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