joswig, *WIREs Comput. Mol. Sci.*, **2**, 456 (2012). Density-Functional Tight Binding-An Approximate Density-Functional Theory Method.
10. P. Koskinen and V. Mäkinen, *Comput. Mater. Sci.*, **47**, 237 (2009). Density-Functional Tight-Binding for Beginners.
11. M. Elstner and G. Seifert, *Philos. Trans. R. Soc. A*, **372**, 20120483 (2014). Density Functional Tight Binding.
12. A. F. Oliveira, G. Seifert, T. Heine, and H. A. A. Duarte, *J. Braz. Chem. Soc.*, **20**, 1193 (2009). Density-Functional Based Tight-Binding: an Approximate DFT Method.
13. M. B. Oviedo, X. Zarate, C. F. Negre, E. Schott, R. Arratia-PÉRez, and C. G. SÁNchez, *J. Phys. Chem. Lett.*, **3**, 2548 (2012). Quantum Dynamical Simulations as a Tool for Predicting Photoinjection Mechanisms in Dye-Sensitized TiO₂ Solar Cells.
14. C. F. A. Negre, V. C. Fuertes, M. B. Oviedo, F. Y. Oliva, and C. G. SÁNchez, *J. Phys. Chem. C*, **116**, 14748 (2012). Quantum Dynamics of Light-Induced Charge Injection in a Model Dye-Nanoparticle Complex.
15. C. F. A. Negre, K. J. Young, M. B. Oviedo, L. J. Allen, C. G. Sánchez, K. N. Jarzemska, J. B. Benedict, R. H. Crabtree, P. Coppens, G. W. Brudvig, et al., *J. Am. Chem. Soc.*, **136**, 16420 (2014). Photoelectrochemical Hole Injection Revealed in Polyoxotitanate Nanocrystals Functionalized with Organic Adsorbates.
16. M. B. Oviedo, C. F. A. Negre, and C. G. Sánchez, *Phys. Chem. Chem. Phys.*, **12**, 6706 (2010). Dynamical Simulation of the Optical Response of Photosynthetic Pigments.
17. M. B. Oviedo and C. G. Sánchez, *J. Phys. Chem. A*, **115**, 12280 (2011). Transition Dipole Moments of the Qy Band in Photosynthetic Pigments.
18. C. R. Medrano, M. B. Oviedo, and C. G. Sánchez, *Phys. Chem. Chem. Phys.*, **18**, 14840 (2016). Photoinduced Charge-Transfer Dynamics Simulations in Noncovalently Bonded Molecular Aggregates.
19. C. Mansilla Wettstein, F. P. Bonafé, M. B. Oviedo, and C. G. Sánchez, *J. Chem. Phys.*, **144**, 224305 (2016). Optical Properties of Graphene Nanoflakes: Shape Matters.
20. E. N. Primo, M. B. Oviedo, C. G. Sánchez, M. D. Rubianes, and G. A. Rivas, *Bioelectrochemistry*, **99**, 8 (2014). Bioelectrochemical Sensing of Promethazine with Bamboo-Type Multiwalled Carbon Nanotubes Dispersed in Calf-Thymus Double Stranded DNA.
21. M. B. Oviedo and B. M. Wong, *J. Chem. Theory Comput.*, **12**, 1862 (2016). Real-Time Quantum Dynamics Reveals Complex, Many-Body Interactions in Solvated Nanodroplets.

22. N. V. Ilawe, M. B. Oviedo, and B. M. Wong, *J. Chem. Theory Comput.*, **13**, 3442 (2017). Real-Time Quantum Dynamics of Long-Range Electronic Excitation Transfer in Plasmonic Nanoantennas.
23. N. V. Ilawe, M. B. Oviedo, and B. M. Wong, *J. Mater. Chem. C*, **6**, 5857 (2018). Effect of Quantum Tunneling on the Efficiency of Excitation Energy Transfer in Plasmonic Nanoparticle Chain Waveguides.
24. B. Aradi, B. Hourahine, and T. Frauenheim, *J. Phys. Chem. A*, **111**, 5678 (2007). DFTB+, A Sparse Matrix-Based Implementation of the DFTB Method.
25. T. Frauenheim, G. Seifert, M. Elstner, Z. Hajnal, G. Jungnickel, D. Porezag, S. Suhai, and R. Scholz, *Phys. Status Solidi B*, **217**, 41 (2000). A Self-Consistent Charge Density-Functional Based Tight-Binding Method for Predictive Materials Simulations in Physics, Chemistry and Biology.
26. W. Liang, S. Yokojima, and G. Chen, *J. Chem. Phys.*, **110**, 1844 (1999). Generalized Linear-Scaling Localized-Density-Matrix Method.
27. K. Yabana and G. F. Bertsch, *Phys. Rev. B*, **54**, 4484 (1996). Time-Dependent Local-Density Approximation in Real Time.
28. R. G. Parr and R. G. Pearson, *J. Am. Chem. Soc.*, **105**, 7512 (1983). Absolute Hardness: Companion Parameter to Absolute Electronegativity.
29. M. B. Oviedo, Ph.D. thesis, Universidad Nacional De Cordoba, Facultad De Ciencias Químicas, Córdoba, Argentina, 2013. Dinámica Cuántica de Sistemas Moleculares Complejos en Tiempo Real.
30. H. Chen, J. M. McMahon, M. A. Ratner, and G. C. Schatz, *J. Phys. Chem. C*, **114**, 14384 (2010). Classical Electrodynamics Coupled to Quantum Mechanics for Calculation of Molecular Optical Properties: a RT-TDDFT/FDTD Approach.
31. D. J. Masiello and G. C. Schatz, *J. Chem. Phys.*, **132**, 064102 (2010). On the Linear Response and Scattering of an Interacting Molecule-Metal System.
32. A. Castro, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, M. A. Marques, E. Gross, and A. Rubio, *Phys. Status Solidi B*, **243**, 2465 (2006). Octopus: A Tool for the Application of Time-Dependent Density Functional Theory.
33. M. Belén Oviedo and C. G. SÁNchez, *ArXiv E-Prints*, ArXiv:1502.00491 (2015). Full Quantum Dynamics of the Electronic Coupling Between Photosynthetic Pigments.
34. A. Boltasseva and H. A. Atwater, *Science*, **331**, 290 (2011). Low-Loss Plasmonic Metamaterials.
35. M. L. Brongersma and V. M. Shalaev, *Science*, **328**, 440 (2010). The Case for Plasmonics.
36. C. M. Soukoulis and M. Wegener, *Science*, **330**, 1633 (2010). Optical Metamaterials—More Bulky and Less Lossy.
37. Y. A. Vlasov, M. O’Boyle, H. F. Hamann, and S. J. McNab, *Nature*, **438**, 65 (2005). Active Control of Slow Light on a Chip with Photonic Crystal Waveguides.

38. D. Solis, A. Paul, J. Olson, L. S. Slaughter, P. Swanglap, W. S. Chang, and S. Link, *Nano Lett.*, **13**, 4779 (2013). Turning the Corner: Efficient Energy Transfer in Bent Plasmonic Nanoparticle Chain Waveguides.
39. M. L. Brongersma, J. W. Hartman, and H. A. Atwater, *Phys. Rev. B*, **62**, R16356 (2000). Electromagnetic Energy Transfer and Switching in Nanoparticle Chain Arrays Below the Diffraction Limit.
40. S. Lal, S. Link, and N. J. Halas, *Nat. Photonics*, **1**, 641 (2007). Nano-Optics From Sensing to Waveguiding.
41. G. D. Scholes, G. R. Fleming, A. Olaya-Castro, and R. Van Grondelle, *Nat. Chem.*, **3**, 763 (2011). Lessons From Nature About Solar Light Harvesting.
42. M. Maiuri, M. B. Oviedo, J. C. Dean, M. Bishop, B. Kudisch, Z. S. D. Toa, B. M. Wong, S. A. McGill, and G. D. Scholes, *J. Phys. Chem. Lett.*, **9**, 5548 (2018). High Magnetic Field Detunes Vibronic Resonances in Photosynthetic Light Harvesting.
43. J. Ma, Z. Wang, and L.-W. Wang, *Nat. Commun.*, **6**, 10107 (2015). Interplay Between Plasmon and Single-Particle Excitations in a Metal Nanocluster.
44. T. LÜNskens, P. Heister, M. ThÄMer, C. A. Walenta, A. Kartouzian, and U. Heiz, *Phys. Chem. Chem. Phys.*, **17**, 17541 (2015). Plasmons in Supported Size-Selected Silver Nanoclusters.
45. M. Quinten, A. Leitner, J. R. Krenn, and F. R. Aussenegg, *Opt. Lett.*, **23**, 1331 (1998). Electromagnetic Energy Transport via Linear Chains of Silver Nanoparticles.
46. B. Willingham and S. Link, *Opt. Express*, **19**, 6450 (2011). Energy Transport in Metal Nanoparticle Chains via Sub-Radiant Plasmon Modes.
47. L. Wu, S. F. Tan, M. Bosman, J. K. Yang, C. A. Nijhuis, and P. Bai, *RSC Adv.*, **6**, 70884 (2016). Charge Transfer Plasmon Resonances Across Silver-Molecule-Silver Junctions: Estimating the Terahertz Conductance of Molecules at Near-Infrared Frequencies.
48. F. Wen, Y. Zhang, S. Gottheim, N. S. King, Y. Zhang, P. Nordlander, and N. J. Halas, *ACS Nano*, **9**, 6428 (2015). Charge Transfer Plasmons: Optical Frequency Conductances and Tunable Infrared Resonances.
49. W. Zhu, R. Esteban, A. G. Borisov, J. J. Baumberg, P. Nordlander, H. J. Lezec, J. Aizpurua, and K. B. Crozier, *Nat. Commun.*, **7**, 11495 (2016). Quantum Mechanical Effects in Plasmonic Structures with Subnanometre Gaps.
50. L. Wu, H. Duan, P. Bai, M. Bosman, J. K. W. Yang, and E. P. Li, *ACS Nano*, **7**, 707 (2012). Fowler-Nordheim Tunneling Induced Charge Transfer Plasmons Between Nearly-Touching Nanoparticles.
51. J. Hutter, M. Iannuzzi, F. Schifffmann, and J. VandeVondele, *WIREs Comput. Mol. Sci.*, **4**, 15 (2014). CP2K: Atomistic Simulations of Condensed Matter Systems.

52. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al., *J. Phys. Condens. Matter*, **21**, 395502 (2009). QUANTUM ESPRESSO: a Modular and Open-Source Software Project for Quantum Simulations of Materials.
53. G. Kresse and J. Hafner, *Phys. Rev. B*, **47**, 558 (1993). Ab Initio Molecular Dynamics for Liquid Metals.
54. J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, *J. Phys. Condens. Matter*, **14**, 2745 (2002). The SIESTA Method for Ab Initio Order-N Materials Simulation.
55. M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, et al., *Comput. Phys. Commun.*, **181**, 1477 (2010). NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations.
56. Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, et al., *Mol. Phys.*, **113**, 184 (2015). Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package.
57. A. V. Akimov and O. V. Prezhdo, *J. Chem. Theory Comput.*, **9**, 4959 (2013). The PYXAID Program for Non-Adiabatic Molecular Dynamics in Condensed Matter Systems.
58. M. Barbatti, G. Granucci, M. Persico, M. Ruckebauer, M. Vazdar, M. Eckert-Maksić, and H. Lischka, *J. Photochem. Photobiol., A*, **190**, 228 (2007). The On-the-Fly Surface-Hopping Program System Newton-X: Application to Ab Initio Simulation of the Nonadiabatic Photodynamics of Benchmark Systems.
59. S. Mai, P. Marquetand, and L. González, *WIREs Comput. Mol. Sci.*, **8**, E1370 (2018). Nonadiabatic Dynamics: the SHARC Approach.
60. F. Zapata, L. Ridder, J. Hidding, C. R. Jacob, I. Infante, and L. Visscher, *J. Chem. Inf. Model.*, **59**, 3191 (2019). QMflows—A Toolkit for Interoperable Parallel Workflows in Quantum Chemistry.
61. A. V. Akimov, *J. Comput. Chem.*, **37**, 1626 (2016). Libra: An Open-Source “Methodology Discovery” Library for Quantum and Classical Dynamics Simulations.
62. S. Hammes-Schiffer and J. C. Tully, *J. Chem. Phys.*, **12**, 4657 (1994). Proton Transfer in Solution: Molecular Dynamics with Quantum Transitions.
63. J. C. Tully, *J. Chem. Phys.*, **93**, 1061 (1990). Molecular Dynamics with Electronic Transitions.
64. H. M. Jaeger, S. Fischer, and O. V. Prezhdo, *J. Chem. Phys.*, **137**, 22A545 (2012). Decoherence-Induced Surface Hopping.
65. N. Shenoi, S. Roy, and J. C. Tully, *J. Chem. Phys.*, **130**, 174107 (2009). Nonadiabatic Dynamics at Metal Surfaces: Independent-Electron Surface Hopping.

66. L. Wang, A. Akimov, and O. V. Prezhdo, *J. Phys. Chem. Lett.*, **7**, 2100 (2016). Recent Progress in Surface Hopping: 2011–2015.
67. M. Barbatti, *WIREs Comput. Mol. Sci.*, **1**, 620 (2011). Nonadiabatic Dynamics with Trajectory Surface Hopping Method.
68. R. Crespo-Otero and M. Barbatti, *Chem. Rev.*, **118**, 7026 (2018). Recent Advances and Perspectives on Nonadiabatic Mixed Quantum–Classical Dynamics.
69. J. E. Subotnik, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi, *Annu. Rev. Phys. Chem.*, **67**, 387 (2016). Understanding the Surface Hopping View of Electronic Transitions and Decoherence.
70. M. Gaus, Q. Cui, and M. Elstner, *J. Chem. Theory Comput.*, **7**, 931 (2011). DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB).
71. C. F. Craig, W. R. Duncan, and O. V. Prezhdo, *Phys. Rev. Lett.*, **95**, 163001 (2005). Trajectory Surface Hopping in the Time-Dependent Kohn-Sham Approach for Electron-Nuclear Dynamics.
72. L. G. Rego and V. S. Batista, *J. Am. Chem. Soc.*, **125**, 7989 (2003). Quantum Dynamics Simulations of Interfacial Electron Transfer in Sensitized TiO₂ Semiconductors.
73. S. G. Abuabara, L. G. C. Rego, and V. S. Batista, *J. Am. Chem. Soc.*, **127**, 18234 (2005). Influence of Thermal Fluctuations on Interfacial Electron Transfer in Functionalized TiO₂ Semiconductors.
74. S. Pal, D. J. Trivedi, A. V. Akimov, B. Aradi, T. Frauenheim, and O. V. Prezhdo, *J. Chem. Theory Comput.*, **12** (4), 1436–1448 (2016). Nonadiabatic Molecular Dynamics for Thousand Atom Systems: A Tight-Binding Approach Toward PYXAID.
75. L. G. Rego, B. C. Hames, K. T. Mazon, and J.-O. Joswig, *J. Phys. Chem. C*, **118**, 126 (2013). Intramolecular Polarization Induces Electron-Hole Charge Separation in Light-Harvesting Molecular Triads.
76. A. Torres, R. S. Oliboni, and L. G. Rego, *J. Phys. Chem. Lett.*, **6**, 4927 (2015). Vibronic and Coherent Effects on Interfacial Electron Transfer Dynamics.
77. B. Space and D. Coker, *J. Chem. Phys.*, **94**, 1976 (1991). Nonadiabatic Dynamics of Excited Excess Electrons in Simple Fluids.
78. N. Yu, C. Margulis, and D. Coker, *J. Phys. Chem. B*, **105**, 6728 (2001). Influence of Solvation Environment on Excited State Avoided Crossings and Photodissociation Dynamics.
79. C. Hu, O. Sugino, H. Hirai, and Y. Tateyama, *Phys. Rev. A*, **82**, 062508 (2010). Nonadiabatic Couplings From the Kohn-Sham Derivative Matrix: Formulation by Time-Dependent Density-Functional Theory and Evaluation in the Pseudopotential Framework.

80. W. R. Duncan, C. F. Craig, and O. V. Prezhdo, *J. Am. Chem. Soc.*, **129**, 8528 (2007). Time-Domain Ab Initio Study of Charge Relaxation and Recombination in Dye-Sensitized TiO₂.
81. A. E. Jailaubekov, A. P. Willard, J. R. Tritsch, W.-L. Chan, N. Sai, R. Gearba, L. G. Kaake, K. J. Williams, K. Leung, P. J. Rossky, et al., *Nat. Mater.*, **12**, 66 (2013). Hot Charge-Transfer Excitons Set the Time Limit for Charge Separation at Donor/Acceptor Interfaces in Organic Photovoltaics.
82. M. Lee, E. Geva, and B. Dunietz, *J. Phys. Chem. Lett.*, **5**, 3810 (2014). Donor-To-Donor Vs. Donor-To-Acceptor Interfacial Charge Transfer States in the Phthalocyanine-Fullerene Organic Photovoltaic System.
83. M. Gaus, A. Goez, and M. Elstner, *J. Chem. Theory Comput.*, **9**, 338 (2013). Parametrization and Benchmark of DFTB3 for Organic Molecules.
84. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, *J. Am. Chem. Soc.*, **114**, 10024 (1992). UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations.
85. T. Liu and A. Troisi, *J. Phys. Chem. C*, **115**, 2406 (2011). Absolute Rate of Charge Separation and Recombination in a Molecular Model of the P3HT/PCBM Interface.
86. Z.-W. Zhao, Q.-Q. Pan, Y. Geng, Y. Wu, L. Zhao, M. Zhang, and Z.-M. Su, *ACS Sustain. Chem. Eng.*, **7**, 19699 (2019). Theoretical Insight Into Multiple Charge-Transfer Mechanisms at the P3HT/Nonfullerenes Interface in Organic Solar Cells.
87. Z. Xu, Y. Zhou, L. Groß, A. De Sio, C. Y. Yam, C. Lienau, T. Frauenheim, and G. Chen, *Nano Lett.*, **19**, 8630 (2019). Coherent Real-Space Charge Transport Across a Donor–Acceptor Interface Mediated by Vibronic Couplings.
88. T. P. Kaloni, P. K. Giesbrecht, G. Schreckenbach, and M. S. Freund, *Chem. Mater.*, **29**, 10248 (2017). Polythiophene: From Fundamental Perspectives to Applications.
89. J. M. Rodríguez-Borbón, A. Kalantar, S. S. R. K. C. Yamijala, M. B. Oviedo, W. Najjar, and B. M. Wong, *J. Chem. Theory Comput.*, **16**, 2085 (2020). Field Programmable Gate Arrays for Enhancing the Speed and Energy Efficiency of Quantum Dynamics Simulations.
90. S. S. R. K. C. Yamijala and P. Huo, *J. Phys. Chem. A.*, **125**, 628 (2021). Direct Nonadiabatic Simulations of the Photoinduced Charge Transfer Dynamics.

3

ADVANCES IN THE MOLECULAR SIMULATION OF MICROPHASE FORMERS

PATRICK CHARBONNEAU¹ and KAI ZHANG²

¹*Department of Chemistry and Department of Physics, Duke University, Durham, NC, USA*

²*Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, Jiangsu, China*

INTRODUCTION

Vapor leaves a boiling liquid, and oil floats above the aqueous fraction in a demixing vinaigrette. Such gas–liquid (or liquid–liquid) *phase* separations are ubiquitous. They occur in systems whatever their constituents may be—atoms, small molecules, proteins, polymer chains, colloids, etc.—as long as these components effectively interact through attractive forces sufficiently strong to compete with thermal noise. As attraction weakens compared to this noise, a critical temperature, T_c , is eventually reached, at which point the two (bulk) phases become indistinguishable and thus form a single homogeneous fluid (Figure 1a).

When attraction between components is frustrated by a longer-ranged repulsion, bulk low-temperature phases are replaced by mesoscale domains organized in various periodic morphologies, such as lamellae and cylinders (Figure 1b).^{1–3} Because these *microphases* also form irrespective of

Reviews in Computational Chemistry, Volume 32, First Edition.

Edited by Abby L. Parrill and Kenny B. Lipkowitz.

© 2022 John Wiley & Sons, Inc. Published 2022 by John Wiley & Sons, Inc.

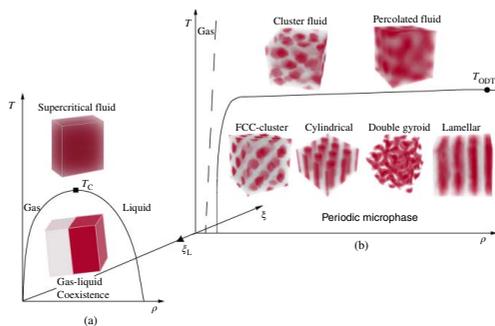


FIGURE 1 Schematic temperature T , density ρ , frustration ξ phase diagram for model microphase formers. (a) At low frustration $\xi < \xi_L$, coexisting gas and liquid phases at low-temperature transform continuously into a homogeneous fluid at the critical point T_c . (b) For frustration beyond the Lifshitz point, $\xi > \xi_L$, T_c becomes a weakly first-order (or weaker) transition at T_{ODT} . The low-temperature gas-liquid coexistence is then replaced by a series of periodic microphases, such as clusters, cylinders, gyroids, and lamellae. Disordered microphases, such as cluster and percolated fluids, are observed at higher temperature. *Source:* Adapted from Ref. [1].

the physical origin of the effective interactions, analogous patterns can be observed in systems as diverse as diblock copolymers,⁴ aggregated colloidal particles,⁵ oil droplets,⁶ magnetic materials,^{7, 8} as well as spin⁹ and charge density waves.¹⁰

The transition from phase to microphase separation upon increasing frustration, ξ , coincides with the gas–liquid critical point (at T_c) transforming into a weakly first-order (or weaker) order–disorder transition (at T_{ODT}) at the *Lifshitz point*, ξ_L .¹¹ At fixed ξ , above T_{ODT} , mesoscopic domains can be observed but remain disordered. On the ξ – T plane, the point (ξ_L, T_L) thus separates three regimes^{12, 13}: a high-temperature disordered phase, gas–liquid coexistence at low temperatures for $\xi < \xi_L$, and periodic microphases for $\xi > \xi_L$ (Figure 1). The frustration, so-called because particles are “frustrated” to choose between several competing states due to its existence, can be of energetic, entropic, and even topological origin, depending on the details of the system. We will see concrete examples in following sections.

Beyond the theoretical and aesthetic fascination with periodic microphases, these structures have also found myriad applications in drug delivery,^{14, 15} patterning^{16–20} and lithography.^{21, 22} In diblock copolymers, for instance, the well-controlled nano-size porous structures can be used to design nanofiltration membranes for water purification.²³ Ordered surfactant micelles can serve as templates for synthesizing mesoporous inorganic molecular sieves.²⁴ Compared with traditional top-down techniques of nanopatterning, the bottom-up self-assembly of microphases enables production of complex three-dimensional patterns with subnanometer precision, often using much simpler equipment. However, the emergence of undesired competing structures often interferes with the formation of targeted ones. External forces such as mechanical shearing,^{25–27} electric,²⁸ and magnetic fields²⁹ have thus been used to enhance and direct periodic microphase assembly. A thorough understanding of the link between particle-level interactions and the complex phase behavior of microphase formers is crucial for designing and controlling these next-generation materials.

In the rest of the introduction, we detail different experimental microphase formers and provide a minimal theoretical framework to present the simulation challenges associated with studying model microphase formers.

Block Copolymers

Block copolymers are by far the most studied microphase formers. Simple diblock copolymers exhibit the whole spectrum of elementary periodic mesoscopic patterns,^{4, 30} and more complex block structures can give rise to structures of seemingly unbounded richness.^{31–33} The competing interactions that give rise to microphase formation in these systems are

effective. In a diblock copolymer, for instance, the monomers of the different blocks, A and B , demix at low temperatures, which forms the basis of the short-range attraction. The tendency to segregate is traditionally quantified by the Flory–Huggins parameter $\chi_{AB} = \frac{z}{k_B T} [\epsilon_{AB} - \frac{1}{2}(\epsilon_{AA} + \epsilon_{BB})]$, where z is the number of nearest neighbors around a monomer, k_B is the Boltzmann constant, and $\epsilon_{ij} < 0$ ($i, j \in \{A, B\}$) is the interaction energy between a pair of monomers. Together with the degree of polymerization N , $\chi_{AB}N$ plays the role of an inverse temperature. Chain connectivity, however, prevents blocks from segregating completely, thus giving rise to long-range frustration. Different mesoscale morphologies then appear as the number fraction f_A (or f_B) of component A (or B) is changed (Figure 2).

Block copolymers are also of particular theoretical interest because the (mean-field) self-consistent field theory of Leibler for f_A vs $\chi_{AB}N$ was solved,³⁵ just as the first mean-field description of a frustrated lattice model was obtained.³⁶

Surfactants and Microemulsions

Surfactants and other small amphiphilic molecules present a hydrophilic head and a hydrophobic tail (containing one or multiple chains). Depending on the relative sizes of the head and tail, the overall effective shape of a surfactant molecule may be cone-like, truncated cone-like, or cylinder-like (Figure 3a–c). Head–head and tail–tail interactions drive segregation, but as for block copolymers, frustration arises from the covalent bonding between the head and the tail.^{2, 39}

Small amounts of surfactants can be dissolved in water, thus forming a homogeneous aqueous solution, but once the *critical micelle concentration* (cmc) is reached supramolecular aggregates assemble instead.⁴⁰ To minimize the contact between the hydrophobic tail and water, surfactants form spherical (or cylindrical) micelles, bicontinuous cubic structures, vesicles, sponge-like phases, planar bilayers, and their inverted counterparts⁴¹ (Figure 3). Note that ordered mesophases formed in amphiphilic–water systems are also known as lyotropic liquid crystals.⁴²

In ternary oil–water–surfactant mixtures, surfactants tend to form monolayers at oil–water interfaces. Micelles swell as they fill with oil, thus forming a thermodynamically stable *microemulsion* of small surfactant-coated oil droplets in water (or water droplets in oil in oil-rich systems). At higher surfactant concentrations, other geometries can assemble, such as alternating layers of oil and water, as well as bicontinuous phases with interpenetrating oil and water domains.⁴¹

Synthetic amphiphiles find notable applications as detergents and lubricants and are also used for drug (and vaccine⁴³) delivery and oil

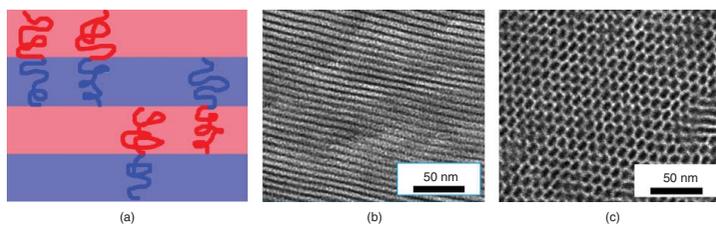


FIGURE 2 (a) Stacking of diblock copolymer chains of two immiscible components *A* (ebook: red, print book: light gray) and *B* (ebook: blue, print book: dark gray) in mesoscopic domains. TEM images of (b) lamellae and (c) hexagonally packed cylinders formed by P3HS-*b*-PDMS block copolymers. *Source:* From Azuma et al. 2018³⁴ / with permission of American Chemical Society.

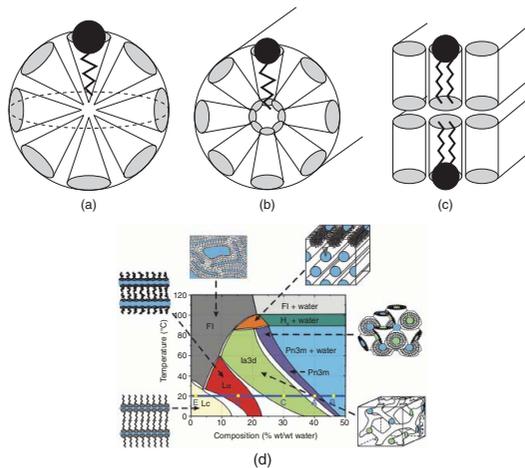


FIGURE 3 Schematic representation of (a) spherical micelles, (b) cylindrical micelles, and (c) planar bilayers formed by cone-like, truncated cone-like, and cylinder-like surfactants, respectively. (d) Temperature–composition phase diagram of monoolein in water, which is the most common lipid system for *in meso* membrane protein crystallization. *Source:* Adapted from Refs. [37, 38].

recovery.^{44–46} Lipids, which are natural amphiphiles, are the foundation of biological membranes. Self-assembled lipid vesicles are indeed primitive models for living cells and principles about surfactant microphases can be used to rationalize the structure of biological membranes^{37, 47} (Figure 3d). Synthetic mesophases are even used as intermediates to bring membrane proteins to crystallize.³⁸ The organization of real cell membranes is, however, much more complex and may contain hundreds of different lipid species mixed with proteins, cholesterols, and other molecules.⁴⁸ Besides three-dimensional microphases, such as gyroid structures in trans-Golgi network,⁴⁹ lipids are also thought to develop mesoscopic domains on the quasi-two-dimensional surface of cell membrane.^{50, 51} These proposed “lipid rafts”, which are concentrated in cholesterols and glycosphingolipids, are believed to strongly influence membrane fluidity and protein trafficking.^{52–55}

Lattice Spin Systems

Sinusoidal magnetic order with stacking of ferromagnetic layers was first observed in rare-earth elements, such as erbium (Er),⁵⁶ and in their compounds, such as cerium antimonide (CeSb).⁵⁷ Because atoms rest on regular crystalline lattices in such systems, their magnetic microphases can be explained by the competing ferromagnetic and antiferromagnetic interactions between spins (Figure 4a,b), which result from the orbital and spin angular momenta of unpaired electrons.⁶⁰ Some high-temperature superconductors also display exotic stripe electronic phases,^{61, 62} which are spin-like,^{63, 64} as the result of the competition between the zero-point kinetic energy of doped holes, which tends to delocalize charges, and the antiferromagnetic interactions among magnetic moments and the Coulomb interactions among charges, which favor their localization⁶⁵ (Figure 4c). More generally, a wide variety of nanoscale charge and spin density waves can develop in transition metal oxides with strongly correlated electronic interactions.^{8, 66}

Colloidal Suspensions

Due to the ease of visualizing micron-scale particles, colloidal suspensions are commonly used as model liquids, crystals, and glasses.^{67, 68} Various attempts have also been made to prepare colloidal microphases. The standard paradigm for synthesizing competing attractive and repulsive effective interactions in colloids combines two separate physical processes. First, the contribution from dispersion forces, which sums up to a strong effective colloidal van der Waals attraction—scaling as $\sim 1/h$ at short interparticle

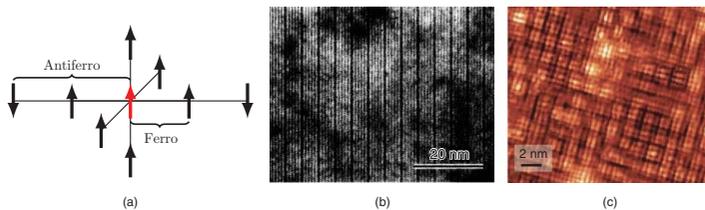


FIGURE 4 (a) A simple cubic lattice model in which a central spin (ebook: red, print book: gray) interacts ferromagnetically with its six nearest neighbors but antiferromagnetically with two next-nearest neighbors along one axial direction. TEM image of (b) lamellar superlattice in Ag-Mg alloy. *Source:* From Fujino et al. 1987⁵⁸ / with permission of American Physical Society, and (c) checkerboard charge-order state in Na-doped cuprates. *Source:* From Hanaguri et al. 2004⁵⁹ / with permission of Springer Nature.

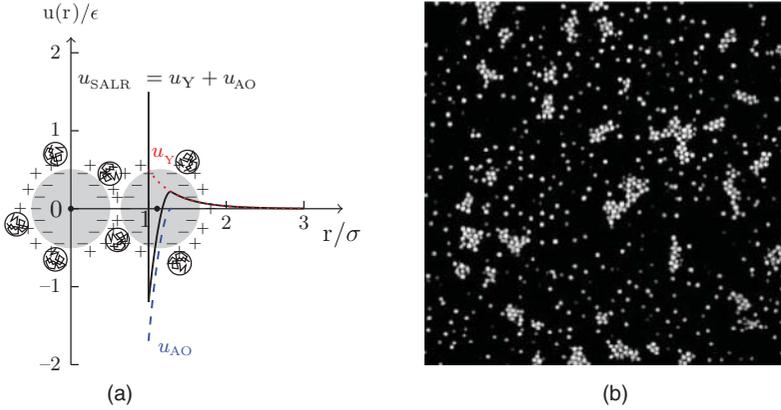


FIGURE 5 (a) Effective Yukawa u_Y (dotted), depletion u_{AO} (dashed) and overall competing u_{SALR} (solid) interactions between a pair of charged colloidal particles of diameter σ (gray), immersed in a suspension of polymers (small circles) of radius of gyration $R_G = 0.15\sigma$. (b) Confocal microscopy image of clusters formed in a colloid–polymer mixture at colloid volume fraction $\phi = 0.086$. *Source:* From Stradner et al. 2004⁵ / with permission of Springer Nature.

surface separations h —and leads to irreversible flocculation of colloidal particles,⁶⁹ must be negated. A common approach is to solvate electrolytes along with charged colloids to obtain a diffuse electric double-layer.^{70, 71} This scheme gives rise to an exponentially decaying repulsion $\sim e^{-\kappa_D h}$ at small h , where κ_D is the inverse Debye screening length, and to a screened electrostatic (Yukawa) repulsion at larger center-to-center distances r (Figure 5a),

$$u_Y(r) \propto \frac{e^{-\kappa_D(r-\sigma)}}{r/\sigma} \quad [1]$$

where $h = r - \sigma$ for particles of diameter σ . Access to the deep energy minimum associated with dispersion forces is then prevented, as described by the Derjaguin–Landau–Verwey–Overbeek (DLVO) theory.⁷² The Debye screening length κ_D^{-1} depends on the Bjerrum length λ_B , which is the distance between two unit charges when their Coulomb energy through a medium with dielectric constant ϵ_r equals the thermal energy, i.e., $k_B T = \frac{e^2}{4\pi\epsilon_0\epsilon_r\lambda_B}$. One can thus generate a fairly strong long-range repulsion by tuning the dielectric constant ϵ_r of the solvent medium.

Second, a weaker attraction regime is obtained and controlled by solvating globular non-adsorbing polymers. The polymer-mediated depletion

attraction thus generated is well described by the Asakura-Oosawa-Vrij (AO) model^{70, 73, 74}

$$u_{\text{AO}}(r) = \begin{cases} -\rho_R k_B T \frac{\pi}{12} (\sigma + 2R_G)^3 \left[2 - 3 \frac{r}{\sigma + 2R_G} + \left(\frac{r}{\sigma + 2R_G} \right)^3 \right], & r < \sigma \\ 0, & \sigma < r < \sigma + 2R_G \\ 0, & r > \sigma + 2R_G \end{cases} \quad [2]$$

where ρ_R and R_G are the number density and radius of gyration of the solvated polymer coils (Figure 5a).

Summing these two model contributions results in a short-range attractive and long-range repulsive (SALR) interaction⁷⁵ (Figure 5a), sometimes described as a “mermaid” potential, owing to its “attractive head” and “repulsive tail”.⁷⁶ Colloidal suspensions of particles with SALR interactions do form various cluster mesophases,⁵ but only disordered ones, such as cluster fluids and gels, have thus far been observed experimentally (Figure 5b).^{77, 78} The difficulty to realize periodic colloidal microphases is thought to result from the non-additivity of the two types of interactions.⁷⁵

Other Examples

In addition to the more common examples described above, a variety of other microphase-forming systems have been reported. The compromise between effective short-range order and long-range frustration can indeed arise in a variety of other physical contexts, such as external rotating magnetic fields⁷⁹ and Janus particles.⁸⁰ Other schemes to form microphases have also been explored. Introducing pairwise repulsion on two different length scales, for instance, gives rise to stripe and cluster mesophases.^{81–83} Particles with very soft repulsive cores,⁸⁴ as in certain dendrimer models,⁸⁵ form multiply-occupied cluster crystals at high density.^{86–88} Purely repulsive interactions in rod-sphere mixtures⁸⁹ and in pear-shaped particles,⁹⁰ can even give rise to entropy-driven microphase formation.

Field Theory of Microphase Formation

It is not the purpose of the current review to discuss how various theoretical tools have been brought to bear on the description of microphases. More complete treatments based on liquid-state theory^{91, 92} or density functional

theory (DFT),^{93, 94} for instance, can be found elsewhere. We here mostly focus on the phenomenological field theory description of the universality of the microphase formation and of the nature of the ODT.

The Ginzburg–Landau theory of gas–liquid-like phase transitions uses a scalar order parameter, the density field $\phi(\mathbf{r})$, and the (nearly) symmetric nature of the pair interactions to express the free energy functional as an even-order power series,

$$F_0[\phi(\mathbf{r})] = \int d\mathbf{r} [a\phi^2(\mathbf{r}) + b\phi^4(\mathbf{r}) + c|\nabla\phi(\mathbf{r})|^2] \quad [3]$$

where the first two terms capture uniform bulk contributions and the gradient term accounts for the cost of forming domain walls or interfaces.⁹⁵ Introducing long-range frustration as a repulsive potential $u_{\text{LR}}(r)$ generically transforms the total free energy functional to⁹⁶

$$F[\phi(\mathbf{r})] = F_0[\phi(\mathbf{r})] + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \phi(\mathbf{r})\phi(\mathbf{r}') u_{\text{LR}}(|\mathbf{r} - \mathbf{r}'|) \quad [4]$$

Minimizing $F[\phi(\mathbf{r})]$ with respect to $\phi(\mathbf{r})$ identifies the stability regime for equilibrium microphases,^{2, 94, 97} whichever the microscopic origin of the attractive and the repulsive contributions may be.

Analysis of this field theory further informs us on the universal nature of the ODT. At the mean-field level, the ODT is second order in nature, albeit in a different universality class than the gas–liquid critical point. In three dimensions, however, the nature of the transition depends on the strength of the frustration, or more precisely on whether the frustration is isotropic or not. Isotropic frustration gives rise to a weakly first-order transition,¹¹ but anisotropic frustration gives rise to a critical ODT transition, albeit of a universality class that depends on the number of frustrated directions. As a result, the universality of the Lifshitz point also depends on the type of frustration⁹⁸.

Molecular Simulations and Challenges

Although existing theoretical descriptions successfully explain the universality of some of the microphase-forming systems, they struggle to provide quantitative estimates for any specific material and to describe the richness of the disordered microphase regime. Molecular simulations are standard recourses in this situation, but have long struggled to characterize microphase formation. Despite extensive efforts at on- and off-lattice

Monte Carlo (MC), molecular dynamics (MD), and other simulation approaches to microphase-forming colloids,^{99–101} surfactants,^{102, 103} block copolymers,^{104–107} and spins^{108, 109} the competing morphologies and occupancies of microphases are challenging to equilibrate. Extracting qualitatively (let alone quantitatively) reliable insight has thus largely been out of computational reach. Direct simulation methods can generate certain microphases and delineate a “dynamical phase diagram”, but the resulting observations typically correspond to kinetically trapped metastable states on the relatively short time scales accessible to molecular simulations. Comparable kinetic hurdles to study simple equilibrium phases, such as gas, liquid, crystal, and even liquid–crystal, have long been surmounted through the development of specialized techniques, such as Frenkel–Ladd thermodynamic integration for crystals, Gibbs ensemble MC for gas–liquid coexistence, finite-size scaling analysis for critical points, etc.¹¹⁰ An equivalent array of treatments for microphases, however, has long been missing.

The numerical challenge is two-fold.

1. Periodic microphases are characterized by both morphology and occupancy (or, equivalently, periodicity), but molecular simulations, which are necessarily finite in size, only allow for the formation of patterns commensurate with the chosen box size. In addition, arbitrarily initialized structures often become kinetically locked in.^{76, 107, 109, 111} Because conventional simulation methods do not include schemes for the lattice occupancy to fluctuate, relaxation to the equilibrium microphase is thus highly unlikely. Even if an equilibrium state is found by chance, a simple simulation scanning over temperatures or densities will still fail to capture phase transitions, because equilibrium morphology and occupancy evolve with the state point. In this context, slow annealing or even parallel tempering are of little help. In light of these challenges, simulation methods that allow mesophase occupancy fluctuation,¹¹² that account for it by free energy calculations,^{113, 114} and that overcome barriers to transform are essential to obtain equilibrium insight into periodic microphase formation.
2. Disordered microphases also present a number of simulation challenges. For instance, achieving equilibration of low-density cluster fluids entails establishing a kinetic balance of inter-cluster particle attachments and detachments, which are slow and rare for they involve crossing high energy barriers and require transport over large distances. High-density fluids can also dynamically arrest in various gel-like structures, and upon approaching the periodic microphase

regime marked dynamical slowdown can emerge. Systems with very strong repulsion further form clusters that arrest in a Wigner-like glass. While slow annealing and parallel tempering do help reach equilibrium in this regime, their computational cost can quickly become out of reach.

In this review, we mainly discuss a free energy-based molecular simulation approach to the equilibrium self-assembly of microphases, which has been successfully applied to lattice spin models,^{115, 116} multiply occupied cluster crystals,^{86, 88} and a schematic SALR colloidal microphase former.^{1, 3, 117} Molecular simulations should thus here be interpreted in the broadest possible sense, in the spirit of the seminal book by Frenkel and Smit.¹¹⁰ We mostly consider very coarse-grained models, in which “molecules” are not to be distinguished. These models might be the only ones that are currently computationally accessible, but there should be no fundamental obstacle to apply the methodology we present to more detailed (molecular) models in the future.

The plan for the rest of this chapter is as follows. In section “Simulating Periodic Microphases”, we first lay down the theoretical and computational foundations for simulating periodic microphases and describe how phase transitions can be determined in this regime. In section “Simulations of Disordered Microphases”, we mostly discuss several classical MC algorithms to enhance the efficiency of simulating disordered microphases as well as describe how different morphological regimes can be distinguished. Section “Microphase Formers Solved by Molecular Simulations” then presents examples of solved model microphase formers, and a brief conclusion follows in Section “Conclusion”.

SIMULATING PERIODIC MICROPHASES

In this section, we describe molecular simulation methods that have been specifically designed to achieve equilibrium in the periodic microphase regime. We first detail the thermodynamic framework and a free energy integration simulation method, followed by a concrete introduction to the ghost particle/cluster switching method. The MC cluster volume moves that accelerate sampling for some of the periodic microphases are then explained. Finally, the approach to extract the order–order and order–disorder phase transitions, including T_{ODT} , from the equilibrium simulation results is discussed.

Expanded Thermodynamics

The Helmholtz free energy $F(N, V, T)$ for a one-component crystal is completely determined by the number of particles N , volume V , and temperature T of that system. In numerical simulations, this phase is typically studied using finite, defectless crystals. Taking into account the contribution of vacancies (and interstitials) to the free energy makes the thermodynamic analysis trickier, because N then need not match the number of lattice sites, N_c . In the thermodynamic limit, the average occupancy, $n_c \equiv N/N_c$, spontaneously equilibrates, but in finite-size simulations microscopic pathways to achieve equilibrium might not be accessible.

To sidestep this challenge, Swope and Anderson introduced a generalized, constrained free energy $F_c(N, V, T, N_c)$ with the differential form¹¹⁸

$$dF_c = -SdT - pdV + \mu dN + \tilde{\mu}_c dN_c \quad [5]$$

where S is the entropy, p the pressure, μ the chemical potential and $\tilde{\mu}_c$ the variable conjugate to N_c , often described as a “site chemical potential”, namely, the free energy cost of inserting one lattice site into the system. At equilibrium, $F(N, V, T) = F_c(N, V, T, N_c^{\text{eq}})$, with $\tilde{\mu}_c = \left(\frac{\partial F_c}{\partial N_c} \right)_{T, V, N} \Big|_{N_c = N_c^{\text{eq}}} = 0$. If the equilibrium state is (nearly) defect-free, as is it for most crystals, then $N_c^{\text{eq}} \approx N$ and the expanded thermodynamics reduces to the conventional form. Certain vacancy-stabilized crystals, such as hard cubes, however, contain a significant fraction of delocalized vacancies at equilibrium.¹¹⁹

The multiple-occupancy (cluster) crystals formed by certain soft-core particles¹²⁰ can also be described by such an expanded thermodynamics.^{86–88} In this context, for a same overall number density ρ , at a given T - ρ phase point, the system may adopt a small cluster occupancy with a small lattice spacing or form larger clusters that are further apart. The per particle free energy differential is then

$$df_c = -sdT - pdv + \mu_c dn_c \quad [6]$$

where $f_c = F_c/N$, $s = S/N$, $v = V/N = 1/\rho$ and $\mu_c = -\tilde{\mu}_c/n_c^2$.

Like cluster crystals, the mesoscopic domains of periodic microphases need not have a fixed occupancy. The only adjustment is that n_c in these systems can also be understood as the line, area, or volume number density of the mesoscopic domain (see section “Thermodynamic Integration for Microphases” for details). In any event, the equilibrium condition remains unaffected; the equilibrium occupancy n_c^{eq} must be such that the constrained free energy is minimized.

The use of an expanded thermodynamics and the state-point-dependent occupancy $n_c^{\text{eq}}(T, \rho)$ is not merely a computational trick to identify the equilibrium state. It also provides physical insight into the anomalous response of these systems, including their heat capacity C_V , bulk modulus B_T , and coefficient of thermal expansion α_T . A response function \mathcal{R} is generically defined as the derivative of a thermal quantity Y (energy, enthalpy, volume, etc.) with respect to a changing parameter x (temperature, pressure, chemical potential, etc.). Because Y is a function of n_c^{eq} , which itself changes with x , at a given state point x_0 , periodic microphases then have¹²¹

$$\mathcal{R}(x_0) \equiv \left. \frac{\partial Y(x, n_c^{\text{eq}}(x))}{\partial x} \right|_{x=x_0} = \underbrace{\left. \frac{\partial Y}{\partial x} \right|_{x=x_0; n_c^{\text{eq}}(x_0)}}_{\mathcal{R}_{\text{vir}}} + \underbrace{\left. \frac{\partial Y}{\partial n_c^{\text{eq}}} \right|_{n_c^{\text{eq}}=n_c^{\text{eq}}(x_0); x_0} \cdot \left. \frac{\partial n_c^{\text{eq}}}{\partial x} \right|_{x=x_0}}_{\mathcal{R}_{\text{domain}}} \quad [7]$$

The “virial” term \mathcal{R}_{vir} is associated with affine transformations of the lattice; and the “domain” term $\mathcal{R}_{\text{domain}}$ is associated with the growth or shrinking of domains, as lattice sites are either eliminated or created. For instance, the bulk modulus of microphases under compression is softened by the growth of the microscale domains. Similarly, microphases under isobaric heating further expand by breaking up these domains. In practice, because a response function can be expressed as fluctuations of thermal quantities, e.g., C_V as energy fluctuation and compressibility $\kappa_T (= B_T^{-1})$ as volume fluctuation,¹²² special attention needs to be paid if n_c is fixed. Fluctuations at fixed n_c only capture the affine transformation of lattices constrained to a specific equilibrium occupancy. The clustering contribution $\mathcal{R}_{\text{domain}}$ must thus be separately calculated.^{86, 88}

Thermodynamic Integration for Microphases

The expanded thermodynamics of microphases suggests a natural approach for pinning down the equilibrium microphase, namely computing the constrained free energy $f_c(T, \rho, n_c)$ for various occupancy n_c (for every morphology), and from these results identifying the conditions that minimize f_c . Free energies, however, cannot be extracted from straightforward ensemble averages, as could energy and magnetization. They are instead obtained by methods that bypass the formidable task of directly counting the number of microstates. The canonical *thermodynamic integration* (TI) method¹²³ is one such method. The challenge is then to establish a reversible integration path between a reference system 0 of known free energy f_0 with the system of interest 1 and thus obtain f_1 .

A common (alchemical) approach to TI entails defining a coupled system with potential energy $U_\gamma = (1 - \gamma)U_0 + \gamma U_1$ as the linear combination of the reference potential energy U_0 and the actual potential energy U_1 and then compute

$$f_1 - f_0 = \frac{1}{N} \int_0^1 d\gamma \left\langle \frac{\partial U_\gamma}{\partial \gamma} \right\rangle_\gamma = \frac{1}{N} \int_0^1 d\gamma \langle U_1 - U_0 \rangle_\gamma \quad [8]$$

where $\langle \dots \rangle_\gamma$ denotes thermal averaging in the coupled system with interaction U_γ .¹¹⁰ Although the choice of a TI path presents considerable freedom, it must nonetheless meet two conditions: (i) f_0 should be analytically (or at least numerically) obtainable; (ii) the integration path along γ should be continuous and reversible, in that it should not suffer from hysteresis like in a first-order phase transition.

In the specific case of periodic mesophases formed by particles with SALR interactions, a possible strategy is to follow a *two-step* TI protocol (Figure 6)¹¹⁷: (a to b) TI from free particles (ideal gas) under an external periodic field, with energy $U_F^{\text{id}}(\mathbf{r}^N)$, to denser purely repulsive particles under the same field, with energy $U_F^{\text{rep}}(\mathbf{r}^N)$; and (b to c) TI from this end point to a system of field-free, fully interacting particles.¹ The key is to choose an appropriate external field, $\mathcal{F}(\mathbf{r})$. To satisfy conditions (i) and (ii), the field symmetry should be akin to the microphase morphology under study, in addition to being sufficiently strong to prevent melting of the mesoscale ordering during the TI path and sufficiently weak for the state of interest to be approached without discontinuity for reasonable integration step sizes. The reference state potential energy is then

$$U_F^{\text{id}}(\mathbf{r}^N) = - \sum_{i=1}^N \mathcal{F}(\mathbf{r}_i) \quad [9]$$

The external field $\mathcal{F}(\mathbf{r})$ for each microphase morphology are summarized in Table 1.

To match the microphase morphology of interest, products of trigonometric functions of wavenumber $k = 2\pi/\lambda$ —to obtain a periodicity λ —have mostly been used. (Including higher harmonics can produce a more sharply confining potential, but relatively slow varying fields accelerate sampling.) For example, for one-dimensional (1D) periodic lamellar phases, a simple sinusoidal function $\mathcal{F}(\mathbf{r}) = \mathcal{F}_0 \cos(kz)$ works, and the number density per area of lamella, $\rho_\lambda = \rho\lambda = 2\pi\rho/k$, can be used as the occupancy n_c parameter. For double gyroids, which are bicontinuous phases emerging

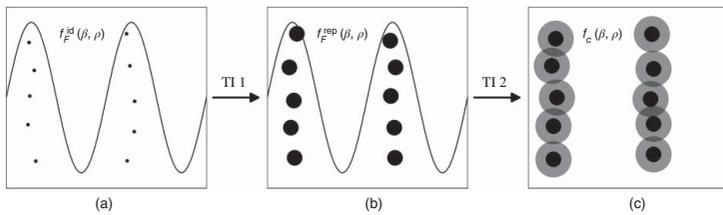
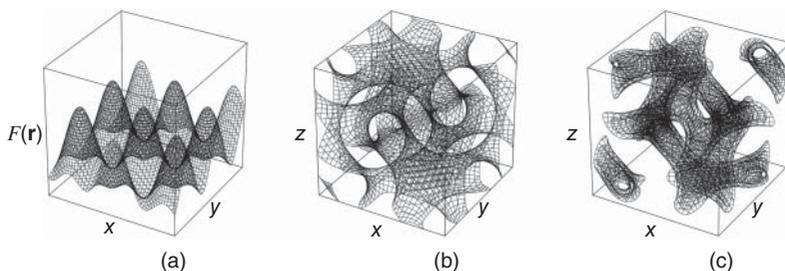


FIGURE 6 Schematic representation of the two-step thermodynamic integration (TI 1 and TI 2) protocol. An external sinusoidal field (solid line) is used to align (a) an ideal gas (small dots) and (b) purely repulsive particles (black disks). (c) If the field is properly chosen, particles assemble into a similar periodic structure once the field is turned off and their SALR interaction is turned on (gray corona).

TABLE 1 Measures of Occupancy n_c and Reference Field $\mathcal{F}(\mathbf{r})$ for Different Microphase Morphologies

| Morphology | n_c | $\mathcal{F}(\mathbf{r})$ |
|----------------------|-----------------------------------|--|
| Lamellar | Area density per lamella | $\mathcal{F}_0 \cos(kz)$ |
| Cylindrical | Line density per cylinder | $\mathcal{F}_0 \cos(kx) \cos \left[k \left(\frac{x}{2} + \frac{\sqrt{3}y}{2} \right) \right] \cos \left[k \left(\frac{x}{2} - \frac{\sqrt{3}y}{2} \right) \right]$ |
| FCC- cluster crystal | Number of particles per cluster | $\mathcal{F}_0 \cos(kx) \cos(ky) \cos(kz)$ |
| BCC- cluster crystal | Number of particles per cluster | $\mathcal{F}_0 \{ \cos[k(x+y)] \cos[k(x-y)] \cos[k(y+z)] \cos[k(y-z)] \cos[k(z+x)] \cos[k(z-x)] \}^{1/2}$ |
| Double gyroid | Number of particles per unit cell | $\mathcal{F}_0 \sin[g(x, y, z)^2]$, where $g(x, y, z) = \cos(kx) \sin(ky) + \cos(ky) \sin(kz) + \cos(kz) \sin(kx)$ |

**FIGURE 7** (a) A two-dimensional field $\mathcal{F}(x, y)$ used for hexagonally packed cylinders with $\mathcal{F}_0 = 6$ and $k = 0.7$. Contour surfaces at (b) $\mathcal{F}(\mathbf{r}) = 400$ and (c) 590 of a three-dimensional double-gyroid field with $\mathcal{F}_0 = 600$ and $k = \pi/5$, as defined in Table 1.

from a minimal surface problem,^{97, 124} an approximate description of the minimal surface in terms of sinusoidal functions¹²⁵ has been shown to work (Figure 7).

The free energy for noninteracting particles at global density ρ and inverse temperature $\beta = (k_B T)^{-1}$ in any external field $\mathcal{F}(\mathbf{r})$ is then

(see Appendix 3.A)

$$f_{\mathcal{F}}^{\text{id}}(\beta, \rho) = f^{\text{id}}(\beta, \rho) - \frac{1}{\beta} \ln \frac{\int_V e^{\beta \mathcal{F}(\mathbf{r})} d\mathbf{r}}{V} \quad [10]$$

where $f^{\text{id}}(\beta, \rho) = \frac{1}{\beta} [\ln(\rho \Lambda^3) - 1]$ is the ideal gas contribution in absence of field, with the thermal de Broglie wavelength Λ set to unity without loss of generality.¹²⁶ For convenience, we define

$$z_{\mathcal{F}}(\beta) = \int_V e^{\beta \mathcal{F}(\mathbf{r})} d\mathbf{r} \quad [11]$$

which reduces to $z_{\mathcal{F}}(\beta) = V$ when $\mathcal{F}(\mathbf{r}) = 0$. Note that because $\mathcal{F}(\mathbf{r})$ is periodic along one, two, or three directions with wavelengths λ_x , λ_y and λ_z , respectively, $z_{\mathcal{F}}(\beta)$ can be computed over a single period.

The endpoint at which the first TI terminates is a system of repulsive particles in a field at the target number density ρ . The total potential energy is then

$$U_{\mathcal{F}}^{\text{rep}}(\mathbf{r}^N) = - \sum_{i=1}^N \mathcal{F}(\mathbf{r}_i) + \sum_{j>i} u^{\text{rep}}(r_{ij}) \quad [12]$$

and hence the integration

$$f_{\mathcal{F}}^{\text{rep}}(\beta, \rho) = f_{\mathcal{F}}^{\text{id}}(\beta, \rho) + \int_0^{\rho} \frac{p(\rho') - \rho' / \beta}{\rho'^2} d\rho' \quad [13]$$

is analogous to obtaining the free energy of a simple liquid in absence of field. The pressure p as a function of ρ can be calculated from the average $\langle \rho \rangle$ under constant NpT simulations and then inverting the function. To preserve the modulation of anisotropic mesophases, however, volume fluctuations then require special care (see Appendix 3.B).

Numerically evaluating the integral in Eq. [13] at low densities can be computationally taxing without specialized simulation methods (see, e.g., section ‘‘Aggregation Volume Biased (AVB) Moves’’). In this regime, one can instead analytically obtain the first two (or more) virial coefficients of the virial series $\frac{\beta p - \rho}{\rho^2} = B_2 + B_3 \rho + \mathcal{O}(\rho^2)$ for repulsive particles in a field^{127, 128} (see Appendix 3.C)

$$B_{2,\mathcal{F}}^{\text{rep}}(\beta) = - \frac{V}{2z_{\mathcal{F}}^2(\beta)} \iint_V d\mathbf{r}_1 d\mathbf{r}_2 f_{12}^{\text{rep}} e^{\beta \mathcal{F}(\mathbf{r}_1)} e^{\beta \mathcal{F}(\mathbf{r}_2)} \quad [14]$$

$$B_{3,\mathcal{F}}^{\text{rep}}(\beta) = -\frac{V^2}{3z_{\mathcal{F}}^3(\beta)} \iiint_V \iiint_V d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 f_{12}^{\text{rep}} f_{13}^{\text{rep}} f_{23}^{\text{rep}} e^{\beta\mathcal{F}(\mathbf{r}_1)} e^{\beta\mathcal{F}(\mathbf{r}_2)} e^{\beta\mathcal{F}(\mathbf{r}_3)} \quad [15]$$

with Mayer function $f_{ij}^{\text{rep}} = e^{-\beta u^{\text{rep}}(r_{ij})} - 1$. (As above, the integral $z_{\mathcal{F}}(\beta)$ can be simplified for specific field symmetries.) The numerical results for the equation of state can then be approximated by polynomial fitting, using $B_{2,\mathcal{F}}^{\text{rep}}$ and $B_{3,\mathcal{F}}^{\text{rep}}$ as relevant coefficients.

The second TI step, as in Eq. [8], brings finite-density repulsive particles in a field to field-free fully interacting SALR particles with

$$U(\mathbf{r}^N) = \sum_{j>i} [u^{\text{rep}}(r_{ij}) + u^{\text{SALR}}(r_{ij})]. \quad [16]$$

Summing the various contributions along the integration path finally gives the constrained free energy for a given microphase morphology and lattice occupancy

$$f_c(\beta, \rho) = f^{\text{id}}(\beta, \rho) - \frac{1}{\beta} \ln \frac{\int_V e^{\beta\mathcal{F}(\mathbf{r})} d\mathbf{r}}{V} + \int_0^\rho \frac{p(\rho') - \rho'/\beta}{\rho'^2} d\rho' + \frac{1}{N} \int_0^1 d\gamma \left\langle \frac{\partial U_\gamma}{\partial \gamma} \right\rangle_\gamma \quad [17]$$

The equilibrium free energy $f(\beta, \rho)$ and microphase occupancy at state point (β, ρ) are then obtained by minimizing Eq. [17] with respect to n_c .

Ghost Particle/Cluster Switching Method

Conventional simulation methods fail to equilibrate periodic structures with fluctuating occupancy because n_c (or, equivalently, N_c) cannot straightforwardly change within a finite-size box. The TI scheme presented above sidesteps this problem by simulating a series of systems with different n_c and locating n_c^{eq} (by free energy minimization) afterward. Even with good initial guesses for n_c^{eq} , however, this approach can be computationally demanding, because multiple simulations must be done for each and every state point of interest. The ghost particle/cluster switching method proposed by Wilding and Sollich¹¹² allows N_c to fluctuate and its equilibrium value to be obtained via histogram reweighting, all within a single (properly tuned) simulation. The method was originally designed for cluster crystals, but generalizing it to microphase-forming SALR particles should be possible, at least for

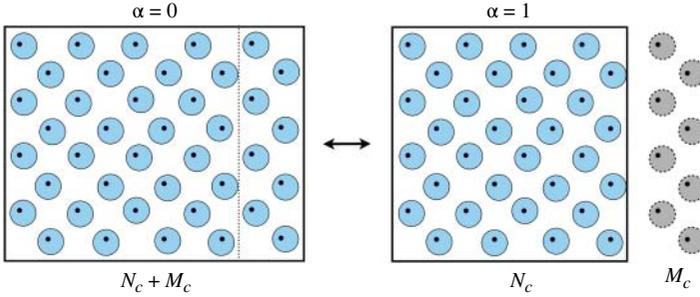


FIGURE 8 Schematic illustration of ghost particle/cluster switching between state $\alpha = 0$ with $N_c + M_c$ sites and state $\alpha = 1$ with N_c real sites plus M_c ghost sites. *Source:* Adapted from Ref. [112]

certain phase symmetries. It does not, however, straightforwardly apply to lattice microphase formers.

The key idea is to construct a reversible sampling path, α , between two states: one with $N_c + M_c$ lattice sites ($\alpha = 0$) and the other with N_c lattice sites ($\alpha = 1$), where $M_c = N_c/m$ is a small fraction of N_c with (large) integer m (Figure 8). Typically, M_c is chosen such that state 0 has one or a few more layers of lattice sites than state 1. At average occupancy n_c , the number of particles is proportional to the number of sites, $N^{(0)} = n_c(N_c + M_c)$ and $N^{(1)} = n_c N_c$, and the relative volume of the states scales similarly, $V^{(1)} = V^{(0)}m/(m + 1)$. For notational convenience, we express the position of particle i , around lattice site $\gamma(i)$ in state α as $\vec{r}_i^{(\alpha)} = \vec{R}_{\gamma(i)}^{(\alpha)} + \vec{u}_i$, where $\vec{R}_{\gamma(i)}^{(\alpha)}$ is the lattice position of site $\gamma(i)$, and \vec{u}_i is the displacement vector from $\gamma(i)$. All particles assigned to a same site γ thus share a same $\vec{R}_{\gamma}^{(\alpha)}$. The first N_c sites of the two states are the same, i.e., $\vec{R}_{\gamma}^{(0)} = \vec{R}_{\gamma}^{(1)}$ for $\gamma = 1, 2, \dots, N_c$, and so are $\vec{r}_i^{(0)} = \vec{r}_i^{(1)}$ for particles associated with these first N_c sites. In state $\alpha = 0$, the additional M_c sites stack regularly on one side of the first N_c sites, but in state $\alpha = 1$ their positions are arbitrary. Upon switching from $\alpha = 0$ to $\alpha = 1$, particles associated with these M_c sites become “ghost” particles that only experience a harmonic confining potential $\phi_g(\vec{u}_i)$. Ghost particles thus only inherit their displacement \vec{u}_i and site association $\gamma(i)$ from their interacting counterpart.

MC simulations can then be performed in the constant- μpT ensemble using standard sampling moves for particle displacements, volume fluctuations, and particle insertions/deletions,¹¹⁰ and including attempts at exchanging the two states, $\alpha \leftrightarrow 1 - \alpha$, which are accepted with probability,

$P_{\text{acc}} = \min \left[1, \frac{V^{(1-\alpha)}}{V^{(\alpha)}} \exp(-\beta(p\Delta V + \Delta E \pm (\mu - \mu_g)N_g)) \right]$, where μ_g is the ghost chemical potential—chosen so as to yield an average n_c to that of real sites—and N_g is the instantaneous number of ghost particles. In practice, such switches suffer low acceptance rates. To enhance the probability of each α , a weight function $\eta_\alpha(v)$ of the instantaneous switch cost $v = \exp(\beta(p\Delta V + \Delta E \pm (\mu - \mu_g)N_g))$, which is obtainable via Wang–Landau sampling,¹²⁹ is added upon P_{acc} to bias the moves.

Note that in standard circumstances, simulating the generalized $\mu p T$ ensemble would be of limited interest, for without specifying an extensive quantity, such as N or V , it leads to unbounded fluctuations.¹³⁰ However, in simulations of crystals (with multiple occupancy), the inaccessibility of pathways that change N_c conveniently bounds the fluctuations of N .¹²¹ The partition function for this ensemble is then

$$\begin{aligned} Z(\mu, p, T, N_c) &= \sum_N e^{\beta\mu N} \int dV e^{-\beta p V} \int dE e^{S(N, V, E, N_c)/k_B} e^{-\beta E} \\ &= e^{\beta\mu N^*} e^{-\beta p V^*} e^{S(N^*, V^*, E^*, N_c)/k_B} e^{-\beta E^*} \\ &= e^{\beta\mu N_c n_c^*} e^{-\beta p N_c v_c^*} e^{N_c s(n_c^*, v_c^*, e_c^*)/k_B} e^{-\beta N_c e_c^*} \end{aligned} \quad [18]$$

which corresponds to the optimal values of the extensive variables (*) in the thermodynamic limit. Using the form with N_c factorized out gives

$$-k_B T \ln Z(\mu, p, T, N_c) = N_c [e_c^* - T s(n_c^*, v_c^*, e_c^*) + p v_c^* - \mu n_c^*] = N_c \tilde{\mu}_c, \quad [19]$$

because the constrained Gibbs free energy $G_c = \mu N + \tilde{\mu}_c N_c = F_c + pV$. Equilibrium is then attained when $\tilde{\mu}_c = 0$, or, in other words, when $Z(\mu, p, T, N_c)$ becomes independent of N_c and any two quantities of the triplet (μ, p, T) determine the third.

If a set (μ, p, T) is chosen arbitrarily for the ghost particle/cluster switching, nonequilibrium states are most likely sampled, but histogram reweighting can correct for this mistake,¹³¹ if it is not too large. By measuring the probability distribution $P(N, V, E | \mu, p, T)$ along the simulation, one can bias the weights η_α and sum over (N, E) to obtain the marginal probability distribution for the volume $P(V | \mu, p, T)$, which should exhibit two peaks corresponding to states $\alpha = 0$ and 1. The area under each peak, $P^{(\alpha)}$, gives the probability that each state α is visited. Because $Z(\mu, p, T, N_c)$ is also the weight with which different N_c is sampled, then $P^{(0)}/P^{(1)} = Z(\mu, p, T, N_c + M_c)/Z(\mu, p, T, N_c)$. At equilibrium, $Z(\mu, p, T, N_c)$ should be independent of

N_c , and one should thus have $P^{(0)}/P^{(1)} = 1$. Using histogram reweighting to explore neighboring (μ', p') to the simulated conditions, equilibrium can be found by making sure that two peaks of $P(V|\mu', p', T)$ have the same area.

In principle, the ghost particle/cluster switching method can be applied to SALR microphase formers, but no explicit demonstration yet exists. We anticipate that a number of technical problems would first need to be addressed. First, SALR particles possess harshly repulsive cores, which can significantly reduce the efficiency of particle insertion. Second, a simple harmonic potential constraining ghost particles to lattice sites cannot prevent particle overlaps, and may thus not be suitable for maintaining large mesoscopic domains, such as lamellae. Third, expanding a three-dimensional network structure like a gyroid, while maintaining its periodicity, would be computationally challenging, because a relatively large state 0 would then have to be sampled.

Cluster Volume Moves

Both of the above simulation schemes for periodic microphases rely on ensembles in which pressure is kept constant, and thus, the system volume is allowed to fluctuate. These fluctuations are traditionally implemented by scaling particle positions affinely along with the box size.¹¹⁰ Because by definition, periodic microphases present high-density mesoscopic domains separated by low-density interstitial spaces, the sampling efficiency of such moves can be extremely low. A small affine compression can cause significant particle overlap within mesodomains, resulting in prohibitively high energy cost, while shrinkage of gaps between mesodomains, which ought to be the natural response to external pressure, is inefficiently sampled. To overcome this problem, a cluster move first proposed by Schultz and Kofke for crystals¹³² has been adapted to the simulation of multiple-occupancy cluster crystals.¹²¹ The volume sampling efficiency is then increased by up to three orders of magnitude.

The key idea of this scheme is to separate the displacement of the clusters from that of the particles within these clusters (Figure 9). In this context, the position of particle i is decomposed as $\mathbf{R}_{\gamma(i)} + \mathbf{u}_i$. The algorithm is then as follows.

1. Propose a volume change from V to $V + \Delta V$ uniformly at random on a logarithmic scale and compute the overall energy cost ΔU .
2. Rescale lattice sites affinely as $\mathbf{R}_{\gamma, \text{new}} = \mathbf{R}_{\gamma, \text{old}}[(V + \Delta V)/V]^{1/3}$.

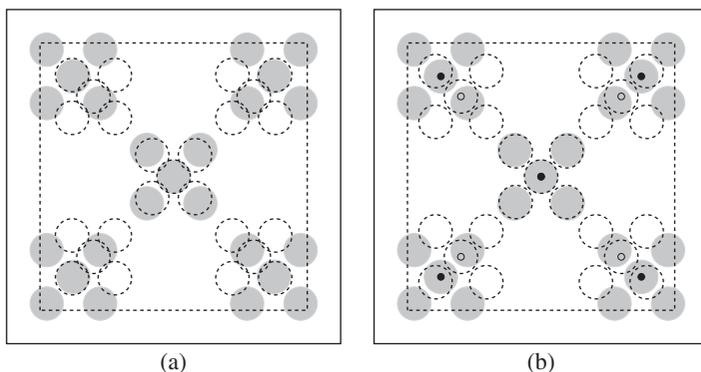


FIGURE 9 Comparison between conventional and cluster volume moves during compression. (a) In conventional volume compression, all particle positions are rescaled affinely with the simulation box from the old (solid square) to the new (dashed square) size, which is likely to cause significant particle overlaps within each cluster (dashed circles). (b) In cluster volume compression, the lattice sites (small black dots) are first rescaled affinely to new positions (small black circles) together with the box, and particles are then moved relative to their own lattice sites. In such a scheme, spacing between clusters can be greatly reduced while overlaps within clusters are prevented.

3. Update the relative displacement vectors as

$$\mathbf{u}_{i,\text{new}} = \mathbf{u}_{i,\text{old}} \left(\frac{V}{V + \Delta V} e^{\beta(p\Delta V + \Delta U_{\text{lattice}})} \right)^{\frac{1}{3(N-1)}} \quad [20]$$

with acceptance probability

$$P_{\text{acc}}(V \rightarrow V + \Delta V, U \rightarrow U + \Delta U) = \min\{1, e^{-\beta(\Delta U - \Delta U_{\text{lattice}})}\} \quad [21]$$

where $\Delta U_{\text{lattice}}$ is the change of lattice energy for a perfect single occupancy crystal.

Decomposing particle displacements in two steps allows lattice sites to be brought together without overcompressing the clusters themselves, and lattice sites to be expanded without breaking bonds between particles in a same cluster. Because this formulation relies on holding the conservation of center of mass fixed, however, special care must be made to track its motion during the simulation. This method can also be extended to (non-percolating) anisotropic periodic microphases, such as layers, cylinders, and clusters,¹¹⁷

but anisotropic volume moves must then be implemented. No version of this scheme for the fully percolating gyroid phase has yet been formulated.

Determining Phase Transitions

Once the free energy of the different mesophases at a sufficient number of state points has been obtained, the first-order transitions between pairs of periodic mesophases as well as between periodic and disordered mesophases can be determined by locating the kink in the equilibrium free energy at fixed density or by standard common tangent construction at fixed temperature.¹³³ The process is akin to that used for the fluid–crystal coexistence of simple fluids and can be complemented by the Gibbs–Duhem construction when appropriate. The Gibbs phase rule is then a useful guide to identify the location of the triple points.

Determining T_{ODT} is more subtle. For systems in which this transition remains critical, standard approaches for locating critical points, such as Binder cumulants,¹³⁴ can be used. For systems in which this transition is weakly first-order, however, standard approaches might not suffice. The free energy kink associated with the transition is then quite difficult to detect numerically, and no length scale critically diverges at the transition. Relying on the analogy with a proper critical point is nonetheless the more promising approach. Examining the decay of the finite-wavevector fluctuations in the periodic microphase regime upon approaching that point, for instance, is expected to exhibit a critical-like decay, before reaching a cutoff (Figure 10). The most effective simulation methodology for detecting this transition, however, remains to be determined.

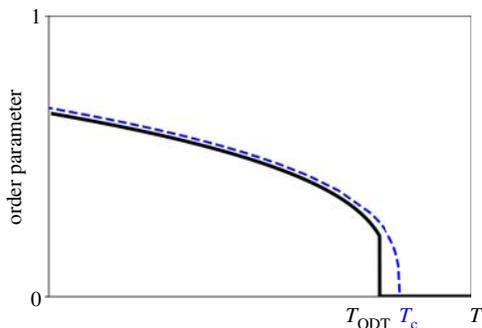


FIGURE 10 A weakly first-order order-disorder transition (ODT) is characterized by a discontinuity of order parameter at T_{ODT} (solid line). By contrast, the order parameter of a standard critical transition vanishes continuously at T_c (dashed line).

SIMULATIONS OF DISORDERED MICROPHASES

Given the appealing morphologies and broad applications of periodic microphases, it is not surprising that physical attention has primarily been given to that materials regime. Understanding disordered microphases, however, is also an essential materials endeavor. The disordered regime is a precursor to periodic mesoscale ordering and its morphological richness is comparable to that of its periodic counterpart. For example, in the language of colloidal microphase formers, upon increasing density, a gas of particles can condense into a gas of clusters, then grow into wormlike micelles that eventually percolate, and inverted structures are expected at higher densities.¹³⁵

Surprisingly, our theoretical understanding of the disordered microphase regime is relatively poor. Some exactly solvable 1D systems exhibit a clustering crossover, but these clusters cannot significantly grow nor percolate; standard mean-field (and density-functional) descriptions miss most of the structural heterogeneity in this regime (see, however, Ref.¹³⁶); and existing liquid-state descriptions are insensitive to many of their rich topological features. For now, this disordered regime is thus best understood using molecular simulations, but this approach is not without challenges.

From a computational standpoint, parallel tempering (PT), or replica exchange, has a long track record of surmounting barriers to rearrangement in frustrated systems of various kinds.^{3,137–141} The genericity of the PT approach, however, does not leverage the structural and dynamical specificities of the various disordered microphase regimes. To surmount the relatively high computational cost of PT, complementing or substituting it with other sampling schemes is thus highly desirable. In this section (see also Ref.¹⁴²), we first describe various simulation methods and their use for different morphological regimes and then describe how to detect changes from one regime to another.

Wolff-Like Cluster Algorithms

Conventional single-particle moves become extremely inefficient upon approaching the standard gas–liquid (paramagnetic–ferromagnetic) critical point T_c . Critical fluctuations and time correlations then grow large, making such moves inefficient at decorrelating configurations.¹⁴³ In spin systems without frustration (i.e., at $\xi = 0$), this critical slowing down can be largely surmounted with the help of cluster algorithms, such as Swendsen–Wang’s¹⁴⁴ and Wolff’s.¹⁴⁵ The key is to group similarly oriented

spins into clusters and to flip them all at once, subject to a selection rule that preserves microscopic reversibility.¹⁴³

The single spin-flip dynamics of systems with $\xi > 0$ also slows down upon approaching the order–disorder transition, irrespective of whether ξ is below or above the Lifshitz point ξ_L . In certain very weakly frustrated models, a generalized Wolff cluster algorithm has found some use.^{115, 116, 146} In a generic Ising-like lattice model, where spins are under nearest-neighbor ferromagnetic interaction of strength $-J$ ($J > 0$) and long-range antiferromagnetic frustration of strength ξJ , the algorithm is as follows.

1. Choose uniformly, at random, a site i to seed the cluster.
2. Add ferromagnetically and antiferromagnetically interacting neighbors of i to the cluster with probability $P_{\text{ferro}} = 1 - e^{-2\beta J}$ and $P_{\text{anti}} = 1 - e^{-2\beta\xi J}$, respectively.
3. Use the spins added to the cluster as new seeds, stopping the cluster growth when all neighbors of all seeds have been considered.
4. Flip the entire cluster with probability one.

In most frustrated systems, however, this generalization does not efficiently sample configurations.¹⁴⁷ Instead, the trial clusters reach a percolation transition at temperature, T_p , well before the critical regime is reached, i.e., $T_p > T_c$. (In the simple Ising model, both transitions coincide.¹⁴⁸) As a result, even for quite small ξ , the efficacy of cluster moves degrades quickly with ξ . Most cluster moves uselessly flip nearly all lattice sites.^{149–151} An elegant algorithmic solution to this problem has not yet been identified, let alone an off-lattice generalization.

Virtual Cluster Moves

In the cluster fluid regime, clusters diffuse more slowly than individual particles, hence the displacement of the latter cannot be relied upon to sample configurations of the former efficiently. Naively displacing whole clusters at once, however, is not microscopically reversible. Cluster–cluster collisions, for instance, are problematic. A virtual cluster move algorithm first proposed by Whitelam and Geissler¹⁵² solves this problem. The key idea is to grow a *pseudocluster* C out of a seed and of particles within its interaction range (Figure 11). The algorithm is as follows.

1. Select uniformly, at random, a seed particle i .
2. Add (interacting) neighboring particles j 's of i to the pseudocluster C with probability P_{ij} .

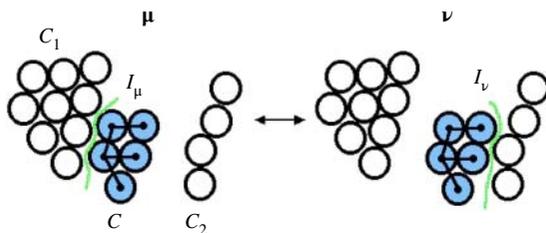


FIGURE 11 A move of a virtual cluster C (shaded) from state μ to ν , which breaks the interface I_μ between C and other particles in the physical cluster C_1 and forms a new interface I_ν with another physical cluster C_2 . A particular iterative linking procedure is reflected in the connecting bonds through C . *Source:* Adapted from Ref. [152]

3. Use the particles added to C as new seeds, stopping the cluster growth when all neighbors of all seeds have been considered. Note that C is usually a subset of a physical cluster.
4. Propose a collective displacement $\mu \rightarrow \nu$ of C . Note that such a move involves breaking the interface between C and the rest of the physical cluster, with energy cost E_μ , and forming an interface between C and a new physical cluster, with energy gain E_ν .
5. Accept this move with probability

$$P_{\text{acc}}(\mu \rightarrow \nu) = \Theta(n_{\text{cut}} - n_C) \min \left\{ 1, e^{-\beta(E_\nu - E_\mu)} \frac{\prod_{I_\nu} [1 - P_{ij}(\nu \rightarrow \mu)]}{\prod_{I_\mu} [1 - P_{ij}(\mu \rightarrow \nu)]} \prod_{\langle ij \rangle} \frac{P_{ij}(\nu \rightarrow \mu)}{P_{ij}(\mu \rightarrow \nu)} \right\}, \quad [22]$$

where $\Theta(x)$ is a Heaviside step function that returns zero if the number of particles in C , n_C , is greater than a specified cutoff n_{cut} . The product \prod_{I_μ} runs over all links across the interface I_μ that do not form during $\mu \rightarrow \nu$. The product $\prod_{\langle ij \rangle}$ runs over all links that realize one way to form a C .

Different choices of $P_{ij}(\mu \rightarrow \nu)$ and n_{cut} result in different pseudocluster sizes and acceptance ratios. They can thus be tuned to optimize the algorithmic efficiency. For instance, one can take P_{ij} to be a function of the

Boltzmann weight of breaking a bond and draw n_{cut} from a power-law distribution during each move.

Although virtual cluster moves accelerate cluster displacements, they also involve a significant computational overhead. MC implementations thus far typically only use a relatively small overall fraction ($\sim 1/N$) of such moves.³ And although this scheme enhances cluster transport, it still largely relies on single-particle transport and cluster collisions for particles to be exchanged. Equilibrating the cluster size distribution can thus remain fairly sluggish.

Aggregation Volume Biased (AVB) Moves

The generic sampling challenge of trafficking particles between different microphase domains—in cluster fluids and elsewhere—must be surmounted for the shape and size distribution of these domains to be equilibrated efficiently. Standard single-particle displacements require either domain collision, or for the bond between a particle and its original domain to be broken and then for that particle to meander in the inter-domain space before reassociating with another domain. Because both the bond-breaking and the diffusion processes are slow with standard local algorithms, nonlocal moves are a natural avenue to consider. The removal of a particle and its reinsertion uniformly at random within the simulation box, however, does not regularly result in a domain surface to domain surface displacement. Aggregation-volume-biased (AVB) MC sampling partly fixes this drawback by biasing the insertion of a particle as the surface of another.¹⁵³

The AVB algorithm proceeds as follows (Figure 12).

1. Select uniformly, at random, a particle i to be displaced.
2. Select uniformly, at random, particle j ($j \neq i$) toward which moving i will be considered.

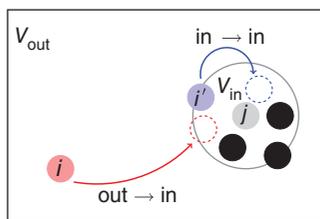


FIGURE 12 An out \rightarrow in AVB move of particle i from the nonbonded region (of volume V_{out}) of the target particle j to its bonded region of volume V_{in} , indicated by j 's interaction range (large solid circle). An in \rightarrow in move of particle i' is also shown.

3. Move i to the bonded region of j with a bias probability P_{bias} (*in* state) or to the nonbonded region of j with probability $1 - P_{\text{bias}}$ (*out* state).
4. Depending on whether i lies originally inside or outside the bonded region of another particle, four types of transitions are possible: $\text{in} \rightarrow \text{in}$, $\text{out} \rightarrow \text{out}$, $\text{in} \rightarrow \text{out}$, and $\text{out} \rightarrow \text{in}$. The first two cases reduce to the standard Metropolis acceptance rule, and the last two cases have an acceptance probability

$$P_{\text{acc}}(\text{out} \rightarrow \text{in}) = \min \left\{ 1, \frac{(1 - P_{\text{bias}})}{P_{\text{bias}}} \frac{V_{\text{in}}}{V_{\text{out}}} e^{-\beta\Delta U} \right\} \quad [23]$$

and

$$P_{\text{acc}}(\text{in} \rightarrow \text{out}) = \min \left\{ 1, \frac{P_{\text{bias}}}{(1 - P_{\text{bias}})} \frac{V_{\text{out}}}{V_{\text{in}}} e^{-\beta\Delta U} \right\} \quad [24]$$

where V_{in} is the volume of the bonded region and $V_{\text{out}} = V - V_{\text{in}}$ is the remainder.

In the context of colloidal microphase formers, the bonded region can be chosen as the attractive shell of the SALR potential. In some systems, it has been observed that using about $\sim 10\%$ AVB moves optimally equilibrates cluster fluids,¹¹⁷ but no systematic parameter tuning has been undertaken in other regimes (see, however, Ref.¹⁴²).

A significant caveat to this approach is that if domains are large, the attractive shell that surrounds a particle is a relatively inefficient region of space toward which to bias sampling. Most particles are then fully surrounded by neighbors, and hence, only a small fraction of the shell volume would allow insertion to proceed. A possible solution could be to bias toward unoccupied shell space alone. Identifying available surface sites can be done fairly efficiently on a lattice, but in continuous space, the geometry is intractable other than through gridding. Either option has yet to be algorithmically explored.

Morphological Crossovers in the Disordered Regime

Because the morphological changes in the disordered microphase regime are not proper phase transitions, their delimitation is subject to a certain degree of arbitrariness. Reasonably robust schemes can nevertheless identify the clustering and the percolation transitions.

The low-density onset of disordered microphase formation entails forming clusters at temperatures of the order of T_{ODT} .^{13, 100, 154–156} When particle concentration exceeds the critical micelle concentration (cmc), or more

generically, the critical cluster density (ccd), the systems goes from being mostly a gas of individual particles to being mostly a gas of particle clusters. This transformation is a crossover, not a phase transition, because it cooperatively involves only a finite number of particles.¹⁵⁷ The structural signature of aggregation could be followed directly in molecular simulations, but the non-universality of the process makes its determination somewhat nontrivial. While surfactant micelles tend to be fairly homogeneous in size and shape, clusters of colloidal microphase formers have a size distribution $P(n)$ that is fairly wide and non-Gaussian.^{158–161} Cluster morphology and size distribution depend sensitively on the details of the interaction potential.¹⁶² For instance, as the effective repulsion strength increases, so does cluster size heterogeneity.¹⁶³ This heterogeneity presents a hurdle to choosing structural order parameters for clustering detection. (It also suppresses the direct assembly of these clusters into a periodic cluster crystal at higher densities,^{117, 164} in contrast to Ref. 99).

Determining the cmc and the ccd from the structure can be avoided by falling back on their definitions as the point of most abrupt change in the system physiochemical properties.^{165, 166} Different schemes have been proposed. A peak in the heat capacity indicates the changing energetics of the system.¹⁶⁷ From an (osmotic) pressure viewpoint, one could also determine ρ_{ccd} or ρ_{cmc} from the intersection of two nearly ideal gas regimes—of single particles and clusters—in the equation of state.¹⁶⁸ Whereas this approach works well for micelle formation, clusters colloidal microphase formers interact strongly and more heterogeneously, which takes them farther from the ideal limit. In those cases, higher-order deviations from the ideal gas equation³ can be considered; For instance, one can use

$$h(\rho; \beta) \equiv \frac{\beta p - \rho}{\rho^2} = B_2(\beta) + B_3(\beta)\rho + B_4(\beta)\rho^2 + \dots \quad [25]$$

where B_n is the n th virial coefficient. The non-monotonicity of this function in cluster-forming systems reflects the competition between attraction that drives aggregation and repulsion between these aggregates. A pronounced valley in $h(\rho; \beta)$ corresponds to a rapid change in pressure response in that density regime and thus provides an estimate of $\rho_{\text{ccd}}(\beta)$.

Increasing the fluid density beyond the onset of clustering commonly extends the roughly spherical clusters into wormlike clusters and then to a disordered percolating network. Once again, this change is not a thermodynamic phase transition, but the morphological change is nonetheless associated with marked changes to the sampling efficiency of certain MC sampling algorithms.^{117, 135} Virtual cluster moves, in particular, are then

essentially irrelevant. The onset of percolation, $\rho_p(T)$, can be reliably identified using the finite-size scaling results from percolation theory.¹⁶⁹ Specifically, the probability $P_p(\rho, T; N)$ that at least one giant cluster network exists spanning the entire system grows with ρ as a sigmoid. As the system size increases, that sigmoid sharpens, hence in the thermodynamic $N \rightarrow \infty$ limit, percolation is sharply defined. Given the mid-point ($\rho = \rho_{\frac{1}{2}}$ and $P_p = \frac{1}{2}$) of this sigmoid, $\rho_p(T) = \rho_{\frac{1}{2}}(T; N \rightarrow \infty)$ can be obtained from critical scaling $\rho_{\frac{1}{2}}(T; N) = N^{-1/d\nu} + \rho_p(T)$, where $d\nu = 2.706$ is the universal scaling exponent for simple percolation in $d = 3$.¹⁶⁹

Other dynamically relevant morphological regimes have been identified within the disordered phase, but their systematic characterization and their onset is not as mature as for clustering and percolation. These regimes are therefore not here further discussed, and we instead refer readers to the original literature on the topic.¹³⁵

MICROPHASE FORMERS SOLVED BY MOLECULAR SIMULATIONS

Although a number of model microphase formers have been studied by molecular simulations, relatively few of these have used free energy-based methods to obtain reliable equilibrium phase information. In this section, we present three models for which quantitative results have been obtained: a 1D, a lattice, and an off-lattice microphase former. These models also illustrate the various tools described in the previous two sections, as well as the diversity of contexts in which microphase formation can be observed.

One-Dimensional Models

1D models with sufficiently short-range interactions cannot undergo thermodynamic phase transitions¹⁷⁰ nor percolate to form a gel, but can exhibit a distinct clustering regime.^{171, 172} Because these models can be solved analytically and are straightforward to simulate, they are also ideal benchmarks.

Both on- and off-lattice 1D models with SALR interactions can be solved by transfer matrix. The former is akin to solving for the Ising chain,¹⁷³ but the latter is a bit more involved. It requires changing absolute particle positions x to relative distances between neighboring particles s , i.e.,

$s_i = x_i$, $s_{i+1} = x_{i+1} - x_i$, etc., and then writing the isothermal-isobaric configurational integral as

$$Z_G(p, \beta) = \int_0^\infty \prod_{i=1}^N ds_i e^{-\beta u_i(s_i, s_{i+1}, \dots) - \beta p s_i} = \text{Tr } M^N = \Lambda_{\max}^N \quad [26]$$

where u_i is the sum of pairwise interactions between particle i and its subsequent neighbors ($i + 1, i + 2, \dots$) along the chain. The last equality only holds in the thermodynamic limit. If the interaction consists of a hard-core of size $\sigma = 1$ and a SALR potential of finite range $\kappa\sigma$, for instance, then only up to $k = \lceil \kappa \rceil - 1$ nearest neighbors can interact at once. Under periodic boundary condition, Z_G can thus be written in terms of the transfer matrix M , with each entry containing the Boltzmann weight for a particular choice of $(s_i, s_{i+1}, \dots, s_{i+k-1})$ along each row and $(s_{i+1}, s_{i+2}, \dots, s_{i+k})$ along each column. Only when a row and a column have matching $(s_{i+1}, s_{i+2}, \dots, s_{i+k-1})$, is the entry nonzero.

Transfer matrix results can be used to benchmark the equation of state and the cluster size distribution extracted from molecular simulations (Figure 13). This comparison reveals that combining single-particle displacements with a 1D analogue of AVB sampling equilibrates the system with remarkable efficiency.¹⁶⁰ This model further illustrates how the ccd can be determined using the equation of state (Figure 13).

Lattice Spin Models

Lattice spin models with frustrated interactions can display a broad array of microphase morphologies. These models are also amenable to various exact theoretical treatments and are thus appealing as theory and simulation benchmarks. They vary in the nature and extent of the frustration of nearest-neighbor, $\langle ij \rangle$, couplings, including interactions between next-nearest neighbors in one or multiple α -axial directions $\langle\langle ij \rangle\rangle^{(\alpha)}$ and between diagonal nearest-neighbors $[ij]$. For simplicity, we here only consider models with no external magnetic field.

The most weakly frustrated, axial next-nearest-neighbor Ising (ANNNI) model has Hamiltonian

$$\mathcal{H}^{\text{ANNNI}} = -J_1 \sum_{\langle ij \rangle} s_i s_j + J_3 \sum_{\langle\langle ij \rangle\rangle^{(\alpha)}} s_i s_j \quad [27]$$

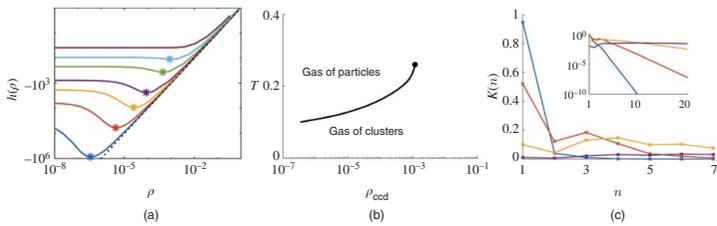


FIGURE 13 (a) Evolution of the equation of state $h(\rho)$ given in Eq. [25] with density ρ , at different temperatures for a one-dimensional microphase former.¹⁶⁰ The minimum of each curve, marked by an asterisk, defines the ccd, and is reported in (b). Note that the ccd disappears at a finite temperature (dot). (c) The probability distribution of cluster size, $K(n)$, at densities below and above the ccd for that same model at temperature $T = 0.2$ evolves smoothly with density. This smoothness makes structural determination of the ccd challenging. The inset presents the same data on a semi-log scale.

for spin variables, $s_i = \pm 1$. More strongly frustrated variants include biaxial (or more) frustration, such as the BNNNI model,

$$\mathcal{H}^{\text{BNNNI}} = -J_1 \sum_{\langle ij \rangle} s_i s_j + J_3 \sum_{\langle\langle ij \rangle\rangle^{(\alpha, \beta)}} s_j s_j \quad [28]$$

or diagonal frustration, such as the DNNNI model,

$$\mathcal{H}^{\text{DNNNI}} = -J_1 \sum_{\langle ij \rangle} s_i s_j + J_2 \sum_{[ij]} s_j s_j \quad [29]$$

Note that on a cubic lattice without external magnetic field the sign of J_1 is not important because the problem can be divided into two independent sublattices. The J_1 - J_2 model,¹⁷⁵ equivalently defined as the sum of two competing antiferromagnetic interactions, is the more common formulation of the DNNNI model. Combining both of these frustration types with $J_2 = 2J_3$ further gives rise to the (simplified) Widom–Wheeler model for surfactant microemulsions

$$\mathcal{H}^{\text{WW}} = -J_1 \sum_{\langle ij \rangle} s_i s_j + J_2 \sum_{[ij]} s_j s_j + J_3 \sum_{\langle\langle ij \rangle\rangle} s_j s_j \quad [30]$$

The ANNNI model is the only of these models for which a reasonably complete simulation phase diagram has been obtained thus far (in three dimensions).^{116, 146} Because frustration in the ANNNI model is only along a single axis, the $T = 0$ phase diagram can be solved analytically,^{60, 176, 177} and only layered mesophases form at low temperatures below the Lifshitz point. The weakness of the frustration also seemingly prevents the generation of large, percolating cluster moves in the generalized Wolff algorithm at temperatures above the critical point (see section “Simulations of Disordered Microphases”). Accurate estimates of the criticality at the Lifshitz point and at stronger frustration are thus computationally accessible through finite-size scaling analysis.^{116, 146, 178}

The resulting T - κ phase diagram, with $\kappa \equiv J_3/J_1$, is divided into several regimes, including a high-temperature paramagnetic state, a ferromagnetic (or antiferromagnetic) ground state at low κ and a modulated phase regime at high κ (Figure 14). The multicritical Lifshitz point, at which three phases (and two types of criticality) meet, is found at $(\kappa_L, T_L) \approx (0.27, 3.75)$.^{146, 179}

The relative free energy of the modulated phases and thus their ordering can be solved by molecular simulations using TI (section “Simulating Periodic Microphases”). These computations confirm the mean-field prediction that lamellae thin as frustration increases¹⁷⁷ as well as the low-temperature

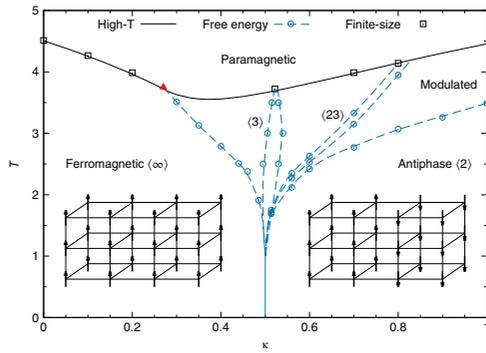


FIGURE 14 Simulation phase diagram of the ANNNI model in the T - k plane.^{116, 146} Paramagnetic, ferromagnetic, and modulated phases meet at the multi-critical Lifshitz point (red triangle). The periodicity of the phases is denoted by integers in brackets, where a symbol of (mn) means m layers of spin up followed by n layers of spin down. The simulation results largely qualitatively agree with the mean-field prediction,³⁶ and the finite-size scaling of the Binder cumulants quantitatively agree with the high-temperature series expansion for the critical line.¹⁷⁴ Simulations can further resolve the ambiguous theoretical predictions for the structure of the periodic microphase regime.

series expansion analysis that an infinite number of lamellar phases with different periodicities spring out at $T = 0$ and $\kappa = 0.5$. Accurate results at intermediate temperatures were long inaccessible to conventional theory and simulations, but molecular simulations using TI provide crisp phase boundaries even in that regime. They have thus revealed that stable regions of lamellar phases like $\langle 3 \rangle$ or $\langle 23 \rangle$ are stripe-like rather than bulging.^{88, 116}

Colloidal Models

Various combinations of attractive and repulsive potential forms have been used to model colloids with SALR interactions, including explicit descriptions of depletion attraction with screened-Yukawa repulsion, coupling Lennard-Jones attraction with Yukawa repulsion, and placing two Yukawa interaction forms in competition.^{13, 70, 99, 180} These models are all expected to exhibit the full microphase phenomenology, but also present marked simulation challenges in configurational sampling. Most studies have struggled to equilibrate the periodic microphase regime, in particular.

For colloid-like systems, molecular simulations using TI (section “Simulating Periodic Microphases”) have thus far only been used for the schematic square-well linear (SWL) model.^{3, 117} The SWL potential, $u_{\text{SWL}}(r) = u_{\text{HS}}(r) + u_{\text{SALR}}(r)$, adds to a hard-core repulsion of diameter σ , a square well (SW) attraction of strength ϵ and a linear repulsive ramp, and hence

$$u_{\text{SALR}}(r) = \begin{cases} -\epsilon & , \quad r < \lambda\sigma \\ \xi\epsilon(\kappa - r/\sigma) & , \quad \lambda\sigma < r < \kappa\sigma \\ 0 & , \quad r > \kappa\sigma \end{cases} \quad [31]$$

where ξ and κ control the repulsion strength and range, respectively.

A sample T - ρ phase diagram for the SWL model with $\xi > \xi_{\text{L}}$ is shown in Figure 15. In the high-temperature disordered microphase regime, a single-particle gas phase at very low densities is replaced by a fluid of clusters as the system crosses the ccd. Further increasing density elongates the clusters, which eventually coalesce into a percolated fluid network. Upon lowering temperature, the disordered microphase regime undergoes a first-order transition into the periodic microphase regime. The highest temperature at which lamellae are stable, T_{ODT} , is a weakly first-order transition. In the periodic regime, the characteristic cluster–cylinder–double gyroid–lamellae morphology sequence is observed as density increases.

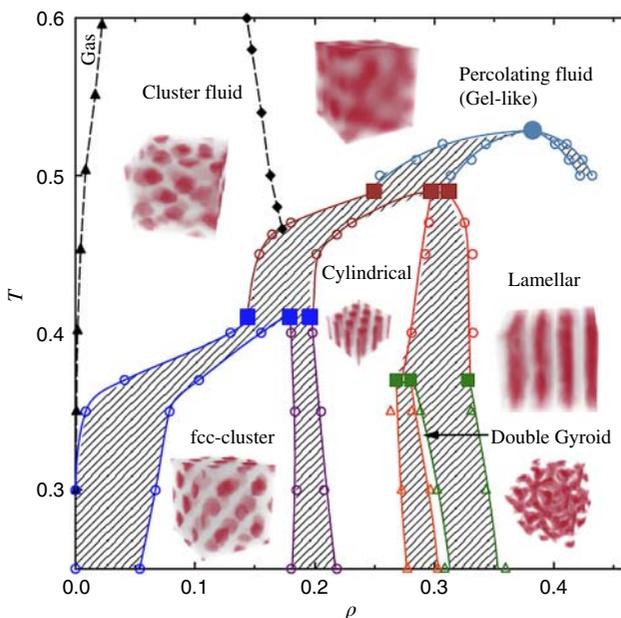


FIGURE 15 T - ρ phase diagram of the SWL model at $\xi = 0.05$, $\kappa = 4$ and $\lambda = 1.5$.¹¹⁷ Shaded areas denote coexistence regimes at first-order transitions and terminate at the maximum periodic microphase stability temperature, T_{ODT} (filled circle). Three triple points (solid squares) can be identified: fluid–fcc cluster–cylindrical coexistence, liquid–cylindrical–lamellar coexistence, and cylindrical–double gyroid–lamellar coexistence. The ccd (triangles) and the percolation line (lozenges) distinguish the gas from the cluster and percolated fluid regimes in the disordered microphase range, respectively.

CONCLUSION

Despite the theoretical and algorithmic advances in the study of microphase-forming models by molecular simulations, many challenges remain. As highlighted throughout the text, further methodological developments are needed to efficiently simulate the disordered microphase regime. Even simple frustrated lattice models obfuscate known cluster algorithms. Simulating the periodic microphase regime of off-lattice models also faces a number of hurdles. Relatively few such morphologies have been considered thus far, and the high-density periodic microphase regime, in which all domains percolate in all directions remains largely unexplored. Enhanced sampling schemes for these phases could dramatically improve

the simulation efficiency. More generally, only a few microphase-forming models have been studied using careful, free energy-based molecular simulations. The generality of our observations and the expectation of physical universality have thus yet to be fully assessed.

Important questions about the assembly of microphases also remain to be answered. If the attraction is very short-ranged, as in many colloidal suspensions, how does the microphase morphology change? Recent advances in complex particle synthesis also open the door for implementing anisotropic competing interactions, which couple liquid–crystal ordering with mesoscale patterns.¹⁸¹

Experimentally, periodic colloidal microphases are still missing. Many possible explanations have been provided such as the complexity of actual interaction potentials,⁷⁵ but no resolution is obviously in sight. Fine-tuning colloidal suspensions to allow the formation of periodic microphases thus remains an open experimental problem.

3.A FREE ENERGY OF AN IDEAL GAS IN A FIELD

The density profile of an ideal gas in a field $-\mathcal{F}(\mathbf{r})$ is

$$\tilde{\rho}(\mathbf{r}) = \frac{Ne^{\beta\mathcal{F}(\mathbf{r})}}{\int e^{\beta\mathcal{F}(\mathbf{r})}d\mathbf{r}} = \frac{\rho V e^{\beta\mathcal{F}(\mathbf{r})}}{z_{\mathcal{F}}} \quad [32]$$

Its free energy can thus be obtained by summing the ideal gas contribution

$$\begin{aligned} \beta F^{\text{id}} &= \int [\tilde{\rho}(\mathbf{r}) \ln(\Lambda^3 \tilde{\rho}(\mathbf{r})) - 1] d\mathbf{r} \\ &= \int \left[\frac{\rho V e^{\beta\mathcal{F}(\mathbf{r})}}{z_{\mathcal{F}}} \ln \left(\Lambda^3 \frac{\rho V e^{\beta\mathcal{F}(\mathbf{r})}}{z_{\mathcal{F}}} \right) - 1 \right] d\mathbf{r} \\ &= \frac{N}{z_{\mathcal{F}}} \left[\int e^{\beta\mathcal{F}(\mathbf{r})} \ln(\Lambda^3 \rho) d\mathbf{r} - \int e^{\beta\mathcal{F}(\mathbf{r})} d\mathbf{r} + \int e^{\beta\mathcal{F}(\mathbf{r})} \ln \frac{V}{z_{\mathcal{F}}} d\mathbf{r} \right. \\ &\quad \left. + \int e^{\beta\mathcal{F}(\mathbf{r})} \beta \mathcal{F}(\mathbf{r}) d\mathbf{r} \right] \quad [33] \\ &= N \ln(\Lambda^3 \rho) - N - N \ln \frac{z_{\mathcal{F}}}{V} + N \beta \frac{\int e^{\beta\mathcal{F}(\mathbf{r})} \mathcal{F}(\mathbf{r}) d\mathbf{r}}{z_{\mathcal{F}}} \end{aligned}$$

to the field contribution to obtain the full free energy,

$$\begin{aligned}
 \beta F_{\mathcal{F}}^{\text{id}} &= \beta F^{\text{id}} + \beta \int \tilde{\rho}(\mathbf{r})(-\mathcal{F}(\mathbf{r}))d\mathbf{r} \\
 &= N \ln(\Lambda^3 \rho) - N - N \ln \frac{z_{\mathcal{F}}}{V} + N \beta \frac{\int e^{\beta \mathcal{F}(\mathbf{r})} \mathcal{F}(\mathbf{r})d\mathbf{r}}{z_{\mathcal{F}}} - \beta \int \tilde{\rho}(\mathbf{r})\mathcal{F}(\mathbf{r})d\mathbf{r} \\
 &= N \ln(\Lambda^3 \rho) - N - N \ln \frac{z_{\mathcal{F}}}{V}
 \end{aligned} \tag{34}$$

and hence the free energy per particle is

$$\begin{aligned}
 f_{\mathcal{F}}^{\text{id}} &= \frac{1}{\beta} [\ln(\Lambda^3 \rho) - 1] - \frac{1}{\beta} \ln \frac{z_{\mathcal{F}}}{V} \\
 &= f^{\text{id}} - \frac{1}{\beta} \ln \frac{z_{\mathcal{F}}}{V}
 \end{aligned} \tag{35}$$

3.B CONSTANT PRESSURE SIMULATIONS OF PARTICLES IN A FIELD

A couple of subtleties arise in constant pressure simulations of particles in a field.

First, the field has to be set relative to the simulation box and changed affinely with it. If the field were to be kept fixed while the box fluctuates, energy would be dissipated as particles are driven by the box to move with respect to the field.

Second, for 1D (e.g., lamellae periodic along z direction) and two-dimensional (e.g., cylinders periodic in the xy plane) periodic structures, a virtual harmonic potential $U_{\alpha} = \frac{1}{2}k_{\alpha}(L_{\alpha} - L_{\alpha,0})^2$ with stiffness k_{α} for $\alpha \in \{x, y, z\}$ should be applied to the box length L_{α} around a targeted value $L_{\alpha,0}$. Because the stress distribution is then spatially inhomogeneous and anisotropic, the inter-domain spacing otherwise shrinks as the domains grow along directions orthogonal to the compression direction. These constraining springs ensure that the sampled configuration preserves the targeted occupancy (periodicity) corresponding to $L_{\alpha,0}$, but do not affect the TI, because the constraining potential, U_{α} , is not felt by particles.

3.C VIRIAL COEFFICIENTS OF PARTICLES IN A FIELD

Expressions for the first few virial coefficients of particles in a field can be obtained as follows. First, recall that the grand canonical partition function

can be expressed as

$$e^{\beta pV} = \Xi(T, V, \lambda) = \sum_{N \geq 0} Q(T, V, N) \lambda^N = 1 + \sum_{N \geq 1} Q_N \lambda^N \quad [36]$$

where $\lambda = e^{\beta \mu}$ and

$$Q_N = Q(T, V, N) = \frac{Z_N}{N! \Lambda^{3N}} \quad [37]$$

is the canonical partial function of N particles with configurational integral

$$Z_N = \int d\mathbf{r}^N e^{-\beta U_N(\mathbf{r}^N)} \quad [38]$$

For particles with pairwise additive interactions, $u(r_{ij})$, under a field, $-\mathcal{F}(\mathbf{r})$, the total potential energy $U_N(\mathbf{r}^N)$ is

$$U_N(\mathbf{r}^N) = \sum_{j>i}^N u(r_{ij}) - \sum_{i=1}^N \mathcal{F}(\mathbf{r}_i). \quad [39]$$

By series expansion and comparison of coefficients,^{127, 128} the second and third virial coefficients are then

$$\begin{aligned} B_{2,\mathcal{F}}(T) &= -\frac{V}{Q_1^2} \left(Q_2 - \frac{1}{2} Q_1^2 \right) = -\frac{V}{2Z_1^2} (Z_2 - Z_1^2) \\ B_{3,\mathcal{F}}(T) &= -\frac{V^2}{3Q_1^3} \left[(6Q_3 - 6Q_1 Q_2 + 2Q_1^3) - \frac{3}{Q_1} (2Q_2 - Q_1^2)^2 \right] \\ &= -\frac{V^2}{3Z_1^3} \left[(Z_3 - 3Z_1 Z_2 + 2Z_1^3) - \frac{3}{Z_1} (Z_2 - Z_1^2)^2 \right] \end{aligned} \quad [40]$$

For a potential energy of the form given in Eq. [39], and defining the Mayer function $f(r) = e^{-\beta u(r)} - 1$, we get

$$\begin{aligned} Z_1 &= \int d\mathbf{r} e^{-\beta U_1(\mathbf{r})} = \int d\mathbf{r} e^{\beta \mathcal{F}(\mathbf{r})} = z_{\mathcal{F}} \\ Z_2 &= \iint d\mathbf{r}_1 d\mathbf{r}_2 e^{-\beta U_2(\mathbf{r}_1, \mathbf{r}_2)} = \iint d\mathbf{r}_1 d\mathbf{r}_2 e^{-\beta u(r_{12})} e^{\beta \mathcal{F}(\mathbf{r}_1)} e^{\beta \mathcal{F}(\mathbf{r}_2)} \\ &= \iint d\mathbf{r}_1 d\mathbf{r}_2 f(r_{12}) e^{\beta \mathcal{F}(\mathbf{r}_1)} e^{\beta \mathcal{F}(\mathbf{r}_2)} + \iint d\mathbf{r}_1 d\mathbf{r}_2 e^{\beta \mathcal{F}(\mathbf{r}_1)} e^{\beta \mathcal{F}(\mathbf{r}_2)} \\ &= \iint d\mathbf{r}_1 d\mathbf{r}_2 f(r_{12}) e^{\beta \mathcal{F}(\mathbf{r}_1)} e^{\beta \mathcal{F}(\mathbf{r}_2)} + z_{\mathcal{F}}^2 \end{aligned} \quad [41]$$

and hence

$$B_{2,F}(T) = -\frac{V}{2z_F^2} \iint d\mathbf{r}_1 d\mathbf{r}_2 f(r_{12}) e^{\beta F(\mathbf{r}_1)} e^{\beta F(\mathbf{r}_2)} \quad [42]$$

Similarly, from the definition for Z_3 ,

$$\begin{aligned} Z_3 &= \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 e^{-\beta U_2(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)} \\ &= \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 e^{-\beta u(r_{12})} e^{-\beta u(r_{13})} e^{-\beta u(r_{23})} e^{\beta F(\mathbf{r}_1)} e^{\beta F(\mathbf{r}_2)} e^{\beta F(\mathbf{r}_3)} \end{aligned} \quad [43]$$

one gets

$$B_{3,F}(T) = -\frac{V^2}{3z_F^3} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 f(r_{12}) f(r_{13}) f(r_{23}) e^{\beta F(\mathbf{r}_1)} e^{\beta F(\mathbf{r}_2)} e^{\beta F(\mathbf{r}_3)} \quad [44]$$

ACKNOWLEDGMENTS

We thank the various colleagues with whom we have had the chance to discuss microphases over the years, including A. Ciach, E. del Gado, J. Oitmaa, J. Pekalski, S. Redner, P. Royall, N. Wilding, and especially our close coworkers on this topic, D. Frenkel, Y. Hu, Y. Liang, B. Mladek, and Y. Zhuang. P.C. acknowledges primary financial support from the National Science Foundation under grant no. DMR-1749374. Kai Zhang acknowledges support from Duke Kunshan startup funding.

REFERENCES

1. Y. Zhuang and P. Charbonneau, *J. Phys. Chem. B*, **120**, 6178–6188 (2016). Equilibrium Phase Behavior of the Square-Well Linear Microphase-Forming Model.
2. A. Ciach, J. Pekalski, and W. Gózdź, *Soft Matter*, **9**, 6301–6308 (2013). Origin of Similarity of Phase Diagrams in Amphiphilic and Colloidal Systems with Competing Interactions.
3. Y. Zhuang and P. Charbonneau, *J. Phys. Chem. B*, **120**, 7775–7782 (2016). Recent Advances in the Theory and Simulation of Model Colloidal Microphase Formers.
4. F. S. Bates and G. Fredrickson, *Phys. Today*, **52**, 32 (2000). Block Copolymers-Designer Soft Materials.

5. A. Stradner, H. Sedgwick, F. Cardinaux, W. C. Poon, S. U. Egelhaaf, and P. Schurtenberger, *Nature*, **432**, 492 (2004). Equilibrium Cluster Formation in Concentrated Protein Solutions and Colloids.
6. P. De Gennes and C. Taupin, *J. Phys. Chem.*, **86**, 2294–2304 (1982). Microemulsions and the Flexibility of Oil/Water Interfaces.
7. O. Portmann, A. Vaterlaus, and D. Pescia, *Nature*, **422**, 701 (2003). An Inverse Transition of Magnetic Domain Patterns in Ultrathin Films.
8. B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, *Science*, **358**, 1155–1160 (2017). Stripe Order in the Underdoped Region of the Two-Dimensional Hubbard Model.
9. J. Orenstein and A. Millis, *Science*, **288**, 468–474 (2000). Advances in the Physics of High-Temperature Superconductivity.
10. R. E. Thorne, *Phys. Today*, **49**, 42–48 (1996). Charge-Density-Wave Conductors.
11. S. Brazovskii, *J. Exp. Theor. Phys.*, **41**, 85 (1975). Phase Transition of an Isotropic System to a Nonuniform State.
12. M. C. Barbosa, *Phys. Rev. E*, **48**, 1744 (1993). Ising Model with Isotropic Competing Interactions in the Presence of a Field: A Tricritical-Lifshitz-Point Realization.
13. A. J. Archer and N. B. Wilding, *Phys. Rev. E*, **76**, 031501 (2007). Phase Behavior of a Fluid with Competing Attractive and Repulsive Interactions.
14. K. Kataoka, A. Harada, and Y. Nagasaki, *Adv. Drug Deliv. Rev.*, **64**, 37–48 (2012). Block Copolymer Micelles for Drug Delivery: Design, Characterization and Biological Significance.
15. A. Rösler, G. W. Vandermeulen, and H.-A. Klok, *Adv. Drug Delivery Rev.*, **64**, 270–279 (2012). Advanced Drug Delivery Devices via Self-Assembly of Amphiphilic Block Copolymers.
16. R. A. Segalman, *Mater. Sci. Eng., R*, **48**, 191–226 (2005). Patterning with Block Copolymer Thin Films.
17. M. Li and C. K. Ober, *Mater. Today*, **9**, 30–39 (2006). Block Copolymer Patterns and Templates.
18. S. Krishnamoorthy, C. Hinderling, and H. Heinzelmann, *Mater. Today*, **9**, 40–47 (2006). Nanoscale Patterning with Block Copolymers.
19. A. T. KG, K. Gotrik, A. Hannon, A. Alexander-Katz, C. Ross, and K. Berggren, *Science*, **336**, 1294–1298 (2012). Templating Three-Dimensional Self-Assembled Structures in Bilayer Block Copolymer Films.
20. B. H. Kim, J. Y. Kim, and S. O. Kim, *Soft Matter*, **9**, 2780–2786 (2013). Directed Self-Assembly of Block Copolymers for Universal Nanopatterning.
21. C. J. Hawker and T. P. Russell, *MRS bulletin*, **30**, 952–966 (2005). Block Copolymer Lithography: Merging “Bottom-Up” with “Top-Down” Processes.

22. C. Tang, E. M. Lennon, G. H. Fredrickson, E. J. Kramer, and C. J. Hawker, *Science*, **322**, 429–432 (2008). Evolution of Block Copolymer Lithography to Highly Ordered Square Arrays.
23. H. Yu, X. Qiu, N. Moreno, Z. Ma, V. M. Calo, S. P. Nunes, and K.-V. Peinemann, *Angew. Chem.*, **54**, 13937–13941 (2015). Self-Assembled Asymmetric Block Copolymer Membranes: Bridging the Gap from Ultra- to Nanofiltration.
24. C. Kresge, M. Leonowicz, W. J. Roth, J. Vartuli, and J. Beck, *Nature*, **359**, 710 (1992). Ordered Mesoporous Molecular Sieves Synthesized by a Liquid-Crystal Template Mechanism.
25. M. Cates and S. T. Milner, *Phys. Rev. Lett.*, **62**, 1856 (1989). Role of Shear in the Isotropic-to-Lamellar Transition.
26. T. Tepe, M. Schulz, J. Zhao, M. Tirrell, F. S. Bates, K. Mortensen, and K. Almdal, *Macromolecules*, **28**, 3008–3011 (1995). Variable Shear-Induced Orientation of a Diblock Copolymer Hexagonal Phase.
27. A. Nikoubashman, R. A. Register, and A. Z. Panagiotopoulos, *Soft Matter*, **9**, 9960–9971 (2013). Simulations of Shear-Induced Morphological Transitions in Block Copolymers.
28. T. Morkved, M. Lu, A. Urbas, E. Ehrichs, H. Jaeger, P. Mansky, and T. Russell, *Science*, **273**, 931–933 (1996). Local Control of Microdomain Orientation in Diblock Copolymer thin Films with Electric Fields.
29. M. Gopinadhan, Y. Choo, K. Kawabata, G. Kaufman, X. Feng, X. Di, Y. Rokhlenko, L. H. Mahajan, D. Ndaya, R. M. Kasi, et al., *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E9437–E9444 (2017). Controlling Orientational Order in Block Copolymers using Low-Intensity Magnetic Fields.
30. C. M. Bates and F. S. Bates, *Macromolecules*, **50**, 3–22 (2017). 50th Anniversary Perspective: Block Polymers–Pure Potential.
31. M. Lee, B.-K. Cho, and W.-C. Zin, *Chem. Rev.*, **101**, 3869–3892 (2001). Supramolecular Structures from Rod–Coil Block Copolymers.
32. K. Hayashida, T. Dotera, A. Takano, and Y. Matsushita, *Phys. Rev. Lett.*, **98**, 195502 (2007). Polymeric Quasicrystal: Mesoscopic Quasicrystalline Tiling in A B C Star Polymers.
33. A. Reddy, M. B. Buckley, A. Arora, F. S. Bates, K. D. Dorfman, and G. M. Grason, *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 10233–10238 (2018). Stable Frank–Kasper Phases of Self-Assembled, Soft Matter Spheres.
34. K. Azuma, J. Sun, Y. Choo, Y. Rokhlenko, J. H. Dwyer, B. Schweitzer, T. Hayakawa, C. O. Osuji, and P. Gopalan, *Macromolecules*, **51**, 6460–6467 (2018). Self-Assembly of an Ultrahigh- χ Block Copolymer with Versatile Etch Selectivity.
35. L. Leibler, *Macromolecules*, **13**, 1602–1617 (1980). Theory of Microphase Separation in Block Copolymers.

36. P. Bak and J. Von Boehm, *Phys. Rev. B*, **21**, 5297 (1980). Ising Model with Solitons, Phasons, and “the Devil’s Staircase”.
37. J. Briggs, H. Chung, and M. Caffrey, *J. Phys. II*, **6**, 723–751 (1996). The Temperature-Composition Phase Diagram and Mesophase Structure Characterization of the Monoolein/Water System.
38. M. Caffrey, *Annu. Rev. Biophys.*, **38**, 29–51 (2009). Crystallizing Membrane Proteins for Structure Determination: Use of Lipidic Mesophases.
39. G. Gompper, *Phase Transitions and Critical Phenomena*, Vol. **16**, Academic Press, New York, 1994.
40. A. Ben-Shaul and W. M. Gelbart, *Micelles, Membranes, Microemulsions, and Monolayers*, Springer, Berlin, 1994, pp. 1–104.
41. J. Israelachvili, *Colloids Surf., A*, **91**, 1–8 (1994). The Science and Applications of Emulsions—an Overview.
42. G. J. Tiddy, *Phys. Rep.*, **57**, 1–46 (1980). Surfactant-Water Liquid Crystal Phases.
43. X. Hou, T. Zaks, R. Langer and Y. Dong, *Nat. Rev. Mater.*, **6**, 1078–1094 (2021). Lipid nanoparticles for mRNA delivery.
44. C. J. Drummond and C. Fong, *Curr. Opin. Colloid Interface Sci.*, **4**, 449–456 (1999). Surfactant Self-Assembly Objects as Novel Drug Delivery Vehicles.
45. C. Mulligan, R. Yong, and B. Gibbs, *Eng. Geol.*, **60**, 371–380 (2001). Surfactant-Enhanced Remediation of Contaminated Soil: a Review.
46. L. L. Schramm, E. N. Stasiuk, and D. G. Marangoni, *Annu. Rep. Prog. Chem. Sect. C: Phys. Chem.*, **99**, 3–48 (2003). 2 Surfactants and Their Applications.
47. J. Israelachvili, S. Marčelja, and R. G. Horn, *Q. Rev. Biophys.*, **13**, 121–200 (1980). Physical Principles of Membrane Organization.
48. K. Simons and J. L. Sampaio, *Cold Spring Harbor Perspect. Biol.*, **3**, a004697 (2011). Membrane Organization and Lipid Rafts.
49. L. Norlén, *J. Invest. Dermatol.*, **117**, 823–829 (2001). Skin Barrier Formation: the Membrane Folding Model.
50. H. M. McConnell, *Annu. Rev. Phys. Chem.*, **42**, 171–195 (1991). Structures and Transitions in Lipid Monolayers at the Air-water Interface.
51. S. L. Veatch and S. L. Keller, *Biophys. J.*, **85**, 3074–3083 (2003). Separation of Liquid Phases in Giant Vesicles of Ternary Mixtures of Phospholipids and Cholesterol.
52. S. Munro, *Cell*, **115**, 377–388 (2003). Lipid Rafts: Elusive or Illusive?
53. K. Jacobson, O. G. Mouritsen, and R. G. Anderson, *Nat. Cell Biol.*, **9**, 7 (2007). Lipid Rafts: At a Crossroad between Cell Biology and Physics.
54. D. Lingwood and K. Simons, *Science*, **327**, 46–50 (2010). Lipid Rafts as a Membrane-Organizing Principle.
55. E. Sezgin, I. Levental, S. Mayor, and C. Eggeling, *Nat. Rev. Mol. Cell Biol.*, **18**,

- 361 (2017). The Mystery of Membrane Organization: Composition, Regulation and Roles of Lipid Rafts.
56. R. Elliott, *Phys. Rev.*, **124**, 346 (1961). Phenomenological discussion of magnetic ordering in the heavy rare-earth metals.
57. J. Rossat-Mignod, P. Burlet, H. Bartholin, O. Vogt, and R. Lagnier, *J. Phys. C: Solid State Phys.*, **13**, 6381 (1980). Specific heat analysis of the magnetic phase diagram of CeSb.
58. Y. Fujino, H. Sato, M. Hirabayashi, E. Aoyagi, and Y. Koyama, *Phys. Rev. Lett.*, **58**, 1012–1015 (1987). In situ Observation of Commensuration Process of Long-Period Superlattice in the Ag-Mg Alloys.
59. T. Hanaguri, C. Lupien, Y. Kohsaka, D.-H. Lee, M. Azuma, M. Takano, H. Takagi, and J. C. Davis, *Nature*, **430**, 1001–1005 (2004). A ‘Checkerboard’ Electronic Crystal State in Lightly Hole-Doped $\text{Ca}_{2-x}\text{Na}_x\text{CuO}_2\text{Cl}_2$.
60. M. E. Fisher and W. Selke, *Phys. Rev. Lett.*, **44**, 1502 (1980). Infinitely Many Commensurate Phases in a Simple Ising Model.
61. J. Lorenzana, G. Seibold, C. Ortix, and M. Grilli, *Phys. Rev. Lett.*, **101**, 186402 (2008). Competing Orders in FeAs Layers.
62. K. A. Moler, *Nature*, **468**, 643 (2010). High-Temperature Superconductivity: How the Cuprates Hid Their Stripes.
63. U. Löw, V. Emery, K. Fabricius, and S. Kivelson, *Phys. Rev. Lett.*, **72**, 1918 (1994). Study of an Ising Model with Competing Long- and Short-Range Interactions.
64. E. Edlund and M. N. Jacobi, *Phys. Rev. Lett.*, **105**, 137203 (2010). Universality of Striped Morphologies.
65. V. Emery, S. Kivelson, and J. Tranquada, *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 8814–8817 (1999). Stripe Phases in High-Temperature Superconductors.
66. E. Dagotto, *Science*, **309**, 257–262 (2005). Complexity in Strongly Correlated Electronic Systems.
67. W. Poon, *Science*, **304**, 830–831 (2004). Colloids as Big Atoms.
68. E. R. Weeks, *ACS Macro Lett.*, **6**, 27–34 (2017). Introduction to the Colloidal Glass Transition.
69. H. C. Hamaker, *Physica*, **4**, 1058–1072 (1937). The London–van der Waals Attraction Between Spherical Particles.
70. A. Vrij, *Pure Appl. Chem.*, **48**, 471–483 (1976). Polymers at Interfaces and the Interactions in Colloidal Dispersions.
71. H. N. W. Lekkerkerker and R. Tuinier, *Colloids and the Depletion Interaction*, Springer, Berlin, 2011.
72. E. J. W. Verwey, *J. Phys. Chem.*, **51**, 631–636 (1947). Theory of the Stability of Lyophobic Colloids.
73. S. Asakura and F. Oosawa, *J. Chem. Phys.*, **22**, 1255–1256 (1954). On Interaction Between Two Bodies Immersed in a Solution of Macromolecules.

74. S. Asakura and F. Oosawa, *J. Polym. Sci.*, **33**, 183–192 (1958). Interaction Between Particles Suspended in Solutions of Macromolecules.
75. C. P. Royall, *Soft Matter*, **14**, 4020–4028 (2018). Hunting Mermaids in Real Space: Known Knowns, Known Unknowns and Unknown Unknowns.
76. E. Lima, P. Pereira, H. Löwen, and S. Apolinario, *J. Phys. Condens. Matter*, **30**, 325101 (2018). Complex Structures Generated by Competing Interactions in Harmonically Confined Colloidal Suspensions.
77. A. I. Campbell, V. J. Anderson, J. S. van Duijneveldt, and P. Bartlett, *Phys. Rev. Lett.*, **94**, 208301 (2005). Dynamical Arrest in Attractive Colloids: The Effect of Long-Range Repulsion.
78. C. L. Klix, C. P. Royall, H. Tanaka, *Phys. Rev. Lett.*, **104**, 165702 (2010). Structural and Dynamical Features of Multiple Metastable Glassy States in a Colloidal System with Competing Interactions.
79. A. T. Pham, Y. Zhuang, P. Detwiler, J. E. Socolar, P. Charbonneau, and B. B. Yellen, *Phys. Rev. E*, **95**, 052607 (2017). Phase Diagram and Aggregation Dynamics of a Monolayer of Paramagnetic Colloids.
80. F. Sciortino, A. Giacometti, and G. Pastore, *Phys. Rev. Lett.*, **103**, 237801 (2009). Phase Diagram of Janus Particles.
81. G. Malescio and G. Pellicane, *Nat. Mater.*, **2**, 97 (2003). Stripe Phases from Isotropic Repulsive Interactions.
82. M. A. Glaser, G. M. Grason, R. D. Kamien, A. Košmrlj, C. D. Santangelo, and P. Ziherl, *EPL*, **78**, 46004 (2007). Soft Spheres Make More Mesophases.
83. H. Pattabhiraman and M. Dijkstra, *Soft Matter*, **13**, 4418–4432 (2017). On the Formation of Stripe, Sigma, and Honeycomb Phases in a Core–Corona System.
84. C. Likos, A. Lang, M. Watzlawek, and H. Löwen, *Phys. Rev. E*, **63**, 031206 (2001). Criterion for Determining Clustering versus Reentrant Melting Behavior for Bounded Interaction Potentials.
85. D. A. Lenz, R. Blaak, C. N. Likos, and B. M. Mladek, *Phys. Rev. Lett.*, **109**, 228301. Microscopically Resolved Simulations Prove the Existence of Soft Cluster Crystals (2012).
86. B. M. Mladek, P. Charbonneau, and D. Frenkel, *Phys. Rev. Lett.*, **99**, 235702 (2007). Phase Coexistence of Cluster Crystals: Beyond the Gibbs Phase Rule.
87. B. M. Mladek, P. Charbonneau, C. N. Likos, D. Frenkel, and G. Kahl, *J. Phys. Condens. Matter*, **20**, 494245 (2008). Multiple Occupancy Crystals Formed by Purely Repulsive Soft Particles.
88. K. Zhang, P. Charbonneau, and B. M. Mladek, *Phys. Rev. Lett.*, **105**, 245701 (2010). Reentrant and Isostructural Transitions in a Cluster–Crystal Former.
89. M. Adams, Z. Dogic, S. L. Keller, and S. Fraden, *Nature*, **393**, 349 (1998). Entropically Driven Microphase Transitions in Mixtures of Colloidal Rods and Spheres.

90. L. Ellison, D. Michel, F. Barmes, and D. Cleaver, *Phys. Rev. Lett.*, **97**, 237801 (2006). Entropy-Driven Formation of the Gyroid Cubic Phase.
91. R. P. Sear and W. M. Gelbart, *J. Chem. Phys.*, **110**, 4582–4588 (1999). Microphase Separation versus the Vapor-Liquid Transition in Systems of Spherical Particles.
92. A. J. Archer, *Phys. Rev. E*, **78**, 031402 (2008). Two-Dimensional Fluid with Competing Interactions Exhibiting Microphase Separation: Theory for Bulk and Interfacial Properties.
93. A. J. Archer, C. Ionescu, D. Pini, and L. Reatto, *J. Phys. Condens. Matter*, **20**, 415106 (2008). Theory for the Phase Behaviour of a Colloidal Fluid with Competing Interactions.
94. A. Ciach and W. Gozdz, *J. Phys. Condens. Matter*, **28**, 244004 (2016). Density Functional Theory for Systems with Mesoscopic Inhomogeneities.
95. P. Hohenberg and A. Krekhov, *Phys. Rep.*, **572**, 1–42 (2015). An Introduction to the Ginzburg–Landau Theory of Phase Transitions and Nonequilibrium Patterns.
96. M. Seul and D. Andelman, *Science*, **267**, 476–483 (1995). Domain Shapes and Patterns: The Phenomenology of Modulated Phases.
97. A. Ciach, *Phys. Rev. E*, **78**, 061505 (2008). Universal Sequence of Ordered Structures Obtained from Mesoscopic Description of Self-Assembly.
98. M. Shpot and H. W. Diehl, *Nucl. Phys. B*, **612**, 340–372 (2001). Two-loop renormalization-group analysis of critical behavior at m -axial Lifshitz points.
99. F. Sciortino, S. Mossa, E. Zaccarelli, and P. Tartaglia, *Phys. Rev. Lett.*, **93**, 055701 (2004). Equilibrium Cluster Phases and Low-Density Arrested Disordered States: the Role of Short-Range Attraction and Long-Range Repulsion.
100. A. Imperio and L. Reatto, *J. Chem. Phys.*, **124**, 164712 (2006). Microphase Separation in Two-Dimensional Systems with Competing Interactions.
101. C. O. Reichhardt, C. Reichhardt, and A. Bishop, *Phys. Rev. E*, **82**, 041502 (2010). Structural Transitions, Melting, and Intermediate Phases for Stripe- and Clump-Forming Systems.
102. B. Smit, P. Hilbers, K. Esselink, L. Rupert, N. Van Os, and A. Schlijper, *Nature*, **348**, 624 (1990). Computer Simulations of a Water/Oil Interface in the Presence of Micelles.
103. D. Wu, D. Chandler, and B. Smit, *J. Phys. Chem.*, **96**, 4077–4083 (1992). Electrostatic Analogy for Surfactant Assemblies.
104. R. D. Groot and T. J. Madden, *J. Chem. Phys.*, **108**, 8713–8724 (1998). Dynamic Simulation of Diblock Copolymer Microphase Separation.
105. F. J. Martínez-Veracochea and F. A. Escobedo, *Macromolecules*, **38**, 8522–8531 (2005). Lattice Monte Carlo Simulations of the Gyroid Phase in Monodisperse and Bidisperse Block Copolymer Systems.

106. F. A. Detcheverry, D. Q. Pike, P. F. Nealey, M. Müller, and J. J. de Pablo, *Phys. Rev. Lett.*, **102**, 197801 (2009). Monte Carlo Simulation of Coarse Grain Polymeric Systems.
107. T. Beardsley and M. W. Matsen, *Eur. Phys. J. E*, **32**, 255–264 (2010). Monte Carlo Phase Diagram for Diblock Copolymer Melts.
108. W. Selke and M. E. Fisher, *Phys. Rev. B*, **20**, 257 (1979). Monte Carlo Study of the Spatially Modulated Phase in an Ising Model.
109. M. Grousson, G. Tarjus, and P. Viot, *Phys. Rev. E*, **64**, 036109 (2001). Monte Carlo Study of the Three-Dimensional Coulomb Frustrated Ising Ferromagnet.
110. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Elsevier, Amsterdam, 2001.
111. E. B. Rasmussen and S. J. K. Jensen, *Phys. Rev. B*, **24**, 2744–2750 (1981). Devil’s-Staircase Behavior of a Simple Spin Model.
112. N. B. Wilding and P. Sollich, *EPL*, **101**, 10004 (2013). A Monte Carlo Method for Chemical Potential Determination in Single and Multiple Occupancy Crystals.
113. A. De Candia, E. Del Gado, A. Fierro, N. Sator, M. Tarzia, and A. Coniglio, *Phys. Rev. E*, **74**, 010403 (2006). Columnar and Lamellar Phases in Attractive Colloidal Systems.
114. M. Müller and K. C. Daoulas, *J. Chem. Phys.*, **128**, 024903 (2008). Calculating the Free Energy of Self-Assembled Structures by Thermodynamic Integration.
115. K. Zhang and P. Charbonneau, *Phys. Rev. Lett.*, **104**, 195703 (2010). Monte Carlo Study of the Axial Next-Nearest-Neighbor Ising Model.
116. K. Zhang and P. Charbonneau, *Phys. Rev. B*, **83**, 214303 (2011). Monte Carlo Approach for Studying Microphases Applied to the Axial Next-Nearest-Neighbor Ising and the Ising-Coulomb Models.
117. Y. Zhuang, K. Zhang, and P. Charbonneau, *Phys. Rev. Lett.*, **116**, 098301 (2016). Equilibrium Phase Behavior of a Continuous-Space Microphase Former.
118. W. C. Swope and H. C. Andersen, *Phys. Rev. A*, **46**, 4539 (1992). Thermodynamics, Statistical Thermodynamics, and Computer Simulation of Crystals with Vacancies and Interstitials.
119. F. Smallenburg, L. Filion, M. Marechal, and M. Dijkstra, *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 17886–17890 (2012). Vacancy-stabilized Crystalline Order in Hard Cubes.
120. B. M. Mladek, D. Gottwald, G. Kahl, M. Neumann, and C. N. Likos, *Phys. Rev. Lett.*, **96**, 045701 (2006). Formation of Polymorphic Cluster Phases for a Class of Models of Purely Repulsive Soft Spheres.
121. K. Zhang and P. Charbonneau, *J. Chem. Phys.*, **136**, 214106 (2012). pT Monte Carlo Simulations of the Cluster-Crystal-Forming Penetrable Sphere Model.

122. T. L. Hill, *Statistical Mechanics: Principles and Selected Applications*, Dover Publications, New York, 1987.
123. J. G. Kirkwood, *J. Chem. Phys.*, **3**, 300–313 (1935). Statistical Mechanics of Fluid Mixtures.
124. E. L. Thomas, D. M. Anderson, C. S. Henkee, and D. Hoffman, *Nature*, **334**, 598 (1988). Periodic Area-Minimizing Surfaces in Block Copolymers.
125. K. Nakazato, K. Oyamatsu, and S. Yamada, *Phys. Rev. Lett.*, **103**, 132501 (2009). Gyroid Phase in Nuclear Pasta.
126. J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids: with Applications to Soft Matter*, Academic Press, Cambridge, MA, 2013.
127. D. A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, CA, 2000.
128. A. Y. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists*, Springer Science & Business Media, Berlin, 2013.
129. F. Wang and D. Landau, *Phys. Rev. Lett.*, **86**, 2050 (2001). Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States.
130. E. Guggenheim, *J. Chem. Phys.*, **7**, 103–107 (1939). Grand Partition Functions and So-Called “Thermodynamic Probability”.
131. A. M. Ferrenberg and D. Landau, *Phys. Rev. B*, **44**, 5081 (1991). Critical Behavior of the Three-Dimensional Ising Model: A High-Resolution Monte Carlo Study.
132. A. J. Schultz and D. A. Kofke, *Phys. Rev. E*, **84**, 046712 (2011). Algorithm for Constant-Pressure Monte Carlo Simulation of Crystalline Solids.
133. M. Dijkstra, R. van Roij, and R. Evans, *Phys. Rev. E*, **59**, 5744 (1999). Phase Diagram of Highly Asymmetric Binary Hard-Sphere Mixtures.
134. K. Binder, *Z. Phys. B Cond. Matt.*, **43**, 119–140 (1981). Finite Size Scaling Analysis of Ising Model Block Distribution Functions.
135. Y. Zhuang and P. Charbonneau, *J. Chem. Phys.*, **147**, 091102 (2017). Communication: Microphase Equilibrium and Assembly Dynamics.
136. P. Charbonneau and M. Tarzia, *J. Chem. Phys.*, **155**, 024501 (2021). Solution of disordered microphases in the Bethe approximation.
137. R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.*, **57**, 2607 (1986). Replica Monte Carlo Simulation of Spin-Glasses.
138. H. G. Katzgraber, S. Trebst, D. A. Huse, and M. Troyer, *J. Stat. Mech.: Theory Exp.*, **2006**, P03018 (2006). Feedback-Optimized Parallel Tempering Monte Carlo.
139. A. Kalz, A. Honecker, S. Fuchs, and T. Pruschke, *Eur. Phys. J. B*, **65**, 533 (2008). Phase Diagram of the Ising Square Lattice with Competing Interactions.
140. S. Jin, A. Sen, W. Guo, and A. W. Sandvik, *Phys. Rev. B*, **87**, 144406 (2013). Phase Transitions in the Frustrated Ising Model on the Square Lattice.

141. A. Imperio and L. Reatto, *J. Phys. Condens. Matter*, **16**, S3769 (2004). A Bidimensional Fluid System with Competing Interactions: Spontaneous and Induced Pattern Formation.
142. M. Zheng and P. Charbonneau, *J. Chem. Phys.*, **154**, 244506 (2021). Characterization and efficient Monte Carlo sampling of disordered microphases.
143. M. Newman and G. Barkema, *Monte Carlo Methods in Statistical Physics*, Oxford University Press, New York, 1999.
144. R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.*, **58**, 86–88 (1987). Nonuniversal Critical Dynamics in Monte Carlo Simulations.
145. U. Wolff, *Phys. Rev. Lett.*, **62**, 361 (1989). Collective Monte Carlo Updating for Spin Systems.
146. N. Pleimling and M. Henkel, *Phys. Rev. Lett.*, **87**, 125702 (2001). Anisotropic Scaling and Generalized Conformal Invariance at Lifshitz Points.
147. W. Janke, *Math. Comput. Simul.*, **47**, 329–346 (1998). Nonlocal Monte Carlo Algorithms for Statistical Physics Applications.
148. A. Coniglio and W. Klein, *J. Phys. A: Math. Gen.*, **13**, 2775–2780 (1980). Clusters and Ising Critical Droplets: a Renormalisation Group Approach.
149. P. W. Leung and C. L. Henley, *Phys. Rev. B*, **43**, 752–759 (1991). Percolation Properties of the Wolff Clusters in Planar Triangular Spin Models.
150. V. Cataudella, G. Franzese, M. Nicodemi, A. Scala, and A. Coniglio, *Phys. Rev. E*, **54**, 175–189 (1996). Percolation and Cluster Monte Carlo Dynamics for Spin Models.
151. H. Fajen, A. K. Hartmann, and A. P. Young, *cond-mat*, 1905.04220 (2019). Percolation of Fortuin-Kasteleyn Clusters for the Random-Bond Ising Model.
152. S. Whitelam and P. L. Geissler, *J. Chem. Phys.*, **127**, 154101 (2007). Avoiding Unphysical Kinetic Traps in Monte Carlo Simulations of Strongly Attractive Particles.
153. B. Chen and J. I. Siepmann, *J. Phys. Chem. B*, **104**, 8725–8734 (2000). A Novel Monte Carlo Algorithm for Simulating Strongly Associating Fluids: Applications to Water, Hydrogen Fluoride, and Acetic Acid.
154. M. Broccio, D. Costa, Y. Liu, and S.-H. Chen, *J. Chem. Phys.*, **124**, 084501 (2006). The Structural Properties of a Two-Yukawa Fluid: Simulation and Analytical Results.
155. J.-M. Bomont, J.-L. Bretonnet, and D. Costa, *J. Chem. Phys.*, **132**, 184508 (2010). Temperature Study of Cluster Formation in Two-Yukawa Fluids.
156. J. A. Bollinger and T. M. Truskett, *J. Chem. Phys.*, **145**, 064902 (2016). Fluids with Competing Interactions. I. Decoding the Structure Factor to Detect and Characterize Self-Limited Clustering.
157. J. N. Israelachvili, *Intermolecular and Surface Forces*, Academic Press, New York, 2015, p 512.

158. E. Mani, W. Lechner, W. K. Kegel, and P. G. Bolhuis, *Soft Matter*, **10**, 4479–4486 (2014). Equilibrium and Non-Equilibrium Cluster Phases in Colloids with Competing Interactions.
159. M. B. Sweatman, R. Fartaria, and L. Lue, *J. Chem. Phys.*, **140**, 124508 (2014). Cluster Formation in Fluids with Competing Short-Range and Long-Range Interactions.
160. Y. Hu and P. Charbonneau, *Soft Matter*, **14**, 4101–4109 (2018). Clustering and Assembly Dynamics of a One-Dimensional Microphase Former.
161. M. B. Sweatman and L. Lue, *Adv. Theory Simul.*, **2**, 1900025 (2019). The Giant SALR Cluster Fluid: A Review.
162. R. Jadrlich, B. Lindquist, and T. Truskett, *J. Chem. Phys.*, **146**, 184103 (2017). Probabilistic Inverse Design for Self-Assembling Materials.
163. P. D. Godfrin, N. E. Valadez-Pérez, R. Castaneda-Priego, N. J. Wagner, and Y. Liu, *Soft Matter*, **10**, 5061–5071 (2014). Generalized Phase Behavior of Cluster Formation in Colloidal Dispersions with Competing Interactions.
164. J. C. F. Toledano, F. Sciortino, and E. Zaccarelli, *Soft Matter*, **5**, 2390–2398 (2009). Colloidal Systems with Competing Interactions: from an Arrested Repulsive Cluster Phase to a Gel.
165. A. P. Santos and A. Z. Panagiotopoulos, *J. Chem. Phys.*, **144**, 044709 (2016). Determination of the Critical Micelle Concentration in Simulations of Surfactant Systems.
166. A. P. Santos, J. Pekalski, and A. Z. Panagiotopoulos, *Soft Matter*, **13**, 8055–8063 (2017). Thermodynamic Signatures and Cluster Properties of Self-Assembly in Systems with Competing Interactions.
167. J. Pekalski, E. Bildanau, and A. Ciach, *Soft Matter*, **15**, 7715 (2019). Self-Assembly of Spiral Patterns in Confined Systems with Competing Interactions.
168. S. Jiao, A. P. Santos, and A. Z. Panagiotopoulos, *Fluid Phase Equilib.*, **470**, 126–133 (2018). Differences in Free Surfactant Concentration and Aggregation Properties for Amphiphiles with the Same Critical Micelle Concentration.
169. D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, CRC Press, London, 1994.
170. L. Van Hove, *Physica*, **16**, 137–143 (1950). Sur l'intégrale de Configuration pour les Systèmes de Particules à une Dimension.
171. J. Pekalski, A. Ciach, and N. Almarza, *J. Chem. Phys.*, **138**, 144903 (2013). Periodic Ordering of Clusters in a One-Dimensional Lattice Model.
172. A. Ciach and J. Pekalski, *Soft Matter*, **13**, 2603–2608 (2017). Exactly Solvable Model for Self-Assembly of Hard Core–Soft Shell Particles at Interfaces.
173. S. Redner, *J. Stat. Phys.*, **25**, 15–23 (1981). One-Dimensional Ising Chain with Competing Interactions: Exact Results and Connection with other Statistical Models.

174. S. Redner and H. Stanley, *Phys. Rev. B*, **16**, 4901 (1977). Helical Order and its Onset at the Lifshitz Point.
175. H. A. Ceccatto, C. J. Gazza, and A. E. Trumper, *Phys. Rev. B*, **45**, 7832–7836 (1992). J_1 - J_2 Model: Energy, Correlations, and Order-Parameter Fluctuations on Finite Lattices.
176. M. E. Fisher, *J. Appl. Phys.*, **52**, 2014–2018 (1981). An Infinity of Commensurate Phases in a Simple Ising System: The ANNNI Model.
177. W. Selke and P. M. Duxbury, *Z. Phys. B: Condens. Matter*, **57**, 49–58 (1984). The Mean Field Theory of the Three-Dimensional ANNNI Model.
178. M. Henkel and M. Pleimling, *Comput. Phys. Commun.*, **147**, 419 (2002). Anisotropic Scaling and Generalized Conformal Invariance at Lifshitz Points.
179. J. Oitmaa, *J. Phys. A: Math. Gen.*, **18**, 365 (1985). A High Temperature Series Study of the ANNNI Model in Two and Three Dimensions.
180. D. Pini and A. Parola, *Soft Matter*, **13**, 9259–9272 (2017). Pattern Formation and Self-Assembly Driven by Competing Interactions.
181. D. Stopper and R. Roth, *Phys. Rev. E*, **96**, 042607 (2017). Phase Behavior and Bulk Structural Properties of a Microphase Former with Anisotropic Competing Interactions: A Density Functional Theory Study.

4

MOLECULAR SIMULATIONS OF DEEP EUTECTIC SOLVENTS: A PERSPECTIVE ON STRUCTURE, DYNAMICS, AND PHYSICAL PROPERTIES

SHALINI J. RUKMANI^{1,2,*}, BRIAN W. DOHERTY^{3,*}, ORLANDO ACEVEDO³ and CORAY M. COLINA^{1,2,4}

¹*Department of Materials Science and Engineering, University of Florida, Gainesville, USA*

²*George and Josephine Butler Polymer Research Laboratory, Department of Chemistry, University of Florida, Gainesville, USA*

³*Department of Chemistry, University of Miami, Coral Gables, USA*

⁴*Department of Chemistry, University of Florida, Gainesville, USA*

INTRODUCTION

Sustainable solvents have attracted significant attention over the last three decades for the synthesis of functional materials in energy conversion, storage, and separation applications.¹⁻⁵ For most applications, conventional organic solvents and inorganic electrolytes have proven to be expensive,

*These authors contributed equally to this work.

Reviews in Computational Chemistry, Volume 32, First Edition.

Edited by Abby L. Parrill and Kenny B. Lipkowitz.

© 2022 John Wiley & Sons, Inc. Published 2022 by John Wiley & Sons, Inc.

energy-intensive, sensitive to moisture, and produce toxic effects harming the environment through formation of residual products and gaseous emissions into the atmosphere.^{5–7} To address these shortcomings, significant research efforts have focused on the development of green and sustainable solvents.^{8–10} Deep eutectic solvents (DESs) have become attractive alternatives, and an increasing research effort to understand the structure–property relation in these solvents for diverse applications have grown since early 2000s.^{11–15} Many research studies have focused on the use of DESs as reaction media for the synthesis of functional materials in electrocatalysis, fuel cells, organic synthesis, biomass, and biodiesel purification.^{2, 16–18} DESs have also been investigated as a possible alternative for CO₂ scrubbing applications such as aqueous amines, aqueous ammonia, and potassium carbonate due to their biodegradability, low cost of production, and low toxicity.^{19, 20} Other important applications of DES include liquid electrolyte alternatives for photovoltaic devices and nanostructured sensors, metal processing such as electrodeposition, metal extraction and processing of metal oxides, and electropolishing.^{21–28}

Several experimental and computational investigations have contributed toward the understanding of their complex structure and interaction between the constituents.^{13, 15, 29} Molecular simulations have played a significant role in elucidating the intricacies present in DESs; in particular, the effect of intermolecular interactions on the observed macroscopic bulk properties.^{30–36} Simulations have also been performed on DES in conjunction with other materials including metal surfaces, proteins, and gas molecules with specific emphasis on the interaction between DES and molecules/surfaces, interfacial properties, and gas sorption.^{37–39} Rather than providing an overview of simulation studies performed to the date this chapter was written on DESs, this review aims to accomplish three goals: (1) Provide the essential background to a novice modeler on the choice of simulation techniques used to model DESs, (2) Describe methods used to obtain important physical, thermodynamic, transport, and structural properties of bulk DES systems including an evaluation of the strengths and drawbacks of the current simulation models, and (3) Discuss future directions for simulating DES-based systems.

In section “Deep Eutectic Solvents”, the molecular structure and types of DESs are described. Sections “Molecular Simulation Methods, Physical Properties, Thermodynamic Properties, Transport Properties, and Deep Eutectic Solvent Structure” provide an overview of atomistic simulation methods used to model DES systems with a discussion as to how physical, thermodynamic, and transport properties are obtained from atomistic molecular simulations. Finally, a summary of the overall performance of current

simulation models is provided that highlights the strengths and drawbacks of each method in representing the structure and properties of DESs. Potential future directions for atomistic simulations of these unique solvent systems are discussed.

DEEP EUTECTIC SOLVENTS

Definition of Deep Eutectic Solvents

The term “deep eutectic solvents” was coined from the decrease observed in the melting temperature of a solvent mixture relative to the melting temperatures of the pure components prior to mixing. The *eutectic temperature* is thus defined as the lowest melting temperature for a given mixture and the corresponding composition is called the *eutectic composition*. DESs collectively denote liquids that are close to this eutectic composition. These solvents consist of large and asymmetric ions with low lattice energies. Abbott et al.¹¹ synthesized the first DES in 2001 by mixing metal chlorides (ZnCl_2 , and/or SnCl_2) and quaternary ammonium salts. DESs are usually obtained by mixing a hydrogen bond acceptor (HBA) (such as quaternary ammonium halide salts) with a hydrogen bond donor (HBD) molecule that has the ability to form a complex with the halide, leading to a depression of the freezing point of the resulting mixture. Figure 1 shows as an example the freezing point curve of a DES system formed from a halide salt and a neutral organic compound, namely, choline chloride (ChCl) and urea.⁴⁰ The freezing point is the lowest (12°C) at 65 mol% urea than the freezing point of the original constituents (ChCl = 302°C and urea = 133°C). In general, the freezing point of most DESs are less than 150°C .^{13, 41}

DES as Ionic Liquid Analogues

The development of DESs by Abbott et al.¹¹ originated from their efforts to overcome the limitations associated with conventional imidazolium-based room temperature ionic liquids (RTILs). For example, the high cost of RTILs for bulk-scale applications and low-moisture stability associated with the use of salts such as aluminum chloride led to replacements featuring the combination of alternative metal chlorides with quaternary ammonium salts. It was observed that low symmetry cations in general led to a decrease in the freezing point, with ChCl showing the lowest freezing point among the systems tested. Subsequently, DESs based on ChCl and carboxylic acids were also synthesized by Abbott et al. and were shown to exhibit a similar depression in the freezing point.⁴² These liquids share similar physical properties

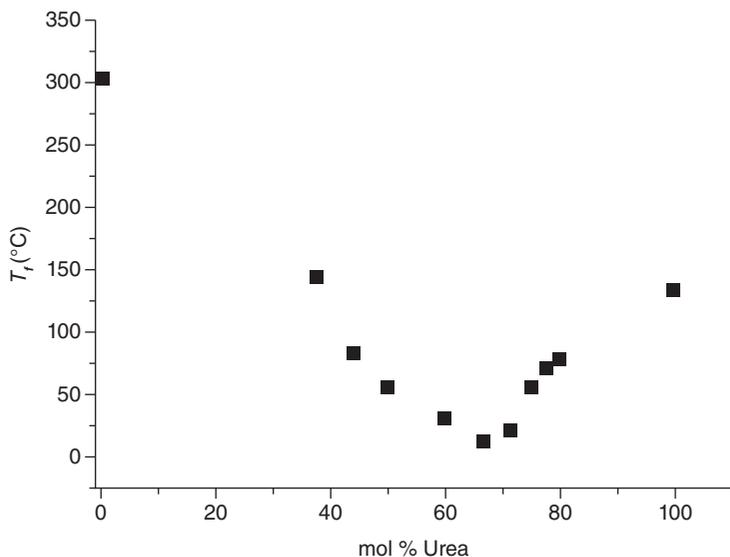


FIGURE 1 Freezing point of choline chloride (ChCl)-urea DES system as a function of composition. *Source:* Abbott et al. 2003⁴⁰.

to RTILs, including high viscosity, large surface tension, low vapor pressure, and nonflammability. Consequently, DESs are often termed as ionic liquid (IL) analogues. Nevertheless, it is important to underscore the differences between DESs and ionic liquids (ILs) with respect to the constituents and molecular interactions that govern their unique properties. One of the important differences between conventional ILs and DESs is that ILs are made from discrete anions and cations, whereas DESs are synthesized by mixing two components that form a eutectic mixture which typically consists of cations, anions, and neutral organic compounds. The differences in the properties exhibited by DESs and ILs arise from the contribution of molecular interactions from different components. While ionic interactions dominate in ILs, molecular interactions have a significant contribution, starting from hydrogen bonding interactions in DESs. Moreover, an ease of synthesis in the pure state, moisture insensitivity, and biodegradability are some of the more attractive properties that differentiate DESs from ILs. Figure 2 shows an example of a DES system, reline, formed from ChCl and urea in a 1 : 2 ratio, respectively, where ChCl is the HBA and urea is the HBD. These components form a eutectic mixture primarily due to contributions from a network of hydrogen bonding interactions.

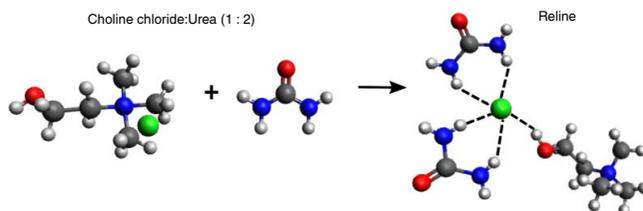


FIGURE 2 Schematic showing choline chloride (ChCl) and urea mixed in the ratio 1 : 2, respectively, to form reline. Elements are colored as ● carbon, ● oxygen, ● chloride, ● nitrogen, ● hydrogen. The dashed lines represent the formation of hydrogen bonds between different pairs including Cl^- and the hydrogen atom of the OH group in Ch^+ , and Cl^- and hydrogen atoms present in urea.

TABLE 1 Estimated Relative Contribution of Hydrogen Bonding Present in CCEtg, CCU, CCGly, and CCMal^a.

| DES ^b | Choline–Cl | Choline–HBD | HBD–Cl | HBD–HBD |
|------------------|------------|-------------|--------|---------|
| CCEtg | 0.15 | 0.05 | 0.57 | 0.23 |
| CCU | 0.08 | 0.08 | 0.32 | 0.52 |
| CCGly | 0.11 | 0.05 | 0.50 | 0.34 |
| CCMal | 0.30 | 0.05 | 0.51 | 0.14 |

^aAveraging was performed over a 1 ns trajectory with 2 ps between each frame.

^bSystem abbreviations are defined in Table 3.

Source: Adapted from Perkins et al. 2014³¹.

Molecular Structure of DESs and Type of Interactions

The properties of DESs are controlled primarily by hydrogen bonding between the different components of the mixture (cation and anion of HBA, and the HBD species). However, appreciable contributions from electrostatic interactions and van der Waals forces on some physical properties such as viscosity have been reported.^{41, 43, 44} Figure 3 illustrates DES formation through a potential complexation of the Cl⁻ anion present in ChCl with urea (HBD)³³ and Table 1 shows the fraction of hydrogen bonds in four ChCl-based DES systems studied by Perkins et al. through atomistic molecular dynamics (MD) simulations.³¹ In three of the systems studied, namely, ethaline, glyceline, and maline, the fraction of hydrogen bonding interactions between the HBD and corresponding anion were found to be the largest. However, in the reline system, the urea–urea interactions were found to be significant. Despite the importance of intermolecular interactions within DESs, quantitatively reproducing the molecular structure of the systems has proven quite challenging for multiple simulation methods including classical MD, *ab initio* MD (AIMD), first-principle MD (FPMD), and mixed quantum and molecular mechanical (QM/MM) with deviations reported for properties such as radial distribution functions (RDFs) and hydrogen bonding behavior.^{34, 35, 45, 46}

Earlier experimental and simulation studies have suggested that negative charge delocalization plays a major role in decreasing the melting point of the individual components, a consequence of hydrogen bonding between the mixture components, especially between the halide ion and the HBD moiety.^{47, 48} However, recent computational investigations based on *ab initio* and molecular mechanics (MM) calculations have shown that there are complex interactions present rather than a simple charge delocalization. For example, charge spreading in ChCl-based DES was investigated by Zahn

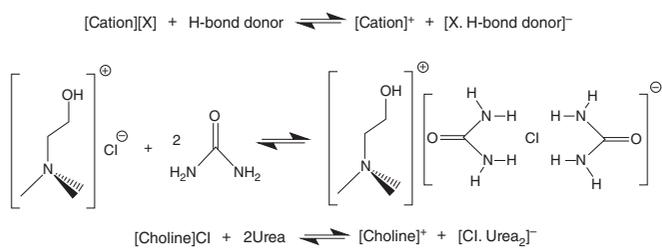


FIGURE 3 Illustration of a possible complex formation in reline DES (ChCl-urea in 1 : 2 molar ratio). *Source:* Ashworth et al. 2016³³.

et al.³⁴ where it was found that hydrogen bonding enhances negative charge spreading from the anion to the HBD, whereas the spreading of positive charge is decreased. However, in the case of ChCl–urea systems, negligible charge spreading was found between the anion and HBD as a result of increased hydrogen bonding between the Ch cation and the Cl anion as compared to Ch–urea. Ashworth et al.³³ also studied ChCl–urea as a model system to understand double ionic hydrogen bond interactions and found that urea forms a H-bonded complex with the cation, namely, urea[choline]⁺, which has been shown to form the strongest H-bond identified between the hydrogen atom of the hydroxyl group of Ch⁺ and the oxygen atom of the urea molecule. This complex is energetically competitive with the [Cl(urea)₂]⁻ complex, which has been associated with the eutectic behavior observed in reline. The negative charge on the anion complex was found to be localized. To summarize, recent investigations have challenged the earlier explanation that negative charge delocalization between the HBD and anion is a predominant factor in decreasing the melting point of the DES mixture relative to the original components.

Types of DES

The general formula used to describe DESs is Cat⁺X^{-z}Y where Cat⁺ denotes the cation, which can be any ammonium, phosphonium, or sulfonium cation, and X is a Lewis base, usually a halide ion such as Cl⁻. Based on the complexing agent, DESs are commonly divided into four types as described in Table 2.

Type I is analogous to the ILs formed using metal-halide/imidazolium salts. The first DES synthesized by Abbott and coworkers^{11, 13} by mixing quaternary ammonium salts and metal chloride salts belong to this category. Type II DESs are formed by mixing hydrated metal halides and ChCl. The inclusion of hydrated metal halides as one of the DES components is

TABLE 2 Classification of DES Based on the General Formula Cat⁺X^{-z}Y.

| Type | General formula |
|------|---|
| I | Cat ⁺ X ^{-z} MCl _x , M = Zn, Sn, Fe, Al, Ga |
| II | Cat ⁺ X ^{-z} MCl _x ·yH ₂ O, M = Cr, Co, Cu, Ni, Fe |
| III | Cat ⁺ X ^{-z} RZ, Z = CONH ₂ , COOH, OH |
| IV | MCl _x + RZ = MCl _{x-1} ⁺ ·RZ + MCl _{x+1} ⁻ , M = Al, Zn and Z = CONH ₂ , OH |

Source: Adapted from Smith et al. 2014¹³.

TABLE 3 Deep Eutectic Solvents Composed of Choline Chloride (ChCl) and a Hydrogen Bond Donor (HBD) at Specific Ratios (e.g., ChCl:HBD of 1 : 1, 1 : 2, 1 : 3).

| Abbreviation | HBD | Name |
|--------------|------------------|-----------|
| CCEtg | Ethylene glycol | Ethaline |
| CCGly | Glycerol | Glyceline |
| CCLev | Levulinic acid | — |
| CCMal | Malonic acid | Maline |
| CCOx | Oxalic acid | Oxaline |
| CCPhe | Phenol | — |
| CCPro | Propylene glycol | Propeline |
| CCU | Urea | Reline |

promising for large-scale production due to lower costs and their ability to withstand moisture.¹³ Following their initial work on metal chloride salts, Abbott et al.^{40, 42} synthesized DESs based on quaternary ammonium salts and molecular HBDs such as urea, alcohols, and carboxylic acids giving rise to type III DESs. DESs formed from inorganic cations constitute type IV eutectics. Transition metal halides such as ZnCl_2 have been able to form eutectics with HBDs such as urea, ethylene glycol, and acetamide.⁴⁹

Type III DESs are one of the most commonly investigated classes, both experimentally and computationally, as they comprise eutectic mixtures formed from a variety of halide salts and neutral HBDs (Figure 4). Type III DESs based on ChCl have been of particular interest to researchers due to several advantages that include simple and versatile preparation from relatively inexpensive components, low toxicity, and biodegradability.^{13, 31, 34, 41, 42, 50–53} Additionally, they have enabled the study of the interactions between the ChCl cation and HBDs, and the effects of molecular interactions on bulk-phase thermodynamic and transport properties in general. Table 3 summarizes the names and composition of the most frequently used type III DESs; for consistency, the DES solvents will be referred to by their abbreviations for the remainder of this chapter.

MOLECULAR SIMULATION METHODS

The first molecular simulation investigations of DESs were performed in 2013,^{31, 54, 55} a decade after DESs were first discovered by Abbott et al.^{11, 40}

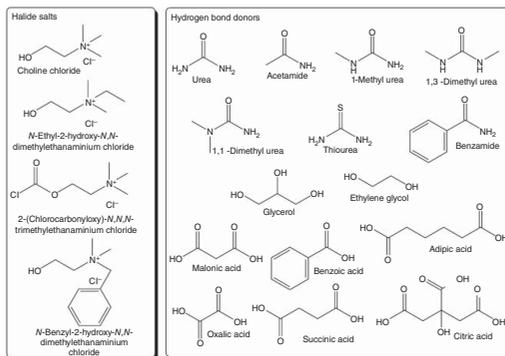


FIGURE 4 Molecular structures of commonly used halide salts and hydrogen bond donors used in the formation of type III DES.
 Source: Smith et al. 2014¹³.

Molecular simulations have played a crucial role in conjunction with experimental investigations in elucidating the structure–property relationships of DESs. Simulation techniques capable of examining different time scales may be necessary depending on the specific property being investigated. In general, simulation studies in the DES field have focused on: (1) understanding negative charge delocalization and charge spreading in DES species to rationalize the lowering of melting temperatures; (2) unravelling complex interactions between different components in the systems (cations, anions, and HBDs) that include contributions from hydrogen bonding and electrostatic interactions, e.g., doubly ionic bonds in different DES systems; and (3) simulating thermodynamic, physical, and transport properties in bulk-phase DES systems.^{20, 37–39, 56} There has also been an increasing number of simulation studies for DES systems in combination with gases, solid interfaces, mixtures, and biomolecules.^{20, 37–39} This chapter describes efforts to reproduce thermodynamic, physical, and structural properties of bulk-phase DES systems using atomistic MD simulations, while also underscoring important conclusions made by *ab initio* methods. In section “An Overview of *Ab Initio* Methods”, a brief overview of the *ab initio* investigations that have played an important role in understanding the molecular structure of DES is provided. Section “Classical Molecular Dynamics at the Atomic Level” discusses atomistic MD simulation methods and section “Nonpolarizable Force Fields Used for DES Simulations” provides a description of nonpolarizable FFs used for DES simulations.

An Overview of *Ab Initio* Methods

One of the major thrust areas of *ab initio* investigations on DESs has been to provide a physical explanation for the observed low melting point in these systems and its effects on their physicochemical properties. García et al.⁵⁷ studied the melting points of ChCl based DESs, where 29 HBDs were optimized at the B3LYP/6-31+G(d) level of theory. The quantitative structure–activity relationship methodology was then utilized to develop a model using a genetic function approximation. Later, similar methods were employed by the same research group³⁰ to rationalize potential correlations between the melting temperatures and the molecular structure for 45 different DES systems, mostly based on choline. A combination of density functional theory (DFT) and a topological analysis of electron density was employed to better understand intermolecular interactions, particularly for hydrogen bonding networks and their effect on the melting point of DESs. Functionals used to perform DFT calculations for DESs must account for dispersion forces given their importance in accurately describing

long-range interactions for these ionic solvents.^{58–61} For example, the DFT investigation by García et al.³⁰ used the B3LYP functional with Grimme's scheme⁶² that accounts for dispersion corrections (B3LYP-D2). Figure 5 provides B3LYP-D2/6-31+G(d,p) optimized structures for four selected DESs with the intermolecular hydrogen-bonded network represented by dotted lines and the Bader cage critical points by points labeled "cp". Cage-like structures were formed by the HBD-chloride hydrogen-bonded interactions and the HBD-cation and anion–cation interactions. This work represents one of the first contributions toward understanding the correlation between macroscopic properties, such as the experimentally observed lowering of melting points, and the molecular structure in terms of hydrogen bond networks.

Wagle et al. also performed DFT calculations to study the mobility of different components in CCGly, in conjunction with quasielastic neutron scattering coupled with selective deuteration.⁶³ In this work, the M06-2X/6-31++G(d,p) level of theory was applied to study the local diffusion dynamics of glyceline's components. The M06-2X DFT functional was chosen as it has been reported to provide accurate descriptions of non-covalent interactions including dispersion effects.^{64–66} The calculations provided a physical explanation, i.e., the competitive nature of hydrogen bonding, for the observed higher local diffusion dynamics of Ch^+ as compared to glycerol in CCGly (Figure 6). This contrasted with the observed slower long-range diffusion dynamics of Ch^+ in comparison to Cl^- . Further *ab initio* calculations were performed on CCU, CCEtg, and CCMal.⁶⁷ Initial geometry optimizations and a subsequent reoptimization were done at the HF/6-311G(d,p) and M06-2X/6-31++G(d,p) levels, respectively. The DES species displayed a cage-like nanostructure due to cooperative H-bonding between HBDs, cations, and anions. A charge distribution analysis indicated higher charge transfer from Ch^+ to the HBD as compared to that occurring from Cl^- to the HBD. The calculated sum of bond orders for $\text{Ch}-\text{Cl}$ interactions correlated directly with the melting point of the DESs.

Zahn et al.³⁴ were the first to perform *ab initio* MD calculations to study charge spreading in the liquid state for CCU, CCGly, and CCOx. The Hirshfield-I partial charge analysis method, previously used for ionic systems,⁶⁸ was applied and it was found that increasing hydrogen bond interactions between the anion and the HBD compound increased the negative charge spreading to the HBD with a reduction of the positive charge spreading for CCGly and CCOx. In contrast, negative charge spreading was negligible on CCU. The results obtained from this study challenged the commonly held notion that negative charge delocalization occurring

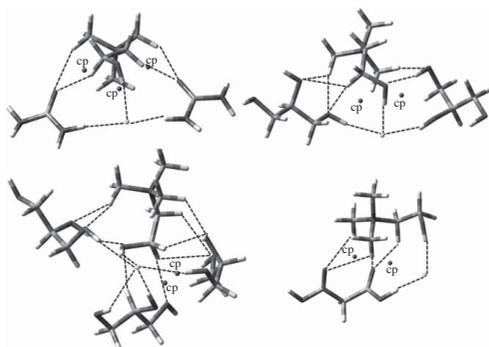


FIGURE 5 Optimized structures for selected DES systems: (top left) ChCl-Urea (1 : 2 molar ratio), (top right) ChCl-Glycerol (1 : 2 molar ratio), (bottom left) ChCl-Glycerol (1 : 3 molar ratio), and (bottom right) ChCl-Malonic acid (1 : 1 molar ratio). Intermolecular hydrogen bonds are represented by ---- and cage critical points are labeled cp. *Source:* García et al. 2015³⁰.

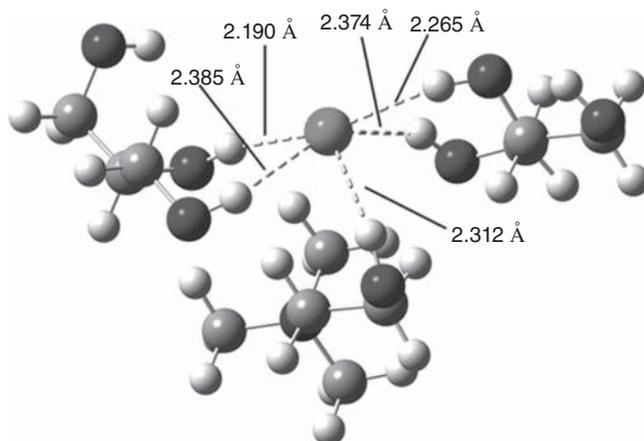


FIGURE 6 *Ab initio* calculations showing hydrogen bonds formed between glycerol molecules and chloride dominate over the ones occurring between the choline cation and chloride. *Source:* Wagle et al. 2015⁶³.

between anion and HBD in DESs was responsible for the observed lowering of their melting points. Follow-up studies further investigated the molecular structure and hydrogen bond dynamics of CCU,⁶⁹ alkali halide crystals,⁷⁰ and doubly ionic bonds.³³

To summarize, *ab initio* simulations have provided a detailed description of the molecular structure of DESs that illustrate the complex nature of component interactions arising from hydrogen bond networks and electrostatic contributions. The charge transfer processes between the cation, anion, and HBD were examined for several ChCl-based DESs. The majority of these investigations have focused on unravelling the effect of molecular interactions on the “deep eutectic” behavior, i.e., a decrease in melting point near the eutectic composition. However, given the large computational cost of these *ab initio* methods in terms of both the time and computer resources required, the expansion of the QM-based methods toward large bulk-phase DES simulations is not feasible at present. Therefore, exploration of existing bulk-phase DES thermodynamic and transport properties and the design of new DES solvents for specific applications necessitates the use of molecular mechanics-based force fields. The subsequent sections provide a description of classical molecular dynamics (MD) simulation methods used to model larger sized, e.g., hundreds to thousands of ionic/molecular components, DES systems.

Classical Molecular Dynamics at the Atomic Level

Classical MD simulations that obey Newton's laws of motion use force fields (FFs) to calculate the potential energy of a system as a function of their atomic coordinates. The choice of FF plays an important role in the prediction of properties from atomistic molecular simulations and must be chosen carefully depending on the type of system to model. Generalized FFs like GAFF⁷¹ and DREIDING⁷² can offer qualitative and quantitative information complementing experimental results for a wide range of small molecule and macromolecular systems. However, the application of these FFs to charged solvents such as ILs and DESs necessitates further refinement and development of new parameter sets to obtain good agreement with experimental data. To reproduce and predict thermodynamic, transport, and structural properties of DESs, it is important that the FF chosen accurately reproduces the molecular geometry, nonbonded interactions, and properly samples the conformational space of these systems. Therefore, any molecular simulations involving a new DES system must involve a thorough validation of the FF selected to provide confidence in the predictive results obtained for the properties of interest.

The majority of the DES simulation studies have used FFs such as GAFF⁷¹ and OPLS-AA,^{73–77} which follow the general “class I” equation [1], with the exception of molecular simulations of CCU^{32, 78} that have employed the Merck molecular force field (MMFF),^{79–81} a class II FF. The potential energy, $U(r)$, in Eq. [1] is represented as the sum of bonded and nonbonded interactions present in a system. The bonded interactions typically comprise U_{bond} , U_{angle} , U_{dihedral} , and U_{improper} terms as described in Eqs. [2]–[5]. The bond stretch interaction is described by a simple harmonic oscillator between atoms i and j as shown in Eq. [2a]. K_{ij}^b denotes the force constant, r_{ij} represents the distance between atoms i and j , and r_{ij}^0 represents the equilibrium bond length. Similarly, U_{angle} can be represented as a harmonic term ($U_{\text{angle},h}$ as described in Eq. [3a]) where K_{ijk}^θ , θ_{ijk} , and θ_{ijk}^0 represent the angle force constant, angle between atoms i , j , and k , and equilibrium angle, respectively. A dihedral energy term is typically expressed as a cosine series as given by Eq. 4, where V_n represents the “barrier height,” n is the periodicity of the potential and γ is the phase angle. Energy contribution from impropers, or out of plane bending, can also be taken explicitly into account, as given by Eq. 5 where K_ω is the force constant and $\omega_{ijk} - \omega_{jkl}^0$ is the out of the plane angle. Impropers can be also implemented using the Wilson wag that is the i - l bond angle with the j - k - l plane. Class II FFs such as COMPASS,⁸² PCFF,⁸³ and MMFF^{79–81} contain higher order force constants (typically cubic and quartic) for bond

and angle terms, and also contain off-diagonal cross-coupling terms such as stretch–bend and bend–bend interactions. For example, in the case of MMFF, cubic-stretch and bend terms are used as shown in Eqs. [2b] and [3b] where $K'_{ij}{}^b$ and $K'_{ijk}{}^\theta$ are cubic-stretch and cubic-bend constants. Also, a stretch–bend cross-term is included as given in Eq. [6], where K^{sb}_{ijk} and K^{sb}_{kji} are the force constants for i – j and k – j stretching coupled to i – j – k bending, r_{ij} and r_{kj} represent the bond lengths between atoms i – j and j – k , and θ_{ijk} denotes the angle between i – j – k atoms.

$$U(r) = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} + U_{\text{improper}} + U_{\text{van der Waals}} + U_{\text{electrostatics}} \quad [1]$$

$$U_{\text{bond},h} = K'_{ij}{}^b (r_{ij} - r_{ij}^0)^2 \quad [2a]$$

$$U_{\text{bond},q} = 143.9325 K'_{ij}{}^b / 2 (r_{ij} - r_{ij}^0) \left(1 + cs (r_{ij} - r_{ij}^0) + \frac{7}{12} (cs^2 (r_{ij} - r_{ij}^0)^2) \right) \quad [2b]$$

$$U_{\text{angle},h} = K'_{ijk}{}^\theta (\theta_{ijk} - \theta_{ijk}^0)^2 \quad [3a]$$

$$U_{\text{angle},c} = 0.043844 K'_{ijk}{}^\theta / 2 (\theta_{ijk} - \theta_{ijk}^0)^2 (1 + cb (\theta_{ijk} - \theta_{ijk}^0)) \quad [3b]$$

$$U_{\text{dihedral}} = \sum_{\text{dihedrals}} \frac{1}{2} V_n (1 + \cos(n\phi - \gamma)) \quad [4]$$

$$U_{\text{improper}} = 0.043844 K_\omega / 2 (\omega_{ijk} - \omega_{jkl}^0)^2 \quad [5]$$

$$U_{\text{stretch-bond}} = 2.51210 (K^{sb}_{ijk} (r_{ij} - r_{ij}^0) + K^{sb}_{kji} (r_{kj} - r_{kj}^0)) (\theta_{ijk} - \theta_{ijk}^0) \quad [6]$$

Nonbonded interactions are composed of van der Waals and electrostatic interactions. The electrostatic energy contribution is computed from the interactions between fixed partial charges in the case of nonpolarizable FFs as given by Eq. [7]. The weak dispersive or van der Waals interactions are typically represented with a 12-6 Lennard-Jones (LJ) potential, but 9-6 and buffered 14-7 LJ potentials can also be used. U_{LJ} in Eq. 8a represents the van der Waals energy contribution, where m and n values equal, for example, 12 and 6 or 9 and 6. ϵ_{ij} and σ represent the potential well depth and collision diameter, respectively. U_{vdW} (Eq. 8b) is another form of the van der Waals term where a buffered 14-7 potential, as used in MMFF, is

shown. The minimum interaction energy distance between atoms i and j is given by r_{ij} (or R_{ij}).

$$U_{\text{coul}} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad [7]$$

$$U_{\text{LJ}} = 4\epsilon_{ij} \left[\left(\frac{\sigma}{r_{ij}} \right)^m - \left(\frac{\sigma}{r_{ij}} \right)^n \right] \quad [8a]$$

$$U_{\text{vdW}} = \epsilon_{ij} \left(\frac{1.07R_{ij}^*}{R_{ij} + 0.07R_{ij}^*} \right)^7 \left(\frac{1.12R_{ij}^{*7}}{R_{ij}^7 + 0.12R_{ij}^{*7}} - 2 \right) \quad [8b]$$

In charged systems, such as DESs and ILs, tuning nonbonded parameters plays an important role in treating polarization implicitly to obtain better agreement with experimental data. The approaches used for implicit treatment for polarization include (1) scaling ionic charges and/or (2) adjusting LJ parameters. One of the earliest investigations to utilize scaled charges was performed by Morrow and Maginn, where overall cation and anion charges of $+0.904e$ and $-0.904e$ (as opposed to integer $\pm 1e$ values) were used for the 1-butyl-3-methylimidazolium hexafluorophosphate [bmim]⁺[PF₆]⁻ IL.⁸⁴ The charges on the fluorine atoms in the anion were found to be asymmetric, which implied polarization of the electron cloud. However, it was found that the anion did not show a preferential orientation close to the nearest cation. In addition, the computed properties did not show a significant difference when using symmetric charges on the anion. Following this work, there were two simulation studies of aqueous-IL interfaces where scaled charges showed better results in terms of interface formation and agreement with experimental values of ILs in humid conditions.^{85, 86} Refined potentials for [bmim]⁺[PF₆]⁻ IL by Bhargava and Balasubramanian⁸⁷ used charges of $\pm 0.8e$ and tuned LJ parameters to match the RDFs obtained from *ab initio* simulations using the Car–Parrinello method. In particular, the emphasis was to reproduce the cation–anion H-bonding behavior. Liu et al. calculated a range of properties including density, heat capacity, thermal conductivity, shear viscosity, and self-diffusion coefficients for six ILs using GAFF parameters with scaled charges ($\pm 0.8e$).⁸⁸ Although other properties showed good agreement, shear viscosities were overestimated by an order of magnitude and self-diffusion coefficients were underestimated by a factor of 2 which has been attributed to the lack of treating polarizability in an explicit way. Consequently, a systematic investigation of different charge

TABLE 4 Evaluation of Different Charge Schemes for ChCl–levulinic Acid DES Employing Isolated Molecule and Cluster Approach.

| Charge assignment method | Cluster | Isolated molecule |
|-----------------------------|---------|-------------------|
| Mulliken | Poor | Poor |
| NPA | Fair | Poor |
| Löwdin | Poor | Poor |
| Mayer | Poor | Poor |
| ChelpG | Good | Fair |
| Merz–Kollman | Good | Fair |
| Atomic Polar Tensor | Fair | Poor |
| Hirshfeld | Fair | Poor |
| Voronoi deformation density | Poor | Poor |
| AIM | Fair | Poor |

Source: García et al. 2015⁹⁴.

scaling factors was completed by Chaban, where an overall scaling factor between 0.7 and 0.8 worked best for imidazolium and pyridinium RTILs.⁸⁹ More recent studies have thoroughly investigated the use, advantages, and disadvantages of employing scaled charges in ILs.^{90–92}

Multiple molecular simulations of DESs have also utilized effective charge scaling to treat polarization implicitly. It is important to note that the choice of charge method, e.g., RESP, ChelpG, AIM, and Merz–Kollman, used during the development of FF partial charges can have a dramatic influence on the accuracy of IL and DES systems.^{93, 94} In addition, the development of atomic charges from either (1) clusters consisting of a 1 : 2 mole ratio, e.g., 1 ChCl and 2 urea in CCU, or (2) isolated molecules/ions can also have a profound effect on the predicted structural arrangement of the system.⁹⁴ Table 4 provides a comparison between different charge models and the use of small clusters versus individual molecules/ions. Overall, the ChelpG and Merz–Kollman charge models coupled to the minimal cluster optimization approach yielded the most accurate reproduction of DES properties. Similar to the IL simulations discussed previously, non-integer charges provided the best results, i.e., charges on Ch⁺ and Cl[−] obtained from ChelpG were +0.8254 e and −0.8392 e, while the Merz–Kollman charges were +0.6849 e and −0.7158 e, respectively.⁹⁴ In addition, Ullah et al. reported a FF for CCLev that possessed non-integer cation and anion ChelpG-derived charges of +0.8254 e and −0.6849 e.³⁷

Tuning of the van der Waals terms for better agreement in structural, thermodynamic, and transport properties in MD simulations has not been

employed as extensively in DES simulations as compared to ILs.⁹⁵ Notably, Doherty and Acevedo adjusted LJ parameters and empirical charges during the development of an OPLS-AA force field for ChCl-based DES systems³⁵ to match RDFs reported by Hammond et al.⁴⁶ A potential drawback of this procedure is the transferability of these parameters to simulations featuring complex systems, e.g., additional species or heterogeneous environments. In the case of DES mixtures, even if one of the cationic/anionic species are common to both components, the chemical environment around the constituent atoms can lead to different charge descriptions.⁹⁶ This emphasizes that in the event of mixing DES systems, or simulating DESs with organic solvents or polymers, scaling of charges and FF parameters in general should be validated prior to the prediction of properties of interest.

Simulations featuring mixtures of DESs and conventional solvents have been reported but are limited in number. For example, Tanner et al. studied the effect of water addition to ILs composed of choline and geranate at varying mole fractions using a combined experimental, atomistic MD, and coarse-grained MD approach.⁹⁷ Fetisov et al.⁴⁵ also performed a FPMD study between CCU and water to study the resultant molecular structure and transport properties for the mixture. Table 5 summarizes the nonpolarizable FFs, with implicit treatment of polarization, used to simulate DES systems through charge scaling and/or LJ tuning and is discussed in greater detail in section “Nonpolarizable Force Fields Used for DES Simulations”.

Nonpolarizable Force Fields used for DES Simulations

A variety of nonpolarizable FFs have been used to study DESs. For example, Perkins et al.⁵⁵ performed molecular simulations on CCU with modified GAFF parameters at several temperatures. Good agreement with experimental densities (1% error) and heat capacities (1.3–1.4% error) was found with a reduced charge model ($\pm 0.8 e$). However, transport properties such as self-diffusion coefficients were underestimated by 25–51% and 29–41% for Ch^+ and urea at 298 K, respectively. Improvement was seen at the higher temperature of 330 K, which reduced the errors to 4–17% and 3–8% for the same species in comparison with experimental values. Subsequent work by the same group³¹ on CCEtg, CCGly, and CCMal also showed good agreement for physical and thermodynamic properties. Once again, the self-diffusion coefficients were difficult to reproduce for CCEtg where values were underestimated by 20–30% at 298 K and 5–25% for simulations at 330 K. For CCGly, the values were underestimated by 14–20% and 17–27% at 298 and 330 K, respectively.

TABLE 5 Comparison of DES Bulk Properties from Atomistic MD Simulations using Nonpolarizable Force Fields.

| Author/force field | Liquid density | Thermal expansion coefficient | Surface tension | Heat capacity | Viscosity | Self-diffusion coefficients | Liquid structure ^a | |
|-----------------------------------|-------------------------|-------------------------------|-----------------|---------------|-----------|-----------------------------|-------------------------------|----|
| CCU | | | | | | | | |
| Perkins et al. ⁵⁵ | GAFF (0.8) ^b | ++ ^c | ++ | na | ++ | na | — ^d | + |
| Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | ++ | na | ++ | + | ++ | — | ++ |
| Shah and Mjalli ⁷⁸ | MMFF | ++ | ++ | na | na | na | — | + |
| García et al. ²⁰ | MDynaMix | ++ | na | Na | na | — | na | + |
| CCEtg | | | | | | | | |
| Perkins et al. ³¹ | GAFF (0.8) | ++ | + | na | + | na | — | + |
| Zhang et al. ³⁶ | GAFF(0.9) | ++ | na | na | na | + | — | ++ |
| Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | ++ | na | ++ | ++ | ++ | — | ++ |
| Ferreira et al. ²⁰ | Mixed ^e | ++ | + | + | na | + | — | + |
| CCGly | | | | | | | | |
| Perkins et al. ³¹ | GAFF (0.8) | ++ | ++ | na | + | na | — | + |
| Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | ++ | na | ++ | + | ++ | — | ++ |
| Mainberger et al. ³² | MMFF | ++ | na | na | ++ | na | — | + |
| García et al. ²⁰ | MDynaMix | ++ | na | na | na | — | na | ++ |
| CCMal | | | | | | | | |
| Perkins et al. ³¹ | GAFF (0.8) | ++ | + | na | na | na | na | + |
| Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | ++ | na | ++ | na | ++ | na | + |
| García et al. ²⁰ | MDynaMix | ++ | na | na | na | — | na | ++ |

TABLE 5 (Continued)

| Author/force field | | Liquid density | Thermal expansion coefficient | Surface tension | Heat capacity | Viscosity | Self-diffusion coefficients | Liquid structure ^d |
|-----------------------------------|---------------|----------------|-------------------------------|-----------------|---------------|-----------|-----------------------------|-------------------------------|
| CCLev | | | | | | | | |
| Ullah et al. ³⁷ | MDynaMix | ++ | ++ | na | na | — | na | ++ |
| Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | ++ | na | na | na | ++ | na | + |
| Mainberger et al. ³² | MMFF | ++ | na | na | na | na | na | + |
| Mainberger et al. ³² | GAFF (0.75) | ++ | na | na | na | na | na | + |
| CCPro | | | | | | | | |
| Ferreira et al. ⁹⁹ | OPLS-AA | ++ | na | na | na | — | — | + |

^aLiquid structure was compared considering the breadth of properties calculated (RDFs, CDFs, SDFs, H-bond analysis, and intermolecular energies) and comparison with *ab initio* simulations or neutron diffraction data.

^bValues in parentheses 0.75, 0.8, and 0.9 indicate charge scaling factor in the given FF.

^c+ and ++ denote reasonable and good agreement with experimental data (deviations <10% and <5%, respectively).

^d— and — denote deviations >10% and >20%, respectively.

^eMixed – GAFF, OPLS-AA, and CHARMM27 were used for Ch⁺, Cl⁻, and ethylene glycol, respectively.

Ferreira et al.⁹⁸ tested various FF combinations for each component in CCEtg including OPLS-AA,^{75,77} GAFF,⁷¹ and CHARMM27^{100,101} with varying charge schemes. Although they obtained good agreement for densities and thermal expansion coefficients using the unscaled charge scheme ($\pm 1e$), self-diffusion coefficients of Ch^+ and ethylene glycol were found to be underestimated by a factor of 8 when compared against experimental values. After scaling the charges by a factor of 0.8, several properties including density, thermal expansion coefficient, enthalpy of vaporization, surface tension, shear viscosity, and structural properties were calculated over a temperature range of 298.15–373.15 K and compared to relevant experimental data. It was reported in this work that the self-diffusion coefficients displayed an improvement of 10% with the refined charges in comparison to the Perkins et al. simulations.^{31,55} Ferreira et al.⁹⁹ applied the same procedure to derive a system-specific FF for CCPro. The authors found success in reproducing experimental densities while combining existing parameters for choline,⁷⁵ chloride,⁷⁶ and propylene glycol.¹⁰² Further refinement was achieved through AIMD simulations of the CCPro system and the restrained electrostatic charge potential (RESP) was utilized to generate new averaged charges on each species. With the newly produced charges, transport properties such as viscosity and diffusion coefficients were improved compared to experiment, and however, the errors associated with each property were still large at 19% and 16%, respectively.

Using the empirical potential structure refinement (EPSR) method that validates the sampling space to neutron diffraction data, Hammond et al.⁴⁶ generated center-of-mass and partial atomic RDFs for CCU along with spatial distribution functions (SDFs) to characterize the solvation environment. To run the refinement simulations, harmonic potentials were used to maintain the geometry for each molecule of interest. A reference potential was either obtained from literature or generated to explore the desired configurational space. The parameters for urea molecules were derived from Soper et al.¹⁰³ and OPLS-AA parameters were used for Ch^+ and Cl^- . The study found a complex H-bonding network consisting of strong HBD (urea)–Cl and Ch–Cl interactions, consistent with other experimental and simulation investigations. Besides being the first work to examine the liquid structure of CCU using neutron diffraction, this work also provided a framework for refining existing atomistic FFs to simulate Ch–Cl DES systems.³⁵

Mainberger et al.³² tested two FFs, GAFF and MMFF, to simulate three ChCl-based DESs containing the HBDs of glycerol, 1,4-butanediol, and levulinic acid. An additional DES containing the zwitterion betaine was also investigated. Two sets of charges derived by the RESP methodology were used with GAFF to investigate the effect of charge scaling. Scaling atomic

charges by 0.75 improved estimation of densities and heat capacities in comparison with experimental data for all systems. However, for the CCGly system where experimental self-diffusion coefficients are available, the predicted values were overestimated by an order of magnitude. The ± 1 charge system underestimated the diffusion coefficients by 16% and 7% for glycerol and Ch^+ , respectively. Simulations with the MMFF did not utilize a charge scaling scheme, but still showed good agreement for densities and heat capacities (within $\sim 3\%$ of experimental values). However, the calculated self-diffusion coefficients for Ch^+ and glycerol were underestimated by 43% and 92%, respectively.

Doherty and Acevedo³⁵ recently developed a set of custom OPLS-AA parameters (called OPLS-DES) to simulate the structural, thermodynamic, and transport properties of ChCl -based DESs. The nonbonded parameters for Ch^+ and Cl^- were adjusted to match the RDFs obtained from the work of Hammond et al.,⁴⁶ and Zahn et al.^{34, 46} Torsional parameters for the Ch cation were originally adjusted in a previous IL study to fit conformational energy minima from LMP2/cc-pVTZ(-f) calculations.⁷⁵ For the HBDs, parameters were taken from OPLS-AA⁷⁷ and nonbonded terms were adjusted to match liquid structure and bulk properties from experimental data and AIMD based calculations.^{34, 46, 69} Physical and thermodynamic properties such as density, shear viscosity, heat capacities, and surface tension showed excellent agreement with experimental data while self-diffusion coefficients showed higher error percentages. The overall mean absolute errors (MAEs) obtained in this work were 1.1%, 1.6%, 5.5%, and 1.5% for the above-mentioned properties, respectively. Self-diffusion coefficients estimated in this work showed error percentages of 31.4% and 78.8% for Ch^+ and urea at 298 K. The percent error values decrease with temperature (28.1% and 0.3%, respectively, at 328 K) similar to the behavior observed in other simulation investigations.^{31, 55, 98} In recent simulations performed by Salehi et al.,¹⁰⁴ the OPLS-AA parameters from Doherty and Acevedo³⁵ and GAFF parameters from Perkins et al.^{31, 55} were used to calculate Hildebrand and Hansen solubility parameters for CCU, CCGly, CCMal, and CCOx. A charge scaling of 0.8 was applied to the simulated systems for both FFs tested. Additional modifications included the removal of intramolecular exclusion terms between hydrogen and oxygen atoms in OH groups of the HBDs and the inclusion of LJ parameters ($\sigma = 0.1 \text{ \AA}$ and $\epsilon = 0.001 \text{ kcal/mol}$) for hydrogen atoms to avoid overlaps. Solubility parameters and enthalpies of vaporization were computed taking into consideration HBD, HBA, and cluster (comprising both HBD and HBA) vaporization. Based on the calculated vaporization enthalpy contributions, the HBD is suggested to vaporize first and the

large values similar to ILs emphasize the polar nature of DESs. As a brief comparison between the GAFF and OPLS-AA parameter sets, in CCU, GAFF overestimated the experimental enthalpies of vaporization by 25 kJ/mol, whereas OPLS-AA showed close agreement.

Beyond generalized FFs that have been modified to simulate DESs, e.g., GAFF and OPLS-AA, additional work has been reported for FFs parameterized specifically to simulate DES systems.^{20, 37, 94} For example, parameterization efforts on CCU, CCGly, and CCMal by García et al.²⁰ developed partial charges by using a minimal cluster approach, where ChCl:HBD clusters (in the ratio 1 : 1, 1 : 2, and 1 : 2 for malonic acid, ethylene glycol, and urea, respectively) were optimized at the B3LYP/6-311+G(d) theory level. Their DFT calculations showed different charge distributions for two urea/ethylene glycol molecules depending on the positions through which the two atoms interact with ChCl. Thus, the two urea/ethylene glycol molecules were assigned different charges. The parameterization effort reported liquid density deviations of 1.68%, 0.29%, and 1.56% for CCU, CCGly, and CCMal, respectively; however, viscosities obtained from these simulations showed large deviations. A similar parameterization procedure was followed by Ullah et al.³⁷ for CCLev where two types of levulinic acid molecules were developed with different ChelpG-based charges. This work showed good agreement with experimental data for density and thermal expansion coefficients (less than 0.8% and 2.8%, respectively) but gave poor shear viscosity reproduction with a deviation of 16.8% compared to the experiment. Experimental data were not available to compare self-diffusion coefficients for CCLev; however, their predictions were comparable to GAFF simulations with no charge scaling.³²

Table 5 provides a summary of bulk properties and liquid structures predicted using the previously discussed nonpolarizable FFs developed for DES systems. Liquid density and thermal expansion coefficient predictions of multiple DES systems showed good agreement with experimental data for all the FFs mentioned in Table 5. In particular, liquid density shows less than 3% deviation from experiments with the exception of CCU (OPLS-AA)³⁵ with 4.0% at 298 K. This has been attributed to the scaling of LJ parameters to reproduce other properties such as heat capacity, surface tension, and molecular interactions. Viscosities calculated by García et al.²⁰ and Ullah et al.³⁷ using the Green–Kubo method showed higher error percentages (ranging from 16.8% to 35.4%) as compared to the nonequilibrium periodic perturbation method,^{35, 98} where error percentages are less than 3%. Surface tension for most systems have only been calculated by OPLS-AA with <2% deviations from experimental data.³⁵ For CCEtg, OPLS-AA parametrized

by Doherty and Acevedo³⁵ displayed better performance than surface tensions computed using a combination of multiple FF parameters, i.e., GAFF (Ch⁺) + OPLS-AA(Cl⁻) + CHARMM27(ethylene glycol).⁹⁸ All nonpolarizable FFs showed large deviations in self-diffusion coefficients at 298 K compared to experimental data (typically >15%, but as high as >50% in several cases). During the development of the nonpolarizable DES FFs, the predicted molecular structures and intermolecular interactions of the solvents were often compared to relevant QM-based calculations and experimental data. For example, the local interactions in CCU predicted by OPLS-AA-based simulations³⁵ were extensively characterized by examining combined distribution functions (CDFs), RDFs, and SDFs, and comparing them in detail to reported AIMD simulations.³⁴ Additional liquid structure properties, such as H-bond residence times, were calculated by García and coworkers.^{20, 37} Detailed comparisons of physical, thermodynamic, and transport properties, and molecular structure and interactions are provided in the subsequent sections “Physical Properties, Thermodynamic Properties, Transport Properties, and Deep Eutectic Solvent Structure”.

From Table 5, it is clear that accurately reproducing transport properties, such as self-diffusion coefficients, was a challenge for all parameterization efforts. The use of polarizable force fields has been advocated for improving agreement with diffusivity experimental data.³⁵ However, to our knowledge, polarizable FFs have not been developed/utilized to simulate bulk properties of DES systems.

PHYSICAL PROPERTIES

Liquid Density

Liquid density is an important physical property that has served as a starting point for the validation of DES FFs at a wide range of temperatures.^{31, 32, 35, 55} It should be noted that density by itself is not sufficient for validating the accuracy of a FF as multiple combinations of parameters can give similar densities within error bars while showing remarkable differences in other predicted thermodynamic, structural, and transport properties. Problematically, experimental solvent characterization data, such as cohesive energy and enthalpy of vaporization, are not readily available for DES systems.¹⁰⁴ Hence, liquid densities have been used for preliminary validation in the literature. Densities obtained from MD simulations have been compared against experimental data^{32, 44, 105–115} for (a) CCU, (b) CCEtg, (c) CCGly, (d) CCMal, and (e) CCLev in Figure 7.

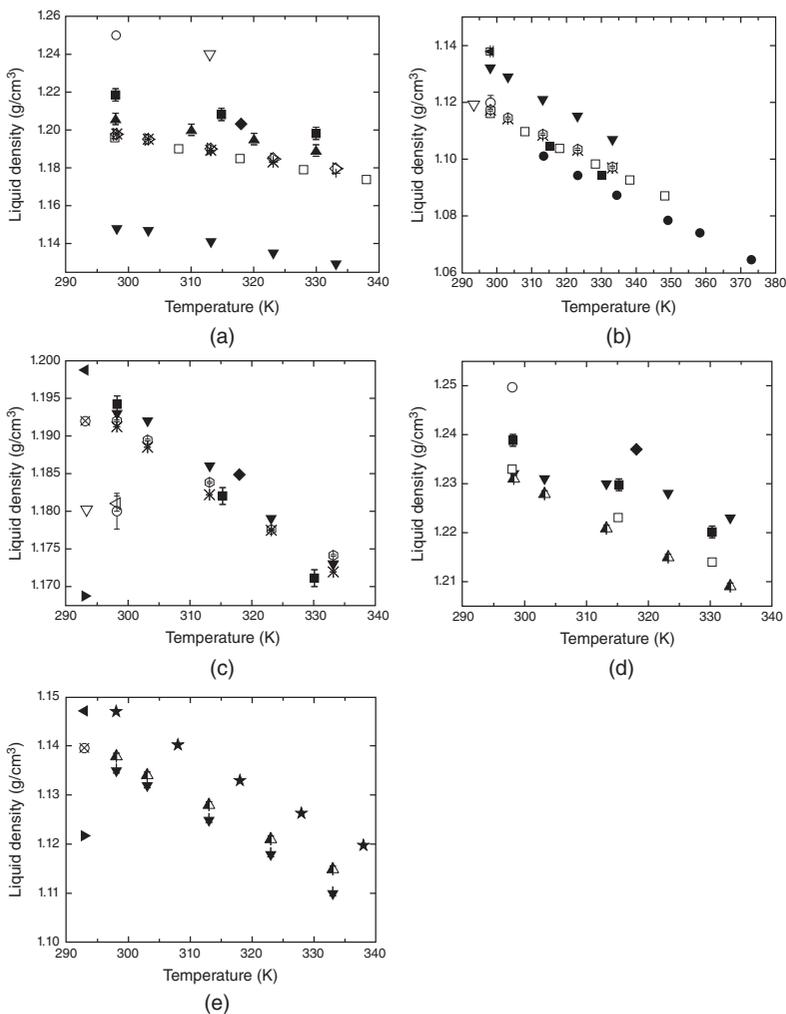


FIGURE 7 Comparison of liquid density data obtained from simulations (solid symbols) and experimental data (half-open symbols) for (a) CCU, (b) CCEtG, (c) CCGly, (d) CCMal, and (e) CCLev. Simulations: ■ Perkins, Painter, and Colina^{31,55}; ▼ Doherty and Acevedo³⁵; ▲ Shah and Mjalli⁷⁸; ● Ferreira et al.⁹⁸; ◆ García, Atilhan, and Aparicio²⁰; ► Mainberger et al. (MMFF)³²; ◀ Mainberger et al. (GAFF)³²; ★ Ullah et al.³⁷. Experimental: □ Ciocirlan et al.¹¹⁵; ○ D'Agostino et al.⁴⁴ + Yadav et al.^{111,116}; × Leron and Li¹⁰⁶; △ Shekaari, Zafarani-Moattar, and Mohammadi¹¹²; ▷ Xie et al.¹⁰⁸; ▽ Abbott et al.¹¹⁷; ◇ Chemat et al.¹⁰⁹; ◊ Shahbaz et al.¹⁰⁵; ◊ Mjalli et al.¹¹⁴; ◊ Mjalli and Abdel Jabbar¹¹³; ▲ Leron, Wong, and Li¹⁰⁷; ⊗ Mainberger et al.³²; ▲ Florindo et al.¹¹⁰

A major caveat is that large deviations are prevalent between the reported experimental densities themselves. These differences have been attributed to several factors that include water absorption by samples and the method of measurement.^{41, 55, 118}

For CCU (Figure 7a), MD simulations by Shah and Mjalli that utilized MMFF showed the closest agreement with experimental densities across a temperature range of 290–330 K.⁷⁸ Reasonably close agreement ($\leq 1\%$) was also found when using GAFF.⁵⁵ However, the OPLS-AA underestimated liquid densities with errors ranging from 3.9–4.2%; this was a consequence of scaling Lennard-Jones parameters to obtain better agreement with other properties that included surface tension, heat capacity, and molecular structure at the expense of increased deviations in density.³⁵ For CCEtg (Figure 7b), GAFF, OPLS-AA, and a mixture of parameters (GAFF(choline) + OPLS-AA(chloride) + CHARMM27(ethylene glycol))⁹⁸ all exhibited good agreement with less than 3% deviation from experimental densities. As a general comparison, GAFF and the mixed parameters underestimated the experimental densities, whereas the OPLS-AA overestimated the values. However, the maximum absolute error for all FF predictions of CCEtg were usually $< 0.01 \text{ g/cm}^3$. Density simulations of CCGly (Figure 7c) reported deviations of less than 1.1% with experiment when utilizing either the GAFF (Perkins, Painter, and Colina³¹) or OPLS-AA³⁵ FFs. However, parameters from Mainberger et al.³² overestimated and underestimated experimental liquid densities at 298 K using GAFF and MMFF, respectively. For CCMal (Figure 7d), both GAFF and OPLS-AA overestimated the densities, except at 298 K where OPLS-AA showed near perfect agreement (0.1% deviation³⁵). Finally, for the CCLev DES (Figure 7e), OPLS-AA³⁵ showed the closest agreement with experimental data (0.2–0.4% error) as compared to GAFF,³² MMFF,³² and MDynaMix.³⁷

An important consideration when developing DES FF parameters is the role of charge scaling to improve agreement in densities. The results presented in Figure 7a–d for the model used by Perkins et al. employed a charge scaling factor of $\pm 0.8 e$ with default GAFF parameters.^{31, 55} This was found to reproduce density better than unscaled charges (i.e., +1 for cations and -1 for anions) or a scaling of $\pm 0.9 e$. Similarly, scaling by $\pm 0.75 e$ was found to yield better results for CCGly and CCLev modeled using GAFF Mainberger et al.³² when compared to unscaled charges. Simulations with MMFF^{32, 78} employed full charges (Figure 7a,c,e) and OPLS-AA³⁵ simulations used a scaling factor of $\pm 0.8 e$. As an alternative, FFs developed by García et al.²⁰ and Ullah et al.³⁷ featured charges calculated from small DES clusters. For example, instead of $\pm 1 e$ for the cation/anion and a neutral

charge for the HBD molecules, charges of +0.8254 e, -0.6849 e, -0.0663 e, and -0.0743 e were used for choline, chloride, and the two levulinic acid molecules, respectively.³⁷ Ferreira et al.⁹⁹ also derived charges from bulk AIMD simulations for the CCPro system that resulted in a scaling of 0.74 and found good agreement (3% error) with a very specific combination of existing FFs.^{75, 76, 102} Finally, Mainberger et al.³² included LJ parameters for hydrogen atoms, but their addition did not show a remarkable improvement in liquid density compared to the other FFs discussed. It is worth mentioning that with respect to DESs, due to a lack of reported experimental densities for many systems, research groups have used alternative methods to predict densities that include empirical group contribution methods and neural network models.^{105, 119}

Volume Expansivity

Volume expansivity is calculated from the slope of the molar volume versus temperature curve as described in Eq. [9]. Table 6 provides a comparison between computed DES volume expansivity values from multiple FFs and experimental measurements. Perkins, Painter, and Colina^{31, 55} computed the volume expansivity for four different DESs using GAFF parameters with $\pm 0.8e$ scaled charges. Their simulations yielded close agreement with most experimental measurements,^{78, 116} with the notable exception of CCGly, which was overestimated. The MMFF⁷⁸ and the MDynaMix³⁷ FFs also yielded excellent volume expansivity predictions. Finally, the use of mixed FF parameters by Ferreira et al.⁹⁸ for CCEtg provided values at the lower and upper bounds of experimental measurements.

$$\alpha_p = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right) \quad [9]$$

Surface Tension

The reproduction of DES surface tension is important for several industrially relevant applications involving separation, distillation, and extraction. Unfortunately, the availability of experimental DES surface tension data is limited due to measurement difficulties; consequently, predictive models have often been employed as an alternative.^{119, 120} Macleod proposed a relation to estimate the surface tension as given by Eq. 10^{121, 122} where ρ_L and ρ_V represent liquid and vapor density, σ_M is the surface tension, and K is a constant. This relation was modified (Eq. [11a])^{121, 123} where the constant K in Macleod's equation was replaced by a new constant known as

TABLE 6 Comparison of Average Volume Expansivity $\alpha_p \times 10^4$ (K⁻¹) for Deep Eutectic Solvents.

| DES | Authors/force field | Simulations | Experiments ^{78, 116} |
|-------|---------------------------------|-------------|--------------------------------|
| CCU | Perkins et al. ⁵⁵ | GAFF (0.8) | 4.378 to 6.0 |
| | Shah and Mijjalli ⁷⁸ | MMFF | |
| CCEtg | Perkins et al. ³¹ | GAFF (0.8) | 5 to 7 |
| | Ferreira et al. ⁹⁸ | Mixed | |
| CCGly | Perkins et al. ³¹ | GAFF (0.8) | 4.6 to 4.7 |
| | | MMFF | |
| CCMal | Perkins et al. ³¹ | GAFF (0.8) | 5 to 6 |
| CCLev | Ullah et al. ³⁷ | MDynaMix | 5.83 to 5.93 |

the parachor (P) expressed in terms of molar quantities. In cases where the vapor density is negligible in comparison with the liquid density, P is given by Eq. [11b] where M_w is the molecular weight and ρ is the liquid density. P can be expressed as a sum of the contributions from its constituents, comprising atoms or groups in the molecule, as shown in Eq. [11b], where P_i denotes individual contributions.¹²⁴ σ_s can be calculated by using the known density of a given compound. The parachor contribution values were later improved for neutral compounds.^{124, 125} Knotts et al. developed a quantitative structure–property relationship for P from data for neutral organic compounds available in the DIPPR database.¹²⁶ This model was then extended to charged systems such as ionic liquids.^{127, 128} The percent error obtained for CCGly and CCEtg using the parachor method at 298 K were 3.74% and 5.91%, respectively.¹²¹

$$\sigma_M = K(\rho_L - \rho_V)^4 \quad [10]$$

$$\sigma_S = [P(\rho_L - \rho_V)]^4 \quad [11a]$$

$$P = \frac{M_w \sigma_s^{1/4}}{\rho} = \sum_i P_i \quad [11b]$$

An alternative approach for estimating surface tension is to employ the Othmer equation (12),¹¹⁹ where the surface tension at a given temperature T can be calculated using the critical temperature (T_C) and a reference surface tension (σ_{ref}) at another temperature (T_{ref}). Deviations between values predicted from the Othmer relation and experimental measurements worsen with increasing temperature.¹¹⁹ Nevertheless, the overall percent error for

nine DES systems was 2.57%, highlighting the empirical Othmer model as a good approach for predicting surface tensions.¹²¹

$$\sigma(T) = \sigma_{\text{ref}} \left[\frac{(T_c - T)}{T_c - T_{\text{ref}}} \right]^{\frac{11}{9}} \quad [12]$$

Employing MD simulations to calculate surface tensions of DESs. Only two simulation studies to the date this chapter was written have been reported. Doherty and Acevedo (OPLS-AA)³⁵ calculated the surface tension for four DES systems: CCU, CCEtg, CCGly, and CCMal. Ferreira et al. (mixed CHARMM27+OPLS-AA+GAFF)⁹⁸ calculated the surface tension for CCEtg. The z -axis of the simulation boxes was elongated by a factor of 3 and 2 by Doherty and Acevedo, and Ferreira et al., respectively. The surface tension was computed from the directional components of the pressure tensor as given by Eq. [13] where σ_{MD} and L_z represent the surface tension and length of the box along the z direction, respectively. P_{ZZ} , P_{XX} , and P_{YY} represent the directional components of the pressure tensor.

$$\sigma_{\text{MD}} = \frac{1}{2} L_z \left[P_{ZZ} - \frac{1}{2} (P_{XX} + P_{YY}) \right] \quad [13]$$

Table 7 provides a comparison of surface tension values for DESs computed using MD simulations and the analogous experimental data/empirical models. Notably, the surface tension computed from OPLS-AA simulations were calculated at 425 K.³⁵ Due to the absence of experimental data at this temperature, MD predictions were compared to surface tensions calculated using the Othmer equation at 425 K utilizing experimental data at 298 K as the reference point.¹¹⁹ The OPLS-AA FF provided excellent performance yielding small deviations ranging from 0.5% to 2.0%. In the work of Ferreira et al.,⁹⁸ surface tensions were directly compared with experimental data available at similar temperatures (298–323 K) and yielded larger deviations of 3.5–5.7%.

THERMODYNAMIC PROPERTIES

Heat Capacity

Heat capacity at constant pressure (C_p) is defined by Eq. [14a], where H represents the enthalpy (Eq. [14b]), U_{inter} and U_{intra} are the intermolecular and intramolecular potential energies, respectively, and KE denotes the kinetic energy. The enthalpy is often expressed as a sum of ideal (H^{id}) and residual

TABLE 7 Comparison of Surface Tension (mN/m) for Deep Eutectic Solvents.

| DES | Authors/force field | Simulation | Experiment ^{98, 119} |
|-----------------------|-------------------------------|------------|-------------------------------|
| CCEtg (298 K) | Ferreira et al. ⁹⁸ | Mixed | 48.91 ± 0.1 |
| CCEtg (313 K) | | | 47.50 ± 0.1 |
| CCEtg (323 K) | | | 46.67 ± 0.1 |
| CCU (425 K) | Doherty and | | |
| Acevedo ³⁵ | OPLS-AA (0.8) | 38.9 | |
| CCEtg (425 K) | | 35.9 | 35.4 |
| CCGly (425 K) | | 43.2 | 44.1 |
| CCMal (425 K) | | 51.3 | 52.3 |

(H^{res}) contributions, as described by Lagache et al.¹²⁹ and Cadena et al.¹³⁰ (Eqs. [15a]–[15c]). Therefore, C_p is written as the sum of ideal and residual contributions (Eqs. [16a]–[16c]).

$$C_p(T, P) = \left(\frac{\partial \langle H \rangle}{\partial T} \right)_P \quad [14a]$$

$$H = U^{\text{inter}} + U^{\text{intra}} + \text{KE} + \text{PV} \quad [14b]$$

$$\langle H \rangle = \langle H^{\text{id}} \rangle + \langle H^{\text{res}} \rangle \quad [15a]$$

$$H^{\text{id}} = U^{\text{intra}} + \text{KE} + Nk_b T \quad [15b]$$

$$H^{\text{res}} = U^{\text{inter}} + \text{PV} - Nk_b T \quad [15c]$$

$$C_p(T, P) = C_p^{\text{id}}(T) + C_p^{\text{res}}(T, P) \quad [16a]$$

$$C_p^{\text{id}}(T) = \left(\frac{\partial \langle H^{\text{id}} \rangle}{\partial T} \right)_P \quad [16b]$$

$$C_p^{\text{res}}(T, P) = \left(\frac{\partial \langle H^{\text{res}} \rangle}{\partial T} \right) \quad [16c]$$

The residual, C_p^{res} , and ideal, C_p^{id} , contributions to the heat capacity are typically obtained from MD simulations and experiments, respectively.^{129, 131} For example, DES C_p^{res} values are typically derived from the slope of a plot featuring H^{res} (Eq. [15c]) at multiple temperatures.^{31, 55} As mentioned, C_p^{id} should be measured experimentally, but when unavailable, *ab initio* calculation-derived values may be substituted.^{31, 55, 88, 130, 131} For example, Perkins, Painter, and Colina^{31, 55} carried out gas-phase DFT calculations to optimize isolated ChCl and HBD moieties and then

performed a vibrational frequency analysis to obtain C_p^{id} . It is important to emphasize that the C_p^{id} values obtained in this manner were significantly overestimated compared to experimental values for multiple systems, including water and DESs.^{130, 132, 133} This can be attributed to the fact that classical FFs use simple harmonic approximations, which overestimate the vibrational energy of molecules.^{131, 133, 134} The need for further refinement of FF parameters to accurately reproduce heat capacities has been acknowledged.^{55, 130, 131} A simple correction is to apply a scaling factor to the computed QM vibrational frequencies that is consistent with the level of theory employed.^{88, 130, 131}

An alternative approach for computing heat capacity is a two-phase model proposed by Lin, Blanco, and Goddard III¹³⁴ where the vibrational density of states (DoS) is calculated to account for QM corrections to the thermodynamic properties of liquids. DoS represents the distribution of the vibrational normal modes of a system, expressed as a function of the normal-mode frequency (ν). The DoS, denoted by $S(\nu)$, can be obtained from the Fourier transform of the mass weighted sum of the atomic velocity autocorrelation functions. The distributions obtained are then normalized to the total number of degrees of freedom in the given system; thermodynamic properties such as heat capacities can then be calculated by assuming each normal mode to be a quantum mechanical oscillator with a frequency ν . This method was employed to compute heat capacities for several IL and DES systems.^{35, 91, 92, 133} The corrected heat capacity is described in Eq. 17a, where C_p^{corr} and C_p^{class} represent the corrected and classical heat capacities at constant pressure. δC_v^{QM} denotes the QM-corrected heat capacity at constant volume and is given by Eq. 17b, where W is a weighting function. In DES simulations by Doherty and Acevedo,³⁵ an additional term $N_c k_b$ was added to the C_p^{class} term to account for neglected contributions due to bond constraints present in the simulations.

$$C_p^{\text{corr}} = C_p^{\text{class}} + \delta C_v^{\text{QM}} \quad [17a]$$

$$\delta C_v^{\text{QM}} = k_b \int_0^\infty (W(\nu) - 1) S(\nu) d\nu \quad [17b]$$

A comparison of heat capacities computed from DES molecular simulations has been provided in Table 8. However, given the multiple methods employed and the differences in temperature reported in the simulations, direct comparison between the force field parameters is not straightforward. Perkins, Painter, and Colina,^{31, 55} reported heat capacities derived from GAFF simulations to be within the range of experimental data

TABLE 8 Comparison of C_p (J/mol K) for Deep Eutectic Solvents.

| DES | Authors/force field | | Temperature (K) | Simulation | Experiment ¹⁰⁶ |
|-------|-----------------------------------|-------------------------|-----------------|---------------|----------------------------|
| CCU | Perkins et al. ⁵⁵ | GAFF (0.8) ^a | 298–330 | 184 | 181.4 ± 0.5 to 186.4 ± 0.5 |
| | Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | 353 | 201.9 | 190.8 ± 0.8 |
| CCEg | Perkins et al. ³¹ | GAFF (0.8) | 298–330 | 209.27 ± 1.55 | 190.8 ± 0.4 to 199.2 ± 0.3 |
| | Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | 353 | 215.8 | 205.6 ± 0.2 |
| CCGly | Perkins et al. ³¹ | GAFF (0.8) | 298–330 | 259.15 ± 2.87 | 237.7 ± 0.6 to 246.9 ± 0.1 |
| | Mainberger et al. ³² | GAFF (0.75) | 303 | 244.0 ± 2.9 | 237.7 ± 0.6 |
| | Mainberger et al. ³² | MMFF | 303 | 246.3 ± 8.6 | 237.7 ± 0.6 |
| | Doherty and Acevedo ³⁵ | OPLS-AA (0.8) | 353 | 240.1 | 254.3 ± 0.4 |

^aValues in parentheses (0.75 or 0.8) indicate charge scaling factor in the given force field.

for CCU; however, 7% errors were computed for the CCEtg and CCGly DESs. Simulations by Mainberger et al.³² using their custom GAFF and MMFF parameters found relatively lower percent errors of 2.6% and 3.6%, respectively, for CCGly. Finally, OPLS-AA simulations by Doherty and Acevedo³⁵ gave error percentages of 5.8%, 5.6%, and 4.9% for CCU, CCEtg, and CCGly, respectively. An alternative to FF predictions was developed by Taherzadeh et al.¹³⁵ using a correlation model trained on 505 C_p values from 28 DES that yielded an absolute average relative deviation of 4.7% for all investigated data points.

Heats of Vaporization

Heats of vaporization (ΔH_{vap}) are calculated according to Eq. 18, where ΔE_{vap} is the difference between the total energies of the gas and liquid phase, T is the temperature, and R is the universal gas constant.

$$\Delta H_{\text{vap}} = \Delta E_{\text{vap}} + RT \quad [18]$$

With respect to DES simulations, calculating ΔH_{vap} is very challenging as the vapor phase composition is experimentally unknown. A thorough investigation by Salehi et al.¹⁰⁴ computed the ΔH_{vap} for CCU, CCEtg, CCGly, CCMal, and CCOx using OPLS-AA parameters by Doherty and Acevedo³⁵ and for CCU using GAFF parameters from Perkins, Painter, and Colina³¹ by utilizing three different vaporization clusters: HBD, HBA, and a cluster from the DES mixture. For the CCU solvent, the ΔH_{vap} derived from the vaporization of urea, ChCl, and a ChCl-urea cluster using OPLS-AA was 82, 165, and 228 kJ/mol, respectively. Comparison to the experimentally estimated ΔH_{vap} values of 46.9 and 79.0 kJ/mol obtained from vapor pressure data of Ravula et al. and Shabaz et al.,^{136, 137} suggests that it is more likely for HBD molecules to vaporize from the DES mixture and dominate the vapor phase. Similar results and agreement with experiment were computed by Salehi et al. for the CCEtg and CCGly DESs using OPLS-AA.¹⁰⁴ However, the GAFF parameters yielded less accurate ΔH_{vap} values for CCU. In separate work, Ferreira et al. utilized their mixed FF parameter set to compute ΔH_{vap} for CCEtg⁹⁸ and CCPro⁹⁹ and found the values ranged from 167.5 ± 0.3 to 179.5 ± 0.5 kJ/mol and 161–210 kJ/mol, respectively. Finally, Ullah et al.³⁷ calculated the ΔH_{vap} value for CCLev using their custom MDynaMix parameters and reported an energy of 52.05 kJ/mol.

Isothermal Compressibility

Isothermal compressibility (κ_T) can be expressed as a change in molar volume (V) with pressure (P) at a given temperature, or, alternatively, in terms of fluid density (ρ) as given in Eq. 19.

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right)_T \quad [19]$$

To compute κ_T using MD simulations, the relationship between V (or ρ) and P could be obtained by fitting an equation of molar volume at varying pressures (at constant temperature). However, a simpler approach was developed by Motakabbir and Berkowitz (Eq. 20),¹³⁸ where a linear isotherm was assumed in the pressure range used to calculate κ_T . Here, ρ_1 and ρ_2 represent the densities at pressure P_1 and P_2 , respectively. The simulations can be divided into smaller blocks in order to obtain multiple values of κ_T at the desired temperature.

$$\kappa_T = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right)_T \approx \frac{\Delta \ln(\rho)}{\Delta P} = \frac{\ln \left(\frac{\rho_2}{\rho_1} \right)}{P_2 - P_1} \quad [20]$$

It is important to emphasize that a linear approximation may not be appropriate for different classes of liquids or within specific pressure/temperature ranges and may lead to inaccurate estimation of κ_T . Alternatively, the Tait equation¹³⁹ can be substituted to fit molar volumes/densities obtained from MD simulations to provide a correlation to pressure.^{140, 141} In this method, κ_T is calculated using the fluctuations formula¹⁴² as given by Eq. 21, where FFs $\langle \delta V^2 \rangle_{\text{NPT}}$ and $\langle V \rangle_{\text{NPT}}$ denote the volume fluctuations and average volume of the simulation box in the NPT ensemble, respectively. The volume fluctuations method has been used widely to compute κ_T for several charged liquid systems including ILs.^{84, 143–146}

$$\kappa_T = \frac{\langle \delta V^2 \rangle_{\text{NPT}}}{\langle V \rangle_{\text{NPT}} k_b T} \quad [21]$$

Both methods, i.e., Eqs. [20] and [21], were compared by computing κ_T values using the SPC/E water model and provided similar κ_T values of 38.47 ± 1.16 and 34.07 ± 5.21 in 10^{-11} Pa^{-1} , respectively.¹³⁸ Simulations at 323 K and 0.98 bar for the IL [bmim][PF₆] performed by Shah et al.¹⁴⁰ also

gave similar error percentages, 40.3% and 41.1%, for the linear approximation and Tait equation methods, respectively. With respect to DES simulations, κ_T values have been reported for CCEtg by Ferreira et al.⁹⁸ at 298, 313, and 323 K using Eq. 21. Multiple FFs were tested, but the mixed parameter set termed 0.8FFM (i.e., GAFF (Ch⁺)³¹ + OPLS-AA/AMBER (Cl⁻)⁷⁶ + OPLS-AA (ethylene glycol)¹⁴⁷), yielded the closest agreement with error percentages of 11%, 1.5%, and 0.1% at 298, 313, and 323 K, respectively. Notably, FF combinations that displayed better agreement for κ_T did not necessarily show good performance for other DES solvent properties including self-diffusion coefficients, surface tension, and viscosity.

TRANSPORT PROPERTIES

Viscosity

Viscosity is an important property for evaluating FF parameters, particularly when gauging the accuracy of computed intermolecular interactions. Fortunately, due to the relative ease of measurement and the importance of DESs in industrial processes,⁵⁷ experimental viscosities are often readily available for comparison. However, the highly viscous nature of DESs⁴¹ often leads to major discrepancies in reported values under the same conditions. For example, viscosities for CCU at ambient conditions have been reported to range from 152 cP⁴⁰ to 527.28 cP.¹¹¹ The cause of such large deviations may stem in part from differences in the experimental methods,¹¹⁰ but more likely arise from the presence of impurities during the preparation process.⁵⁷ Many classes of DESs are highly hygroscopic, and water has been found to have a dramatic effect on the viscosities of these solvents. For example, the viscosity of CCU was shown to decrease by approximately 60% when shifting from pure DES to 0.1 mol fraction of water.¹¹¹ Additionally, Florindo et al.¹¹⁰ found that CCOx has a considerable affinity for atmospheric water, citing a 19.40% water content that dropped the viscosity from 5363 to 44.49 cP at 303.15 K. The cause of such dramatic decreases in viscosity is believed to be from the disruption of the complex hydrogen bond network that is attributed to DES's highly viscous nature.^{41, 110}

Along with impurities, increasing the molar ratio of HBDs present in a DES can also decrease the viscosity as a consequence of disrupting the hydrogen bond network of the solvent. For example, the viscosities for CCPhe at 1 : 2, 1 : 3, and 1 : 6 molar ratios are reported as 90.33, 44.64,

and 21.43 cP, respectively.^{41, 148} As more phenol was added, it is suggested that the HBD eventually acts as an organic solvent which disrupts the Coulomb interaction between the cation and anion, leading to the same effect seen when water is introduced to the system. However, for CCGly the opposite effect was observed, as the molar ratio increased from 1 : 2 to 1 : 3 to 1 : 4 the reported viscosity values (at 293.15 K) also increased, i.e., 376, 450, and 503 cP, respectively.^{41, 148, 149} This phenomena is attributed to the strong cohesive energy between glycerol molecules that generates a strong hydrogen bond network limiting ion mobility and thus increasing the viscosity.⁴¹ Considering these examples, it is important to consider the ratios used during the construction of DESs when studying their viscosities.

Temperature also has a significant effect on the viscosity of DESs.^{110, 150} For example, substantial decreases in viscosity going from 298 to 328 K were reported for CCU (1 : 2 molar ratio): 750 to 95 cP, CCGly (1 : 2 molar ratio): 259 to 52 cP, and CCMal (1 : 1 molar ratio): 1124 to 161 cP, respectively.^{44, 110} A strong relationship between the temperature dependence of the DES viscosity and the strength of the ion-HBD intermolecular forces has been suggested based on fittings to an Arrhenius model which allows for activation energies (E_a) to be calculated.¹¹⁰ In this correlation, DESs that exhibit low viscosities have a low E_a , whereas systems with high viscosity values have a relatively higher E_a . For example, CCEtg (1 : 2 molar ratio) which has a reported viscosity of 39.7 cP at 298.15 K has an E_a of -11.26 kJ/mol, while CCOx (1 : 1 molar ratio) which has a reported viscosity of 208.3 at 348.15 K has an E_a of -65.20 kJ/mol.^{41, 57} This trend further highlights the dramatic effect of ion-HBD interactions upon the overall viscosity.

When calculating the viscosity of a system computationally, the type of simulation falls under two categories: equilibrium MD and nonequilibrium molecular dynamics (NEMD). MD is desirable because the viscosity can be calculated from either pressure or momentum fluctuations in an equilibrium trajectory via a single simulation. A common approach to relating pressure fluctuations to viscosity uses the Green–Kubo formula (Eq. 22).¹⁵¹

$$\eta = \frac{V}{k_B T} \int_0^\infty \langle P_{xz}(t_0) P_{xz}(t_0 + t) \rangle_{t_0} dt \quad [22]$$

Here, V is the volume, T is the temperature, k_b is the Boltzmann constant, and P_{xz} is the off-diagonal element of the stress tensor. To reduce uncertainties arising from large pressure fluctuations, the integral is often fit to a pressure tensor autocorrelation function. However, due to the slow dynamics and relaxation times of the highly viscous solvents, the Green–Kubo method

struggles to accurately sample pressure tensors unless very long simulations are performed.^{152–154}

Alternatively, NEMD methods have been shown to properly treat highly viscous solvents^{91, 92, 155, 156} by applying an external force to the solvent and relating the resulting flux back to the viscosity. As a result, additional simulations are required beyond the equilibrium trajectory. One method that has been used extensively is the periodic perturbation method.¹⁵⁷ In this method, an external force of a chosen amplitude is applied in the x direction, a_x , to three-dimensional periodic cells to create a velocity field u . The velocity field can then be described using the Navier–Stokes equation (Eq. 23) where a_y and a_z are equal to zero, resulting in velocity fields in the y and z direction to also be zero. The equation for the velocity field then becomes

$$\rho \frac{\delta u_x(z)}{\delta t} = \rho a_x(z) + \eta \frac{\delta^2 u_x(z)}{\delta z^2} \quad [23]$$

where ρ is the mass density. The velocity field is easily calculated throughout the simulation with the use of a velocity profile, V . The velocity profile is then related to the viscosity of the system using Eqs. [24] and [25], where l_z is the height of the box and Λ is the acceleration amplitude of the external force, $a_x(z)$ (Eq. 26).

$$\eta = \frac{\Lambda \rho}{V k^2} \quad [24]$$

$$k = \frac{2\pi}{l_z} \quad [25]$$

To ensure a smooth velocity profile with small local shear rates, the external force is controlled with a cosine function.

$$a_x(z) = \Lambda \cos(kz) \quad [26]$$

The selection of a proper acceleration amplitude is crucial, as it should be large enough to properly probe the system, but small enough so that the equilibrium of the system is not completely destroyed. Multiple simulations are then required at varying amplitudes, typically ranging between 0.02 and 0.25 nm/ps², in order to get point viscosities at each amplitude. Extrapolation to an undisturbed system where, $\Lambda = 0$, is taken as the viscosity of the system.

An alternative NEMD, called the D-base method, was recently tested on a highly viscous ionic liquid system [Bmim][Tf₂N].¹⁵⁸ This method utilizes finite-size effects of self-diffusion coefficients to calculate the viscosity of a

bulk system through the equation of Yeh and Hummer (Eq. 27).¹⁵⁹ Multiple simulations are required to provide diffusion coefficients at various system sizes.

$$D_{\text{self}}^{\text{MD}} = \left(\frac{1}{\eta} \right) \left(-\frac{\xi k_b T}{6\pi L} \right) + D_{\text{self}}^{\infty} \quad [27]$$

From Eq. 27, ξ is a dimensionless constant equal to 2.837297, and L is the length of the simulation box. Eq. 27 is in a linear form, where $D_{\text{self}}^{\text{MD}}$ is the independent variable and $-\xi k_b T/6\pi L$ is the dependent variable. D_{self}^{∞} is the intercept of the linear function, representing the thermodynamic limit in which $L \rightarrow \infty$. The inverse slope of the line is then taken as the viscosity of the system. In order to incorporate all species in the system, $D_{\text{self}}^{\text{MD}}$ is replaced by an average of each self-diffusion coefficient weighted by their corresponding mole fraction (Eq. 28).

$$D_{\text{avg}} = \lim_{t \rightarrow \infty} \frac{1}{6t} \frac{1}{N} \left\langle \sum_{i=1}^n \sum_{j=1}^{N_i} (r_{j,i}(t) - r_{j,i}(0))^2 \right\rangle \quad [28]$$

Here n and N are the total number of species and molecules in the mixture, respectively. When used on the [Bmim][Tf₂N] ionic liquid system at increasing temperatures, the predicted viscosities matched well with Green–Kubo data reported by Zhang et al.¹⁶⁰ Although the D-base method has not been applied to DESs to the date this chapter was written, it does have considerable potential for future use with its ability to handle viscous mixtures possessing more than one molecular species.

Another potential avenue for computing the viscosities of DESs is the Müller-Plathe method, which has been successfully applied to the ionic liquid system [Emim][Tf₂N].^{161, 162} The Müller-Plathe method uses a reverse nonequilibrium molecular dynamics approach (RNEMD), where a momentum flux causes the corresponding external field that is related to the viscosity using Eq. [29]. This differs from the NEMD approach, where an external field elicits a flux within the system which is then related to the viscosity by the Navier–Stokes equation.

$$j_y = -\eta(\dot{\gamma}) \frac{\delta v_x}{\delta y} \quad [29]$$

In Eq. 29, j_y is the momentum flux that is imposed, and $\dot{\gamma}$ is the velocity gradient which can be calculated using a velocity profile throughout a simulation. The momentum flux is arbitrarily chosen by dividing the simulation box into an N number of “slabs” in the y direction and then exchanging

the largest momentum difference, $P_{x,n_c} - P_{x,n_1}$, in the x direction between two atoms from differing slabs. The total momentum exchange, P_{total} , and resulting flux at a time t is calculated with Eqs. [30] and [31].

$$P_{\text{total}} = \sum (P_{x,n_c} - P_{x,n_1}) \quad [30]$$

$$j_y(P_x) = \frac{P_{\text{total}}}{2tL_xL_z} \quad [31]$$

L_x and L_z are the lengths of the simulation boxes in the x and z direction. The resulting velocity gradients, which are calculated by the velocity profiles of the simulation, are then related back to Eq. [29] to obtain the viscosity. Calculated viscosities for the IL system [Emim][Tf₂N] gave a root mean squared error of 15% when compared to experimental values.¹⁶² Again, while this method has yet to be applied to DES systems, it may present a viable option for future studies.

Errors between MD and NEMD methods have been highlighted in the DES system CCLev (1 : 2 molar ratio). Using the Green–Kubo method, a calculated viscosity of 265 cP at 298 K was overestimated compared to the weighted experimental value of 226.8 cP, a 16.8% error.³⁷ Comparatively, the periodic perturbation method proved to be more accurate yielding a viscosity of 220.8 cP, a 2.6% error.³⁵ Additional simulated viscosities featuring the Green–Kubo method have been reported for CCU (1 : 2 molar ratio), CCGly (1 : 2 molar ratio), and CCMal (1 : 1 molar ratio) at 318 K with percent errors of 35.4%, 26.8%, and 31.8% respectively.²⁰ Simulated viscosities using the periodic perturbation method for these same systems gave percent errors of 1.1%, 3.3%, and 1.9%, respectively, at 303 and 348 K.³⁵ Altamash et al. calculated the viscosity of the DES system choline chloride phenylacetic acid (1 : 2 molar ratio) using the Green–Kubo method and found at 298 K the simulations performed modestly with percent error of 16.4%.¹⁶³ However, when the temperature increased to 345 K, the percent errors became as high as 80.1% illustrating the importance of studying DESs systems within a wide range of temperatures. Alternatively, Zhang et al.³⁶ found using the time decomposition method¹⁶⁰ of the Green–Kubo theory for CCEtg (1 : 2) while utilizing the parameters of Perkins et al.³¹ performed better at elevated temperatures where calculated viscosities were overestimated by 5–8 cP. The periodic perturbation method when coupled to the OPLS-AA FF developed by Doherty and Acevedo³⁵ found that temperature had little effect on the accuracy of each prediction, where a MAE of 14 data points was calculated to be 1.6% (Table 9). However, there are cases where the percent error did increase marginally as the temperature was

TABLE 9 Calculated and Experimental Viscosities (cP) at Various Temperatures.

| DES | Simulation | Experiment | % error |
|-----------------|------------|------------|---------|
| 298.15 K | | | |
| CCEtg | 39.5 | 39.7 | 0.6 |
| CCGly | 258.8 | 259.0 | 0.1 |
| CCLev | 220.8 | 226.8 | 2.6 |
| CCPhe (1 : 2) | 89.1 | 90.3 | 1.3 |
| CCPhe (1 : 3) | 44.4 | 44.6 | 0.4 |
| CCU | 753.1 | 749.9 | 0.4 |
| 303.15 K | | | |
| CCEtg | 35.0 | 35.0 | 0.0 |
| CCGly | 246.8 | 238.9 | 3.3 |
| CCLev | 164.0 | 164.5 | 0.3 |
| CCPhe (1 : 2) | 64.7 | 68.4 | 5.4 |
| CCPhe (1 : 3) | 36.5 | 35.2 | 3.7 |
| CCU | 520.5 | 514.8 | 1.1 |
| 348.15 K | | | |
| CCMal | 94.9 | 96.7 | 1.9 |
| CCOx | 205.2 | 208.3 | 1.5 |
| MAE (%) | | | 1.6 |

^aWeighted experimental averages were computed at various temperatures where each weight was determined by the inverse of its reported uncertainty.

Source: Doherty et al. 2018³⁵.

raised by 5 K, and further studies may be needed at higher temperatures for a full evaluation.

Along with the calculated viscosity method, the quality of the FF can influence bulk-phase properties significantly. A list of FFs that have been used to simulate DES viscosities is given in Table 10. The importance of selecting a proper force field is emphasized by Ferreira et al. who tested eight combinations of nonpolarizable FF parameters from the literature for the different components of CCEtg (1 : 2 molar ratio).⁹⁸ This includes choline parameters from Sambasivarao and Acevedo,⁷⁵ Perkins et al.⁵⁵ and OPLS-AA intramolecular parameters.⁷⁷ Chloride parameters were taken either directly from OPLS-AA⁷⁷ or from Canongia Lopes et al.⁷⁶ which were developed for IL simulations. For ethylene glycol, the HBD parameters were taken from either OPLS-AA,⁷⁷ Szcfczyk and Cordeiro,¹⁴⁷ or Gorny et al.¹⁰⁰ In addition, each combination was evaluated with integer charges ($\pm 1 e$) and scaled charges ($\pm 0.8 e$). Substantial improvement for self-diffusion coefficients were observed when the charges were scaled.

TABLE 10 Viscosity Calculation Methods Utilized in DES Simulations.

| Simulation by | DES studied | Charge assignment | Method |
|-----------------------------------|--|-------------------|-----------------------|
| Ullah et al. ³⁷ | CCLev (1 : 2) | Scaled | Green–Kubo |
| Doherty and Acevedo ³⁵ | CCEtg (1 : 2), CCGly (1 : 2), CCLev (1 : 2), CCMal (1 : 1), CCOx (1 : 1), CCPhe (1 : 2), CCPhe (1 : 3), CCU (1 : 2) | Scaled | Periodic Perturbation |
| García et al. ²⁰ | CCGly (1 : 2), CCMal (1 : 1), CCU (1 : 2) | Scaled | Green–Kubo |
| Altamash et al. ¹⁶³ | CCPhOAc (1 : 2) | Scaled | Green–Kubo |
| Zhang et al. ³⁶ | CCEtg (1 : 2) | Scaled | Green–Kubo |
| Ferreira et al. ⁹⁸ | CCEtg (1 : 2) | Unscaled, scaled | Periodic perturbation |
| Ferreira et al. ⁹⁹ | CCPro (1 : 2) | Unscaled, scaled | Periodic perturbation |

While using the periodic perturbation method, five FF combinations struggled to match experimental viscosity values shown in Figure 8, where measurements were underestimated by a factor of 2.5 on average. However, at 298 K some FFs performed considerably better than others, stemming from the treatment of short-range interactions and hydrogen bonding that are governed by the quality of the parameters. Also worth noting is the ability to capture the temperature dependence of viscosity measurements, where calculated values normalized by the 298 K measurement are in very good agreement with experimental data. Overall, great care should be taken when selecting (1) FF parameters and (2) viscosity calculation methods, as viscosity predictions derived from atomistic DES simulations have been shown to be highly sensitive to both choices.

Aside from MD simulations, the viscosity of DES systems can also be calculated with equations of state models in combination with various friction theories which has been outlined in the review by González de Catilla et al.¹⁶⁴ For example, Haghbakhsh et al.¹⁶⁵ used friction theory along with the cubic plus association and perturbed chain-statistical associating fluid theory to calculate the viscosity of 27 different DES systems and found an average relative deviation from experimental values of 4.4% for both models. A follow-up version of the models were used by the same

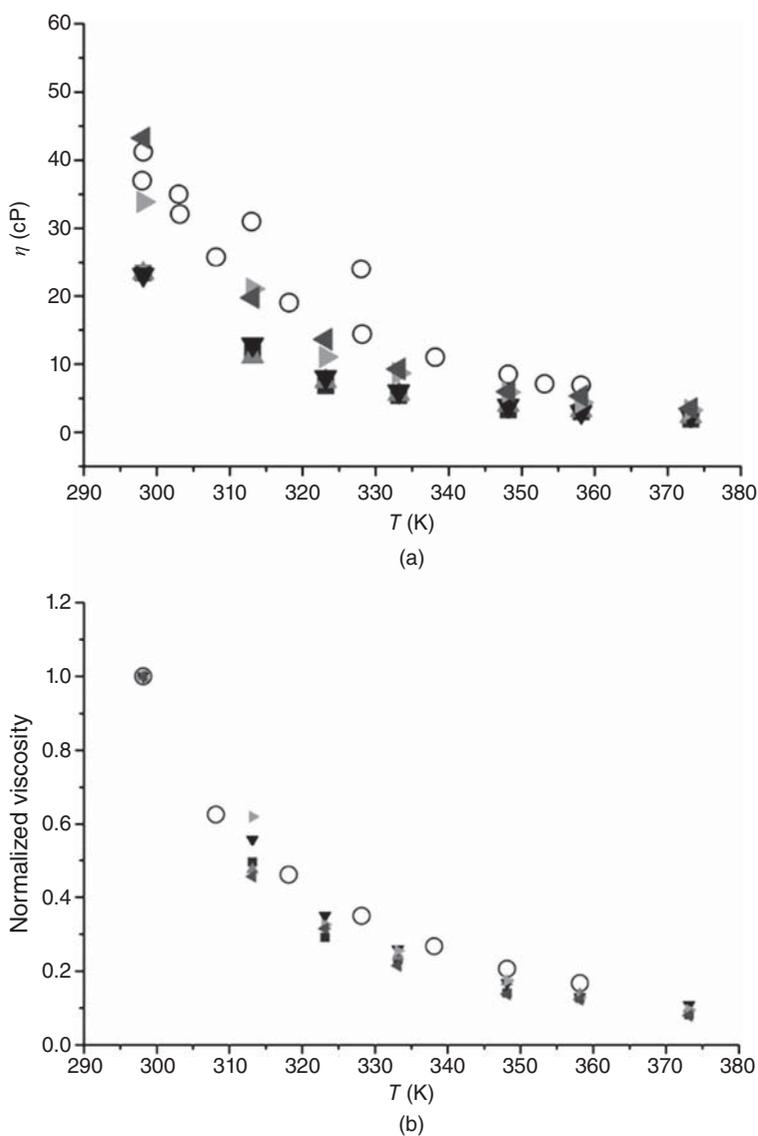


FIGURE 8 (a) Experimental and simulated viscosities from multiple FF combinations for CCEtg at 1 bar and (b) Experimental and simulated normalized viscosity by the temperature 298.15 K. Experimental values are depicted as empty black circles while simulated values are filled shapes. *Source:* Ferreira et al. 2016⁹⁸.

group where results were then improved to an average deviation of 2.7%.¹⁶⁶ Additionally, Lloret et al.¹⁶⁷ used a soft-SAFT with free volume theory for quaternary ammonium chloride containing DES systems and found good agreement with experimental trends.

Diffusion Coefficients

The importance of understanding the mechanism of how ions/molecules diffuse in DES systems has been discussed in the literature and tested experimentally.^{44, 57, 168} Originally, DESs were thought to move similarly to ILs which have been characterized by Zhao et al. using a modified hole theory.¹⁶⁹ In the modified hole theory, ions diffuse moving from one vacancy to another when an ion's hole size is smaller than the one adjacent to it. While it is still believed that the mobility of the holes are the dominant factor in an ion's diffusivity resulting in a jumping mechanism,⁴⁴ the hydrogen bond network between the ions and HBD should also be considered as it can hinder the mobility of each component. To our knowledge, only D'Agostino et al.⁴⁴ and Abbott et al.¹⁶⁸ have reported diffusion coefficients with the use of pulse field gradient nuclear magnetic resonance (PFG-NMR) for DES systems. D'Agostino et al.⁴⁴ found the HBD interactions to be important for CCMal, where it is believed that the carboxylic acid functional groups of maline create a dimerization through hydrogen bonding that leads to long chains and hinders the mobility of the ions within the system. Abbott et al.¹⁶⁸ also found that when more choline chloride was added to a CCGly system, the choline chloride would break up the intermolecular forces between glycerol molecules and increase their diffusivity, highlighting the importance of the hydrogen bond network established as well as the molar ratio in each DES system. Experimentally, cations are found to diffuse slower than HBDs for urea, glycerol, and ethylene glycol which can be explained by the hole theory. However, the opposite trend is seen with malonic acid due to its dimerization that was explained previously. The importance of temperature has also been stated in literature,⁴⁴ where results show an Arrhenius-type behavior. The temperature dependence is also crucial when considering simulated diffusion coefficients which is explained later.

For a simulated system, the diffusion coefficient is calculated by applying the Einstein relation and the average mean square displacement for each ion/molecular center of mass (Eq. [32]).¹⁴²

$$D_s = \frac{1}{6} \lim_{t \rightarrow \infty} \frac{d}{dt} \frac{1}{N} \sum_{i=1}^N \langle |r_i(t) - r_i(0)|^2 \rangle \quad [32]$$

Here, $r_i(t)$ is the center of mass of species i at time t , and N is the number of individual species. An important aspect of using the Einstein relation is that the system is studied within a proper diffusive regime where ions/molecules are moving freely. A way to monitor the diffusive regime is with the calculation of the beta-parameter (β), which has been previously discussed by Del Popolo and Voth¹⁷⁰ and applied to DES systems by Perkins et al.,^{31, 55} Doherty and Acevedo,³⁵ Ferreira et al.,⁹⁸ and Mainberger et al.³²

$$\beta(t) = \frac{d \log_{10} \langle \Delta r(t)^2 \rangle}{d \log_{10} t} \quad [33]$$

Here, $\langle \Delta r(t)^2 \rangle$ is the mean square displacement and t is the time. The beta-parameter can therefore be plotted versus time and when $\beta < 1$, the system is considered to be in the sub-diffusive regime. When $\beta = 1$, the system is then considered in the diffusive regime and the diffusion coefficient can be properly calculated with Eq. [33].

Similar to viscosity calculations, diffusion coefficients are very sensitive to the treatment of charges assigned to each molecule/ion. For simulations that utilize a nonpolarizable FF, the use of integer charges ($\pm 1 e$) for the ions has shown to significantly underestimate the diffusion coefficients. For example, Mainberger et al.³² used MMFF parameters with unscaled charges and found that for CCGly the calculated diffusion coefficients at 328 K had percent errors as high as 92%. As an alternative, GAFF parameters in combination with RESP charges derived from a minimal cluster of ChCl/HBDs (1 : 2 ratio) were also tested. Due to charge transfer effects, the species Ch^+ , Cl^- , and glycerol had scaled point charges of 0.7615 e, $-0.6527 e$, and $-0.0544 e$, respectively. Diffusion coefficient results improved dramatically to 15.6% and 6.5% errors for choline cation and glycerol, respectively. Ferreira et al.⁹⁸ observed the same improvement when charges were scaled for CCEtG and CCPro. For example, using unscaled charges by Perkins et al.⁵⁵ for CCEtG combined with parameters from OPLS-AA,⁷⁷ calculated diffusion coefficients at 313.15 K had errors of 90.13% and 87.88% for choline and ethylene glycol, respectively. However, scaling the charges by a factor of 0.8, improved the errors to 11.8% and 2.8%, respectively. For CCPro, system specific charges resulted in a scaling factor of 0.74 that did not perform as well with errors of 17% and 15% for choline and propylene glycol, respectively. However, this was a major improvement over the unscaled systems that never reached the diffusive regime. Calculations of self-diffusivity using a variety of FFs are presented in Table 11 for multiple DESs.

Another important factor to consider is the temperature of the system. Typically, at room temperature (298.15 K) systems exhibit a sub-diffusive character where $\beta < 1$ due to the strongly correlated hydrogen bonding that

TABLE 11 Calculated Self-Diffusion Coefficients (D^+ and D^{HBD} at $10^{-11} \text{ m}^2 \text{ s}^{-1}$) at 298.15 and 330.15 K for Deep Eutectic Solvents.

| DES | Force field | D^+ | D^{HBD} | % error D^+ | % error D^{HBD} |
|-------|-------------------------------------|-------|------------------|---------------|--------------------------|
| CCU | | | | 298.15 K | |
| | Doherty and Acevedo ³⁵ | 0.46 | 0.35 | 31.4 | 47.0 |
| | Perkins et al. ⁵⁵ | 0.17 | 0.39 | 51.4 | 40.9 |
| | Experiment ⁴⁴ | 0.35 | 0.66 | — | — |
| | | | | 330.15 K | |
| | Doherty and Acevedo ^{35,a} | 1.51 | 3.56 | 28.1 | 0.3 |
| CCEtg | | | | 298.15 K | |
| | Perkins et al. ³¹ | 1.81 | 3.75 | 30.9 | 21.4 |
| | Ferreira et al. ⁹⁸ | 2.08 | 4.15 | 20.6 | 13.0 |
| | Experiment ⁴⁴ | 2.62 | 4.77 | — | — |
| | | | | 330.15 K | |
| | Perkins et al. ³¹ | 7.44 | 15.3 | 24.0 | 6.7 |
| CCGly | | | | 298.15 K | |
| | Perkins et al. ³¹ | 0.30 | 0.45 | 21.1 | 13.5 |
| | Experiment ⁴⁴ | 0.38 | 0.52 | — | — |
| | | | | 330.15 K | |
| | Perkins et al. ³¹ | 2.11 | 3.12 | 17.9 | 27.3 |
| | Mainberger et al. ^{32,a} | 2.07 | 2.61 | 15.6 | 6.5 |
| CCPro | Experiment ⁴⁴ | 1.79 | 2.45 | — | — |
| | | | | 298.15 K | |
| | Ferreira et al. ⁹⁹ | 1.37 | 2.11 | 17.0 | 15.0 |
| | Experiment ⁹⁹ | 1.66 | 2.49 | — | — |

^aMeasured at 328.15 K

occurs between the cation and anion creating a cage that the ions cannot escape until the temperature is raised.⁴⁴ Although scaling charges has shown to improve results, simulations at room temperature still struggle as shown by Perkins et al.^{31,55} where they analyzed diffusion coefficients at both 298 and 330 K for CCU, CCEtg, and CCGly using the GAFF FF with RESP derived charges on isolated ions/molecules. Most noticeably for the CCU system, results at room temperature gave large errors of 51.4% and 40.9% for choline and urea, respectively. Results improved considerably however when the temperature was raised to 330 K, with calculated diffusion coefficients for choline and urea having 3.8% and 3.4% errors, respectively. Although improvement was seen when the temperature

was raised to 330 K for every DES system, reported percent errors were still as high as 27.3%. Despite the errors, diffusive trends were adequately captured for all simulated results reported. Shown in Figure 9, Perkins et al.⁵⁵ found that urea diffuses faster than the heavier and larger choline ion for CCU which is consistent with experimental findings.

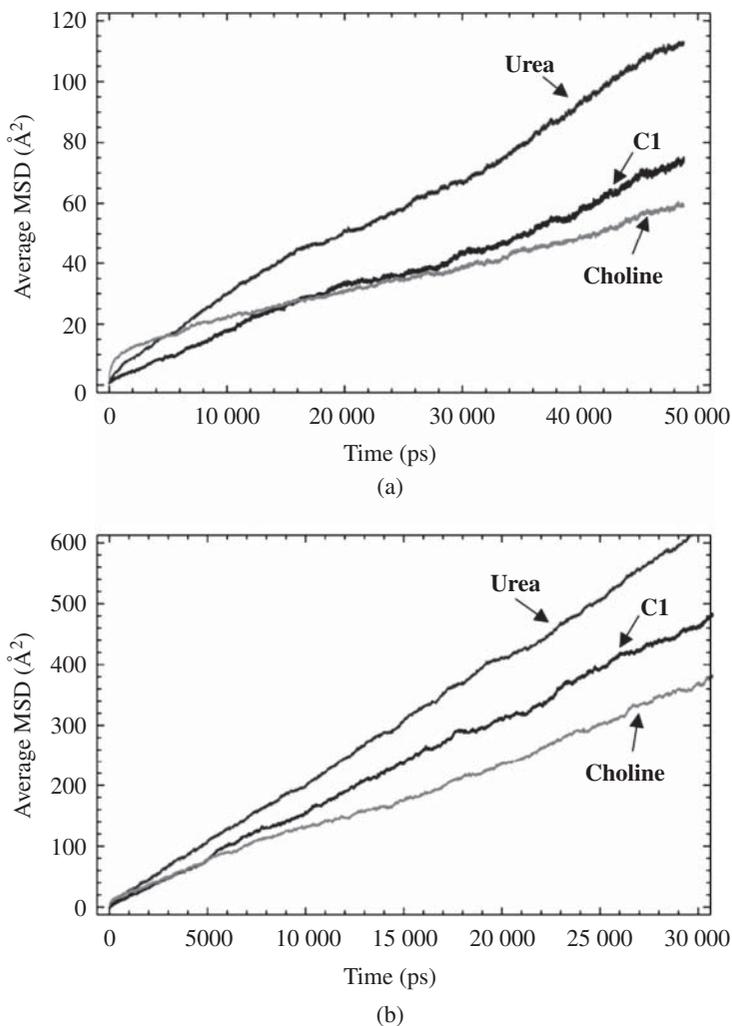


FIGURE 9 Average mean squared displacement as a function of time at (a) 298 K and (b) 330 K. *Source:* Perkins et al. 2013⁵⁴.

As with any simulation, refinement of FF parameters can lead to more accurate predictions of bulk-phase and transport properties. Modest improvement was accomplished by Ferreira et al.⁹⁸ by testing various combinations of FF parameters for each component of CCETg. After scaling the charges by 0.8 to the best performing parameter combination, predicted diffusion coefficients improved to an average error of 10% for the temperature range of 298.15–323.15 K. However, the results were inconsistent and for all combinations tested there was a parabola-like temperature

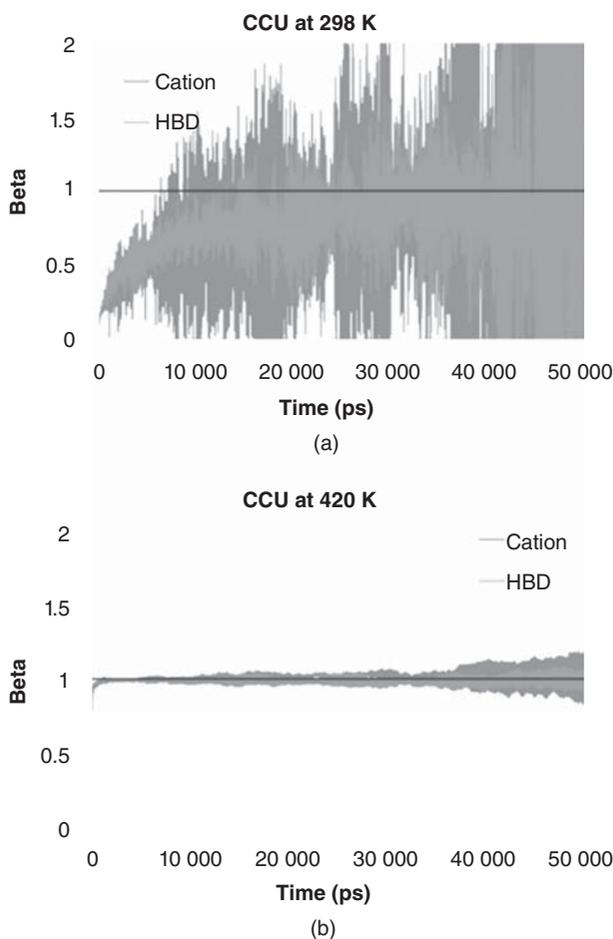


FIGURE 10 Calculated beta-parameter as a function of time for CCU at (a) 298 K and (b) 420 K. *Source:* Based on Doherty et al. 2018³⁵.

dependence for each diffusion coefficient where percent errors more than doubled between temperatures. A similar trend was reported by Doherty and Acevedo³⁵ who developed OPLS-AA parameters utilizing a 0.8 charge scaling for eight different DES systems. At temperatures between 298.15 and 328.15 K, a sub-diffusive regime was observed. Illustrated in Figure 10 are calculated beta-parameters for the choline chloride urea system at both 298.15 and 420.15 K, clearly showing the temperature dependence where $\beta < 1$ is observed for the majority of the 298.15 K simulations and $\beta = 1$ at 420.15 K. To compensate, simulations were performed at higher temperature (400.15–500.15 K) and calculated diffusion coefficients were extrapolated to room temperature. Results varied substantially for the CCU system, where errors were reported as 31.4%, 0.0%, and 23.2% for 298.15, 308.15, and 323.15 K, respectively.³⁵

Although calculated self-diffusion coefficients can seem accurate at specific temperatures, there exists an inconsistency for all nonpolarizable FFs when tested over a range of temperatures. At minimum, simulations should be run at higher temperatures with adequate lengths in order to ensure a diffusive regime is properly sampled. Polarizable FFs are likely required to adequately capture the charge fluctuations that have a large effect on transport properties.

DEEP EUTECTIC SOLVENT STRUCTURE

Radial Distribution Functions

The driving force for DES melting point depression has often been linked to the complex hydrogen bond network formed between both the ions and HBDs.^{44, 54, 168, 171} In addition, physical properties such as viscosity and diffusion coefficients can be directly related to the DES local structure and interaction strengths between each component.^{36, 150} To clarify the liquid structuring of various DES systems including CCU (1 : 2 molar ratio), CCOx (1 : 1 molar ratio), CCEtg (1 : 2 molar ratio), and CCGly (1 : 1 molar ratio), neutron diffraction^{36, 46, 63, 172} and simulation studies^{31, 32, 34–36, 45, 55, 69, 78, 150, 173} have been performed. The existence of a H-bond network is well supported by HOESY NMR,⁴⁰ FT-IR,^{55, 174, 175} PFG-NMR,⁴⁴ and Quasi-elastic neutron scattering.⁶³ The use of ND/EP SR has been shown to provide specific atomic site–site interactions at an accurate level.^{46, 172, 176} From the ND/EP SR data, comparisons can be made to high-level computational methods such as FPMD,⁴⁵ AIMD,^{34, 36, 69, 173} and QM/MD simulations.¹⁵⁰ MD simulations utilizing nonpolarizable FFs have also been used to study the structure of CCU (1 : 2 molar

ratio),^{20, 35, 54, 55, 177, 178} CCEtg (1 : 2 molar ratio),^{31, 35, 36} CCLev (1 : 2 molar ratio),^{37, 94} CCMal (1 : 2 molar ratio),²⁰ and CCGly (1 : 2 molar ratio)^{20, 31, 32, 35} and compared to the previously mentioned methods when applicable.

Analyzing the center-of-mass (COM) RDFs provides insight into how each component of the DES system is coordinated to one another by integrating the first peak of each interaction. For the system CCU (1 : 2 molar ratio), the COM RDFs were computed using ND/EPSR⁴⁶ and MD^{35, 54} at room temperature, as well as FPMD⁴⁵ at 333 K. Table 12 shows the COM RDF peak distances and coordination numbers computed from ND/EPSR, indicating the strongest interactions occurred between choline–chloride, urea–chloride, and urea–urea near 4 Å. Integration of the urea–chloride peak resulted in a coordination number of two urea molecules per chloride, which is expected due to the 2 : 1 urea:chloride ratio used to construct this particular DES. The choline–chloride peak shows a distinct shoulder around 5 Å implying that there exists multiple interaction sites within the first solvation shell in which a chloride ion can oscillate between. Meanwhile, the urea molecules are not only interacting with the chloride ions, but also with surrounding urea molecules with a coordination value of 6.77. These results suggest that the HBDs in DESs are highly involved with the structural ordering of the system and that a complex hydrogen bond network exists between both the ion pair and HBDs. The importance of the HBD was also examined by Sun et al.⁵⁴ who studied the effect of increasing mole percentages (0%, 25%, 67.7%, and 75%) of urea into choline chloride using MD. Pure ChCl exhibited strong long-range ordering between the ions, but when urea was added to the system, the COM RDFs showed a gradual increase in the ion interaction distance and the second solvation layer decreasing. This is the result of the chloride ions interacting urea molecules that have inserted themselves within the ionic lattice and disrupting the long-range interactions.

Also provided in Table 12 are the coordination numbers calculated using FPMD and classical MD. Good agreement between all three methods was observed, with all coordination numbers within error of each other. Minimum and maximum distances are also in good agreement for each peak. It should be noted that for the MD simulations, the FF was fit specifically to reproduce the ND/EPSR data while maintaining accurate bulk property predictions. Significant tailoring of both the Lennard-Jones terms and charges were required to obtain accurate results and would be a necessary procedure for future DES systems.³⁵

TABLE 12 Average Coordination Number (N_{coord}) and Position (\AA) of the First Maximum and Minimum in Center-of-Mass RDFs between Choline Cation (Ch), Chloride Anion (Cl), and Urea.

| Center | Shell | Classical MD (303 K) ³⁵ | | | ND/ESPR (303 K) ⁴⁶ | | | FPMD (333 K) ⁴⁵ | | |
|--------|-------|---------------------------------------|------------------|--------------------|----------------------------------|------------------|--------------------|-------------------------------|------------------|--------------------|
| | | r_{max} | r_{min} | N_{coord} | r_{max} | r_{min} | N_{coord} | r_{max} | r_{min} | N_{coord} |
| Urea | Cl | 4.3 | 5.4 | 1.90 | 4.0 | 5.5 | 2.08 ± 1.01 | 4.1 | 5.3 | 1.9 ± 0.4 |
| Ch | Cl | 4.1 | 6.4 | 3.49 | 4.2 | 6.7 | 4.35 ± 1.30 | 4.2 | 6.5 | 3.1 ± 0.6 |
| Ch | Urea | 4.7 | 7.2 | 8.76 | 5.4 | 6.9 | 5.91 ± 2.84 | 5.1 | 7.1 | 8.6 ± 0.7 |
| Ch | Ch | 6.5 | 8.2 | 5.41 | 6.3 | 8.0 | 6.74 ± 2.16 | — | — | — |
| Urea | Urea | 4.8 | 6.6 | 6.00 | 4.3 | 6.1 | 6.77 ± 3.05 | 4.7 | 6.3 | 4.9 ± 0.5 |

Source: Based on Doherty et al. 2018³⁵.

TABLE 13 Position of the First Peak in the COM Radial Distribution Functions for CCGly at 1:1 and 1:2 Molar Ratios.

| RDF | Peak position (\AA) | |
|-------------------------|--------------------------------|--------------|
| | CCU (1 : 2) | CCOx (1 : 1) |
| Choline–choline | 6.4 | 6.3 |
| Choline–Cl ⁻ | 4.6 | 4.5 |
| Choline–HBD | 5.4 | 5.7 |
| HBD–HBD | 4.3 | 5.0 |
| HBD–Cl ⁻ | 3.7 | 3.7 |

Source: Gilmore et al. 2018¹⁷².

COM RDFs have also been computed for CCU (1 : 2 molar ratio) and CCOx (1 : 1 molar ratio) at 338 K using ND/ESPR (Table 13).¹⁷² When comparing the CCU (1 : 2 molar ratio) results to the room temperature ND/ESPR, the maximum peak positions were very similar, varying only by an average of 0.2 \AA . There is a noticeable difference in the choline–chloride interaction where the shoulder indicating the presence of multiple binding motifs is now a singular broad peak, suggesting that the elevated temperature favors interactions with primarily the hydroxyl group and trimethylammonium region of choline. For the CCOx (1 : 1 molar ratio) system, the same peak distances are observed as the CCU (1 : 2 molar ratio) system. Similar to the CCU system, the choline–HBD occurs at a shorter distance than the choline–choline interaction, which indicates an intercalation of oxalic acid within the ionic

TABLE 14 Position of the First Peak in the COM Radial Distribution Functions for CCGly at 1 : 1 and 1 : 2 Molar Ratios.

| RDF | Peak position (Å) | |
|--------------------------|-----------------------------|-----------------------------|
| | $\chi_{\text{chcl}} = 0.33$ | $\chi_{\text{chcl}} = 0.50$ |
| Choline–choline | 6.5 | 6.3 |
| Choline–Cl ⁻ | 4.1 | 4.1 |
| Choline–glycerol | 5.9 | 5.7 |
| Glycerol–Cl ⁻ | 4.1 | 4.1 |
| Glycerol–glycerol | 5.3 | 5.5 |

Source: Turner et al. 2019¹⁷⁶.

lattice. A prominent peak for the choline–chloride interaction over the oxalic acid–chloride interaction is worth noting, as an AIMD study performed at 375 K by Zahn et al.³⁴ found the opposite trend where preference was given to the HBD–chloride interaction. This could be a result of elevated temperatures but should be examined in future studies.

The importance of the HBD molar ratio has also been recently studied with the use of ND/EP SR for CCGly.¹⁷⁶ Looking at both the 1 : 1 and 1 : 2 ratios, Turner and Holbrey found through COM RDFs that there was no significant changes in the local structuring when choline chloride was added to the system (Table 14). However, by observing the partial site–site RDFs and the corresponding coordination numbers, it was found that the hydrogen bond network formed between glycerol molecules was disrupted as choline chloride was added. This may be the cause of the increase in distance seen in the glycerol–glycerol COM RDF from 5.3 to 5.5 Å. Site–site RDFs also showed that as the ratio of ions increases from 1 : 2 to 1 : 1, there was a significant reorganization that occurred to compensate for the excess choline chloride to the point where the system may be considered more of a choline chloride ionic liquid environment with glycerol clusters dispersed throughout.

While COM RDFs provided insight into the general coordination between each component of a DES system, partial site–site RDFs allowed for further investigation on specific interactions between atoms of each species. From the COM RDFs for CCU (1 : 2 molar ratio), it is understood that within the first solvation shell there are close range interactions between the cation–anion, anion–HBD, and HBD–HBD creating a complex hydrogen bond network. From the ND/EP SR results of Hammond et al.,⁴⁶ there was a clear preference for the chloride to interact with the hydroxyl

hydrogen of choline for the ion pair. The interactions with the other hydrogens of choline showed a lower correlation with chloride, but still have significant peak heights which could be the reason why the shoulder was seen for the choline–chloride interaction in the COM RDFs. As for the anion–HBD and HBD–HBD interactions, differentiation between the hydrogens *cis* and *trans* to the urea oxygen is apparent when interacting with the chloride anion. The ND/EPSR data suggest that the chloride ion is more likely to interact with the *cis* hydrogens while the *trans* hydrogens have a stabilizing effect in interacting with surrounding urea molecules. This same trend is observed by Doherty and Acevedo³⁵ using a refined FF fitted to ND/EPSR data. Alternatively, both AIMD⁶⁹ and MD⁵⁵ studies have observed the opposite phenomena, where site–site RDFs show the *trans* hydrogens prefer interactions with the oxygen of surrounding urea molecules and the *cis* hydrogens bond with the chloride anion.

Consistent among all DES systems is that the majority of the cation–anion interactions occurred between the hydroxyl hydrogen of choline and chloride, while the anion–HBD interactions were dominated by the hydrogen bonding groups, i.e., OH or NH. Zahn et al.³⁴ studied the influence of the HBD group with AIMD when the hydrogen bonding ability of the organic compound varied from urea to ethylene glycol and oxalic acid (Figure 11). Site–site RDFs showed that as the HBD shifted from an amine to hydroxyl and eventually a carboxylic acid, the anion–HBD interaction became stronger as indicated by the increase of the peak height as well as the decrease in the hydrogen bond distance. These results match the partial charge analysis performed using the Hirshfield-I charge partitioning scheme,¹⁷⁹ where charge transfer was much more significant to oxalic acid when compared to urea indicating a stronger hydrogen bond to the carboxylic acid group. Interestingly, Ullah et al.³⁷ found through MD simulations that the cation–anion interaction through the hydroxyl group of choline was still the dominating interaction when compared to the anion–HBD interaction between Cl[−] and the carboxylate containing levulinic acid through site–site RDFs. This is likely due to the fact that oxalic acid has two sites where the chloride can hydrogen bond to as opposed to levulinic acid’s single site.

The importance of the HBD is also showcased in the AIMD simulation of CCGly (1 : 1 molar ratio).¹⁷³ The strongest hydrogen bonding occurred between the OH groups of glycerol and Cl[−], with the middle hydroxyl group being the most dominant while the peripheral OH groups hydrogen bonded with surrounding glycerol molecules. Considerable interactions were also

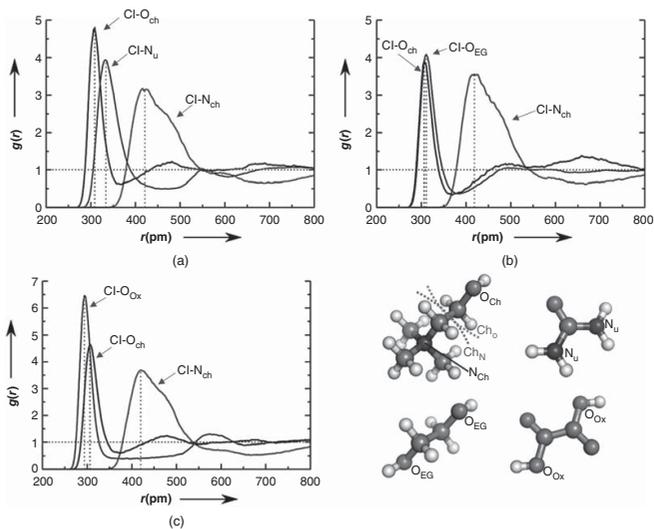


FIGURE 11 Radial distribution functions of selective interactions for (a) urea, (b) ethylene glycol, and (c) oxalic acid for choline chloride deep eutectic solvents. *Source:* Zahn et al. 2016⁵⁴.

observed between choline and glycerol through the hydroxyl groups of both moieties.

A comprehensive look into the structure of CCEtg (1 : 2 molar ratio) was recently studied by Zhang et al. where a combination of classical MD, AIMD, and neutron scattering experiments were utilized.³⁶ The FF used by Perkins et al.³¹ was validated as a means to simulate the solvation environment by matching experimental neutron scattering structure factors, $S(q)$, as well as computed $S(q)$ through AIMD. RDFs were also compared to AIMD results at 400 K, and although there were discrepancies in peak height and position, the overall structure features were captured. Further analysis was provided by the MD simulations, where calculated coordination number probabilities showed that on average, the chloride anion is surrounded by either 1 or 2 ethylene glycol molecules while the choline acts more as an observer. However, once the chloride anion interacts with the hydroxyl group of the choline, calculated hydrogen bond lifetimes and peak heights in the RDFs indicate that the cation–anion interaction is the strongest.

The use of nonpolarizable FFs to recreate the site–site RDFs from AIMD data has proven to be a difficult task. While Doherty and Acevedo³⁵ were able to match RDFs of ND/EPSR data for CCU,⁴⁶ the systems CCEtg, CCOx, and CCGly were less successful. Mainberger et al.³² also reported problems in obtaining consistent site–site RDFs when comparing the GAFF and MMFF with scaled charges for CCEtg and CCLev. The source of this difficulty likely arises from the charge assignment given to each atom in the deep eutectic system. García et al.⁹⁴ found that using a variety of different charge partitioning schemes resulted in considerably different site–site RDFs for CCLev, and the atoms involved in hydrogen bonding were particularly sensitive to the charges assigned. Zahn et al.³⁴ suggested that depending on the HBD involved, different scaling factors should be used for nonpolarizable FFs, and strongly recommended the use of polarizable FFs as an alternative.

Hydrogen Bond Analysis

In order to perform a hydrogen bond analysis, distances and angles characteristic to DES donors and acceptors need to be defined, e.g., donor–acceptor distances that range between 2.95 and 3.5 Å and X–H–Y angles ranging between a starting point of 130°–150° (Table 15). An additional criterion was proposed by Fetisov et al.⁴⁵ and Wernet et al.¹⁸⁰ who defined hydrogen bonding in DES through a CDF resembling an ellipsoid as shown in Figure 13. This ellipsoid criterion has also been applied to CCU in MD simulations by Doherty and Acevedo.³⁵

TABLE 15 List of Hydrogen Bond Criteria for Various DES Systems.

| | System | Simulation method | Donor-acceptor distance (Å) | X-H-Y angle cutoff |
|-----------------------------------|--------|------------------------------------|-----------------------------|--------------------|
| Perkins et al. ³¹ | CCEtg | Molecular dynamics | 3.5 | 150 |
| García et al. ⁹⁴ | CCLev | Molecular dynamics | 3.0 | 130 |
| Ullah et al. ³⁷ | CCLev | Molecular dynamics | 3.0 | 130 |
| García et al. ²⁰ | CCGly | Molecular dynamics | 3.0 | 130 |
| Perkins et al. ³¹ | CCGly | Molecular dynamics | 3.5 | 150 |
| Turner and Holbrey ¹⁷⁶ | CCGly | ND/EPSR | 3.4 | 135 |
| García et al. ²⁰ | CCMal | Molecular dynamics | 3.0 | 130 |
| Perkins et al. ³¹ | CCMal | Molecular dynamics | 3.5 | 150 |
| García et al. ²⁰ | CCU | Molecular dynamics | 3.0 | 130 |
| Fetisov et al. ⁴⁵ | CCU | First principle molecular dynamics | 3.5 | 150 |
| Sun et al. ⁵⁴ | CCU | Molecular dynamics | 2.95 | 150 |
| Perkins et al. ³¹ | CCU | Molecular dynamics | 3.5 | 150 |

Using FPMD, Fetisov et al.⁴⁵ studied the effects of water in CCU (1 : 2 molar ratio). Hydrogen bonds were monitored with a combination of RDFs and CDFs, while also taking into account the average hydrogen bond types at each frame of the trajectory. For the pure CCU system, strong interactions between urea and chloride were observed through both the RDFs and average fraction of hydrogen bonds where the *trans* hydrogens of urea were found to prefer bonding with Cl⁻. Conversely, urea-urea hydrogen bonding of the NH-O=C type was primarily dominated by the *cis* hydrogens of urea. This same trend was also observed by Perkins et al.^{31,55} when the average fraction of hydrogen bonds were calculated between urea and chloride throughout an MD trajectory (Figure 12). Hydrogen bonds between choline and chloride were also observed through RDFs, however due to the

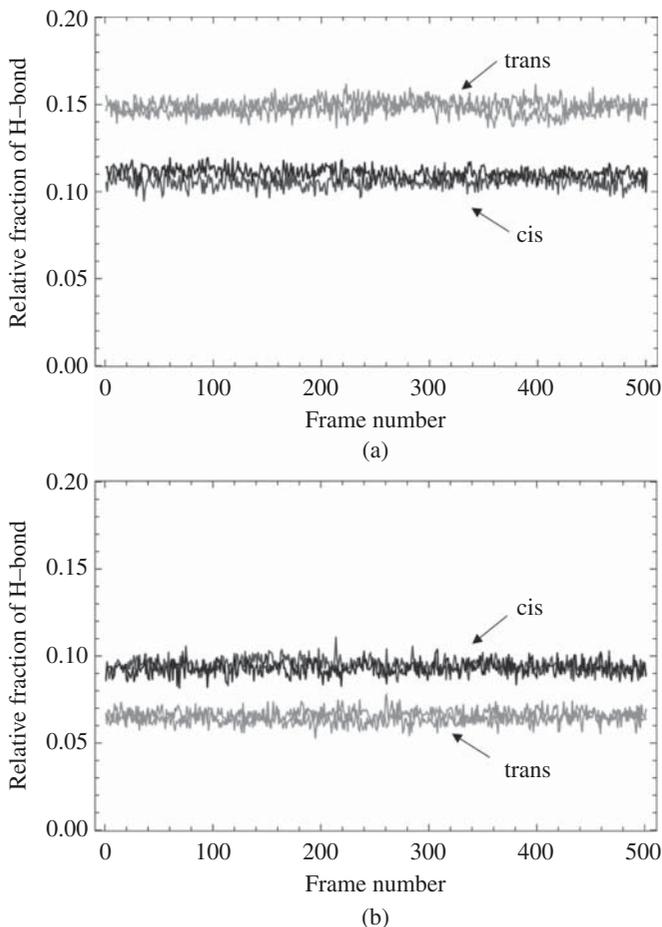


FIGURE 12 Average fraction of H-bonds of the type (a) NH-O=C and (b) NH-Cl^- . *Source:* Perkins et al. 2013⁵⁵.

1 : 2 molar ratio the majority of hydrogen bonds involving Cl^- included interactions with urea. Raman spectroscopy and DFT calculations by Silva et al. also found hydrogen bonding between chloride and urea to be the most important interaction present in CCU.¹⁸¹ When water is introduced into the system, Cl-urea interactions become weaker as water begins to form new intermolecular interactions with each DES component. Figure 13 illustrates

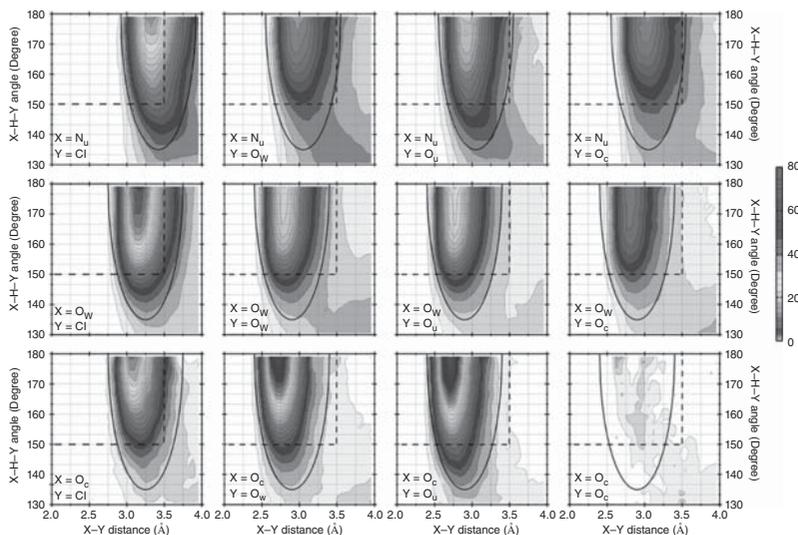


FIGURE 13 CDF to examine hydrogen bonding interactions in the aqueous reline system. Dashed lines show rectangular boundaries for the hydrogen bond criteria of Perkins, Painter, and Colina⁵⁵ while the solid lines show the elliptical boundaries as defined by Wernet et al.¹⁸⁰ *Source:* Fetisov et al. 2018⁴⁵.

the CDFs of the X–H–Y interactions between choline, urea, water, and Cl^- that showcase the ellipsoid criterion that has been proposed by Wernet et al.¹⁸⁰ The ellipsoid criterion has also been used by Doherty and Acevedo³⁵ who used a nonpolarizable FF to study a pure CCU (1 : 2 molar ratio) system. Shown in Figure 14, strong hydrogen bonding is evident between the choline–chloride and urea–chloride, whereas choline and urea rarely interacted through hydrogen bonding.

Perkins et al.^{31,55} also studied the systems CCEtG (1 : 2 molar ratio), CCGly (1 : 2 molar ratio), and CCMal (1 : 2 molar ratio) using the same methods as their CCU (1 : 2 molar ratio) simulations. Following the relative contributions of hydrogen bonds for a given hydrogen bond type, it was observed that the largest fraction of hydrogen bonds were between the HBD and the anion (see Figure 14). CCEtG and CCGly exhibited similar trends with the exception of HBD–HBD interactions due to the extra hydroxyl group of glyceline resulting in a higher fraction. Of these systems, CCMal has the highest viscosity, perhaps a consequence of the strong cation–anion and HBD–anion interactions present, which indicates a very stable hydrogen bond network and limited mobility within the system.

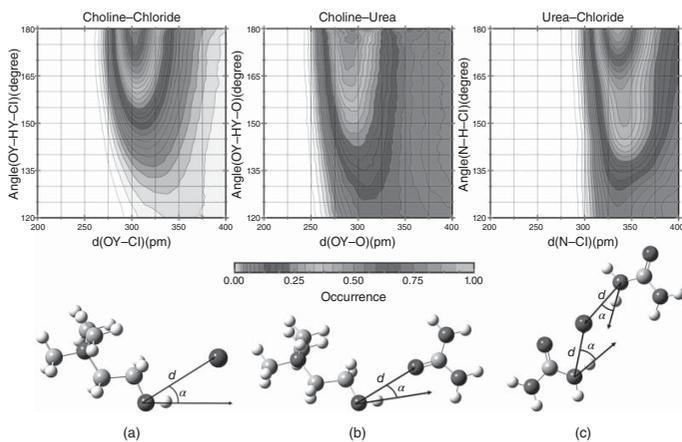


FIGURE 14 CDFs for selected interactions in reline. *Source:* Doherty et al. 2018³⁵.

Additional hydrogen bond analysis can be observed through residence times calculated by the following autocorrelation functional (Eq. [34]).

$$\tau_{\text{HB}} = 2 \cdot \int \frac{\langle h_i(t)h_i(0) \rangle}{\langle h_i(0) \rangle} dt \quad [34]$$

where $h_i(0)$ and $h_i(t)$ are hydrogen bonds at time 0 or time t when given a set criteria. Zahn focused on CCU (1 : 2 molar ratio) using AIMD⁶⁹ and found that the cation–anion hydrogen bonding occurring between the hydroxyl hydrogen and chloride contributed the longest hydrogen bond residence time of 10.3 ps. Interestingly, the second longest residence time was also between the hydroxyl hydrogen of choline and the oxygen of urea with a time of 6.4 ps. These results contradicted reports by Hammond et al.⁴⁶ where the rotation of the OH group in choline was dynamic thus preventing rigid hydrogen bonding. Instead, it was discovered that choline preferred the *gauche* conformation, leading to longer residence times between the cation and both anion and HBD. Sun et al.⁵⁴ also discovered the longest hydrogen bond residence time in reline to be between choline and chloride (12.6 ps), while hydrogen bonding between urea and chloride was about five times shorter (2.4 ps).

CCLev (1 : 2 molar ratio) hydrogen bond lifetimes were also monitored by García et al.⁹⁴ and Ullah et al.³⁷ through MD simulations. Ullah extended the hydrogen bonding criteria to the second solvation shell at a maximum distance of 6.0 Å and found that the cation–anion (Hc-Cl) interaction had the longest lifetime followed by cation–HBD (H1-Oc) and anion–HBD (H1-O11). All other hydrogen bonds in the system were similar and ranged between 35 and 45 ps (Figure 15). García et al. found that the hydrogen bond lifetimes were highly dependent on the charge partitioning scheme to assign charges in the MD simulation.⁹⁴ When considering the top performing charge partitions, however, it was discovered that the cation–anion interaction still remained as the longest hydrogen bond lifetime.

Hydrogen bond lifetimes were also calculated by Zhang et al.³⁶ for CCEtg (1 : 2 molar ratio) using the same FF parameters as Perkins et al.³¹ Interactions between chloride and choline were found to have considerably longer lifetimes compared to the studies mentioned above, with the hydroxyl group of Ch^+ and Cl^- having a lifetime of 1462.4 ± 56.0 ps. The second longest lifetime calculated was between ethylene glycol and chloride with a value of 972.4 ± 27.9 ps. All other hydrogen bond interactions fell into the same range, i.e., 24.1–80.5 ps, as the previous studies mentioned. These results match the same trend seen in other systems, where the cation–anion interactions provide the longest lifetimes.

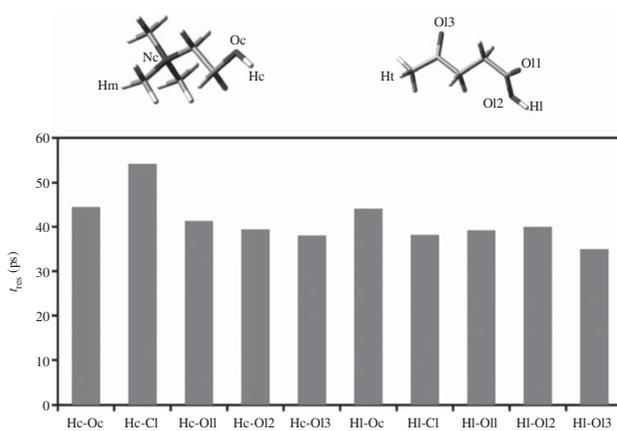


FIGURE 15 Hydrogen bond lifetimes, t_{res} , for selected atoms in CCLev system. *Source:* Ullah et al. 2015³⁷/Royal Society of Chemistry /CC BY 3.0.

Further detail into the dynamics of the CCEtg system was also provided by fitting of the molecular dipole moment correlation function for choline and ethylene glycol to the fractional kinetic Mittag–Leffler model.¹⁸² By using the fractional kinetic model, the dipole relaxation can be separated into a fast mode and slow mode. The fast time process calculated was attributed to the weaker hydrogen bond interactions between choline and ethylene glycol which matches the hydrogen bond lifetimes, while the slower modes were partially attributed to the longer hydrogen bond lifetimes in the interactions with nearby chlorides. These results further connect the considerable influence that the hydrogen bond network has upon system dynamics in DESs.

Spatial Distribution Functions

General structuring of DES systems have been monitored with the help of SDFs which provide a three-dimensional visualization of the distribution of the nearest neighbor to a reference molecule. Systems studied include CCU (1 : 2 molar ratio),^{20, 35, 45, 46, 69, 178} CCLev (1 : 2 molar ratio),^{37, 94} CCEtg (1 : 2 molar ratio), CCGly (1 : 2 molar ratio),^{35, 176} CCPhe (1 : 2 and 1 : 3 molar ratio), CCMal (1 : 2 molar ratio), and CCOx (1 : 2 molar ratio).³⁵ As expected from the RDFs, the chlorine anion resided near the hydrogen bond donating groups of both the cation and HBD, while the cation–HBD interactions resided in the remaining space surrounding the anion (Figure 16). Maintaining this ordering via favorable electrostatic interactions has been suggested to be the driving force for deep eutectic solvent formation.⁴⁶

APPLICATION OF DES THROUGH SIMULATION

Gas Sorption Studies on DES

The rapid increase of atmospheric pollutants through the combustion of fossil fuels has been linked to health effects as well as climate change, leading to the need for environmentally friendly gas capturing techniques.¹⁸³ DESs have been proposed as carbon capture sorbents due to their tunability and natural affinity for gas compounds such as CO₂.^{12, 57, 184} Computational studies have been performed for the choline chloride phenylacetic acid system to investigate the binding modes between the DES components and CO₂ at both the vacuum surface and bulk liquid phase.¹⁶³ Through DFT calculations, CO₂ was found to interact strongly with the COOH group of phenylacetic acid as well as Cl⁻, while the choline cation stabilized the chloride anion through hydrogen bonding with its hydroxyl group. This binding

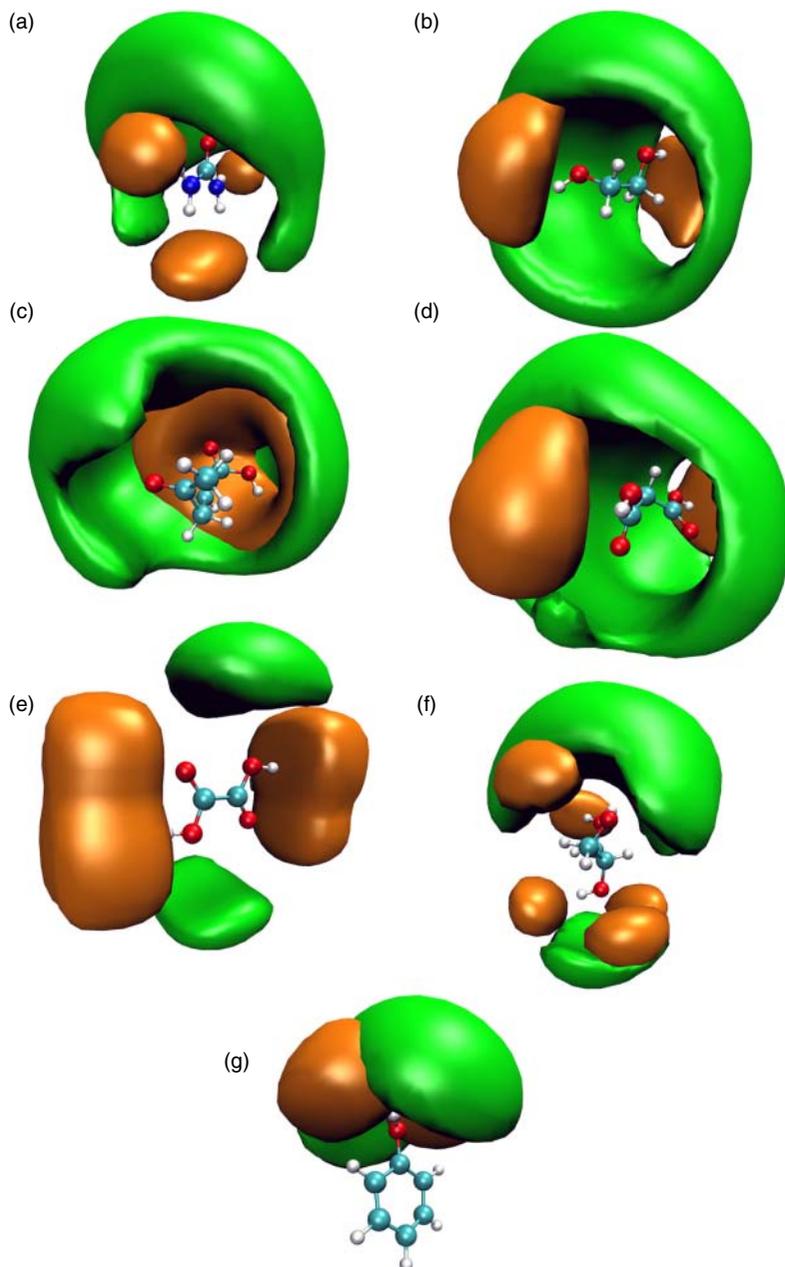


FIGURE 16 Spatial distribution functions for (a) urea, (b) ethylene glycol, (c) levulinic acid, (d) malonic acid, (e) oxalic acid, (f) glycerol, (g) phenol in choline chloride deep eutectic solvents. Orange denotes the position of the chloride anion, and the green depicts the choline cation. *Source:* Doherty et al. 2018³⁵.

motif matched previous DFT studies by the same group for CCLev (1 : 2 molar ratio) and CO₂.³⁷ In the bulk phase, MD simulations were performed at the DES interface where flue gas molecules (N₂, H₂O, CO₂, and O₂) were placed to fill the vacuum to monitor diffusion of each gas into the liquid.¹⁶³ Interestingly, CO₂ absorption was found to occur with very little volume expansion indicating minor rearrangement of the DES components was necessary to maintain the hydrogen bond network. Additionally, water was readily absorbed into the first layer and hindered the diffusion of CO₂ into the bulk region; future designs of DESs for CO₂ capture should take this observation into consideration.

The DES-based capture of SO₂ has also been studied both experimentally and computationally. Experimental studies for the eutectic mixtures of choline chloride with glycerol,¹⁸⁵ levulinic acid,¹⁸⁶ urea, thiourea, malonic acid, and ethylene glycol¹⁸⁷ have all resulted in absorption capacities similar to ILs where the absorption process is reversible and showed no signs of decreasing capacity throughout the absorption–desorption cycle. Interactions between SO₂ and CCGly (1 : 1 molar ratio) have been studied using DFT¹⁸⁸ as well as AIMD¹⁷³ where both studies confirmed that the chloride anion binds to SO₂ which disrupted the OH–anion interactions with choline and glycerol causing the system to become more fluid. Weaker interactions between cation–glycerol and glycerol–glycerol were also disrupted with the addition of SO₂ due to dispersion-like interactions between the gas and the nonpolar groups of choline and glycerol.¹⁷³ Although these interactions are not as significant as the cation–anion and glycerol–anion interactions, Korotkevich et al.¹⁷³ suggested that future design of DESs tuned for SO₂ absorption may be able take advantage of this interaction by expanding the nonpolar regions.

DES Interactions at Metal Surfaces

An attractive property that DESs have over traditional organic solvents is their high conductivity that can lead to catalytic behavior when involved in the electrodeposition of metals.^{24, 26–28, 189, 190} Understanding how DESs nucleate at the surface of metals is crucial for the design of DES-based materials and technology. Thus, studies have emerged focusing upon how DES species orient themselves around metal surfaces.^{191, 192} Experimentally, CCEtg has been studied on the surface of glassy carbon (GC) with the use of polarization modulation infrared reflection absorption spectroscopy.¹⁹³ This spectroscopic study suggested that decreasing the surface potential to -0.6 V caused the choline cation to absorb vertically to the surface by means of the N⁺(CH₃)₃ group, which resulted in a decrease

of the molecular dipole moment. Conversely, when the potential was increased to $E > +0.4$ V, the choline cation was replaced by the chloride anion which formed an adlayer on the GC. In this case, the choline cation was still vertically oriented to maximize the electrostatic interaction with the anion. The electrodeposition of Cu^{2+} onto GC and Pt surfaces has also been investigated in a separate study by Vukmirovic et al.²⁴ Cyclic voltammetry studies revealed that nucleation rates were sluggish on GC electrodes in relation to Pt and are likely due to the nucleation structure on each surface. Although deposition of Cu^{2+} onto Pt in the CCEtg solvent medium exhibited faster kinetics relative to GC, a comparative study was performed replacing CCEtg with a 3 M aqueous solution of NaCl to increase the chloride concentration that resulted in faster mass and charge transfer. This study highlights the importance that although DESs provide a large reduction potential that will ultimately increase cell voltage and energy density in energy storage devices, the kinetic limitations of charge and mass transfer at the electrode surface must first be overcome to compete with current aqueous electrolytes.

The absorption of choline into the (100) surface of a metal was also studied for CCEv on Ag, Al, and Pt using MD methods.⁵⁶ Two distinct absorption layers within 10 Å of the uncharged surfaces was computed for all three metals, where the first layer is primarily composed of levulinic acid and choline cations oriented in a parallel fashion to the surface and the second consisted of excess chloride anions and levulinic acid molecules. Within the absorbed layer, diffusion rates of the ions were significantly lower than that of the bulk solution further emphasizing the strong interactions present between the DES and the metal surface.

Proteins in DES

DESs have been utilized as cosolvents in enzymatic catalysis for completely green chemical processes. For example, DES systems have been shown to drive regioselectivity when combined with the potato epoxide hydrolase StEH1 for the hydrolysis of chiral (1,2)-*trans*-2-methylstyrene oxide.¹⁹⁴ Additionally, eight different DES systems have been shown to stabilize *Candida antarctica* lipase B (CALB) and *C. antarctica* lipase A.¹⁹⁵ The DESs/enzymes provided a 30% enhancement to the production yield of α -monobenzoate glycerol through the esterification between benzoate and glycerol compared to commercially available biocatalysts.¹⁹⁶ Interestingly, CALB has been shown to lose its activity by nearly 70% when pretreated with 10M urea due to protein denaturing. However, when pretreated with CCU, the activity loss was <1%.¹⁹⁵ The origin of why CALB remained

active in a solvent that contained 66% urea was investigated by Monhemi et al.³⁹ through classical MD simulations. The CALB in an 8 M urea simulation found that urea molecules rapidly diffused into the α -Helix5 active site disrupting hydrogen bonding, which resulted in a denaturing process. In the CCU environment, the urea molecules preferred to interact with the choline and Cl ions, allowing the α -Helix5 site to retain its intramolecular hydrogen bonds and remain active. Enzyme stability was also observed in the CCU mixture as the chloride anions formed hydrogen bonds with surface residues of the enzyme leading to the phenomenon known as “enzyme immobilization.” This technique has also been reported for CALB in CCGly where no loss in enzyme activity was observed for up to 14 days.¹⁹⁶ The combination of DESs and biocatalysts is a rapidly developing and exciting field. However, many unanswered molecular level questions remain necessitating the future development and application of novel computational tools.

SUMMARY

Provided in this chapter is a comprehensive overview of DESs and the methods used to study these systems through simulation. The majority of simulations have been performed on Type III DESs that contain choline chloride as the salt and a corresponding organic HBD at specific molar ratios. Ideally, due to the strong polarization present in each of the systems, *ab initio* methods such as DFT, FPMD, and AIMD would be utilized because of their explicit treatment of polarization and many body effects. While *ab initio* methods can provide valuable information such as solvation structure, charge transfer, and at times reactivity, limitations in their trajectory lengths and system sizes call for more computationally affordable methods. Considerable efforts in the development of nonpolarizable FFs for classical MD simulations have been reported that have provided in depth analysis into various DES properties such as density, thermal expansion coefficient, surface tension, heat capacity, enthalpy of vaporization, isothermal compressibility, viscosity, and self-diffusion coefficients. Due to the importance of parameter validation, this chapter provided an overview of how prominent DES FFs performed in reproducing bulk properties and liquid structures measured experimentally or computed using higher-theory *ab initio* methods.

Generally, most published nonpolarizable FFs performed well in reproducing the bulk properties of DESs as outlined in Table 5. Significant improvement was noted when scaling the charges for the various DES systems to mimic charge transfer effects and polarization. In most cases,

a uniform scaling factor between 0.7 and 0.9 was chosen due to reported success in previous applications to ILs. Alternatively, some groups have developed system-specific charge models based on DFT calculation of small DES clusters or charge analysis from AIMD simulations. Success in reproducing bulk properties in MD simulations has also elucidated the structure–property relationship of these solvents by highlighting the importance of the hydrogen bond network that is formed between each component. The hydrogen bond network greatly influences properties such as viscosity and diffusion coefficients, which can limit the application of some DESs as alternative solvents. A major challenge for nonpolarizable FFs was the poor reproduction of self-diffusion coefficients; explicit treatment of polarization effects may be required to improve agreement. Additionally, a drawback of employing a scaled charge model is the treatment of additives that may alter the magnitude of charge transfer. Due to these concerns, a general and completely transferable nonpolarizable FF may be difficult or even impossible to develop for DESs.

ACKNOWLEDGMENTS

Gratitude is expressed to the National Science Foundation (CHE-2102038 for B.D. and O.A., SI2-1613155 for C.M.C and S.J.R.) and the University of Florida Preeminence Initiative for support of this research.

REFERENCES

1. C. J. Clarke, W.-C. Tu, O. Levers, A. Bröhl, and J. P. Hallett, *Chem. Rev.*, **118**, 747–800 (2018). Green and Sustainable Solvents in Chemical Processes.
2. D. V. Wagle, H. Zhao, and G. A. Baker, *Acc. Chem. Res.*, **47**, 2299–2308 (2014). Deep Eutectic Solvents: Sustainable Media for Nanoscale and Functional Materials.
3. R. L. Vekariya, *J. Mol. Liq.*, **227**, 44–60 (2017). A Review of Ionic Liquids: Applications Towards Catalytic Organic Transformations.
4. M. Espino, M. de los Ángeles Fernández, F. J. V. Gomez, and M. F. Silva, *Trends Anal. Chem.*, **76**, 126–136 (2016). Natural Designer Solvents for Greening Analytical Chemistry.
5. M. C. Bubalo, S. Vidović, I. R. Redovniković, and S. Jokić, *J. Chem. Technol. Biotechnol.*, **90**, 1631–1639 (2015). Green Solvents for Green Technologies.
6. H. Zhao and G. A. Baker, *J. Chem. Technol. Biotechnol.*, **88**, 3–12 (2012). Ionic Liquids and Deep Eutectic Solvents for Biodiesel Synthesis: A Review.

7. J. H. Clark, H. F. Sneddon, L. Summerton, and L. C. Jones, *Green and Sustainable Medicinal Chemistry*, Royal Society of Chemistry, 2016.
8. Y. Medina-Gonzalez, S. Camy, and J.-S. Condoret, *ACS Sustain. Chem. Eng.*, **2**, 2623–2636 (2014). ScCO_2 /Green Solvents: Biphasic Promising Systems for Cleaner Chemicals Manufacturing.
9. J. I. García, H. García-Marín, and E. Pires, *Green Chem.*, **16**, 1007–1033 (2014). Glycerol Based Solvents: Synthesis, Properties and Applications.
10. E. Knipping, C. Aucher, G. Guirado, and L. Aubouy, *New J. Chem.*, **42**, 4693–4699 (2018). Room Temperature Ionic Liquids Versus Organic Solvents as Lithium–Oxygen Battery Electrolytes.
11. A. P. Abbott, G. Capper, D. L. Davies, H. L. Munro, R. K. Rasheed, and V. Tambyrajah, *Chem. Commun.*, 2010–2011 (2001). Preparation of Novel, Moisture-Stable, Lewis-Acidic Ionic Liquids Containing Quaternary Ammonium Salts with Functional Side Chains.
12. A. Paiva, R. Craveiro, I. Aroso, M. Martins, R. L. Reis, and A. R. C. Duarte, *ACS Sustain. Chem. Eng.*, **2**, 1063–1071 (2014). Natural Deep Eutectic Solvents – Solvents for the 21st Century.
13. E. L. Smith, A. P. Abbott, and K. S. Ryder, *Chem. Rev.*, **114**, 11060–11082 (2014). Deep Eutectic Solvents (Dess) and Their Applications.
14. T. Altamash, M. S. Nasser, Y. Elhamarnah, M. Magzoub, R. Ullah, B. Anaya, S. Aparicio, and M. Atilhan, *ChemistrySelect*, **2**, 7278–7295 (2017). Gas Solubility and Rheological Behavior of Natural Deep Eutectic Solvents (Nades) Via Combined Experimental and Molecular Simulation Techniques.
15. D. V. Wagle, L. Adhikari, and G. A. Baker, *Fluid Phase Equilib.*, **448**, 50–58 (2017). Computational Perspectives on Structure, Dynamics, Gas Sorption, and Bio-Interactions in Deep Eutectic Solvents.
16. P. Liu, J.-W. Hao, L.-P. Mo, and Z.-H. Zhang, *RSC Adv.*, **5**, 48675–48704 (2015). Recent Advances in the Application of Deep Eutectic Solvents as Sustainable Media as Well as Catalysts in Organic Reactions.
17. D. Z. Troter, Z. B. Todorović, D. R. Đokić-Stojanović, O. S. Stamenković, and V. B. Veljković, *Renew. Sustain. Energy Rev.*, **61**, 473–500 (2016). Application of Ionic Liquids and Deep Eutectic Solvents in Biodiesel Production: A Review.
18. L. Wei, Y.-J. Fan, H.-H. Wang, N. Tian, Z.-Y. Zhou, and S.-G. Sun, *Electrochim. Acta*, **76**, 468–474 (2012). Electrochemically Shape-Controlled Synthesis in Deep Eutectic Solvents of Pt Nanoflowers with Enhanced Activity for Ethanol Oxidation.
19. S. Sarmad, J.-P. Mikkola, and X. Ji, *ChemSusChem*, **10**, 324–352 (2017). Carbon Dioxide Capture with Ionic Liquids and Deep Eutectic Solvents: A New Generation of Sorbents.

20. G. García, M. Atilhan, and S. Aparicio, *J. Phys. Chem. C*, **119**, 21413–21425 (2015). Interfacial Properties of Deep Eutectic Solvents Regarding to CO₂ Capture.
21. A. P. Abbott, A. Ballantyne, R. C. Harris, J. A. Juma, K. S. Ryder, and G. Forrest, *Electrochim. Acta*, **176**, 718–726 (2015). A Comparative Study of Nickel Electrodeposition Using Deep Eutectic Solvents and Aqueous Solutions.
22. A. J. Goddard, R. C. Harris, S. Saleem, M. Azam, C. Hood, D. Clark, J. Satchwell, and K. S. Ryder, *Trans. IMF*, **95**, 137–146 (2017). Electropolishing and Electrolytic Etching of Ni-Based Hip Consolidated Aerospace Forms: A Comparison between Deep Eutectic Solvents and Aqueous Electrolytes.
23. C. Li, J. Zhang, Z. Li, J. Yin, Y. Cui, Y. Liu, and G. Yang, *Green Chem.*, **18**, 3789–3795 (2016). Extraction Desulfurization of Fuels with ‘Metal Ions’ Based Deep Eutectic Solvents (MDESS).
24. M. B. Vukmirovic, R. R. Adzic, and R. Akolkar, *J. Phys. Chem. B*, **124**, 5465–5475 (2020). Copper Electrodeposition from Deep Eutectic Solvents-Voltammetric Studies Providing Insights into the Role of Substrate: Platinum Vs Glassy Carbon.
25. X. Shen, N. Sinclair, J. Wainright, R. Akolkar, and R. F. Savinell, *J. Electrochem. Soc.*, **167**, 086509 (2020). Evaluating and Developing a Reliable Reference Electrode for Choline Chloride Based Deep Eutectic Solvents.
26. N. A. Shaheen, I. Mahesh, M. B. Vukmirovic, and R. Akolkar, *Electrochem. Commun.*, **115**, 106721 (2020). Hysteresis Effects and Roughness Suppression Efficacy of Polyethylenimine Additive in Cu Electrodeposition in Ethaline.
27. D. Shen, M. B. Vukmirovic, and R. Akolkar, *J. Electrochem. Soc.*, **166**, E526–E532 (2019). Understanding the Role of Complexation in the Charge-Transfer Kinetics of the $\text{Cu}^{2+} + e \leftrightarrow \text{Cu}^{1+}$ Redox Reaction in Ethaline Deep Eutectic Solvent.
28. D. Shen, K. Steinberg, and R. Akolkar, *J. Electrochem. Soc.*, **165**, E808–E815 (2018). Avoiding Pitfalls in the Determination of Reliable Electrochemical Kinetics Parameters for the $\text{Cu}^{2+} \rightarrow \text{Cu}^{1+}$ Reduction Reaction in Deep Eutectic Solvents.
29. I. I. I. Alkhatib, D. Bahamon, F. Llovel, M. R. M. Abu-Zahra, and L. F. Vega, *J. Mol. Liq.*, **298**, 112183 (2020). Perspectives and Guidelines on Thermodynamic Modelling of Deep Eutectic Solvents.
30. G. García, M. Atilhan, and S. Aparicio, *Chem. Phys. Lett.*, **634**, 151–155 (2015). An Approach for the Rationalization of Melting Temperature for Deep Eutectic Solvents from DFT.
31. S. L. Perkins, P. Painter, and C. M. Colina, *J. Chem. Eng. Data*, **59**, 3652–3662 (2014). Experimental and Computational Studies of Choline Chloride-Based Deep Eutectic Solvents.
32. S. Mainberger, M. Kindlein, F. Bezold, E. Elts, M. Minceva, and H. Briesen, *Mol. Phys.*, **115**, 1309–1321 (2017). Deep Eutectic Solvent Formation: A

- Structural View Using Molecular Dynamics Simulations with Classical Force Fields.
33. C. R. Ashworth, R. P. Matthews, T. Welton, and P. A. Hunt, *Phys. Chem. Chem. Phys.*, **18**, 18145–18160 (2016). Doubly Ionic Hydrogen Bond Interactions within the Choline Chloride–Urea Deep Eutectic Solvent.
 34. S. Zahn, B. Kirchner, and D. Mollenhauer, *ChemPhysChem*, **17**, 3354–3358 (2016). Charge Spreading in Deep Eutectic Solvents.
 35. B. Doherty and O. Acevedo, *J. Phys. Chem. B*, **122**, 9982–9993 (2018). OPLS Force Field for Choline Chloride-Based Deep Eutectic Solvents.
 36. Y. Zhang, D. Poe, L. Heroux, H. Squire, B. W. Doherty, Z. Long, M. Dadmun, B. Gurkan, M. E. Tuckerman, and E. J. Maginn, *J. Phys. Chem. B*, **124**, 5251–5264 (2020). Liquid Structure and Transport Properties of the Deep Eutectic Solvent Ethaline.
 37. R. Ullah, M. Atilhan, B. Anaya, M. Khraisheh, G. García, A. ElKhattat, M. Tariq, and S. Aparicio, *Phys. Chem. Chem. Phys.*, **17**, 20941–20960 (2015). A Detailed Study of Cholinium Chloride and Levulinic Acid Deep Eutectic Solvent System for CO₂ Capture Via Experimental and Molecular Simulation Approaches.
 38. J. M. Rimsza and L. René Corrales, *Comput. Theor. Chem.*, **987**, 57–61 (2012). Adsorption Complexes of Copper and Copper Oxide in the Deep Eutectic Solvent 2:1 Urea–Choline Chloride.
 39. H. Monhemi, M. R. Housaindokht, A. A. Moosavi-Movahedi, and M. R. Bozorgmehr, *Phys. Chem. Chem. Phys.*, **16**, 14882–14893 (2014). How a Protein Can Remain Stable in a Solvent with High Content of Urea: Insights from Molecular Dynamics Simulation of Candida Antarctica Lipase B in Urea: Choline Chloride Deep Eutectic Solvent.
 40. A. P. Abbott, G. Capper, D. L. Davies, R. K. Rasheed, and V. Tambyrajah, *Chem. Commun.*, 70–71 (2003). Novel Solvent Properties of Choline Chloride/Urea Mixtures.
 41. Q. Zhang, K. De Oliveira Vigier, S. Royer, and F. Jérôme, *Chem. Soc. Rev.*, **41**, 7108–7146 (2012). Deep Eutectic Solvents: Syntheses, Properties and Applications.
 42. A. P. Abbott, D. Boothby, G. Capper, D. L. Davies, and R. K. Rasheed, *J. Am. Chem. Soc.*, **126**, 9142–9147 (2004). Deep Eutectic Solvents Formed between Choline Chloride and Carboxylic Acids: Versatile Alternatives to Ionic Liquids.
 43. Y. Hou, Y. Gu, S. Zhang, F. Yang, H. Ding, and Y. Shan, *J. Mol. Liq.*, **143**, 154–159 (2008). Novel Binary Eutectic Mixtures Based on Imidazole.
 44. C. D’Agostino, R. C. Harris, A. P. Abbott, L. F. Gladden, and M. D. Mantle, *Phys. Chem. Chem. Phys.*, **13**, 21383–21391 (2011). Molecular Motion and Ion Diffusion in Choline Chloride Based Deep Eutectic Solvents Studied by 1 H Pulsed Field Gradient NMR Spectroscopy.

45. E. O. Fetisov, D. B. Harwood, I. F. W. Kuo, S. E. E. Warrag, M. C. Kroon, C. J. Peters, and J. I. Siepmann, *J. Phys. Chem. B*, **122**, 1245–1254 (2018). First-Principles Molecular Dynamics Study of a Deep Eutectic Solvent: Choline Chloride/Urea and Its Mixture with Water.
46. O. S. Hammond, D. T. Bowron, and K. J. Edler, *Green Chem.*, **18**, 2736–2744 (2016). Liquid Structure of the Choline Chloride-Urea Deep Eutectic Solvent (Reline) from Neutron Diffraction and Atomistic Modelling.
47. C. Ruß and B. König, *Green Chem.*, **14**, 2969 (2012). Low Melting Mixtures in Organic Synthesis – an Alternative to Ionic Liquids?
48. D. Carriazo, M. C. Serrano, M. C. Gutierrez, M. L. Ferrer, and F. del Monte, *ChemInform*, **43**, 4996–5019 (2012). Cheminform Abstract: Deep-Eutectic Solvents Playing Multiple Roles in the Synthesis of Polymers and Related Materials.
49. A. P. Abbott, J. C. Barron, K. S. Ryder, and D. Wilson, *Chem. Eur. J.*, **13**, 6495–6501 (2007). Eutectic-Based Ionic Liquids with Metal-Containing Anions and Cations.
50. L. I. N. Tomé, V. Baião, W. da Silva, and C. M. A. Brett, *Appl. Mater. Today*, **10**, 30–50 (2018). Deep Eutectic Solvents for the Production and Application of New Materials.
51. B. Gurkan, H. Squire, and E. Pentzer, *J. Phys. Chem. Lett.*, **10**, 7956–7964 (2019). Metal-Free Deep Eutectic Solvents: Preparation, Physical Properties, and Significance.
52. B. Chen, S. Mitchell, N. Sinclair, J. Wainright, E. Pentzer, and B. Gurkan, *Mol. Syst. Des. Eng.*, **5**, 1147–1157 (2020). Feasibility of Tempo-Functionalized Imidazolium, Ammonium and Pyridinium Salts as Redox-Active Carriers in Ethaline Deep Eutectic Solvent for Energy Storage.
53. J. M. Klein, H. Squire, W. Dean, and B. E. Gurkan, *J. Phys. Chem. B*, **124**(29), 6348–6357 (2020). From Salt in Solution to Solely Ions - Solvation of Methyl Viologen in Deep Eutectic Solvents and Ionic Liquids.
54. H. Sun, Y. Li, X. Wu, and G. Li, *J. Mol. Model.*, **19**, 2433–2441 (2013). Theoretical Study on the Structures and Properties of Mixtures of Urea and Choline Chloride.
55. S. L. Perkins, P. Painter, and C. M. Colina, *J. Phys. Chem. B*, **117**, 10250–10260 (2013). Molecular Dynamic Simulations and Vibrational Analysis of an Ionic Liquid Analogue.
56. M. Atilhan and S. Aparicio, *J. Phys. Chem. C*, **120**, 10400–10409 (2016). Deep Eutectic Solvents on the Surface of Face Centered Cubic Metals.
57. G. García, S. Aparicio, R. Ullah, and M. Atilhan, *Energy Fuels*, **29**, 2616–2644 (2015). Deep Eutectic Solvents: Physicochemical Properties and Gas Separation Applications.

58. S. Kristyán and P. Pulay, *Chem. Phys. Lett.*, **229**, 175–180 (1994). Can (Semi)Local Density Functional Theory Account for the London Dispersion Forces?
59. J. Pérez-Jordá and A. D. Becke, *Chem. Phys. Lett.*, **233**, 134–137 (1995). A Density-Functional Study of Van Der Waals Forces: Rare Gas Diatomics.
60. P. Hobza, J. Spöner, and T. Reschel, *J. Comput. Chem.*, **16**, 1315–1325 (1995). Density Functional Theory and Molecular Clusters.
61. S. Grimme, *WIREs Comput. Mol. Sci.*, **1**, 211–228 (2011). Density Functional Theory with London Dispersion Corrections.
62. S. Grimme and S. Ehrlich, L. Goerigk, *J. Comput. Chem.*, **32**, 1456–1465 (2011). Effect of the Damping Function in Dispersion Corrected Density Functional Theory.
63. D. V. Wagle, G. A. Baker, and E. Mamontov, *J. Phys. Chem. Lett.*, **6**, 2924–2928 (2015). Differential Microscopic Mobility of Components within a Deep Eutectic Solvent.
64. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, **120**, 215–241 (2008). The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals.
65. Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.*, **41**, 157–167 (2008). Density Functionals with Broad Applicability in Chemistry.
66. M. Walker, A. J. A. Harvey, A. Sen, and C. E. H. Dessent, *J. Phys. Chem. A*, **117**, 12590–12600 (2013). Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better Than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions.
67. D. V. Wagle, C. A. Deakyne, and G. A. Baker, *J. Phys. Chem. B*, **120**, 6739–6746 (2016). Quantum Chemical Insight into the Interactions and Thermodynamics Present in Choline Chloride Based Deep Eutectic Solvents.
68. P. Bultinck, C. Van Alsenoy, P. W. Ayers, and R. Carbó-Dorca, *J. Chem. Phys.*, **126**, 144111 (2007). Critical Analysis and Extension of the Hirshfeld Atoms in Molecules.
69. S. Zahn, *Phys. Chem. Chem. Phys.*, **19**, 4041–4047 (2017). Deep Eutectic Solvents: Similia Similibus Solvuntur?
70. C. J. Martin and D. A. O'Connor, *J. Phys. C: Solid State Phys.*, **10**, 3521–3526 (1977). An Experimental Test of Lindemann's Melting Law.
71. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, *J. Comput. Chem.*, **25**, 1157–1174 (2004). Development and Testing of a General Amber Force Field.

72. S. L. Mayo, B. D. Olafson, and W. A. Goddard, *J. Phys. Chem.*, **94**, 8897–8909 (1990). Dreiding: A Generic Force Field for Molecular Simulations.
73. S. W. I. Siu, K. Pluhackova, and R. A. Böckmann, *J. Chem. Theory Comput.*, **8**, 1459–1470 (2012). Optimization of the OPLS-AA Force Field for Long Hydrocarbons.
74. D. Kony, W. Damm, S. Stoll, and W. F. Van Gunsteren, *J. Comput. Chem.*, **23**, 1416–1429 (2002). An Improved OPLS-AA Force Field for Carbohydrates.
75. S. V. Sambasivarao and O. Acevedo, *J. Chem. Theory Comput.*, **5**, 1038–1050 (2009). Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids.
76. J. N. Canongia Lopes, J. Deschamps, and A. A. H. Pádua, *J. Phys. Chem. B*, **108**, 2038–2047 (2004). Modeling Ionic Liquids Using a Systematic All-Atom Force Field.
77. W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.*, **118**, 11225–11236 (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids.
78. D. Shah and F. S. Mjalli, *Phys. Chem. Chem. Phys.*, **16**, 23900–23907 (2014). Effect of Water on the Thermo-Physical Properties of Reline: An Experimental and Molecular Simulation Based Approach.
79. T. A. Halgren, *J. Comput. Chem.*, **17**, 490–519 (1996). Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94.
80. T. A. Halgren, *J. Comput. Chem.*, **17**, 520–552 (1996). Merck Molecular Force Field. II. MMFF94 Van Der Waals and Electrostatic Parameters for Intermolecular Interactions.
81. T. A. Halgren, *J. Comput. Chem.*, **17**, 553–586 (1996). Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94.
82. H. Sun, *J. Phys. Chem. B*, **102**, 7338–7364 (1998). Compass: An Ab Initio Force-Field Optimized for Condensed-Phase Applications overview with Details on Alkane and Benzene Compounds.
83. H. Sun, *J. Comput. Chem.*, **15**, 752–768 (1994). Force Field for Computation of Conformational Energies, Structures, and Vibrational Frequencies of Aromatic Polyesters.
84. T. I. Morrow and E. J. Maginn, *J. Phys. Chem. B*, **107**, 9160–9160 (2003). Molecular Dynamics Study of the Ionic Liquid 1-N-Butyl-3-Methylimidazolium Hexafluorophosphate.
85. A. Chaumont, R. Schurhammer, and G. Wipff, *J. Phys. Chem. B*, **109**, 18964–18973 (2005). Aqueous Interfaces with Hydrophobic Room-Temperature Ionic Liquids: A Molecular Dynamics Study.
86. G. Chevrot, R. Schurhammer, and G. Wipff, *Phys. Chem. Chem. Phys.*, **8**, 4166–4174 (2006). Molecular Dynamics Simulations of the Aqueous Interface with the [BMI][PF₆] Ionic Liquid: Comparison of Different Solvent Models.

87. B. L. Bhargava and S. Balasubramanian, *J. Chem. Phys.*, **127**, 114510 (2007). Refined Potential Model for Atomistic Simulations of Ionic Liquid [BMIM][PF₆].
88. H. Liu, E. Maginn, A. E. Visser, N. J. Bridges, and E. B. Fox, *Ind. Eng. Chem. Res.*, **51**, 7242–7254 (2012). Thermal and Transport Properties of Six Ionic Liquids: An Experimental and Molecular Dynamics Study.
89. V. Chaban, *Phys. Chem. Chem. Phys.*, **13**, 16055–16062 (2011). Polarizability Versus Mobility: Atomistic Force Field for Ionic Liquids.
90. K. Cui, A. Yethiraj, and J. R. Schmidt, *J. Phys. Chem. B*, **123**, 9222–9229 (2019). Influence of Charge Scaling on the Solvation Properties of Ionic Liquid Solutions.
91. B. Doherty, X. Zhong, S. Gathiaka, B. Li, and O. Acevedo, *J. Chem. Theory Comput.*, **13**, 6131–6145 (2017). Revisiting OPLS Force Field Parameters for Ionic Liquid Simulations.
92. B. Doherty, X. Zhong, and O. Acevedo, *J. Phys. Chem. B*, **122**, 2962–2974 (2018). Virtual Site OPLS Force Field for Imidazolium-Based Ionic Liquids.
93. J. Rigby and E. I. Izgorodina, *Phys. Chem. Chem. Phys.*, **15**, 1632–1646 (2013). Assessment of Atomic Partial Charge Schemes for Polarisation and Charge Transfer Effects in Ionic Liquids.
94. G. García, M. Atilhan, and S. Aparicio, *J. Mol. Liq.*, **211**, 506–514 (2015). The Impact of Charges in Force Field Parameterization for Molecular Dynamics Simulations of Deep Eutectic Solvents.
95. D. Bedrov, J.-P. Piquemal, O. Borodin, A. D. MacKerell, Jr., B. Roux, and C. Schröder, *Chem. Rev.*, **119**, 7940–7995 (2019). Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields.
96. M. Atilhan and S. Aparicio, *J. Mol. Liq.*, **283**, 147–154 (2019). Molecular Dynamics Simulations of Mixed Deep Eutectic Solvents and Their Interaction with Nanomaterials.
97. E. E. L. Tanner, K. M. Piston, H. Ma, K. N. Ibsen, S. Nangia, and S. Mitragotri, *ACS Biomater. Sci. Eng.*, **5**, 3645–3653 (2019). The Influence of Water on Choline-Based Ionic Liquids.
98. E. S. C. Ferreira, I. V. Voroshylova, C. M. Pereira, and M. N. D. S. Cordeiro, *J. Phys. Chem. B*, **120**, 10124–10137 (2016). Improved Force Field Model for the Deep Eutectic Solvent Ethaline: Reliable Physicochemical Properties.
99. E. S. C. Ferreira, I. V. Voroshylova, N. M. Figueiredo, C. M. Pereira, and M. N. D. S. Cordeiro, *J. Mol. Liq.*, **298**, 111978 (2020). Computational and Experimental Study of Propeline: A Choline Chloride Based Deep Eutectic Solvent.
100. K. Górný, Z. Dendzik, M. Pabiszczak, and Z. Gburski, *J. Non-Cryst. Solids*, **364**, 28–33 (2013). Non-Debye Dipolar Relaxation of Ethylene Glycol Embedded in ZSM-5 Zeolite Host Matrix — Computer Simulation Study.

101. A. D. Mackerell, Jr., M. Feig, and C. L. Brooks, 3rd, *J. Comput. Chem.*, **25**, 1400–1415 (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.
102. E. S. C. Ferreira, I. V. Voroshylova, V. A. Koverga, C. M. Pereira, and M. N. D. S. Cordeiro, *J. Phys. Chem. B*, **121**, 10906–10921 (2017). New Force Field Model for Propylene Glycol: Insight to Local Structure and Dynamics.
103. A. K. Soper, E. W. Castner, and A. Luzar, *Biophys. Chem.*, **105**, 649–666 (2003). Impact of Urea on Water Structure: A Clue to Its Properties as a Denaturant?
104. H. S. Salehi, M. Ramdin, O. A. Moutos, and T. J. H. Vlugt, *Fluid Phase Equilib.*, **497**, 10–18 (2019). Computing Solubility Parameters of Deep Eutectic Solvents from Molecular Dynamics Simulations.
105. K. Shahbaz, S. Baroutian, F. S. Mjalli, M. A. Hashim, and I. M. AlNashef, *Thermochim. Acta*, **527**, 59–66 (2012). Densities of Ammonium and Phosphonium Based Deep Eutectic Solvents: Prediction Using Artificial Intelligence and Group Contribution Techniques.
106. R. B. Leron and M.-H. Li, *Thermochim. Acta*, **530**, 52–57 (2012). Molar Heat Capacities of Choline Chloride-Based Deep Eutectic Solvents and Their Binary Mixtures with Water.
107. R. B. Leron, D. S. H. Wong, and M.-H. Li, *Fluid Phase Equilib.*, **335**, 32–38 (2012). Densities of a Deep Eutectic Solvent Based on Choline Chloride and Glycerol and Its Aqueous Mixtures at Elevated Pressures.
108. Y. Xie, H. Dong, S. Zhang, X. Lu, and X. Ji, *J. Chem. Eng. Data*, **59**, 3344–3352 (2014). Effect of Water on the Density, Viscosity, and CO₂ solubility in Choline Chloride/Urea.
109. F. Chemat, H. Anjum, A. M. Shariff, P. Kumar, and T. Murugesan, *J. Mol. Liq.*, **218**, 301–308 (2016). Thermal and Physical Properties of (Choline Chloride Urea L-Arginine) Deep Eutectic Solvents.
110. C. Florindo, F. S. Oliveira, L. P. N. Rebelo, A. M. Fernandes, and I. M. Marrucho, *ACS Sustain. Chem. Eng.*, **2**, 2416–2425 (2014). Insights into the Synthesis and Properties of Deep Eutectic Solvents Based on Cholinium Chloride and Carboxylic Acids.
111. A. Yadav and S. Pandey, *J. Chem. Eng. Data*, **59**, 2221–2229 (2014). Densities and Viscosities of (Choline Chloride + Urea) Deep Eutectic Solvent and Its Aqueous Mixtures in the Temperature Range 293.15 K to 363.15 K.
112. H. Shekaari, M. T. Zafarani-Moattar, and B. Mohammadi, *J. Mol. Liq.*, **243**, 451–461 (2017). Thermophysical Characterization of Aqueous Deep Eutectic Solvent (Choline Chloride/Urea) Solutions in Full Ranges of Concentration at T = (293.15–323.15) K.
113. F. S. Mjalli and N. M. Abdel Jabbar, *Fluid Phase Equilib.*, **381**, 71–76 (2014). Acoustic Investigation of Choline Chloride Based Ionic Liquids Analogs.

114. F. S. Mjalli, G. Vakili-Nezhaad, K. Shahbaz, and I. M. AlNashef, *Thermochim. Acta*, **575**, 40–44 (2014). Application of the Eötvös and Guggenheim Empirical Rules for Predicting the Density and Surface Tension of Ionic Liquids Analogues.
115. O. Ciocirlan, O. Iulian, and O. Croitoru, *Rev. Chim.*, **62**, 721–723 (2010). Effect of Temperature on the Physico-Chemical Properties of Three Ionic Liquids Containing Choline.
116. A. Yadav, J. R. Kar, M. Verma, S. Naqvi, and S. Pandey, *Thermochim. Acta*, **600**, 95–101 (2015). Densities of Aqueous Mixtures of (Choline Chloride Ethylene Glycol) and (Choline Chloride Malonic Acid) Deep Eutectic Solvents in Temperature Range 283.15–363.15 k.
117. A. P. Abbott, G. Capper, D. L. Davies, K. J. McKenzie, and S. U. Obi, *J. Chem. Eng. Data*, **51**, 1280–1282 (2006). Solubility of Metal Oxides in Deep Eutectic Solvents Based on Choline Chloride.
118. M. E. Senko and D. H. Templeton, *Acta Crystallogr.*, **13**, 281–285 (1960). Unit Cells of Choline Halides and Structure of Choline Chloride.
119. K. Shahbaz, F. S. Mjalli, M. A. Hashim, and I. M. AlNashef, *Thermochim. Acta*, **515**, 67–72 (2011). Prediction of Deep Eutectic Solvents Densities at Different Temperatures.
120. R. Haghbakhsh, M. Taherzadeh, A. R. C. Duarte, and S. Raeissi, *J. Mol. Liq.*, **307**, 112972 (2020). A General Model for the Surface Tensions of Deep Eutectic Solvents.
121. K. Shahbaz, F. S. Mjalli, M. A. Hashim, and I. M. AlNashef, *Fluid Phase Equilib.*, **319**, 48–54 (2012). Prediction of the Surface Tension of Deep Eutectic Solvents.
122. D. B. Macleod, *J. Chem. Soc., Faraday Trans.*, **19**, 38 (1923). On a Relation between Surface Tension and Density.
123. S. Sugden, *J. Chem. Soc. Trans.*, **125**, 32–41 (1924). Vi.—The Variation of Surface Tension with Temperature and Some Related Functions.
124. O. R. Quayle, *Chem. Rev.*, **53**, 439–589 (1953). The Parachors of Organic Compounds. An Interpretation and Catalogue.
125. S. A. Mumford and J. W. C. Phillips, *J. Chem. Soc.*, 2112–2133 (1929). CCLXXIV.—The Evaluation and Interpretation of Parachors.
126. T. A. Knotts, W. Vincent Wilding, J. L. Oscarson, and R. L. Rowley, *J. Chem. Eng. Data*, **46**, 1007–1012 (2001). Use of the Dippr Database for Development of Qspr Correlations: Surface Tension†.
127. M. Deetlefs, K. R. Seddon, and M. Shara, *Phys. Chem. Chem. Phys.*, **8**, 642–649 (2006). Predicting Physical Properties of Ionic Liquids.
128. R. L. Gardas and J. A. P. Coutinho, *Fluid Phase Equilib.*, **265**, 57–65 (2008). Applying a QSPR Correlation to the Prediction of Surface Tensions of Ionic Liquids.

129. M. Lagache, P. Ungerer, A. Boutin, and A. H. Fuchs, *Phys. Chem. Chem. Phys.*, **3**, 4333–4339 (2001). Prediction of Thermodynamic Derivative Properties of Fluids by Monte Carlo Simulation.
130. C. Cadena, Q. Zhao, R. Q. Snurr, and E. J. Maginn, *J. Phys. Chem. B*, **110**, 2821–2832 (2006). Molecular Modeling and Experimental Studies of the Thermodynamic and Transport Properties of Pyridinium-Based Ionic Liquids.
131. K. E. Gutowski, B. Gurkan, and E. J. Maginn, *Pure Appl. Chem.*, **81**, 1799–1828 (2009). Force Field for the Atomistic Simulation of the Properties of Hydrazine, Organic Hydrazine Derivatives, and Energetic Hydrazinium Ionic Liquids.
132. P. H. Berens, D. H. J. Mackay, G. M. White, and K. R. Wilson, *J. Chem. Phys.*, **79**, 2375–2389 (1983). Thermodynamics and Quantum Corrections from Molecular Dynamics for Liquid Water.
133. M. Chen, R. Pendrill, G. Widmalm, J. W. Brady, and J. Wohlert, *J. Chem. Theory Comput.*, **10**, 4465–4479 (2014). Molecular Dynamics Simulations of the Ionic Liquid 1-N-Butyl-3-Methylimidazolium Chloride and Its Binary Mixtures with Ethanol.
134. S.-T. Lin, M. Blanco, and W. A. Goddard, *J. Chem. Phys.*, **119**, 11792–11805 (2003). The Two-Phase Model for Calculating Thermodynamic Properties of Liquids from Molecular Dynamics: Validation for the Phase Diagram of Lennard-Jones Fluids.
135. M. Taherzadeh, R. Haghbakhsh, A. R. C. Duarte, and S. Raeissi, *J. Mol. Liq.*, **307**, 112940 (2020). Estimation of the Heat Capacities of Deep Eutectic Solvents.
136. S. Ravula, N. E. Larm, M. A. Mottaleb, M. P. Heitz, and G. A. Baker, *ChemEngineering*, **3**, 42 (2019). Vapor Pressure Mapping of Ionic Liquids and Low-Volatility Fluids Using Graded Isothermal Thermogravimetric Analysis.
137. K. Shahbaz, F. S. Mjalli, G. Vakili-Nezhaad, I. M. AlNashef, A. Asadov, and M. M. Farid, *J. Mol. Liq.*, **222**, 61–66 (2016). Thermogravimetric Measurement of Deep Eutectic Solvents Vapor Pressure.
138. K. A. Motakabbir and M. Berkowitz, *J. Phys. Chem.*, **94**, 8359–8362 (1990). Isothermal Compressibility of SPC/E Water.
139. J. H. Dymond and R. Malhotra, *Int. J. Thermophys.*, **9**, 941–951 (1988). The Tait Equation: 100 Years On.
140. J. K. Shah, J. F. Brennecke, and E. J. Maginn, *Green Chem.*, **4**, 112–118 (2002). Thermodynamic Properties of the Ionic Liquid 1-N-Butyl-3-Methylimidazolium Hexafluorophosphate from Monte Carlo Simulations.
141. S. M. Hosseini, S. Aparicio, M. M. Alavianmehr, and R. Khalifeh, *J. Mol. Liq.*, **266**, 751–761 (2018). On the Volumetric Properties of 2-Hydroxy Ethylammonium Formate Ionic Liquid under High-Pressures: Measurement and Molecular Dynamics.

142. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, 1989.
143. C. L. Diel and V. F. Cabral, *J. Mol. Liq.*, **282**, 226–234 (2019). Equilibrium Properties of Protic Ionic Liquids Based on Metil-2-Hydroxyethylammonium Cation.
144. Z. Pouramini, A. Mohebbi, and M. H. Kowsari, *J. Mol. Liq.*, **246**, 39–47 (2017). Atomistic Insights into the Thermodynamics, Structure, and Dynamics of Ionic Liquid 1-Hexyl-3-Methylimidazolium Hexafluorophosphate Via Molecular Dynamics Study.
145. H. Sakuma and M. Ichiki, *Geofluids*, **16**, 89–102 (2016). Density and Isothermal Compressibility of Supercritical H₂O–NaCl Fluid: Molecular Dynamics Study from 673 to 2000 K, 0.2 to 2 Gpa, and 0 to 22 Wt% NaCl Concentrations.
146. E. S. Minina, E. S. Pyanzina, E. V. Novak, and S. S. Kantorovich, *Eur. Phys. J. E Soft Matter*, **41**, 67 (2018). Compressibility of Ferrofluids: Towards a Better Understanding of Structural Properties.
147. B. Szeferczyk and M. N. D. S. Cordeiro, *J. Phys. Chem. B*, **115**, 3013–3019 (2011). Physical Properties at the Base for the Development of an All-Atom Force Field for Ethylene Glycol.
148. W. Guo, Y. Hou, S. Ren, S. Tian, and W. Wu, *J. Chem. Eng. Data*, **58**, 866–872 (2013). Formation of Deep Eutectic Solvents by Phenols and Choline Chloride and Their Physical Properties.
149. A. P. Abbott, R. C. Harris, and K. S. Ryder, *J. Phys. Chem. B*, **111**, 4910–4913 (2007). Application of Hole Theory to Define Ionic Liquids by Their Transport Properties.
150. R. Stefanovic, M. Ludwig, G. B. Webber, R. Atkin, and A. J. Page, *Phys. Chem. Chem. Phys.*, **19**, 3297–3306 (2017). Nanostructure, Hydrogen Bonding and Rheology in Choline Chloride Deep Eutectic Solvents as a Function of the Hydrogen Bond Donor.
151. R. Zwanzig, *Annu. Rev. Phys. Chem.*, **16**, 67–102 (1965). Time-Correlation Functions and Transport Coefficients in Statistical Mechanics.
152. C. Rey-Castro and L. F. Vega, *J. Phys. Chem. B*, **110**, 14426–14435 (2006). Transport Properties of the Ionic Liquid 1-Ethyl-3-Methylimidazolium Chloride from Equilibrium Molecular Dynamics Simulation. The Effect of Temperature.
153. M. H. Kowsari, S. Alavi, M. Ashrafizaadeh, and B. Najafi, *J. Chem. Phys.*, **130**, 014703 (2009). Molecular Dynamics Simulation of Imidazolium-Based Ionic Liquids. II. Transport Coefficients.
154. M. Atilhan, J. Jacquemin, D. Rooney, M. Khraisheh, and S. Aparicio, *Ind. Eng. Chem. Res.*, **52**, 16774–16785 (2013). Viscous Behavior of Imidazolium-Based Ionic Liquids.

155. L. Zhao, X. Wang, and L. Wang, H. Sun, *Fluid Phase Equilib.*, **260**, 212–217 (2007). Prediction of Shear Viscosities Using Periodic Perturbation Method and OPLS Force Field.
156. J. Ma, Z. Zhang, Y. Xiang, F. Cao, and H. Sun, *Mol. Simul.*, **43**, 1502–1512 (2017). On the Prediction of Transport Properties of Ionic Liquid Using 1-N-Butylmethylpyridinium Tetrafluoroborate as an Example.
157. B. Hess, *J. Chem. Phys.*, **116**, 209 (2002). Determining the Shear Viscosity of Model Liquids from Molecular Dynamics Simulations.
158. S. H. Jamali, R. Hartkamp, C. Bardas, J. Söhl, T. J. H. Vlugt, and O. A. Moutos, *J. Chem. Theory Comput.*, **14**, 5959–5968 (2018). Shear Viscosity Computed from the Finite-Size Effects of Self-Diffusivity in Equilibrium Molecular Dynamics.
159. I.-C. Yeh and G. Hummer, *J. Phys. Chem. B*, **108**, 15873–15879 (2004). System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions.
160. Y. Zhang, A. Otani, and E. J. Maginn, *J. Chem. Theory Comput.*, **11**, 3537–3546 (2015). Reliable Viscosity Calculation from Equilibrium Molecular Dynamics Simulations: A Time Decomposition Method.
161. F. Müller-Plathe, *Phys. Rev. E*, **59**, 4894 (1999). Reversing the Perturbation in Nonequilibrium Molecular Dynamics: An Easy Way to Calculate the Shear Viscosity of Fluids.
162. M. S. Kelkar and E. J. Maginn, *J. Phys. Chem. B*, **111**, 4867–4876 (2007). Effect of Temperature and Water Content on the Shear Viscosity of the Ionic Liquid 1-Ethyl-3-Methylimidazolium Bis(Trifluoromethanesulfonyl)Imide as Studied by Atomistic Simulations.
163. T. Altamash, M. Atilhan, A. Aliyan, R. Ullah, G. García, and S. Aparicio, *RSC Adv.*, **6**, 109201–109210 (2016). Insights into Choline Chloride–Phenylacetic Acid Deep Eutectic Solvent for CO₂ Absorption.
164. A. González de Castilla, J. P. Bittner, S. Müller, S. Jakobtorweihen, and I. Smirnova, *J. Chem. Eng. Data*, (2019). Thermodynamic and Transport Properties Modeling of Deep Eutectic Solvents: A Review on G E -Models, Equations of State, and Molecular Dynamics.
165. R. Haghbakhsh, S. Raeissi, K. Parvaneh, and A. Shariati, *J. Mol. Liq.*, **249**, 554–561 (2018). The Friction Theory for Modeling the Viscosities of Deep Eutectic Solvents Using the Cpa and Pc-Saft Equations of State.
166. R. Haghbakhsh, K. Parvaneh, S. Raeissi, and A. Shariati, *Fluid Phase Equilib.*, **470**, 193–202 (2018). A General Viscosity Model for Deep Eutectic Solvents: The Free Volume Theory Coupled with Association Equations of State.
167. J. O. Lloret, L. F. Vega, and F. Llovel, *Fluid Phase Equilib.*, **448**, 81–93 (2017). Accurate Description of Thermophysical Properties of Tetraalkylammonium Chloride Deep Eutectic Solvents with the Soft-Saft Equation of State.

168. A. P. Abbott, R. C. Harris, K. S. Ryder, C. D'Agostino, L. F. Gladden, and M. D. Mantle, *Green Chem.*, **13**, 82–90 (2011). Glycerol Eutectics as Sustainable Solvent Systems.
169. H. Zhao, Z.-C. Liang, and F. Li, *J. Mol. Liq.*, **149**, 55–59 (2009). An Improved Model for the Conductivity of Room-Temperature Ionic Liquids Based on Hole Theory.
170. M. G. Del Pópolo and G. A. Voth, *J. Phys. Chem. B*, **108**, 1744–1752 (2004). On the Structure and Dynamics of Ionic Liquids.
171. M. Francisco, A. van den Bruinhorst, and M. C. Kroon, *Angew. Chem. Int. Ed.*, **52**, 3074–3085 (2013). Low-Transition-Temperature Mixtures (LTTMs): A New Generation of Designer Solvents.
172. M. Gilmore, L. M. Moura, A. H. Turner, M. Swadźba-Kwaśny, S. K. Callear, J. A. McCune, O. A. Scherman, and J. D. Holbrey, *J. Chem. Phys.*, **148**, 193823 (2018). A Comparison of Choline:Urea and Choline:Oxalic Acid Deep Eutectic Solvents at 338 K.
173. A. Korotkevich, D. S. Firaha, and A. A. H. Padua, B. Kirchner, *Fluid Phase Equilib.*, **448**, 59–68 (2017). Ab Initio Molecular Dynamics Simulations of SO₂ Solvation in Choline Chloride/Glycerol Deep Eutectic Solvent.
174. Y. Dai, J. van Spronsen, G.-J. Witkamp, R. Verpoorte, and Y. H. Choi, *Anal. Chim. Acta*, **766**, 61–68 (2013). Natural Deep Eutectic Solvents as New Potential Media for Green Technology.
175. R. Ahmadi, B. Hemmateenejad, A. Safavi, Z. Shojaeifard, A. Shahsavari, A. Mohajeri, M. Heydari Dokoohaki, and A. R. Zolghadr, *Phys. Chem. Chem. Phys.*, **20**, 18463–18473 (2018). Deep Eutectic-Water Binary Solvent Associations Investigated by Vibrational Spectroscopy and Chemometrics.
176. A. H. Turner and J. D. Holbrey, *Phys. Chem. Chem. Phys.*, **21**, 21782–21789 (2019). Investigation of Glycerol Hydrogen-Bonding Networks in Choline Chloride/Glycerol Eutectic-Forming Liquids Using Neutron Diffraction.
177. S. Kaur, S. Sharma, and H. K. Kashyap, *J. Chem. Phys.*, **147**, 194507 (2017). Bulk and Interfacial Structures of Reline Deep Eutectic Solvent: A Molecular Dynamics Study.
178. V. Migliorati, F. Sessa, and P. D'Angelo, *Chem. Phys. Lett.: X*, **2**, 100001 (2019). Deep Eutectic Solvents: A Structural Point of View on the Role of the Cation.
179. F. L. Hirshfeld, *Theor. Chem. Acc.*, **44**, 129–138 (1977). Bonded-Atom Fragments for Describing Molecular Charge Densities.
180. P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. A. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, and A. Nilsson, *Science*, **304**, 995–999 (2004). The Structure of the First Coordination Shell in Liquid Water.
181. L. P. Silva, C. F. Araújo, D. O. Abranches, M. Melle-Franco, M. A. R. Martins, M. M. Nolasco, P. J. A. Ribeiro-Claro, S. P. Pinho, and J. A. P. Coutinho, *Phys.*

- Chem. Chem. Phys.*, **21**, 18278–18289 (2019). What a Difference a Methyl Group Makes – Probing Choline–Urea Molecular Interactions through Urea Structure Modification.
182. M. E. Tuckerman, A. Chandra, and D. Marx, *J. Chem. Phys.*, **133**, 124108 (2010). A Statistical Mechanical Theory of Proton Transport Kinetics in Hydrogen-Bonded Networks Based on Population Correlation Functions with Applications to Acids and Bases.
183. F. Perera, *Int. J. Environ. Res. Public Health*, **15**, (2017). Pollution from Fossil-Fuel Combustion Is the Leading Environmental Threat to Global Pediatric Health and Equity: Solutions Exist.
184. T. J. Trivedi, J. H. Lee, H. J. Lee, Y. K. Jeong, and J. W. Choi, *Green Chem.*, **18**, 2834–2842 (2016). Deep Eutectic Solvents as Attractive Media for CO₂ Capture.
185. D. Yang, M. Hou, H. Ning, J. Zhang, J. Ma, G. Yang, and B. Han, *Green Chem.*, **15**, 2261 (2013). Efficient SO₂ Absorption by Renewable Choline Chloride–Glycerol Deep Eutectic Solvents.
186. D. Deng, G. Han, and Y. Jiang, *New J. Chem.*, **39**, 8158–8164 (2015). Investigation of a Deep Eutectic Solvent Formed by Levulinic Acid with Quaternary Ammonium Salt as an Efficient SO₂ Absorbent.
187. S. Sun, Y. Niu, Q. Xu, Z. Sun, and X. Wei, *Ind. Eng. Chem. Res.*, **54**, 8019–8024 (2015). Efficient SO₂ Absorptions by Four Kinds of Deep Eutectic Solvents Based on Choline Chloride.
188. H. Li, Y. Chang, W. Zhu, C. Wang, C. Wang, S. Yin, M. Zhang, and H. Li, *Phys. Chem. Chem. Phys.*, **17**, 28729–28742 (2015). Theoretical Evidence of Charge Transfer Interaction between SO₂ and Deep Eutectic Solvents Formed by Choline Chloride and Glycerol.
189. O. S. Hammond, K. J. Edler, D. T. Bowron, and L. Torrente-Murciano, *Nat. Commun.*, **8**, 14150 (2017). Deep Eutectic-Solvothermal Synthesis of Nanostructured Ceria.
190. L. Wei, Z.-Y. Zhou, S.-P. Chen, C.-D. Xu, D. Su, M. E. Schuster, and S.-G. Sun, *Chem. Commun.*, **49**, 11152–11154 (2013). Electrochemically Shape-Controlled Synthesis in Deep Eutectic Solvents: Triambic Icosahedral Platinum Nanocrystals with High-Index Facets and Their Enhanced Catalytic Activity.
191. A. P. Abbott, J. C. Barron, G. Frisch, K. S. Ryder, and A. F. Silva, *Electrochim. Acta*, **56**, 5272–5279 (2011). The Effect of Additives on Zinc Electrodeposition from Deep Eutectic Solvents.
192. N. M. Pereira, P. M. V. Fernandes, C. M. Pereira, and A. Fernando Silva, *J. Electrochem. Soc.*, **159**, D501–D506 (2012). Electrodeposition of Zinc from Choline Chloride–Ethylene Glycol Deep Eutectic Solvent: Effect of the Tartrate Ion.

193. L. Vieira, R. Schennach, and B. Gollas, *Phys. Chem. Chem. Phys.*, **17**, 12870–12880 (2015). In Situ Pm-Irras of a Glassy Carbon Electrode/Deep Eutectic Solvent Interface.
194. D. Lindberg, M. de la Fuente Revenga, and M. Widersten, *J. Biotechnol.*, **147**, 169–171 (2010). Deep Eutectic Solvents (Dess) Are Viable Cosolvents for Enzyme-Catalyzed Epoxide Hydrolysis.
195. J. T. Gorke, F. Sreenc, and R. J. Kazlauskas, *Chem. Commun.*, 1235–1237 (2008). Hydrolase-Catalyzed Biotransformations in Deep Eutectic Solvents.
196. N. Guajardo, K. Ahumada, P. Domínguez de María, and R. A. Schreiber, *Biocatal. Biotransform.*, **37**, 106–114 (2019). Remarkable Stability of Candida Antarctica Lipase B Immobilized Via Cross-Linking Aggregates (Clea) in Deep Eutectic Solvents.

INDEX

- ab initio* methods, 145–148
- aggregation volume biased (AVB) moves, 99, 109–110
- Asakura-Oosawa–Vrij (AO) model, 90
- axial next-nearest-neighbor Ising (ANNNI) model, 113, 115, 116

- barrier crossing, 15–19, 24
- basin hopping (BH), 11
- Bjerrum length, 89
- block copolymers, 83–84
- Boltzmann constant, 11, 84, 171
- Born-Oppenheimer molecular dynamics, 60

- Car–Parrinello molecular dynamics, 60
- charge transfer plasmon (CTP), 57
- classical MD simulations, 149–153
- classical path approximation (CPA), 62–63

- cluster structure optimization, 29
- collision and dissociation detection (CD/DD), 24
- combined distribution functions (CDFs), 159
- constant pressure, 120, 164, 166
- Coulomb energy, 89
- critical cluster density (ccd), 111
- critical micellar concentration (cmc), 84
- crossover-style information exchange, 24

- D-base method, 172
- decoherence induced surface hopping (DISH), 63
- deep eutectic solvents (DESs)
 - ab initio* methods, 145–148
 - classical MD simulations, 149–153
 - definition, 137
 - gas sorption studies on, 196–198

Reviews in Computational Chemistry, Volume 32, First Edition.

Edited by Abby L. Parrill and Kenny B. Lipkowitz.

© 2022 John Wiley & Sons, Inc. Published 2022 by John Wiley & Sons, Inc.

- deep eutectic solvents (DESs)
 (*Continued*)
 hydrogen bond analysis, 189–196
 interactions at metal surfaces,
 198–199
 as ionic liquid analogues, 137–140
 molecular structure and type of
 interactions, 140–142
 nonpolarizable force fields,
 153–159
 physical properties, 159–164
 proteins in, 199–200
 radial distribution functions,
 183–189
 spatial distribution functions, 196
 thermodynamic properties, 164–170
 transport properties, 170–183
 types of, 142–143
 deep learning, 19
 degrees of freedom (DOF), 3
 density functional theory (DFT), 25,
 90–91, 145
 density functional tight binding (DFTB)
 computational details, 67–68
 example, 68–72
 nonadiabatic electron dynamics,
 59–67
 RT-TDDFTB (*see* real-time
 time-dependent DFTB
 (RT-TDDFTB))
 Derjaguin–Landau–Verwey–Overbeek
 (DLVO) theory, 89
 deterministic global optimization
 (DGO) methods, 8
 diffusion coefficients, DES, 178–183
 disordered microphases, 92–93

 electronic couplings, 56–57
 empirical potential structure refinement
 (EPSR) method, 156
 equilibrium MD, 171
 eutectic composition, 137
 eutectic temperature, 137
 excitation energy transfer (EET), 47,
 53–56

 fewest-switches surface-hopping
 (FSSH) method, 61
 Flory–Huggins parameter, 84
 force fields (FF), 16
 fourth-order Runge–Kutta (RK4)
 method, 65
 free energy
 calculations, 16, 92
 Helmholtz, 94
 of ideal gas in field, 119–120
 freezing point, 137

 ghost particle/cluster switching method,
 100–103
 Gibbs–Duham construction, 105
 Ginzburg–Landau theory, 91
 Green–Kubo formula, 158, 171–174

 heat capacity, 164–168
 heats of vaporization, 168
 Hellmann–Feynman theorem, 62
 Helmholtz free energy, 94
 high-performance computing (HPC), 27
 hydrogen bond acceptor (HBA), 137
 hydrogen bond analysis, 189–196
 hydrogen bond donor (HBD), 137

 Independent Electron Surface Hopping
 (IESH), 63
 infrared (IR) spectrum, 28
 ionic liquid
 analogues, 137–140
 room temperature ionic liquids, 137
 isothermal compressibility, 169–170

 lattice spin models, 87, 113–117
 Lennard–Jones (LJ) potential, 150
 Lifshitz point, 83, 91, 107
 Liouville–von Neumann equation, 47
 liquid density, DES, 159–162

 machine learning (ML), 19
 Macleod’s equation, 162
 message-passing interface (MPI)
 paradigms, 27

- microphases, 81
 - aggregation volume biased (AVB) moves, 99, 109–110
 - block copolymers, 83–84
 - cluster volume moves, 103–105
 - colloidal models, 117–118
 - colloidal suspensions, 87–90
 - determining phase transitions, 105
 - disordered, 92–93
 - expanded thermodynamics, 94–95
 - field theory of, 90–91
 - ghost particle/cluster switching method, 100–103
 - lattice spin models, 87, 113–117
 - molecular simulations and challenges, 91–93
 - morphological crossovers in, 110–112
 - one-dimensional models, 112–113
 - other examples, 90
 - periodic, 92
 - surfactants and microemulsions, 84–87
 - thermodynamic integration, 95–100
 - virtual cluster moves, 107–109
 - Wolff-like cluster algorithms, 106–107
- Mittag–Leffler model, 196
- mixed quantum-classical (MQC) approaches, 61
- Møller-Plesset perturbation theory of second order (MP2), 25
- Monte Carlo with minimization (MCM), 11
- Müller-Plathe method, 173
- Mulliken approximation, 48
- naphthalene
 - absorption spectrum for, 49–51
 - with laser-type perturbation, 51
 - molecule, electron dynamics for, 49
- Navier–Stokes equation, 172, 173
- neural networks, 19
- no free lunch theorem (NFLT), 13
- nonadiabatic electron dynamics
 - adiabatic *vs* nonadiabatic dynamics, 59–61
 - classical path approximation, 62–63
 - CPA-FSSH-DFTB implementation, 65–66
 - equations governing, 61–62
 - post-processing tools, 67
 - surface hopping and fewest switches criterion, 63–65
- nonadiabatic molecular dynamics (NAMD), 61
- non-deterministic global optimization (NDGO)
 - algorithm comparisons, 14–15
 - barrier crossing, 15–19
 - deterministic *vs* non-deterministic search, 5–8
 - fundamental algorithm ideas, 21–22
 - fundamental take-home lessons, 8
 - method design choices, 22–28
 - need for structural optimization, 2–3
 - no free lunch (NFL) theorem, 8, 11–14
 - old *vs* new machine learning, 19–20
 - search space, 3–5
 - take-home lessons for, 20–21
 - things to do, and pitfalls to avoid, 31–32
 - tips for absolute beginners, 28–31
- nonequilibrium molecular dynamics (NEMD), 171, 172
- nonpolarizable force fields, 153–159
- OPLS-DES, 157
- parallel tempering (PT), 106
- periodic microphases, 92
- phenotype representation, 23
- physical properties, DES
 - liquid density, 159–162
 - surface tension, 162–164
 - volume expansivity, 162

- plasmonic excitations, 57–59
- potential energy surface (PES), 16, 59
- pulse field gradient nuclear magnetic resonance (PFG-NMR), 178
- quantitative structure–activity relationship, 145
- quantum tunneling, 17–19, 21, 57
- radial distribution functions, 183–189
- real-time time-dependent DFTB (RT-TDDFTB)
- naphthalene, absorption spectrum for, 49–51
 - naphthalene molecule, electron dynamics for, 49
 - naphthalene with laser-type perturbation, 51
 - realistic large systems, electron dynamics of, 51–59
 - theory and methodology, 46–49
- restrained electrostatic charge potential (RESP), 156
- reverse nonequilibrium molecular dynamics approach (RNEMD), 173
- room temperature ionic liquids (RTILs), 137
- short-range attractive and long-range repulsive (SALR) interaction, 90
- single plasmonic nanoparticle, 53
- spatial distribution functions (SDFs), 156, 196
- staircase transformation, 16, 19, 24
- surface hopping (SH)
- and fewest switches criterion, 63–65
 - method, 61
- surface tension, DES, 162–164
- thermal de Broglie wavelength, 99
- thermal energy, 89
- thermodynamic integration (TI)
- method, 95–100
- thermodynamic properties, DES
- heat capacity, 164–168
 - heats of vaporization, 168
 - isothermal compressibility, 169–170
- time-dependent Kohn-Sham (TDKS)
- approximation, 65
- time-independent Schrödinger equation (TISE), 59, 61
- transport properties, DES
- diffusion coefficients, 178–183
 - viscosity, 170–178
- van der Waals attraction, 87
- virial coefficients, 99, 111, 120–122
- virtual cluster moves, 107–109
- viscosity, DES, 170–178
- volume expansivity, DES, 162
- Wang–Landau sampling, 102
- weak no free lunch theorem (wNFLT), 13
- Widom–Wheeler model, 115
- Wolff-like cluster algorithms, 106–107