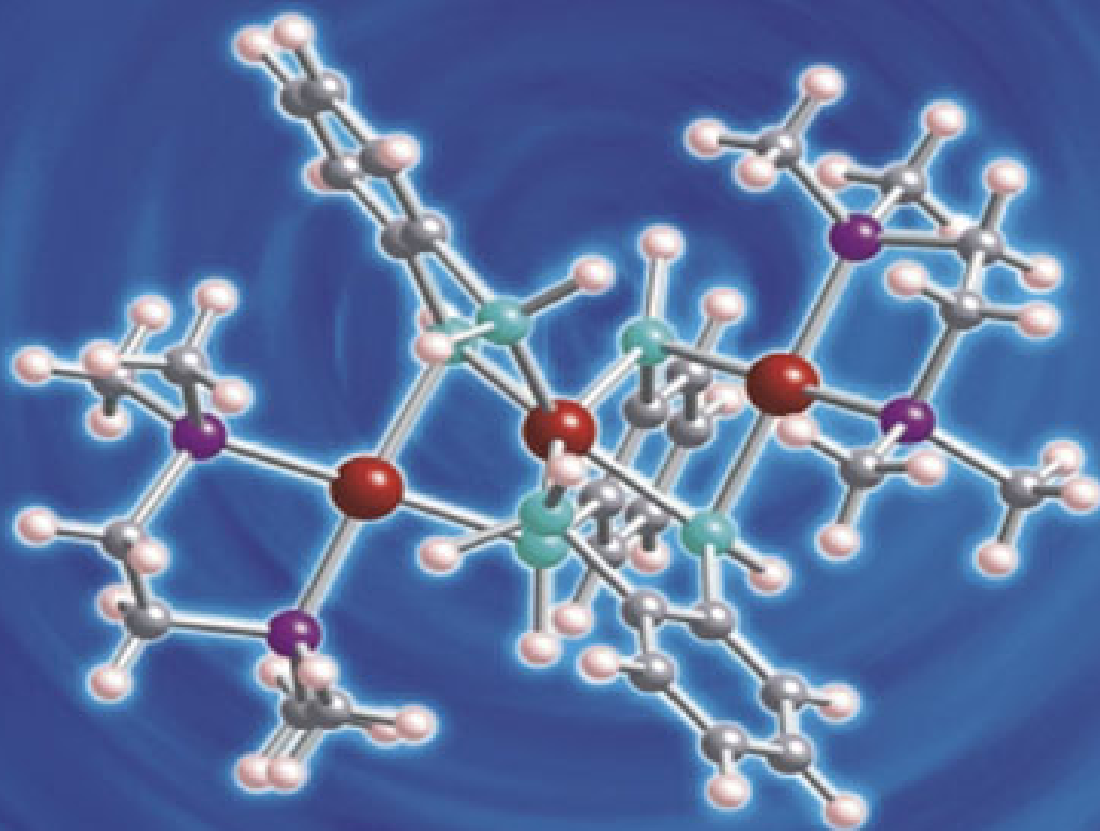


Essentials of
**Computational
Chemistry** Second Edition

Theories and Models

Christopher J Cramer



 WILEY

Contents

[Preface to the First Edition](#)

[Preface to the Second Edition](#)

[Acknowledgments](#)

[1 What are Theory, Computation, and Modeling?](#)

[1.1 Definition of Terms](#)

[1.2 Quantum Mechanics](#)

[1.3 Computable Quantities](#)

[1.4 Cost and Efficiency](#)

[1.5 Note on Units](#)

[Bibliography and Suggested Additional Reading](#)
[References](#)

[2 Molecular Mechanics](#)

[2.1 History and Fundamental Assumptions](#)

[2.2 Potential Energy Functional Forms](#)

[2.3 Force-field Energies and Thermodynamics](#)

[2.4 Geometry Optimization](#)

[2.5 Menagerie of Modern Force Fields](#)

[2.6 Force Fields and Docking](#)

[2.7 Case Study: \(2R*4S*\)-1-Hydroxy-2,4-dimethylhex-5-ene](#)

[Bibliography and Suggested Additional Reading](#)
[References](#)

3 Simulations of Molecular Ensembles

[3.1 Relationship Between MM Optima and Real Systems](#)

[3.2 Phase Space and Trajectories](#)

[3.3 Molecular Dynamics](#)

[3.4 Monte Carlo](#)

[3.5 Ensemble and Dynamical Property Examples](#)

[3.6 Key Details in Formalism](#)

[3.7 Force Field Performance in Simulations](#)

[3.8 Case Study: Silica Sodalite](#)

[Bibliography and Suggested Additional Reading](#)
[References](#)

4 Foundations of Molecular Orbital Theory

[4.1 Quantum Mechanics and the Wave Function](#)

[4.2 The Hamiltonian Operator](#)

[4.3 Construction of Trial Wave Functions](#)

[4.4 Hückel Theory](#)

[4.5 Many-electron Wave Functions](#)

[Bibliography and Suggested Additional Reading](#)
[References](#)

5 Semiempirical Implementations of Molecular Orbital Theory

[5.1 Semiempirical Philosophy](#)

[5.2 Extended Hückel Theory](#)

[5.3 CNDO Formalism](#)

[5.4 INDO Formalism](#)

[5.5 Basic NDDO Formalism](#)

[5.6 General Performance Overview of Basic NDDO Models](#)

[5.7 Ongoing Developments in Semiempirical MO Theory](#)

[5.8 Case Study: Asymmetric Alkylation of Benzaldehyde](#)

[Bibliography and Suggested Additional Reading References](#)

6 *Ab Initio* Implementations of Hartree-Fock Molecular Orbital Theory

[6.1 *Ab Initio* Philosophy](#)

[6.2 Basis Sets](#)

[6.3 Key Technical and Practical Points of Hartree-Fock Theory](#)

[6.4 General Performance Overview of *Ab Initio* HF Theory](#)

[6.5 Case Study: Polymerization of 4-Substituted Aromatic Enynes](#)

[Bibliography and Suggested Additional Reading References](#)

7 Including Electron Correlation in Molecular Orbital Theory

[7.1 Dynamical vs. Non-dynamical Electron Correlation](#)

[7.2 Multiconfiguration Self-Consistent Field Theory](#)
[7.3 Configuration Interaction](#)
[7.4 Perturbation Theory](#)
[7.5 Coupled-cluster Theory](#)
[7.6 Practical Issues in Application](#)
[7.7 Parameterized Methods](#)
[7.8 Case Study: Ethylenedione Radical Anion Bibliography and Suggested Additional Reading](#)
[References](#)

8 Density Functional Theory

[8.1 Theoretical Motivation](#)
[8.2 Rigorous Foundation](#)
[8.3 Kohn-Sham Self-consistent Field Methodology](#)
[8.4 Exchange-correlation Functionals](#)
[8.5 Advantages and Disadvantages of DFT Compared to MO Theory](#)
[8.6 General Performance Overview of DFT](#)
[8.7 Case Study: Transition-Metal Catalyzed Carbonylation of Methanol Bibliography and Suggested Additional Reading](#)
[References](#)

9 Charge Distribution and Spectroscopic Properties

[9.1 Properties Related to Charge Distribution](#)
[9.2 Ionization Potentials and Electron Affinities](#)
[9.3 Spectroscopy of Nuclear Motion](#)

[9.4 NMR Spectral Properties](#)

[9.5 Case Study: Matrix Isolation of Perfluorinated *p*-Benzyne](#)

[Bibliography and Suggested Additional Reading References](#)

10 Thermodynamic Properties

[10.1 Microscopic-macroscopic Connection](#)

[10.2 Zero-point Vibrational Energy](#)

[10.3 Ensemble Properties and Basic Statistical Mechanics](#)

[10.4 Standard-state Heats and Free Energies of Formation and Reaction](#)

[10.5 Technical Caveats](#)

[10.6 Case Study: Heat of Formation of H₂NOH](#)

[Bibliography and Suggested Additional Reading References](#)

11 Implicit Models for Condensed Phases

[11.1 Condensed-phase Effects on Structure and Reactivity](#)

[11.2 Electrostatic Interactions with a Continuum](#)

[11.3 Continuum Models for Non-electrostatic Interactions](#)

[11.4 Strengths and Weaknesses of Continuum Solvation Models](#)

[11.5 Case Study: Aqueous Reductive Dechlorination of Hexachloroethane](#)

[Bibliography and Suggested Additional Reading References](#)

12 Explicit Models for Condensed Phases

[12.1 Motivation](#)

[12.2 Computing Free-energy Differences](#)

[12.3 Other Thermodynamic Properties](#)

[12.4 Solvent Models](#)

[12.5 Relative Merits of Explicit and Implicit Solvent Models](#)

[12.6 Case Study: Binding of Biotin Analogs to Avidin](#)

[Bibliography and Suggested Additional Reading References](#)

13 Hybrid Quantal/Classical Models

[13.1 Motivation](#)

[13.2 Boundaries Through Space](#)

[13.3 Boundaries Through Bonds](#)

[13.4 Empirical Valence Bond Methods](#)

[13.5 Case Study: Catalytic Mechanism of Yeast Enolase](#)

[Bibliography and Suggested Additional Reading References](#)

14 Excited Electronic States

[14.1 Determinantal/Configurational Representation of Excited States](#)

[14.2 Singly Excited States](#)

[14.3 General Excited State Methods](#)

[14.4 Sum and Projection Methods](#)

[14.5 Transition Probabilities](#)

[14.6 Solvatochromism](#)

[14.7 Case Study: Organic Light Emitting Diode](#)

[Alq₃](#)

[Bibliography and Suggested Additional Reading](#)

[References](#)

[15 Adiabatic Reaction Dynamics](#)

[15.1 Reaction Kinetics and Rate Constants](#)

[15.2 Reaction Paths and Transition States](#)

[15.3 Transition-state Theory](#)

[15.4 Condensed-phase Dynamics](#)

[15.5 Non-adiabatic Dynamics](#)

[15.6 Case Study: Isomerization of Propylene](#)

[Oxide](#)

[Bibliography and Suggested Additional Reading](#)

[References](#)

[Appendix A Acronym Glossary](#)

[Appendix B Symmetry and Group Theory](#)

[B.1 Symmetry Elements](#)

[B.2 Molecular Point Groups and Irreducible Representations](#)

[B.3 Assigning Electronic State Symmetries](#)

[B.4 Symmetry in the Evaluation of Integrals and Partition Functions](#)

[**Appendix C Spin Algebra**](#)

[C.1 Spin Operators](#)

[C.2 Pure- and Mixed-spin Wave Functions](#)

[C.3 UHF Wave Functions](#)

[C.4 Spin Projection/Annihilation](#)

[Reference](#)

[**Appendix D Orbital Localization**](#)

[D.1 Orbitals as Empirical Constructs](#)

[D.2 Natural Bond Orbital Analysis](#)

[References](#)

[**Index**](#)

Essentials of Computational Chemistry

Theories and Models

Second Edition

Christopher J. Cramer

*Department of Chemistry and Supercomputing Institute,
University of Minnesota, USA*



John Wiley & Sons, Ltd

Copyright © 2004 John Wiley & Sons Ltd, The Atrium,
Southern Gate, Chichester, West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk

Visit our Home Page on www.wileyeurope.com or
www.wiley.com

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, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

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

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-
1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Cramer, Christopher J., 1961-

Essentials of computational chemistry : theories and models /

Christopher J. Cramer. - 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-470-09181-9 (cloth: alk. paper) - ISBN 0-470-09182-7 (pbk.: alk. paper)

1. Chemistry, Physical and theoretical - Data processing. 2.

Chemistry, Physical and theoretical - Mathematical models.

I. Title.

QD455.3.E4C73 2004

541'.0285 - dc22

2004015537

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-09181-9 (cased)

ISBN 0-470-09182-7 (pbk)

For Katherine

Preface to the First Edition

Computational chemistry, alternatively sometimes called theoretical chemistry or molecular modeling (reflecting a certain factionalization amongst practitioners), is a field that can be said to be both old and young. It is old in the sense that its foundation was laid with the development of quantum mechanics in the early part of the twentieth century. It is young, however, insofar as arguably no technology in human history has developed at the pace that digital computers have over the last 35 years or so. The digital computer being the 'instrument' of the computational chemist, workers in the field have taken advantage of this progress to develop and apply new theoretical methodologies at a similarly astonishing pace.

The evidence of this progress and its impact on Chemistry in general can be assessed in various ways. Boyd and Lipkowitz, in their book series *Reviews in Computational Chemistry*, have periodically examined such quantifiable indicators as numbers of computational papers published, citations to computational chemistry software packages, and citation rankings of computational chemists. While such metrics need not necessarily be correlated with 'importance', the exponential growth rates they document are noteworthy. My own personal (and somewhat more whimsical) metric is the staggering increase in the percentage of exposition floor space occupied by computational chemistry software vendors at various chemistry meetings worldwide - *someone* must be buying those products!

Importantly, the need for at least a cursory understanding of theory/computation/modeling is by no means restricted to practitioners of the art. Because of the broad array of theoretical tools now available, it is a rare problem of

interest that does not occupy the attention of both experimental *and* theoretical chemists. Indeed, the synergy between theory and experiment has vastly accelerated progress in any number of areas (as one example, it is hard to imagine a modern paper on the matrix isolation of a reactive intermediate and its identification by infrared spectroscopy not making a comparison of the experimental spectrum to one obtained from theory/calculation). To take advantage of readily accessible theoretical tools, and to understand the results reported by theoretical collaborators (or competitors), even the wettest of wet chemists can benefit from some familiarity with theoretical chemistry. My objective in this book is to provide a survey of computational chemistry – its underpinnings, its jargon, its strengths and weaknesses – that will be accessible to both the experimental and theoretical communities. The level of the presentation assumes exposure to quantum and statistical mechanics; particular topics/examples span the range of inorganic, organic, and biological chemistry. As such, this text could be used in a course populated by senior undergraduates and/or beginning graduate students without regard to specialization.

The scope of theoretical methodologies presented in the text reflects my judgment of the degree to which these methodologies impact on a broad range of chemical problems, i.e., the degree to which a practicing chemist may expect to encounter them repeatedly in the literature and thus should understand their applicability (or lack thereof). In some instances, methodologies that do not find much modern use are discussed because they help to illustrate in an intuitive fashion how more contemporary models developed their current form. Indeed, one of my central goals in this book is to render less opaque the fundamental natures of the various theoretical models. By understanding the assumptions implicit in a theoretical model, and the

concomitant limitations imposed by those assumptions, one can make informed judgments about the trustworthiness of theoretical results (and economically sound choices of models to apply, if one is about to embark on a computational project).

With no wish to be divisive, it must be acknowledged: there are some chemists who are not fond of advanced mathematics. Unfortunately, it is simply not possible to describe computational chemistry without resort to a fairly hefty number of equations, and, particularly for modern electronic-structure theories, some of those equations are fantastically daunting in the absence of a detailed knowledge of the field. That being said, I offer a promise to present no equation without an effort to provide an intuitive explanation for its form and the various terms within it. In those instances where I don't think such an explanation *can* be offered (of which there are, admittedly, a few), I will provide a qualitative discussion of the area and point to some useful references for those inclined to learn more.

In terms of layout, it might be preferable from a historic sense to start with quantum theories and then develop classical theories as an approximation to the more rigorous formulation. However, I think it is more pedagogically straightforward (and far easier on the student) to begin with classical models, which are in the widest use by experimentalists and tend to feel very intuitive to the modern chemist, and move from there to increasingly more complex theories. In that same vein, early emphasis will be on single-molecule (gas-phase) calculations followed by a discussion of extensions to include condensed-phase effects. While the book focuses primarily on the calculation of equilibrium properties, excited states and reaction dynamics are dealt with as advanced subjects in later chapters.

The quality of a theory is necessarily judged by its comparison to (accurate) physical measurements. Thus, careful attention is paid to offering comparisons between theory and experiment for a broad array of physical observables (the first chapter is devoted in part to enumerating these). In addition, there *is* some utility in the computation of things which cannot be observed (e.g., partial atomic charges), and these will also be discussed with respect to the performance of different levels of theory. However, the best way to develop a feeling for the scope and utility of various theories is to apply them, and instructors are encouraged to develop computational problem sets for their students. To assist in that regard, case studies appear at the end of most chapters illustrating the employ of one or more of the models most recently presented. The studies are drawn from the chemical literature; depending on the level of instruction, reading and discussing the original papers as part of the class may well be worthwhile, since any synopsis necessarily does away with some of the original content.

Perversely, perhaps, I do not include in this book specific problems. Indeed, I provide almost no discussion of such nuts and bolts issues as, for example, how to enter a molecular geometry into a given program. The reason I eschew these undertakings is not that I think them unimportant, but that computational chemistry software is not particularly well standardized, and I would like neither to tie the book to a particular code or codes nor to recapitulate material found in users' manuals. Furthermore, the hardware and software available in different venues varies widely, so individual instructors are best equipped to handle technical issues themselves. With respect to illustrative problems for students, there *are* reasonably good archives of such exercises provided either by software vendors as part of their particular package or developed for

computational chemistry courses around the world. Chemistry 8021 at the University of Minnesota, for example, has several years worth of problem sets (with answers) available at pollux.chem.umn.edu/8021. Given the pace of computational chemistry development and of modern publishing, such archives are expected to offer a more timely range of challenges in any case.

A brief summary of the mathematical notation adopted throughout this text is in order. Scalar quantities, whether constants or variables, are represented by italic characters. Vectors and matrices are represented by boldface characters (individual matrix *elements* are scalar, however, and thus are represented by italic characters that are indexed by subscript(s) identifying the particular element). Quantum mechanical operators are represented by italic characters if they have scalar expectation values and boldface characters if their expectation values are vectors or matrices (or if they are typically *constructed* as matrices for computational purposes). The only deliberate exception to the above rules is that quantities represented by Greek characters typically are made neither italic nor boldface, irrespective of their scalar or vector/matrix nature.

Finally, as with most textbooks, the total content encompassed herein is such that only the most masochistic of classes would attempt to go through this book cover to cover in the context of a typical, semester-long course. My intent in coverage is not to act as a firehose, but to offer a reasonable degree of flexibility to the instructor in terms of optional topics. Thus, for instance, Chapters 3 and 11-13 could readily be skipped in courses whose focus is primarily on the modeling of small- and medium-sized molecular systems. Similarly, courses with a focus on macromolecular modeling could easily choose to ignore the more advanced levels of quantum mechanical modeling. And, clearly, time constraints in a typical course are unlikely to allow the

inclusion of more than one of the last two chapters. These practical points having been made, one can always hope that the eager student, riveted by the content, will take time to read the rest of the book him- or herself!

Christopher J. Cramer

September 2001

Preface to the Second Edition

Since publication of the first edition I have become increasingly, painfully aware of just how short the half-life of certain 'Essentials' can be in a field growing as quickly as is computational chemistry. While I utterly disavow any hubris on my part and indeed blithely assign all blame for this text's title to my editor, that does not detract from my satisfaction at having brought the text up from the ancient history of 2001 to the present of 2004. Hopefully, readers too will be satisfied with what's new and improved.

So, what *is* new and improved? In a nutshell, *new* material includes discussion of docking, principal components analysis, force field validation in dynamics simulations, first-order perturbation theory for relativistic effects, tight-binding density functional theory, electronegativity equalization charge models, standard-state equilibrium constants, computation of pK_a values and redox potentials, molecular dynamics with implicit solvent, and direct dynamics. With respect to *improved* material, the menagerie of modern force fields has been restocked to account for the latest in new and ongoing developments and a *new* menagerie of density functionals has been assembled to help the computational innocent navigate the forest of acronyms (in this last regard, the acronym glossary of Appendix A has also been expanded with an additional 64 entries). In addition, newly developed basis sets for electronic structure calculations are discussed, as are methods to scale various theories to infinite-basis-set limits, and new thermochemical methods. The performances of various more recent methods for the prediction of nuclear magnetic resonance chemical shifts are summarized, and discussion of the generation of condensed-phase potentials of mean force from simulation is expanded.

As developments in semiempirical molecular orbital theory, density functional theory, and continuum solvation models have proceeded at a particularly breakneck pace over the last three years, Chapters 5, 8, and 11 have been substantially reworked and contain much fresh material. In addition, I have tried wherever possible to update discussions and, while so doing, to add the most modern references available so as to improve the text's connection with the primary literature. This effort poses something of a challenge, as I definitely do not want to cross the line from writing a text to writing instead an outrageously lengthy review article - I leave it to the reader to assess my success in that regard. Lastly, the few remaining errors, typographical and otherwise, left over from the second printing of the first edition have been corrected - I accept full responsibility for all of them (with particular apologies to any descendants of Leopold Kronecker) and I thank those readers who called some of them to my attention.

As for important things that have *not* changed, with the exception of Chapter 10 I have chosen to continue to use all of the existing case studies. I consider them still to be sufficiently illustrative of modern application that they remain useful as a basis for thought/discussion, and instructors will inevitably have their own particular favorites that they may discuss 'off-text' in any case. The thorough nature of the index has also, hopefully, not changed, nor I hope the deliberate and careful explanation of all equations, tables, and figures.

Finally, in spite of the somewhat greater corpulence of the second edition compared to the first, I have done my best to maintain the text's liveliness - at least to the extent that a scientific tome can be said to possess that quality. After all, to what end science without humor?

Christopher J. Cramer

July 2004

Acknowledgments

It is a pleasure to recognize the extent to which conversations with my computationally minded colleagues at the University of Minnesota – Jiali Gao, Steven Kass, Ilja Siepmann, Don Truhlar, Darrin York, and the late Jan Almlöf – contributed to this project. As a longtime friend and collaborator, Don in particular has been an invaluable source of knowledge and inspiration. So, too, this book, and particularly the second edition, has been improved based on the input of graduate students either in my research group or taking Computational Chemistry as part of their coursework. Of these, Ed Sherer deserves special mention for having offered detailed and helpful comments on the book when it was in manuscript form. In addition, my colleague Bill Tolman provided inspirational assistance in the preparation of the cover art, and I am grateful to Sheryl Frankel for exceedingly efficient executive assistance. Finally, the editorial staff at Wiley have been consummate professionals with whom it has been a pleasure to work.

Most of the first edition of this book was written during a sabbatical year spent working with Modesto Orozco and Javier Luque at the University of Barcelona. Two more gracious hosts are unlikely to exist anywhere (particularly with respect to ignoring the vast amounts of time the moonlighting author spent writing a book). Support for that sabbatical year derived from the John Simon Guggenheim Foundation, the Spanish Ministry of Education and Culture, the Foundation BBV, and the University of Minnesota, and the generosity of those agencies is gratefully acknowledged. The writing of the second edition was shoehorned into whatever free moments presented themselves, and I thank the members of my research group for not complaining

about my assiduous efforts to hide myself from them over the course of a long Minnesota winter.

Finally, if it were not for the heroic efforts of my wife Katherine and the (relative) patience of my children William, Matthew, and Allison – all of whom allowed me to spend a ridiculous number of hours hunched over a keyboard in a non-communicative trance – I most certainly could never have accomplished anything.

1

What are Theory, Computation, and Modeling?

1.1 Definition of Terms

A clear definition of terms is critical to the success of all communication. Particularly in the area of computational chemistry, there is a need to be careful in the nomenclature used to describe predictive tools, since this often helps clarify what approximations have been made in the course of a modeling ‘experiment’. For the purposes of this textbook, we will adopt a specific convention for what distinguishes theory, computation, and modeling.

In general, ‘theory’ is a word with which most scientists are entirely comfortable. A theory is one or more rules that are postulated to govern the behavior of physical systems. Often, in science at least, such rules are quantitative in nature and expressed in the form of a mathematical equation. Thus, for example, one has the theory of Einstein that the energy of a particle, E , is equal to its relativistic mass, m , times the speed of light in a vacuum, c , squared,

$$(1.1) \ E = mc^2$$

The quantitative nature of scientific theories allows them to be tested by experiment. This testing is the means by which the applicable range of a theory is elucidated. Thus, for instance, many theories of classical mechanics prove applicable to macroscopic systems but break down for very small systems, where one must instead resort to quantum

mechanics. The observation that a theory has limits in its applicability might, at first glance, seem a sufficient flaw to warrant discarding it. However, if a sufficiently large number of ‘interesting’ systems falls within the range of the theory, practical reasons tend to motivate its continued use. Of course, such a situation tends to inspire efforts to find a more *general* theory that is not subject to the limitations of the original. Thus, for example, classical mechanics can be viewed as a special case of the more general quantum mechanics in which the presence of macroscopic masses and velocities leads to a simplification of the governing equations (and concepts).

Such simplifications of general theories under special circumstances can be key to getting anything useful done! One would certainly *not* want to design the pendulum for a mechanical clock using the fairly complicated mathematics of quantal theories, for instance, although the process would ultimately lead to the same result as that obtained from the simpler equations of the more restricted classical theories. Furthermore, at least at the start of the twenty-first century, a generalized ‘theory of everything’ does not yet exist. For instance, efforts to link theories of quantum electromagnetics and theories of gravity continue to be pursued.

Occasionally, a theory has proven so robust over time, even if only within a limited range of applicability, that it is called a ‘law’. For instance, Coulomb’s law specifies that the energy of interaction (in arbitrary units) between two point charges is given by

$$(1.2) \quad E = \frac{q_1 q_2}{\epsilon r_{12}}$$

where q is a charge, ϵ is the dielectric constant of a homogeneous medium (possibly vacuum) in which the charges are embedded, and r_{12} is the distance between them. However, the term ‘law’ is best regarded as honorific

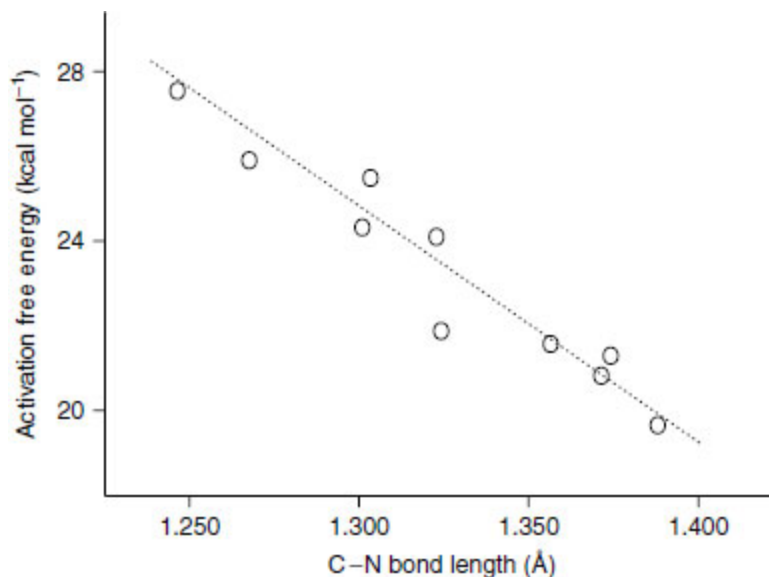
- indeed, one might regard it as hubris to imply that experimentalists *can* discern the laws of the universe within a finite span of time.

Theory behind us, let us now move on to 'model'. The difference between a theory and a model tends to be rather subtle, and largely a matter of intent. Thus, the goal of a theory tends to be to achieve as great a generality as possible, irrespective of the practical consequences. Quantum theory, for instance, has breathtaking generality, but the practical consequence is that the equations that govern quantum theory are intractable for all but the most ideal of systems. A model, on the other hand, typically involves the deliberate introduction of simplifying approximations into a more general theory so as to extend its practical utility. Indeed, the approximations sometimes go to the extreme of rendering the model deliberately qualitative. Thus, one can regard the valence-shell-electron-pair repulsion (VSEPR; an acronym glossary is provided as Appendix A of this text) model familiar to most students of inorganic chemistry as a drastic simplification of quantum mechanics to permit discrete choices for preferred conformations of inorganic complexes. (While serious theoreticians may shudder at the empiricism that often governs such drastic simplifications, and mutter gloomily about lack of 'rigor', the value of a model is not in its intrinsic beauty, of course, but in its ability to solve practical problems; for a delightful cartoon capturing the hubris of theoretical dogmatism, see Ghosh 2003.)

Another feature sometimes characteristic of a *quantitative* 'model' is that it incorporates certain constants that are derived wholly from experimental data, i.e., they are empirically determined. Again, the degree to which this distinguishes a model from a theory can be subtle. The speed of light and the charge of the electron are fundamental constants of the universe that appear either

explicitly or implicitly in [Eqs. \(1.1\)](#) and [\(1.2\)](#), and we know these values only through experimental measurement. So, again, the issue tends to be intent. A model is often designed to apply specifically to a restricted volume of what we might call chemical space. For instance, we might imagine developing a model that would predict the free energy of activation for the hydrolysis of substituted β -lactams in water. Our motivation, obviously, would be the therapeutic utility of these species as antibiotics. Because we are limiting ourselves to consideration of only very specific kinds of bond-making and bond-breaking, we may be able to construct a model that takes advantage of a few experimentally known free energies of activation and correlates them with some other measured or predicted quantity. For example, we might find from comparison with X-ray crystallography that there is a linear correlation between the aqueous free energy of activation, ΔG^\ddagger , and the length of the lactam C-N bond in the crystal, r_{CN} ([Figure 1.1](#)). Our 'model' would then be

Figure 1.1 Correlation between activation free energy for aqueous hydrolysis of β -lactams and lactam C-N bond lengths as determined from X-ray crystallography (data entirely fictitious)



$$(1.3) \quad \Delta G^\ddagger = a r_{\text{CN}} + b$$

where a would be the slope (in units of energy per length) and b the intercept (in units of energy) for the empirically determined correlation.

[Equation \(1.3\)](#) represents a very simple model, and that simplicity derives, presumably, from the small volume of chemical space over which it appears to hold. As it is hard to imagine deriving [Eq. \(1.3\)](#) from the fundamental equations of quantum mechanics, it might be more descriptive to refer to it as a 'relationship' rather than a 'model'. That is, we make some attempt to distinguish between correlation and causality. For the moment, we will not parse the terms too closely.

An interesting question that arises with respect to [Eq. \(1.3\)](#) is whether it may be more broadly applicable. For instance, might the model be useful for predicting the free energies of activation for the hydrolysis of γ -lactams? What about amides in general? What about imides? In a statistical sense, these chemical questions are analogous to asking about the degree to which a correlation may be trusted for extrapolation vs. interpolation. One might say that we have derived a correlation involving two axes of multi-dimensional chemical space, activation free energy for β -

lactam hydrolysis and β -lactam C-N bond length. Like any correlation, our model is expected to be most robust when used in an interpolative sense, i.e., when applied to newly measured β -lactam C-N bonds with lengths that fall within the range of the data used to derive the correlation. Increasingly less certain will be application of [Eq. \(1.3\)](#) to β -lactam bond lengths that are *outside* the range used to derive the correlation, or assumption that other chemical axes, albeit qualitatively similar (like γ -lactam C-N bond lengths), will be coincident with the abscissa.

Thus, a key question in one's mind when evaluating any application of a theoretical model should be, 'How similar is the system being studied to systems that were employed in the development of the model?' The generality of a given model can only be established by comparison to experiment for a wider and wider variety of systems. This point will be emphasized repeatedly throughout this text.

Finally, there is the definition of 'computation'. While theories and models like those represented by [Eqs. \(1.1\)](#), [\(1.2\)](#), and [\(1.3\)](#), are not particularly taxing in terms of their mathematics, many others can only be efficiently put to use with the assistance of a digital computer. Indeed, there is a certain synergy between the development of chemical theories and the development of computational hardware, software, etc. If a theory cannot be tested, say because solution of the relevant equations lies outside the scope of practical possibility, then its utility cannot be determined. Similarly, advances in computational technology can permit existing theories to be applied to increasingly complex systems to better gauge the degree to which they are robust. These points are expanded upon in Section 1.4. Here we simply close with the concise statement that 'computation' is the use of digital technology to solve the mathematical equations defining a particular theory or model.