Christophe Chipot
Andrew Pohorille

Editors

Free Energy Calculations

Theory and Applications
in Chemistry and Biology

Springer

Springer Series in
CHEMICAL PHYSICS 86

Springer Series in
CHEMICAL PHYSICS

Series Editors: A. W. Castleman, Jr. J. P. Toennies K. Yamanouch W. Zinth

The purpose of this series is to provide comprehensive up-to-date monographs in both well established disciplines and emerging research areas within the broad fields of chemical physics and physical chemistry. The books deal with both fundamental science and applications, and may have either a theoretical or an experimental emphasis. They are aimed primarily at researchers and graduate students in chemical physics and related fields.

Ch. Chipot   A. Pohorille   (Eds.)

Free Energy Calculations

Theory and Applications
in Chemistry and Biology

With 85 Figures, 11 in Color and 2 Tables

Springer

Christophe Chipot
Equipe de Dynamiques des Assemblages Membranaires
UMR CNRS/UHP No 7565
Université Henri Poincaré, BP 239
54506 Vandœvre-lès-Nancy cedex, France
E-Mail: Christophe.Chipot@edam.uhp-nancy.fr

Andrew Pohorille
University of California San Francisco
Department of Pharmaceutical Chemistry
600 16th Street
San Francisco, CA 94143, USA
and
NASA Ames Research Center
MS 239-4
Moffett Field, CA 94035, USA
E-Mail: pohorill@max.arc.nasa.gov

Series Editors:
Professor A. W. Castleman, Jr.
Department of Chemistry, The Pennsylvania State University
152 Davey Laboratory, University Park, PA 16802, USA

Professor J.P. Toennies
Max-Planck-Institut für Strömungsforschung, Bunsenstrasse 10
37073 Göttingen, Germany

Professor K. Yamanouchi
University of Tokyo, Department of Chemistry
Hongo 7-3-1, 113-0033 Tokyo, Japan

Professor W. Zinth
Universität München, Institut für Medizinische Optik
Öttingerstr. 67, 80538 München, Germany

Foreword

Andrew Pohorille and Christophe Chipot

In recent years, impressive advances have been made in the calculation of free energies in chemical and biological systems. Whereas some of this be ascribed to a rapid increase in computational power, progress has been facilitated primarily by the emergence of a wide variety of methods that have greatly improved both the efficiency and accuracy of free energy calculations. This progress has, however, come at a price: it is increasingly difficult for researchers to find their way through the maze of available computational techniques. Why are there so many methods? Are they conceptually related? Do they differ in efficiency and accuracy? Why do methods that appear to be very similar carry different names? Which method is the best for a specific problem? These questions leave not only most novices but also many experts in the field confused and desperately looking for guidance.

As a response, we attempt to present in this book a coherent account of the concepts that underlie the different approaches devised for the determination of free energies. Our guiding principle is that most of these approaches are rooted in a few basic ideas, which have been known for quite some time. These original ideas were contributed by such pioneers in the field as John Kirkwood [1, 2], Robert Zwanzig [3], Benjamin Widom [4], John Valleau [5] and Charles Bennett [6]. With a few exceptions, recent developments are not so much due to the discovery of ground-breaking, new fundamental principles, but rather to astute and ingenious ways of applying those already known. This statement is not meant as a slight on the researchers who have contributed to these developments. In fact, they have produced a considerable body of beautiful theoretical work, based on increasingly deep insights into statistical mechanics, numerical methods and their applications to chemistry and biology. We hope, instead, that this view will help to introduce order into the seemingly chaotic field of free energy calculations.

The present book is aimed at a relatively broad readership that includes advanced undergraduate and graduate students of chemistry, physics and engineering, post-doctoral associates and specialists from both academia and industry who carry out research in the fields that require molecular modeling and numerical simulations. This book will also be particularly useful to students in biochemistry, structural

biology, bioengineering, bioinformatics, pharmaceutical chemistry, as well as other related areas, who have an interest in molecular-level computational techniques.

To benefit fully from this book readers should be familiar with the fundamentals of statistical mechanics at the level of a solid undergraduate course, or an introductory graduate course. It is also assumed that the reader is acquainted with basic computer simulation techniques, in particular the molecular dynamics (MD) and Monte Carlo (MC) methods. Several very good books are available to learn about these methodologies, such as that of Allen and Tildesley [7], or Frenkel and Smit [8]. In the case of Chaps. 4 and 11, a basic knowledge of classical and quantum mechanics, respectively, is a prerequisite. The mathematics required is at the level typically taught to undergraduates of science and engineering, although occasionally more advanced techniques are used.

The book consists of 14 chapters, in which we attempt to summarize the current state of the art in the field. We also offer a look into the future by including descriptions of several methods that hold great promise, but are not yet widely employed. The first six chapters form the core of the book. In Chap. 1, we define the context of the book by recounting briefly the history of free energy calculations and presenting the necessary statistical mechanics background material utilized in the subsequent chapters.

The next three chapters deal with the most widely used classes of methods: free energy perturbation (FEP) [3], methods based on probability distributions and histograms, and thermodynamic integration (TI) [1, 2]. These chapters represent a mix of traditional material that has already been well covered, as well as the description of new techniques that have been developed only recently. The common thread followed here is that different methods share the same underlying principles. Chapter 5 is dedicated to a relatively new class of methods, based on calculating free energies from nonequilibrium dynamics. In Chap. 6, we discuss an important topic that has not received, so far, sufficient attention – the analysis of errors in free energy calculations, especially those based on perturbative and nonequilibrium approaches.

In the next three chapters, we cover methods that do not fall neatly into the four groups of approaches described in Chaps. 2–5, but still have similar conceptual underpinnings. Chapter 7 is devoted to path sampling techniques. They have been, so far, used primarily for chemical kinetics, but recently have become the object of increased interest in the context of free energy calculations. In Chap. 8, we discuss a variety of methods targeted at improving the sampling of phase space. Here, readers will find the description of techniques such as multi-canonical sampling, Tsallis sampling and parallel tempering or replica exchange. The main topic of Chap. 9 is the potential distribution theorem (PDT). Some readers might be surprised that this important theorem comes so late in the book, considering that it forms the theoretical basis, although often not spelled out explicitly, of many methods for free energy calculations. This is, however, not by accident. The chapter contains not only relatively well-known material, such as the particle insertion method [4], but also a generalized formulation of the potential distribution theorem followed by an outline of the quasi-chemical theory and its applications, which may be unfamiliar to many readers.

Chapters 10 and 11 cover methods that apply to systems different from those discussed so far. First, the techniques for calculating chemical potentials in the grand canonical ensemble are discussed. Even though much of this chapter is focused on phase equilibria, the reader will discover that most of the methodology introduced in Chap. 3 can be easily adapted to these systems. Next, we will provide a brief presentation of the methods devised for calculating free energies in quantum systems. Again, it will be shown that many techniques described previously for classical systems, such as PDT, FEP and TI, can be profitably applied when quantum effects are taken into account explicitly.

In Chap. 12, we discuss approximate methods for calculating free energies. These methods are of particular interest to those who are interested in computer-aided drug design and in silico genetic engineering. Chapter 13 provides a brief and necessarily incomplete review of significant, current and future applications of free energy calculations to systems of both chemical and biological interest. One objective of this chapter is to establish the connection between the quantities obtained from computer simulations and from experiments. The book closes with a short summary that includes recommendations on how the different methods presented here should be chosen for several specific classes of problems. Although the book contains no exercises, most chapters provide examples and pseudo-codes to illustrate how the different free energy methods work.

Each chapter is written by one or several authors, who are specialists in the area covered by the chapter. In spite of considerable efforts, this arrangement does not guarantee the level of consistency that could be attained if the book were written by a single or a small number of authors. The reader, however, gets something in return. By recruiting experts in different areas to write individual chapters, it is possible to achieve the depth in the treatment of each subject matter, that would otherwise be very hard to reach.

The material of this book is presented with greater rigor and at a higher level of detail than is customary in general reviews and book chapters on the same subject. We hope that theorists who are actively involved in research on free energy calculations, or want to gain depth in the field, will find it beneficial. Those who do not need this level of detail, but are simply interested in effective applications of existing methods, should not feel discouraged. Instead of following all the mathematical developments, they may wish to focus on the final formulas, their intuitive explanations, and some examples of their applications. Although the chapters are not truly self-contained, they may, nevertheless, be read individually, or in small clusters, especially by those with sufficient background knowledge in the field.

Several interesting topics have been excluded, perhaps somewhat arbitrarily, from the scope of this book. Specifically, we do not discuss analytical theories, mostly based on the integral equation formalism, even though they have contributed importantly to the field. In addition, we do not discuss coarse-grained, and, in particular, lattice and off-lattice approaches. At the opposite end of the wide spectrum of methods, we do not deal with purely quantum mechanical systems consisting of a small number of atoms.

On several occasions, the reader will notice a direct connection between the topics covered in the book and other, related areas of statistical mechanics, such as the methodology of computer simulations, nonequilibrium dynamics or chemical kinetics. This is hardly a surprise because free energy calculations are at the nexus of statistical mechanics of condensed phases.

Acknowledgments

References

1. Kirkwood, J. G., Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **1935**, *3*, 300–313
2. Kirkwood, J. G., in *Theory of Liquids*, Alder, B. J., Ed., Gordon and Breach, New York, 1968
3. Zwanzig, R. W., High-temperature equation of state by a perturbation method. I. Nonpolar gases, *J. Chem. Phys.* **1954**, *22*, 1420–1426
4. Widom, B., Some topics in the theory of fluids, *J. Chem. Phys.* **1963**, *39*, 2808–2812
5. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free energy estimation: umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199
6. Bennett, C. H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comp. Phys.* **1976**, *22*, 245–268
7. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon, Oxford, 1987
8. Frenkel, D.; Smit, B., *Understanding Molecular Simulations: From Algorithms to Applications*, Academic, San Diego, 1996

Contents

6 Understanding and Improving Free Energy Calculations in Molecular Simulations: Error Analysis and Reduction Methods

Nandou Lu and Thomas B. Woolf 199

7 Transition Path Sampling and the Calculation of Free Energies

Christoph Dellago 249

List of Contributors

Ioan Andricioaei
Department of Chemistry
Center for Computational Medicine and
Biology
University of Michigan
Ann Arbor, Michigan 48109–1055
andricio@umich.edu

Dilip Asthagiri
Theoretical Division
Los Alamos National Laboratory
Los Alamos, New Mexico 87545
dilipa@lanl.gov

Thomas L. Beck
Departments of Chemistry and Physics
University of Cincinnati
Cincinnati, Ohio 45221–0172
thomas.beck@uc.edu

Christophe Chipot
Equipe de Dynamique des Assemblages
Membranaires
UMR CNRS/UHP 7565
Université Henri Poincaré, BP 239
54506 Vandœuvre–lès–Nancy cedex
France
Christophe.Chipot@edam.
uhp-nancy.fr

Eric Darve
Mechanical Engineering Department
Stanford University
Stanford, California 94305
darve@stanford.edu

Christoph Dellago
Faculty of Physics
University of Vienna
Boltzmanngasse 5, 1090 Vienna, Austria
Christoph.Dellago@univie.ac.at

Gerhard Hummer
Laboratory of Chemical Physics
National Institute of Diabetes and
Digestive and Kidney Diseases
National Institutes of Health
Building 5, Room 132
Bethesda, Maryland 20892–0520
gerhard.hummer@nih.gov

Nandou Lu
Departments of Physiology and of
Biophysics and Biophysical Chemistry
School of Medicine
Johns Hopkins University
Baltimore, Maryland 21205
nlu@groucho.med.jhmi.edu

Alan E. Mark
Institute for Molecular Bioscience
The University of Queensland
Brisbane QLD 4072 Australia
a.mark@uq.edu.au

Athanassios Z. Panagiotopoulos
Department of Chemical Engineering
Princeton University
Princeton, New Jersey 08544
azp@princeton.edu

Vijay S. Pande
Departments of Chemistry
and of Structural Biology
Stanford University, Stanford
California 94305
pande@stanford.edu

Andrew Pohorille
University of California San Francisco,
Department of Pharmaceutical
Chemistry,
600 16th Street,
San Francisco, CA 94143, USA
and
NASA Ames Research Center
MS 239–4
Moffett Field, CA 94035, USA
pohorill@max.arc.nasa.gov

Lawrence R. Pratt
Theoretical Division
Los Alamos National Laboratory
Los Alamos, New Mexico 87545
lrp@lanl.gov

M. Scott Shell
Department of Pharmaceutical
Chemistry
University of California San Francisco
600 16th Street, Box 2240
San Francisco, California 94143
shell@maxwell.ucsf.edu

Thomas Simonson
Laboratoire de Biochimie
UMR CNRS 7654
Department of Biology
Ecole Polytechnique
91128 Palaiseau, France
Thomas.Simonson@polytechnique.fr

Thomas B. Woolf
Departments of Physiology and of
Biophysics and Biophysical Chemistry
School of Medicine
Johns Hopkins University
Baltimore, Maryland 21205
woolf@groucho.med.jhmi.edu

1

Introduction

Christophe Chipot, M. Scott Shell and Andrew Pohorille

1.1 Historical Backdrop

To understand fully the vast majority of chemical processes, it is often necessary to examine their underlying free energy behavior. This is the case, for instance, in protein–ligand binding and drug partitioning across the cell membrane. These processes, which are of paramount importance in the field of computer-aided, rational drug design, cannot be predicted reliably without knowledge of the associated free energy changes.

The reliable determination of free energy changes using numerical simulations based on the fundamental principles of statistical mechanics is now within reach. Developments on the methodological front in conjunction with the continuous increase in computational power have contributed to bringing free energy calculations to the level of robust and well-characterized modeling tools, while widening their field of applications.

1.1.1 The Pioneers of Free Energy Calculations

The theory underlying free energy calculations and several different approximations to its rigorous formulation were developed a long time ago. Yet, due to computational limitations at the time when this methodology was introduced, numerical applications of this theory remained very limited. In many respects, John Kirkwood laid the foundations for what would become standard methods for estimating free energy differences – perturbation theory and thermodynamic integration (TI) [1, 2]. Reconciling statistical mechanics and the concept of degree of evolution of a chemical reaction, put forth by Théophile De Donder [3] in his work on chemical affinity, Kirkwood introduced in his derivation of integral equations for liquid-state theory the notion of the order parameter, or generalized extent parameter, and used it to infer the free energy difference between two well-defined thermodynamic states [1, 2].

Almost 20 years later, Robert Zwanzig [4] followed a perturbative route to free energy calculations, showing how physical properties of a hard-core molecule change upon adding a rudimentary form of an attractive potential. The high-temperature

expansions that he established for simple, nonpolar gases form the theoretical basis of the popular free energy perturbation (FEP) method, widely employed for determining free energy differences. However, the significance of FEP was appreciated much earlier. In fact, Lev Landau [5] included a simple derivation of the thermodynamic perturbation formula in the first edition of his widely read textbook on statistical mechanics as early as 1938.

Nearly 10 years after Zwanzig published his perturbation method, Benjamin Widom [6] formulated the potential distribution theorem (PDT). He further suggested an elegant application of PDT to estimate the excess chemical potential – i.e., the chemical potential of a system in excess of that of an ideal, noninteracting system at the same density – on the basis of the random insertion of a test particle. In essence, the particle insertion method proposed by Widom may be viewed as a special case of the perturbative theory, in which the addition of a single particle is handled as a one-step perturbation of the liquid.

1.1.2 Escaping from Boltzmann Sampling

Central to the accurate determination of free energy differences between two systems – viz. target and reference – is to explore the configurational space of the reference system such that relevant, low-energy states of the target system are adequately sampled. It has long been recognized, however, that direct applications of conventional computer simulations methods, such as molecular dynamics (MD) or Monte Carlo (MC), are not successful in this respect [7]. In the late 1960s and in the 1970s a number of remarkable strategies were developed to circumvent this difficulty by generating effective non-Boltzmann sampling. The basic ideas behind these strategies have been broadly exploited in most subsequent theoretical developments.

One of the most influential ideas was the energy distribution formalism, in which free energy difference was represented in terms of a one-dimensional integral over the distribution of potential energy differences between the target and reference states weighted by the unbiased or biased Boltzmann factor. This idea was proposed and applied to calculate thermodynamic properties of Lennard-Jones fluids by McDonald and Konrad Singer [8, 9] as early as 1967. In subsequent developments it formed the conceptual basis for some of the best techniques for estimating free energies.

Returning to the concept of a generalized extent parameter, John Valleau and Damon Card [10] devised so-called multistage sampling, which relies on the construction of a chain of configurational energies that bridge the reference and the target states whenever their low-energy regions overlap poorly. The basic idea of this stratification method is to split the total free energy difference into a sum of free energy differences between intermediate states that overlap considerably better than the initial and final states.

Finding the best estimate of the free energy difference between two canonical ensembles on the same configurational space, for which finite samples are available, is a nontrivial problem. Charles Bennett [11] addressed this problem by developing the acceptance ratio estimator, which corresponds to the minimum statistical

variance. He further showed that the efficiency of this estimator is proportional to the extent to which the two ensembles overlap. A remarkable feature of Bennett's method is that, once data are collected for the two ensembles, good estimates of the free energy difference can be obtained even if the overlap between the ensembles is poor.

Another approach to improving the efficiency of free energy calculations is to sample the reference ensemble sufficiently broadly that adequate statistics about low-energy configurations of the target ensemble can be acquired. In 1977, Glenn Torrie and John Valleau [12] devised such an approach by introducing non-Boltzmann weighting function that can subsequently be removed to yield unbiased probability distribution. This method became widely known as umbrella sampling (US). It is interesting to note that an embryonic form of the US scheme had been laid 10 years earlier in the pioneering computational study of McDonald and Konrad Singer [8].

The seminal work on stratification and sampling opened new vistas for the accurate determination of free energy profiles. Both approaches are still widely used to tackle a variety of problems of physical, chemical, and biological relevance. Perhaps because they are most efficient when used in combination the distinction between them has often been lost. At present, the name 'umbrella sampling' is commonly used to describe simulations in which an order parameter connecting the initial and final ensembles is divided into mutually overlapping regions, or 'windows,' which are sampled using non-Boltzmann weights.

1.1.3 Early Successes and Failures of Free Energy Calculations

As we have already pointed out, the theoretical basis of free energy calculations were laid a long time ago [1, 4, 5], but, quite understandably, had to wait for sufficient computational capabilities to be applied to molecular systems of interest to the chemist, the physicist, and the biologist. In the meantime, these calculations were the domain of analytical theories. The most useful in practice were perturbation theories of dense liquids. In the Barker–Henderson theory [13], the reference state was chosen to be a hard-sphere fluid. The subsequent Weeks–Chandler–Andersen theory [14] differed from the Barker–Henderson approach by dividing the intermolecular potential such that its unperturbed and perturbed parts were associated with repulsive and attractive forces, respectively. This division yields slower variation of the perturbation term with intermolecular separation and, consequently, faster convergence of the perturbation series than the division employed by Barker and Henderson.

Analytical perturbation theories led to a host of important, nontrivial predictions, which were subsequently probed by and confirmed in numerical simulations. The elegant theory devised by Lawrence Pratt and David Chandler [15] to explain the hydrophobic effect constitutes a noteworthy example of such predictions.

As more computational power became accessible and confidence in the potential energy functions developed for statistical simulations increased, applications of free energy calculations to systems of chemical, physical, and biological interest began to flourish. The excellent agreement between theory and experiment reported in pioneering application studies encouraged attempts to employ similar methods to increasingly complex molecular assemblies.

Most of the earliest free energy calculations were based on MC simulations. Initial applications to Lennard-Jones fluids [8] were extended to study atomic clusters [16] and hydration of ions by a small number of water molecules [17]. Atomic clusters were also studied in one of the first applications of MD to free energy calculations [18]. All these calculations were based on the thermodynamic integration method originally proposed by Kirkwood [1]. The thermodynamic integration approach was also used by Mihaly Mezei et al. [19, 20] to calculate the free energy of liquid water. Using a different approach, based on multistage [10] and US [12] numerical schemes, Gren Patey and John Valleau [21] further extended the range of free energy calculations by deriving a free energy profile characterizing the interaction of an ion pair dissolved in a dipolar fluid.

In 1979, two studies appeared that addressed the nature of the hydrophobic effect through free energy calculations. Susumu Okazaki et al. [22] used MC simulations to estimate the free energy of hydrophobic hydration. They found that, consistent with the conventional picture of the hydrophobic effect, hydrophobic hydration is accompanied by a decrease in internal energy and a large entropy loss. In the second study, Bruce Berne and coworkers [23] adopted a multistage strategy to investigate a model system formed by two Lennard-Jones spheres in a bath of 214 water molecules. They successfully recovered the features of hydrophobic interactions predicted by Pratt and Chandler [15]. Subsequent results based on more-accurate potential energy functions and markedly extended sampling further fully confirmed these predictions – see for instance [24]. Two years later, Postma et al. [25] further contributed to our understanding of the hydrophobic effect by investigating the solvation of noble gases and estimated the reversible work required to form a cavity in water.

In the early 1980s, free energy calculations were extended in several new directions in ways that were not possible only a few years earlier. In 1980, Chyuan-Yin Lee and Larry Scott [26] estimated the interfacial free energy of water from MC simulations. In this work, they also derived and applied for the first time a useful technique that is currently often called simple overlap sampling (SOS). Two years later, Quirke and Jacucci [27] calculated the free energy of liquid nitrogen from MC simulations, Shing and Gubbins [28] used US combined with the particle insertion method to determine chemical potentials, focusing sampling on cavity volumes sufficiently large to accommodate a solute molecule, and Arieh Warshel [29] calculated the contribution of the solvation free energy to electron and proton transfer reactions, using a rudimentary hard-sphere model of the donor and acceptor, and a dipolar representation of water. The same year, Scott Northrup et al. [30] applied US simulations to examine the free energy changes in a biologically relevant system. Isomerization of a tyrosine residue in bovine pancreatic trypsin inhibitor (BPTI) was studied by rotating the aromatic ring in sequentially overlapping windows. From the resulting free energy profile, the authors inferred the rate constant for the ring-flipping reaction.

In 1984, using a very rudimentary model, Tembe and McCammon [31] demonstrated that the FEP machinery could be applied successfully to model ligand–receptor assemblies. In 1985, Jorgensen and Ravimohan [32] followed the same perturbative route to estimate the relativesolvation free energy of methanol and

ethane. To reach their goal, they elaborated an elegant paradigm, in which a common topology was shared by the reference and the target states of the transformation. Employing a similar strategy, William Jorgensen and coworkers [33, 34] pioneered the estimation of the pK_a values of simple organic solute in aqueous environments. These pioneering efforts, which initially met with only moderate enthusiasm, constitute what might be considered today the turning point for free energy calculations on chemically relevant systems, paving the way for extensions to far more complex molecular assemblies.

In early studies, complete free energy profiles along a chosen order parameter were obtained by combining US and stratification strategies—e.g. Chandrasekhar et al. investigated the S_N2 reaction of $Cl^- + CH_3Cl$, both in the gas phase and in aqueous solution [35], thus, laying the ground for the forthcoming hybrid quantum mechanical/molecular mechanical (QM/MM) calculations. In 1987, Douglas Tobias and Charles Brooks III showed that the same information could be extracted from thermodynamic perturbation theory. They did so by constructing the free energy profile for separating two tagged argon atoms in liquid argon [36].

The same year, Peter Kollman and coworkers published three papers that opened new horizons for in silico modeling of site-directed mutagenesis. Employing the FEP methodology, they estimated the free energy changes associated with point mutations of the side chains of naturally occurring amino acids [37]. They used the same approach for computing the relative binding free energies in protein–inhibitor complexes of thermolysin [38] and substilisin [39]. The same year, they also explored an alternative route to the costly FEP calculations, in which perturbation was carried out using very minute increments of the general extent, or coupling parameter [40]. It is worth mentioning, however, that this so-called 'slow-growth' (SG) strategy had to wait for 10 years and the work of Christopher Jarzynski [41] to find a rigorous theoretical formulation. Yet, during that period, a number of ambitious problems were tackled employing SG simulations, including a heroic effort to understand structural modifications in deoxyribonucleic acid (DNA) [42].

Considering that the chemical transformations attempted hitherto involved only one or two atoms, the series of articles from the group of Peter Kollman appeared to represent a quantum leap forward. It was soon recognized, however, that these calculations were evidently too short and probably not converged. They demonstrated, nonetheless, that modeling biologically relevant systems was a realistic goal for the computational chemist.

Also back in 1987, Fleischman and Brooks [43] devised an efficient approach to the estimation of enthalpy and entropy differences. They concluded that the errors associated with the calculated enthalpies and entropies were about one order of magnitude larger than those of the corresponding free energies. Only recently did Lu et al. [44] revisit this issue, proposing an attractive scheme to improve the accuracy of enthalpy and entropy calculations. Wilfred van Gunsteren and coworkers [45] further concluded that reasonably accurate estimates of entropy differences might be obtained through the TI approach, in which several copies of the solute of interest are desolvated. It is fair to acknowledge that, although several improvements to the original approaches for extracting enthalpic and entropic contributions to free energies

have recently been put forth, the conclusions drawn by Fleischman and Brooks remain qualitatively correct.

In contrast to FEP and US, TI was not widely applied in the late 1970s and early 1980s. Only in the late 1980s, did TI regain its well-deserved position as one of the most useful techniques to obtain free energies from computer simulations. In 1988, Tjerk Straatsma and Herman Berendsen [46] used this technique to study the free energy of ionic hydration by performing the mutation of neon into sodium. Three years later, Wang et al. [47] used TI to construct the free energy profile describing interactions between two hydrophobic solutes – viz. a pair of neon atoms in a bath of water. Today, TI remains one of the favorite methods for free energy calculations.

Several research groups paved the way for future progress through innovative applications of free energy methods to physical and organic chemistry, as well as structural biology. An exhaustive account of the plethora of articles published in the early years of free energy calculations falls beyond the scope of this introduction. The reader is referred to the review articles by William Jorgensen [48, 49], David Beveridge and Frank DiCapua [50, 51] and Peter Kollman [52], for summaries of these efforts.

1.1.4 Characterizing, Understanding, and Improving Free Energy Calculations

After the initial enthusiasm ignited by pioneering studies, which often reported excellent agreement between computed and experimentally determined free energy differences, it was progressively realized that the some of the published, highly promising results reflected good fortune rather than actual accuracy of computer simulations. For example, in many instances, it was observed that the calculated free energy differences showed a tendency to depart from the experimental target value as more sampling was accumulated. It became widely appreciated that many free energy calculations were plagued by an inherent slow convergence, sometimes to such extent that, for all practical purposes, systems under study appeared nonergodic. These observations clearly indicated that improved sampling and analysis techniques were needed. Efforts were thus expended, with excellent results, to address these issues. It was further discovered that several aspects of early calculations had not been treated with sufficient care to theoretical details. In the subsequent years, the underlying methodological problems received considerable attention and at present most of them have been solved. Along different lines, much work was devoted to large-scale free energy calculations, especially in the biological domain, in which improved efficiency was achieved by relaxing theoretical rigor through a series of well-motivated approximations. Below, we outline some of the main advances of the last 15 years. A more complete account of these advances is given in the subsequent chapters.

A large body of methodological work is devoted to clarifying and improving the basic strategies for determining free energy – stratification, US, FEP, and TI methods. A common class of problems involves calculating free energy along an order parameter – e.g., the reaction coordinate, based on a combination of US and

stratification. The efficiency of these methods relies on designing biases that improve the uniformity of sampling. Intuitive guesses of such biases may turn out to be very difficult, especially for qualitatively new problems. Improperly set biasing potentials could result in highly nonuniform probability distributions and a paucity of data at some values of the order parameter. To improve accuracy, additional simulations with revised biases are required. This raises a question: what is the optimal scheme for combining the data acquired at different ranges of the order parameter and using different biases?

Recasting the Ferrenberg–Swendsen multiple histogram equations [53], Kumar et al. [54] answered this question by devising the weighted histogram analysis method (WHAM). WHAM rapidly superseded previously used ad hoc methods and became the basic tool for constructing free energy profiles from distributions derived through stratification.

Four years later, Christian Bartels and Martin Karplus [55] used the WHAM equations as the core of their adaptive US approach, in which the efficiency of free energy calculations was improved through refinement of the biasing potentials as the simulation progressed. Efforts to develop adaptive US techniques had, however, started even before WHAM was developed. They were pioneered by Mihaly Mezei [56], who used a self-consistent procedure to refine non-Boltzmann biases.

Observing that stratification strategies, which rely on breaking the path connecting the reference and the target states into intermediate states, often led to singularities and numerical instabilities at the end points of the transformation, Beutler et al. [57] suggested that introducing a soft-core potential might alleviate end-point catastrophes. This simple technical trick turned out to be a highly successful approach to estimate solvation free energies in computationally challenging systems, involving, for example, the creation or annihilation of chemical groups.

Another technical problem that plagued early estimations of free energy was their strong dependence on system size whenever significant electrostatic interactions were present [46]. Once long-range corrections using Ewald lattice summation or the reaction field are included in molecular simulations, size effects in neutral systems decrease markedly. The problem, however, persists in charged systems, for example in determining the free energy of charging a neutral species in solution. Hummer et al. [58] showed that system-size dependence could be largely eliminated in these cases by careful treatment of the self-interaction term, which is associated with interactions of charged particles with their periodic images and a uniform neutralizing charge background. Surprisingly, they found that it was possible to calculate accurately the hydration energy of the sodium ion using only 16 water molecules if self-interactions were properly taken into account.

The determination of the character and location of phase transitions has been an active area of research from the early days of computer simulation, all the way back to the 1953 Metropolis et al. [59] MC paper. Within a two-phase coexistence region, small systems simulated under periodic boundary conditions show regions of apparent thermodynamic instability [60]; simulations in the presence of an explicit interface eliminate this at some cost in system size and equilibration time. The determination of precise coexistence boundaries was usually done indirectly, through the

use of a method to determine the free energies of the coexisting phases, such as TI or the particle insertion method [61, 62]. A notable advance emerged with the Gibbs ensemble approach [63], in which two phases were simulated directly without an interface by coupling separate simulation boxes via particle and volume fluctuations. In the last 10 years, however, the preferred approach to fluid-phase coexistence has become histogram reweighting methods, which offer greater control over simulation errors and enable more precise determination of critical points than the Gibbs ensemble [64]. For equilibria involving dense fluid or solid phases – for which attempted particle insertions are infrequently accepted – the approach of tracing phase coexistence lines by Gibbs–Duhem integration [65] remains a primary technique.

An aspect of free energy calculations that caused considerable, and somewhat surprising difficulties is the treatment of holonomic constraints. In numerical simulations these are often used to remove high-frequency vibrations, and by doing so allow the equations of motion to be integrated with larger time steps. In the early years of free energy calculations, the effect of frozen internal degrees of freedom on the generated ensemble was essentially ignored [66]. It was shown, however, that hard constraints might alter the accessible volume of phase space, and, consequently, might significantly influence the computed free energy differences. Stefan Boresch and Martin Karplus [67] pointed out the importance of metric tensor corrections in free energy calculations, and showed that, in a number of instances, these corrections could be evaluated analytically. To a large extent, the foundations for the treatment of constrained internal degrees of freedom may be found in the articles of Marshall Fixman [68] and Nubuhiro Gō and Harold Scheraga [69], published some 20 years earlier.

Holonomic constraints also appear in the determination of free energy profiles along a chosen order parameter, ξ, using TI. In this framework, the thermodynamic force – i.e., the first derivative of the free energy with respect to the order parameter – is calculated at fixed values of the parameter and subsequently integrated to recover the free energy profile along ξ. Wilfred van Gunsteren [70] hypothesized that the thermodynamic force was equal to the constraint force acting along ξ. It, however, soon became apparent that this conjecture was incorrect whenever ξ was a nonlinear function of the Cartesian coordinates. A rigorous framework for handling holonomic constraints in the simulation of rare events was proposed the very same year by Carter et al. [71]. The complete treatment of such constraints in free energy calculations that involved other rigid constraints was proposed nearly another decade later by Wouter den Otter and Wim Briels [72], and further extended to the multidimensional case [73].

Almost immediately, it was realized that keeping the system at fixed values of the order parameter was not a prerequisite to calculating the thermodynamic force. Following a different route than den Otter and Briels, Eric Darve, and Andrew Pohorille derived the formulas for this force in both constrained and unconstrained simulations. They further showed how the latter could be used to combine TI and US into a highly efficient scheme that yielded uniform sampling of the order parameter. They called this approach the adaptive biasing force (ABF) method [74]. Gains in efficiency of ABF, compared to the previous adaptive US schemes based on probability distribution functions, are due to the fact that forces, in contrast to probabilities,

are local properties and, therefore, they can be readily estimated without the need to sample broad ranges of ξ. The efficiency of this approach in the treatment of complex systems has been demonstrated by Jérôme Hénin and Christophe Chipot [75, 76].

ABF is an example of a strategy in which nearly optimal sampling of a low-dimensional configurational space is achieved even in the presence of high free energy barriers. In recent years, other strategies aimed at the same goal have been proposed. In 2002, Allessandro Laio and Michele Parrinello [77] introduced a meta-dynamics approach for exploring free energy surfaces that relied on the definition of collective degrees of freedom to which coarse-grained, non-Markovian dynamics was applied. A memory kernel guarantees that, as the simulation progresses, the visited minima of the free energy landscape are continuously filled, ensuring that, in the long run, exploration of the system is uniform.

Some of the most efficient techniques for sampling configurational space were developed in association with the MC method rather than MD. In 1992, Berg and Neuhaus [78] devised a multicanonical method in which weighting factors that yield equiprobable distributions of order parameters are determined through an iterative procedure. A similar underlying idea is at the origin of the method proposed by Fugao Wang and David Landau [79]. In their algorithm, independent random walks are performed over different ranges of the order parameter – e.g., the energy. The derived density of states is then updated in a continuous fashion, eventually yielding flat probability distributions. This method, originally designed for discrete lattice systems, was later adapted to continuum fluids by Shell et al. [80] and Yan et al. [81]. A somewhat different approach was taken by Smith and Bruce [82, 83] in their transition matrix method. Instead of estimating probabilities of visiting different states of the system, they calculated transition probabilities between macrostates. This method proved to generate excellent estimates of thermodynamic functions with a high statistical accuracy. Another multicanonical strategy devised by John Valleau allows a range of both densities and temperatures to be spanned in a single simulation, thus giving access to accurate free energies and other ensemble averages [84, 85].

In comparison with MC-based methods, US-based molecular dynamics appeared to be limited by the fact that order parameters had to be dynamical variables, for which equations of motion existed. This limitation was removed by introducing to free energy calculations the extended ensemble formalism. In 1996, Xianjun Kong and Charles Brooks III [86] adopted an extended Hamiltonian approach, which allowed general order parameters to be treated as dynamical variables, to follow a pathway along which the free energy is always minimal. The same idea forms the basis of an algorithm recently put forth by Bitetti-Putzer et al. [87]. The authors observed that using the generalized ensemble helped to cross free energy barriers and to overcome kinetic traps. An extended ensemble formalism is also an inherent part of the previously discussed method proposed by Laio and Parrinello [77].

In the early 1990s, another approach was developed for improving the efficiency of free energy calculations through non-Boltzmann sampling [88–91]. Its basic idea is to construct simultaneously a series of MD trajectories or MC walks that are characterized by different values of an order parameter. The method is effective if the probability of visiting different states of the system varies significantly for the target

value of the parameter, but becomes progressively smoother as the parameter increases, or decreases. Occasionally, one attempts to update the simulations by swapping configurations between the systems characterized by the consecutive values of the parameter, and accepting this modification according to the Metropolis criterion. The result is that the rugged nature of the probability density function at the target value of the parameter is *tempered* by exchanging configurations with those sampled from smoother probability distributions. For this reasons, the approach is called *parallel tempering*, although versions of this method are also known under different names, such as replica exchange and J-walking. A suitable and most frequently used parameter that increases smoothness of the probability distribution and efficiency of sampling is the temperature, although other choices are possible and occasionally employed. In recent years, the method has gained considerable popularity as a successful approach to problems that involve high-energy barriers between different states of the system.

Also in the early 1990s, a somewhat related method for calculating free energy differences was proposed by Ron Elber and coworkers [92, 93]. It relies on simulating multiple, noninteracting replicas that differ only locally. As a result, the method is applicable to systems that undergo only local modification – e.g., point mutations in proteins. For this reason, it has been called the locally enhanced sampling (LES) technique.

In contrast to the FEP, US, and TI methods, which provided general routes to calculating free energy, methods based on the PDT had only limited applications. Their standard formulation, the particle insertion method, was successful only if the cavities formed spontaneously due to thermal fluctuations in the solvent were sufficiently large to accommodate solvent molecules. These methods, however, proved to be of considerable conceptual importance, especially in improving our understanding of the hydrophobic effect. To this end, particularly influential was the work of Hummer et al. [94]. Building on the earlier studies of Pratt and Pohorille [95, 96], they connected information theory with statistical mechanics to model the probability distribution of solvent centers in a given cavity volume. This approach was not only able to describe the primitive hydrophobic effects that drive cavity formation in water and association of nonpolar solutes but also provided a convenient framework for investigating other hydrophobic phenomena, such as the conformational equilibria in alkanes and nonpolar peptide chains, and the effects of temperature and pressure on protein folding.

Recently, Lawrence Pratt and coworkers applied the generalized form of the PDT, which included averaging not only over particle positions but also over molecular orientations and conformations, in a new context. They developed a quasichemical theory for the evaluation of solution free energies [97] and applied it to several challenging problems, such as the hydration free energy of ions – viz. H^+, Li^+, Na^+ and HO^- [98]. They further argued that the PDT forms the basis for approaches to calculating free energies that are as general and practical as other, widely used methods.

One of the most important theoretical developments of the last decade is due to Chris Jarzynski, who established a remarkably simple relationship between the equilibrium free energy difference and an ensemble of properly constructed irreversible

transformations linking the initial and final state of the system [41, 99]. Jarzynski's identity laid the foundations for a new, general class of methods for estimating free energies, which is applicable to phenomena that are either irreversible or clearly driven out of equilibrium. Not surprisingly, this work stimulated further theoretical developments [100], and applications on both the experimental [101] and computational [102] fronts.

In one of the most advanced applications of the nonequilibrium method, Klaus Schulten and coworkers [103] coupled steered MD simulations with the Jarzynski identity to derive the free energy profile that characterizes glycerol conduction in the aquaglyceroporin GlpF [104]. This computationally challenging study, which required MD simulations of a system composed of approximately 106,000 atoms, provided theoretical support for the proposed mechanism of glycerol transport by identifying potential binding sites, energy barriers, and a vestibular low-energy region conducive to glycerol uptake within the channel.

Further improvements to Jarzynski's method were proposed in 2004 by Marty Ytreberg and Daniel Zuckerman [105], who combined it with a path sampling scheme. Transition path sampling was used to refine in an iterative fashion the reaction pathway along which the nonequilibrium work was evaluated. Compared to standard calculations relying on the Jarzynski identity, this approach appears to be substantially more effective, because it favors rare events involving small works, and focuses sampling on regions that truly contribute to the free energy change.

Until recently, advances in calculating the free energy were not accompanied by comparable progress in rigorous error analysis and reduction. Although a variety of methods to estimate the error in calculated free energies were proposed [32, 106], they were usually somewhat heuristic or involved approximations that were not always sufficiently well supported. Only recently, considerable progress has been made on this front, in particular by Daniel Zuckerman and Thomas Woolf [107].

An interesting approach for eliminating the systematic sampling bias caused by the exponential averaging in FEP calculations has been proposed by Lu et al. [108]. In a nutshell, it relies on a combination of the forward and reverse transformations between a reference and target state, employing Bennett's acceptance ratio [11] for the optimal averaging of these simulations in terms of overlap sampling. The merit of the scheme devised by Lu et al. lies in the reconciliation of two techniques that have been employed widely, albeit always independently and for different purposes – i.e., running forward and reverse simulations, usually to infer some estimate of the statistical error associated with the free energy difference [32], and the long-known, elegant method put forth by Charles Bennett back in 1976. Amazingly enough, the connection between these two commonly adopted sampling strategies had to wait almost 20 years to be clearly articulated. The latter illustrates that concepts once popular may become dormant, until they are rediscovered years later and used in a computationally more attractive version.

Realizing that practical application of free energy calculations outside the purely academic environment, in particular in the pharmaceutical industry, required significant cost reductions, much effort was invested towards developing faster and cheaper methods for estimating free energy differences in complex systems. The goal

for this line of research, primarily aimed at drug design applications, was quite ambitious: to make approximate methods sufficiently efficient and reliable that they would provide answers faster than laboratory experiments [109].

In this context, of particular interest are protein–ligand associations, which are typically accompanied by significant conformational changes. Since these changes occur on time scales that make direct, atomic-level simulation of these processes impractical, alternative, simplified strategies had to be devised. One such strategy was proposed by Åqvist. He assumed that the change in the binding free energy due to the mutation of a ligand associated with a protein obeyed the linear response theory [110]. Empirical parameters that appeared in his formulation were determined from training sets of protein–ligand complexes and were subsequently applied to predict the binding affinities of new ligands.

Another computational strategy relied on simultaneous in silico creation of the ligand in the free and the bound states. The term *creation*, that will be discussed in detail in Chap. 2, refers to the progressive scaling of parameters that describe interaction of the ligand with its environment. Andrew McCammon and coworkers [111] laid the statistical–mechanical foundations for deriving protein–ligand association constants, showing, in particular, how the double-creation scheme should be modified to obtain rigorous binding free energies. In a related work, Jan Hermans and Lu Wang [112] proposed a complete treatment of the binding free energy, which included the so-called *cratic* term arising from the loss of rotational and translational entropy upon association.

In 2000, Erin Duffy and William Jorgensen [113] simulated a set of 200 organic solutes of potential pharmaceutical interest in aqueous solution. Using an automated procedure, they inferred solvation free energies on the basis of configurationally averaged descriptors obtained through linear regression. Noting that the estimated free energies were sensitive to the choice of the net atomic charges on the solutes, they proposed that specific corrections be included in the regression equations for poorly described functional groups. With the increase of computational power, William Jorgensen showed how lead optimization could be guided employing FEP calculations to design new, very potent anti-HIV-1 agents [114]. To find a compromise between accurate but low-throughput free energy calculations and inexpensive but generally poor-scoring function-based schemes, David Pearlman and Paul Charifson [115] suggested that one-step FEP simulations on a grid surrounding the solute of interest represented a promising tradeoff for high-throughput determination of protein–ligand binding constants.

Paul Smith and Wilfred van Gunsteren [116, 117] suggested an approach to inferring a set of free energy differences based on a single simulation of the initial state. Herman Berendsen and coworkers [118] developed another strategy, which was based on the potential energy distribution function. Using a quasi-Gaussian entropy theory, the free energy and entropy changes were expressed in terms of the potential energy moments. This approach was shown to reproduce accurately the free energy of water and methanol over an appreciable range of temperatures.

New horizons for treating computationally challenging problems opened with the emergence of reliable implicit solvation models. For example, Simonson et al. [119]

showed that a continuum treatment of long-range interactions could be used in free energy calculations without sacrificing accuracy, which led to significant reductions in the cost of atomistic simulations. More recently, the application of an implicit solvation scheme to the calculation of association free energies was revisited by Andrew McCammon and coworkers [120]. Employing a molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) model, they successfully tackled the difficult problem of estimating changes in the conformational free energy upon binding of a ligand to its receptor. In general, application of implicit solvent to protein–ligand assemblies, in which solvent molecules do not contribute directly to the association, is a possible answer to the need for high-throughput de novo drug design in industrial settings.

The vast majority of free energy theory/calculation approaches originate from a classical statistical–mechanical underpinning. This assumption is appropriate for a wide range of ion and molecule solvation problems. Even in the early stages of the development of free energy methodology, however, emphasis was placed on quantum aspects of free energies. These early developments followed two general lines. In the first, Eugene Wigner and John Kirkwood, as early as the 1930s, derived an expansion for the free energy in powers of $\hbar$; the first term in the series is the classical free energy, and subsequent terms yield increasingly accurate quantum corrections. In addition, an effective potential can be derived which allows for a classically based simulation moving on a quantum-modified potential. The Wigner–Kirkwood and thermodynamic perturbation theory approaches are described thoroughly in reference [5]. The second line in the development of approximate quantum free energy methods was the discovery of variational approaches pioneered by Richard Feynman [121], Albert Hibbs [122], and Hagen Kleinert [123]. Starting from a path integral description of the quantum system, and integrating out the path modes, effective potentials were derived which ensure that the computed free energies were above the exact result. More recently, the PDT has been extended to the quantum domain using Feynman path integral methods [124, 125], and these ideas have found utility in modeling quantum behavior in fluids [126, 127].

The ideas mentioned in this section, and many others, will be discussed in detail in subsequent chapters. As we have already stressed, the goal of this section is not to be exhaustive. Instead, the guiding idea has been to show how developments in the field were motivated by the theoretical and practical challenges arising as both the computational power and the popularity of free energy calculations increased. The reader interested in learning more about the history of free energy calculations is referred to the previously mentioned articles by William Jorgensen [48, 49], David Beveridge and Frank DiCapua [50, 51], and Peter Kollman [52] from the late 1980s and early 1990s, and to more recent reviews by Thomas Simonson et al. [128], Christophe Chipot and David Pearlman [129], Bruce Berne and John Straub [130], as well as Tomas Rodinger and Régis Pomès [131].

1.2 The Density of States

In the remainder of this chapter, we review the fundamentals that underlie the theoretical developments in this book. We outline, in sequence, the concept of density of states and partition function, the most basic approaches to calculating free energies and the essential strategies for improving the efficiency of these calculations. The ideas discussed here are, most likely, known to the reader. They can also be found in classical books on statistical mechanics [132–134] and molecular simulations [135, 136]. Thus, we do not attempt to be exhaustive. On the contrary, we present the material in a way that is most directly relevant to the topics covered in the book.

The density of states is the central function in statistical thermodynamics, and provides the key link between the microscopic states of a system and its macroscopic, observable properties. In systems with continuous degrees of freedom, the correct treatment of this function is not as straightforward as in lattice systems – we, therefore, present a brief discussion of its subtleties later. The section closes with a short description of the microcanonical MC simulation method, which demonstrates the properties of continuum density of states functions.

1.2.1 Mathematical Formalism

We begin by considering the density of states, Ω, or microcanonical partition function for a single-component, structureless fluid of N particles – although the extension to structured, or multicomponent systems is rather straightforward. Our use of the notation Ω refers to the energy *density* of microstates, and not the integrated phase space volume [137, 138]. Although there has been some debate about which is appropriate to the microcanonical entropy, the former is tied to histograms, as discussed in Chap. 3, and, hence, it is our focus here. For an in-depth mathematical treatment of these issues, the reader is referred to [137–139].

For discrete systems such as the Ising model, the density of states counts the number of microstate configurations of the system consistent with each macrostate – e.g., $\Omega(\mathscr{E})$ gives the number of microstates with energy $\mathscr{E}$. In a system with continuous degrees of freedom, this 'counting' is ill-defined because the number of configurations is infinite. In contrast, for our fluid, we consider the entire $3N$-dimensional space defined by all the coordinates of the particles, and let the 'number' of configurations of a given potential energy be proportional to the $(3N - 1)$-dimensional area of the associated energy hypersurface. In mathematical terms, this translates to:

$$\Omega_{\mathrm{con}} \propto \frac{1}{N!} \int_{V^N} \delta[U(\mathbf{q}) - \mathscr{E}] \, \mathrm{d}\mathbf{q}. \qquad (1.1)$$

Here, δ is the Dirac delta function, U is the potential energy function, and $\mathbf{q}$ represents the $3N$ coordinates. In this expression, the integral is performed over the entire configuration space – each coordinate runs over the volume of the simulation box, and the delta function 'selects' only those configurations of energy $\mathscr{E}$. The $N!$ term factors out the identical configurations which differ only by particle permutation. It is worth noting that the density of states is an implicit function of N and V,

which define the dimensionality and boundaries of the hypersurface $\mathscr{E}$, respectively. We have also used here the annotation "con," because this integral depends on the configurational coordinates and the potential energy alone. The complete density of states depends on the total Hamiltonian of the system, and is expressed as follows:

$$\Omega_{\text{tot}}(N, V, \mathscr{E}) = \frac{1}{h^{3N} N!} \int \int_{V^N} \delta[\mathscr{H}(\mathbf{q}, \mathbf{p}_x) - \mathscr{E}] \, \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_x \qquad (1.2)$$

where $\mathbf{p}_x$ are the $3N$ conjugate momenta. Here, we have introduced $1/h^{3N}$ as the factor of proportionality, which is necessary to retrieve the correct correspondence with the high-temperature quantum-mechanical prediction. For a detailed discussion of this proportionality, see for instance [132]. The interpretation of the density of states in this classical, continuum setting is that the quantity $\Omega(\mathscr{E})\mathrm{d}\mathscr{E}$ measures the volume of microstates of energy $\mathscr{E} \pm \mathrm{d}\mathscr{E}/2$. Although this definition may seem vague in physical terms, the important result is that relative values of the density of states have a clear significance. This is to say, if $\Omega(\mathscr{E}_1)$ is twice $\Omega(\mathscr{E}_2)$, then there are twice as many microstates at energy level 1 than 2, even though we may not have a clear way of counting their absolute number at $\mathscr{E}_1$ or $\mathscr{E}_2$. Ultimately, at a classical level, we need only know the density of states to a multiplicative constant, since this will not change the relative measures at different energy levels – or volumes, or even particle numbers.

The connection between the multiplicative insensitivity of Ω and thermodynamics is actually rather intuitive: classically, we are normally only concerned with entropy differences, not absolute entropy values. Along these lines, if we examine Boltzmann's equation, $S = k_{\text{B}} \ln \Omega$, where k_{B} is the Boltzmann constant, we see that a multiplicative uncertainty in the density of states translates to an additive uncertainty in the entropy. From a simulation perspective, this implies that we need not converge to an absolute density of states. Typically, however, one implements a heuristic rule which *defines* the minimum value of the working density of states to be one.

As suggested previously, the density of states has a direct connection to the entropy, and, hence, to thermodynamics, via Boltzmann's equation. Alternately, we can consider the free energy analogue, using the Laplace transform of the density of states – the canonical partition function:

$$Q(N, V, T) = \int \exp(-\beta \mathscr{E}) \, \Omega_{\text{tot}}(N, V, \mathscr{E}) \, \mathrm{d}\mathscr{E} \qquad (1.3)$$

$\beta = (k_{\text{B}} T)^{-1}$. In this expression, the macrostate probabilities at a given temperature are easy to identify – the probability that each energy will be visited is proportional to the integrand.

We now return to the issue of configurational density of states. In the simulation of molecular systems, we are interested only in the calculation of their configurational properties, or more explicitly, the configurational contribution to their partition functions. This is because the kinetic component is analytic, and, hence, there is no need to measure it via simulation. For conventional MC simulations in the

canonical N, V, T ensemble, for example, we readily integrate out these kinetic degrees of freedom, which are simply factored out of the total partition function [135]. The situation in the microcanonical ensemble is somewhat more intricate [139]. Since the kinetic and the potential energies are additive in the Hamiltonian, one can rewrite the single δ-function in (1.2) as a convolution integral involving two δ-functions of each energy term:

$$
\begin{aligned}
\Omega_{\text{tot}}(N, V, \mathscr{E}) &= \frac{1}{h^{3N} N!} \int \int_{V^N} \delta[U(\mathbf{q}) + K(\mathbf{p}_x) - \mathscr{E}] \, \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_x \\
&= \frac{1}{h^{3N} N!} \int \left[\int \delta[K(\mathbf{p}_x) - \mathscr{E}'] \, \mathrm{d}\mathbf{p}_x \right] \\
&\quad \times \left[\int_{V^N} \delta[U(\mathbf{q}) - \mathscr{E} + \mathscr{E}'] \, \mathrm{d}\mathbf{q} \right] \, \mathrm{d}\mathscr{E}' \\
&= \int \Omega_{\text{ig}}(N, V, \mathscr{E}') \Omega_{\text{ex}}(N, V, \mathscr{E} - \mathscr{E}') \, \mathrm{d}\mathscr{E}' \quad (1.4)
\end{aligned}
$$

where the ideal gas and excess density of states in the last lines are defined by

$$
\begin{aligned}
\Omega_{\text{ig}}(N, V, \mathscr{E}) &= \frac{V^N}{h^{3N} N!} \int \delta[K(\mathbf{p}_x) - \mathscr{E}] \, \mathrm{d}\mathbf{p}_x \\
&= \left[\frac{(2\pi m \mathscr{E})^{3/2} V}{h^3} \right]^N \frac{\mathscr{E}^{-1}}{N! \Gamma\left(\frac{3}{2} N\right)} \quad (1.5)
\end{aligned}
$$

and

$$
\Omega_{\text{ex}}(N, V, \mathscr{E}) = \frac{1}{V^N} \int_{V^N} \delta[U(\mathbf{q}) - \mathscr{E}] \, \mathrm{d}\mathbf{q}. \quad (1.6)
$$

Here m is the mass of a particle and Γ is the Γ function. In (1.5), we have determined the explicit ideal gas density of states. This is possible since the kinetic energy is a quadratic function of the momentum, $K = \sum \mathbf{p}^2 / 2m$, which allows us to switch to hyper-spherical coordinates for the treatment of the δ-function. The important fact is that the kinetic contribution to the total, microcanonical partition function is analytical, whereas the excess quantity is the subject of our simulation. This should not cause any confusion, since the excess and the configurational density of states differ only by a simple factor:

$$
\Omega_{\text{ex}}(N, V, \mathscr{E}) = \frac{N!}{V^N} \Omega_{\text{con}}(N, V, \mathscr{E}). \quad (1.7)
$$

The simulation algorithms presented in Chap. 3, for example, may be formulated in such a way that one is calculating either the excess or the configurational density of states, the only distinction being whether the functionality of the multiplicative term on the right-hand side of (1.7) is absorbed into Ω or introduced into the reweighting of results. The use of Ω_{ex} might be mathematically more aesthetic, in that it has natural dimensions. It should, however, be emphasized that it is the configurational

quantity which retains the physical significance of a density of states. In other words, $\Omega_{\mathrm{con}}(N, V, \mathscr{E})$ remains proportional to the number of microstates with given $N, V, \mathscr{E}$.

The excess density of states figures straightforwardly into the canonical partition function. Substituting the convolution in (1.4) into (1.3) and making the substitution $\mathscr{E}'' = \mathscr{E} - \mathscr{E}'$, it follows that

$$Q(N, V, T) = \left[\int e^{-\beta\mathscr{E}'} \Omega_{\mathrm{ig}}(N, V, \mathscr{E}') \, d\mathscr{E}' \right] \left[\int e^{-\beta\mathscr{E}''} \Omega_{\mathrm{ex}}(N, V, \mathscr{E}'') \, d\mathscr{E}'' \right]$$

$$= \frac{V^N}{\Lambda(\beta)^{3N} N!} \int e^{-\beta\mathscr{E}} \Omega_{\mathrm{ex}}(N, V, \mathscr{E}) \, d\mathscr{E}$$

$$= \frac{1}{\Lambda(\beta)^{3N}} \int e^{-\beta\mathscr{E}} \Omega_{\mathrm{con}}(N, V, \mathscr{E}) \, d\mathscr{E} \tag{1.8}$$

In the second line, we have carried the integral over the ideal gas part, which results in the temperature-dependent de Broglie wavelength, Λ. The final expression is similar to the familiar casting of the canonical partition function,

$$Q(N, V, T) = \frac{1}{N! h^{3N}} \int \exp\left[-\beta\mathscr{H}\left(\mathbf{q}, \mathbf{p}_x\right)\right] \, d\mathbf{q} \, d\mathbf{p}_x$$

$$= \frac{1}{\Lambda^{3N} N!} Z(N, V, T) \tag{1.9}$$

except that the multidimensional integral over coordinates is now replaced by a one-dimensional integral over energy. In (1.9), $Z(N, V, T)$ is the configurational integral defined by:

$$Z(N, V, T) = \int \exp(-\beta U(\mathbf{q})) \, d\mathbf{q} \tag{1.10}$$

where $U(\mathbf{q})$ is the potential energy of the system.

In this chapter and in others of the present book, we will often drop the subscript "con" from the configurational density of states, which will simply be denoted by Ω. Any other quantity, such as the total and excess density of states, will retain its subscript.

1.2.2 Application: MC Simulation in the Microcanonical Ensemble

A working example will help illustrate some of the mathematical properties of the density of states and its connection to the microcanonical ensemble. It is possible to perform a MC simulation in a microcanonical setting (constant total energy, kinetic plus potential) using the previous arguments. This method was developed by John Ray [140] and later by Rolf Lustig [141], and though it is not frequently used, its derivation is instructive. As with any MC simulation, the first concern is the ensemble of interest, which specifies the relevant underlying partition function and, importantly, the probability with which configurations should be visited or sampled. In this case, we extract these probabilities with a simple manipulation of the density of states. Starting with the analytically evaluated ideal gas density of states in

(1.5), we substitute this contribution back into the convolution integral determining the total microcanonical partition function in (1.4):

$$
\begin{aligned}
\Omega_{\text{tot}}(N, V, \mathscr{E}) &= \frac{1}{N!\,\Gamma\left(\frac{3}{2}N\right)} \left[\frac{(2\pi m)^{3/2}}{h^3}\right]^N \\
&\quad \times \int \int_{V^N} \delta\left[U(\mathbf{q}) - \mathscr{E} + \mathscr{E}'\right] (\mathscr{E}')^{3N/2-1}\, \mathrm{d}\mathbf{q}\, \mathrm{d}\mathscr{E}' \\
&= \frac{1}{N!\,\Gamma\left(\frac{3}{2}N\right)} \left[\frac{(2\pi m)^{3/2}}{h^3}\right]^N \\
&\quad \times \int_{V^N} \left[\mathscr{E} - U(\mathbf{q})\right]^{3N/2-1} \theta\left[\mathscr{E} - U(\mathbf{q})\right]\, \mathrm{d}\mathbf{q}
\end{aligned}
\tag{1.11}
$$

where θ is the Heaviside step function. In going to the last line in this expression, we have switched the order of integration and performed the integral over $\mathscr{E}'$ to remove the delta function. The final expression gives a clear significance of the microstate probabilities in the ensemble and has a nice correspondence with the canonical configurational partition function. Compare this result to that of a constant-temperature simulation, in the NVT ensemble. There we must specify the temperature, the partition function is Q, and the state probabilities follow the Boltzmann factor. Similarly, in the microcanonical simulation we must specify a total energy, the partition function is Ω, and the weight each configuration should carry is:

$$
\boxed{\mathscr{P}(\mathbf{q}) \propto \left[\mathscr{E} - U(\mathbf{q})\right]^{3N/2-1} \theta\left[\mathscr{E} - U(\mathbf{q})\right].}
\tag{1.12}
$$

Based on (1.12), we can implement any complement of MC moves and formulate appropriate acceptance criteria such that the progression of configurations satisfies this distribution. For simple moves in which the proposal probability equals that of its inverse – symmetric moves, such as single-particle displacements – the Metropolis acceptance criterion then reads [141]:

$$
\mathscr{P}_{\text{acc}}(U_0 \to U_n) = \min\left[1, \left(\frac{\mathscr{E} - U_n}{\mathscr{E} - U_0}\right)^{3N/2-1} \theta\left(\mathscr{E} - U_n\right)\right]
\tag{1.13}
$$

where it is assumed that the initial energy, U_0, is less than $\mathscr{E}$. Similar arguments can be used to adapt (1.13) in the presence of additional constraints, such as nonspherical rigid molecules or fixed total momentum [141].

1.3 Free Energy

1.3.1 Basic Approaches to Free Energy Calculations

The Helmholtz free energy, A, which is the thermodynamic potential, the natural independent variables of which are those of the canonical ensemble, can be expressed in terms of the partition function:

$$\boxed{A = -\beta^{-1} \ln Q(N, V, T).}$$

(1.14)

This equation forms the fundamental connection between thermodynamics and statistical mechanics in the canonical ensemble, from which it follows that calculating A is equivalent to estimating the value of Q. In general, evaluating Q is a very difficult undertaking. In both experiments and calculations, however, we are interested in free energy *differences*, ΔA, between two systems or states of a system, say 0 and 1, described by the partition functions Q_0 and Q_1, respectively – the arguments N, V, T have been dropped to simplify the notation:

$$\Delta A = -\beta^{-1} \ln Q_1 / Q_0$$

(1.15)

If the quantity of interest is the excess Helmholtz free energy, as is almost always the case, or if the masses of particles in systems 0 and 1 are the same, (1.15) can be rewritten in terms of the configurational integrals Z_0 and Z_1

$$\Delta A = -\beta^{-1} \ln Z_1 / Z_0.$$

(1.16)

Almost all problems that require knowledge of free energies are naturally formulated or can be framed in terms of (1.15) or (1.16). Systems 0 and 1 may differ in several ways. For example, they may be characterized by different values of a macroscopic parameter, such as the temperature. Alternatively, they may be defined by two different Hamiltonians, $\mathscr{H}_0$ and $\mathscr{H}_1$, as is the case in studies of free energy changes upon point mutation of one or several amino acids in a protein. Finally, the definitions of 0 and 1 can be naturally extended to describe two different, well-defined macroscopic states of the same system. Then, Q_0 is defined as:

$$Q_0 = \frac{1}{N! h^{3N}} \int_{\Gamma_0} \exp\left[-\beta \mathscr{H}\left(\mathbf{x}, \mathbf{p}_x\right)\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x$$

(1.17)

where Γ_0 is the volume in the phase space accessible to the system in state 0. Q_1 can be defined in a similar manner. The macroscopic states defined by Γ_0 and Γ_1 may correspond to different conformations of a flexible molecule, or the bound and unbound structures of a protein–ligand complex.

Calculating free energies in these three types of systems requires slightly different theoretical treatments, but the underlying ideas remain the same. For this reason, we will draw a distinction between these systems only when it is necessary for theoretical developments. If treatments of different types of systems are essentially identical, yet require somewhat different notations, we will often limit our discussion to only one case, leaving the exercise of changing the notation to the reader.

Equation (1.15) indicates that our ultimate focus in calculating ΔA is on determining the ratio Q_1/Q_0 – or equivalently Z_1/Z_0 – rather than on individual partition functions. On the basis of computer simulations, this can be done in several ways. One approach consists in transforming (1.16) as follows:

$$\Delta A = -\beta^{-1} \ln \frac{\int \exp\left[-\beta U_0\left(\mathbf{x}\right)\right] d\mathbf{x}}{\int \exp\left[-\beta U_0\left(\mathbf{x}\right)\right] d\mathbf{x}} \tag{1.18}$$

$$= -\beta^{-1} \ln \exp\left\{-\beta\left[U_1(\mathbf{x}) - U_0(\mathbf{x})\right]\right\} P_0\left(\mathbf{x}\right)$$

$$= -\beta^{-1} \ln\langle\exp\left\{-\beta\left[U_1(\mathbf{x}) - U_0(\mathbf{x})\right]\right\}\rangle_0$$

Here, the systems 0 and 1 are described by the potential energy functions, $U_0(\mathbf{x})$, and $U_1(\mathbf{x})$, respectively. Generalization to conditions in which systems 0 and 1 are at two different temperatures is straightforward. β_0 and β_1 are equal to $(k_B T_0)^{-1}$ and $(k_B T_1)^{-1}$, respectively. $P_0(\mathbf{x})$ is the probability density function of finding system 0 in the microstate defined by positions $\mathbf{x}$ of the particles:

$$P_0\left(\mathbf{x}\right) = \frac{\exp\left[-\beta_0 U_0\left(\mathbf{x}\right)\right]}{Z_0} \tag{1.19}$$

An interesting feature of (1.18) is that ΔA is estimated from a simulation of system 0 only. During such a simulation, a sample of the value of $\beta_1 U_1 - \beta_0 U_0$ needs to be collected which is sufficient to estimate with the desired accuracy the average exponential in (1.18). Using one system as the reference and focusing on energy differences is reminiscent of perturbation methods. Not surprisingly, this general approach is called the FEP method. This method will be discussed in detail in Chaps. 2 and 6.

Another approach to calculating ΔA relies on estimating the appropriate probability density functions. The connection between the probabilities of different states and the partition function is natural in statistical mechanics. Equation (1.19) is a reflection of this connection. Similarly, the probability of observing the potential energy of the system being equal to U is:

$$P(U) = \frac{\exp\left(-\beta U\right)\Omega\left(U\right)}{Z} \tag{1.20}$$

where, again, the arguments N, V, T have been omitted for simplicity.

Let us assume that system 0 can be transformed to system 1 through the continuous change of some parameter λ defined such that λ_0 and λ_1 correspond to systems 0 and 1, respectively. This parameter could be a macroscopic variable – viz. the temperature, a parameter that transforms $\mathscr{H}_0$ to $\mathscr{H}_1$, or a generalized coordinate (e.g., a torsional angle or an intermolecular distance) that allows the different structural states of the system to be distinguished. It follows that:

$$P_0 = P\left(\lambda_0\right) = \frac{\int \exp\left(-\beta\mathscr{H}\right)\delta\left(\lambda - \lambda_0\right)\,d\mathbf{x}\,d\mathbf{p}_x}{\mathscr{N}} = \frac{Q_0}{\mathscr{N}} \tag{1.21}$$

where $\mathscr{N}$ is a normalization constant. Here, β, $\mathscr{H}$ or $\mathbf{x}, \mathbf{p}_x$ could be functions of λ. P_1 can be obtained in the same way, by substituting subscript 1 for 0. Combining (1.15) and (1.21) leads to:

$$\Delta A = -\beta^{-1} \ln \frac{P_1}{P_0} \tag{1.22}$$

This equation provides a prescription for calculating ΔA. The probability distribution function, $P(\lambda)$, for the range of λ comprised between λ_0 and λ_1 is obtained from computer simulations, usually as a histogram. The ratio P_1/P_0 is then estimated. This generic idea has been implemented in various, creative ways, yielding a class of techniques called probability distribution or histogram methods. These methods are discussed in Chap. 3.

In the third approach, one calculates $\mathrm{d}\Delta A/\mathrm{d}\lambda$ rather than ΔA directly. Differentiating (1.14) yields:

$$\frac{\mathrm{d}\beta A}{\mathrm{d}\lambda} = -\frac{1}{Q(\lambda)} \frac{\partial Q}{\partial \lambda} \tag{1.23}$$

If λ is a parameter in the Hamiltonian, we obtain:

$$\boxed{\frac{\mathrm{d}A}{\mathrm{d}\lambda} = \frac{\displaystyle\int \frac{\partial \mathscr{H}}{\partial \lambda} \exp\left(-\beta\mathscr{H}\right) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x}{\displaystyle\int \exp\left(-\beta\mathscr{H}\right) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x} = \left\langle \frac{\partial \mathscr{H}}{\partial \lambda} \right\rangle_\lambda} \tag{1.24}$$

and the free energy difference between system 0 and system 1 is evaluated by integrating the average derivative of the Hamiltonian with respect to λ, which is in units of the force, in the range extending from λ_0 to λ_1. For this reason, the method is called thermodynamic integration. If λ is a function of the positions of the particles, derivation of the formula for $\mathrm{d}A/\mathrm{d}\lambda$ is more intricate, but the quantity that needs to be averaged remains the same. Details are given in Chap. 4.

Conceptually, the three methods outlined above are closely connected. For example, one can derive the TI formula from (1.18) by assuming that the transformation from system 0 to system 1 proceeds through a sequential series of small perturbations, in which λ changes by an increment $\Delta\lambda$, and then taking the limit of $\Delta\lambda \rightarrow 0$. Even though the methods are related, the distinction between them is useful, because the developments of advanced techniques for each of them is often markedly different.

As we will see further in the book, almost all methods for calculating free energies in chemical and biological problems by means of computer simulations of equilibrium systems rely on one of the three approaches that we have just outlined, or on their possible combination. These methods can be applied not only in the context of the canonical ensemble, but also in other ensembles. As will be discussed in Chap. 5, ΔA can be also estimated from nonequilibrium simulations, to such extent that FEP and TI methods can be considered as limiting cases of this approach.

1.4 Ergodicity, Quasi-nonergodicity and Enhanced Sampling

Central to many developments in this book is the concept of *ergodicity*. Let us consider a physical system consisting of N particles. Its time evolution can be described as a path, or trajectory, in phase space. If the system was initially in the state

$\{\mathbf{p_0}, \mathbf{q_0}\}$, the time average – if it exists – of any property, f, observed over time T would be equal to

$$\overline{f(\mathbf{q_0}, \mathbf{p_o})} = \frac{1}{T} \int_0^T f\left[\mathbf{q(t)}, \mathbf{p(t)}\right] \, \mathrm{d}t \qquad (1.25)$$

Similarly, we can define the ensemble average:

$$\boxed{\langle f \rangle = \int f P\left(\mathbf{x}, \mathbf{p}_x\right) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x,} \qquad (1.26)$$

where $P(\mathbf{x}, \mathbf{p}_x)$ is the time-independent probability that measures the fraction of systems that are in the state $\{\mathbf{x}, \mathbf{p}_x\}$.

For ergodic systems, the probability of visiting the neighborhood of each point in phase space converges to a unique limiting value as $T \to \infty$, such that the time average of f is equal to its ensemble average

$$\lim_{T \to \infty} \overline{f(\mathbf{q_0}, \mathbf{p_0})} = \langle f \rangle. \qquad (1.27)$$

There are two important consequences of this equality for computer simulations of many-body systems. First, it means that statistically averaged properties of these systems are accessible from simulations that are aimed at generating trajectories – e.g., molecular dynamics, or ensemble averages such as Monte Carlo. Furthermore, for sufficiently long trajectories, the time-averaged properties become independent of the initial conditions. Stated differently, it means that for almost all values of $\{\mathbf{q_0}, \mathbf{p_0}\}$, the system will pass arbitrarily close to any point $\{\mathbf{x}, \mathbf{p}_x\}$ in phase space at some later time.

The assumption that (1.27) holds, i.e., that time averages of macroscopic variables can be replaced by their ensemble averages is called the *ergodic hypothesis*. It is equivalent to the statement that a system assumes, in the long run, all conceivable microstates that are compatible with the conservation laws, and, therefore, lies at the foundation of statistical mechanics developed by Boltzmann and Maxwell. From our perspective, it is clear that the theoretical outline given in the previous two sections would not be appropriate for nonergodic systems. Moreover, for such systems, it is not expected that different computer simulations of the same system, no matter how long, would yield the same estimates of the free energy.

Although it is usually very difficult to prove ergodicity, it is strongly believed that almost all many-body systems are ergodic. There are, however, a few known examples of nonergodic systems. Perhaps the best known are completely integrable systems – i.e., systems for which the number of degrees of freedom is equal to the number of constants of motion. This was proven in the famous Kolmogorov–Arnold–Moser (KAM) theorem [142]. Fortunately, systems known to be nonergodic are usually not of interest in chemistry and biology.

Even if a system is formally ergodic, its behavior during computer simulations may resemble those of nonergodic systems. This means that the system does not properly explore phase space, and, therefore, the calculated statistical averages might

exhibit strong dependence on the initial conditions. This phenomenon is called quasi-nonergodicity. It may occur because the system diffuses very slowly, to the extent that the volume in phase space covered during the simulation is insufficient to estimate reliably statistical averages of properties of interest. More often, the appearance of nonergodicity is caused by high energy barriers separating different volumes of phase space. It follows that transitions between these volumes constitute rare events that might never happen during a computer simulation, or that occur so infrequently that accurate estimates of statistical averages cannot be achieved in practice. Even if the volumes are connected by low-energy regions, but these regions are very narrow – viz. so-called 'entropy bottlenecks,' and hence rarely sampled, the appearance of nonergodicity persists.

Quasi-nonergodicity is a common phenomenon in complex chemical and biological systems. If this is the case, direct application of the methods outlined in the previous section might not yield correct estimates of free energies. To improve these estimates, more-advanced strategies that allow relevant rare events to be sampled are needed. These strategies are called enhanced sampling methods. Most of them are also used in other fields of science, but under a different name – viz. variance-reduction methods. The connection between these two names is fairly obvious. The primary goal in applying enhanced sampling methods is to explore efficiently the regions in phase space that are important for calculating free energy, and, by doing so, reduce the variance of the estimates of this quantity.

Two enhanced sampling strategies have proved to be particularly effective in dealing with quasi-nonergodicity, namely stratification and importance sampling. In fact, almost all techniques used to improve the efficiency of free energy calculations rely on one of these strategies, or their combination. Their thoughtful and creative implementation often makes the difference between successful and unreliable simulations.

Stratification, sometimes also called multistage sampling [10], is a strategy for distributing samples so that all parts of the function are adequately sampled. In an unstratified process, all the samples are generated from the same probability distribution function, $P(x)$, which might vary greatly in the domain Ω. In a stratified method, this domain is first partitioned into a number of disjoint regions Ω_i, called strata, such that their union covers the whole domain. In the region Ω_i, x_i is sampled according to $P_i(x_i)$, equal to $P(x)$ in this region. In the process, every strata is sampled, even if it is associated with a very low $P(x)$, and, as a consequence, is unlikely to be visited in an unstratified sampling. The end result of stratification is a lowered variance on the estimate of any function $f(x)$ averaged over Ω with the probability measure $P(x)$.

To illustrate how stratification works in the context of free energy calculations, let us consider the transformation of state 0 into state 1 described by the parameter λ. We further assume that these two states are separated by a high-energy barrier that corresponds to a value of λ between λ_0 and λ_1. Transitions between 0 and 1 are then rare and the free energy estimated from unstratified computer simulations would converge very slowly to its limiting value, irrespective of the initial conditions. If, however, the full range of λ is partitioned into a number of smaller intervals, and

each of these intervals is sampled independently, it is possible to recover the complete $P(\lambda)$ and estimate ΔA from (1.22), with great savings of computer time.

Importance sampling is another, highly successful variance-reduction technique [143]. The idea behind it is that certain regions in phase space are important for estimating the quantity of interest, even though these regions might have low probability of being visited. It is thus advantageous to choose a sampling distribution from which these 'important' regions are sampled more frequently than they would be from the true distribution. If this approach were applied directly in a simulation, it would yield a biased estimator. The results of the simulation obtained using the modified distribution can, however, be properly weighted to ensure that the estimator is unbiased. The weight is given by the likelihood ratio of the true distribution to the biased simulation distribution.

The basic idea of importance sampling can be illustrated simply in the example of the transformation from 0 to 1 along λ, as described above. In lieu of sampling from the true probability distribution, $P(\lambda)$, we design simulations in which λ is sampled according to $P'(\lambda)$. The latter probability should be chosen so that it is more uniform than $P(\lambda)$. The relation between the two probabilities may then be expressed as follows:

$$P'(\lambda) = P(\lambda) \exp\left[\beta\eta(\lambda)\right] \tag{1.28}$$

where $\eta(\lambda)$ is the weight factor that depends on the value of λ. Next, ΔA in (1.22) can be expressed in terms of $P'(\lambda_0)$ and $P'(\lambda_1)$ derived from the biased simulation:

$$\Delta A = -\beta^{-1} \ln \frac{P(\lambda_1)}{P(\lambda_0)} = -\beta^{-1} \ln \frac{P'(\lambda_1)}{P'(\lambda_0)} + \eta(\lambda_1) - \eta(\lambda_0) \tag{1.29}$$

The fundamental issue in implementing importance sampling in simulations is the proper choice of the biased distribution, or, equivalently, the weighting factor, η. A variety of ingenious techniques that lead to great improvement in the efficiency and accuracy of free energy calculations have been developed for this purpose. They will be mentioned frequently throughout this book.

References

1. Kirkwood, J. G., Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **1935**, *3*, 300–313
2. Kirkwood, J. G., in *Theory of Liquids*, Alder, B. J., Ed., Gordon and Breach: New York, 1968
3. De Donder, T., *L'affinité*, Gauthier-Villars: Paris, 1927
4. Zwanzig, R. W., High-temperature equation of state by a perturbation method. I. Nonpolar gases, *J. Chem. Phys.* **1954**, *22*, 1420–1426
5. Landau, L. D., *Statistical Physics*, Clarendon: Oxford, 1938
6. Widom, B., Some topics in the theory of fluids, *J. Chem. Phys.* **1963**, *39*, 2808–2812
7. Owicki, J. C.; Scheraga, H. A., Monte Carlo calculations in the isothermal–isobaric ensemble. 1. Liquid water, *J. Am. Chem. Soc.* **1977**, *99*, 7403–7412

8. McDonald, I. R.; Singer, K., Machine calculation of thermodynamic properties of a simple fluid at supercritical temperatures, *J. Chem. Phys.* **1967**, *47*, 4766–4772

9. McDonald, I. R.; Singer, K., Calculation of thermodynamic properties of liquid argon from Lennard-Jones parameters by a Monte Carlo method, *Discuss. Faraday Soc.* **1967**, *43*, 40–49

10. Valleau, J. P.; Card, D. N., Monte Carlo estimation of the free energy by multistage sampling, *J. Chem. Phys.* **1972**, *57*, 5457–5462

11. Bennett, C. H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **1976**, *22*, 245–268

12. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free energy estimation: Umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199

13. Barker, J. A.; Henderson, D., Perturbation theory and equation of state for fluids: the square-well potential, *J. Chem. Phys.* **1967**, *47*, 2856–2861

14. Weeks, J. D.; Chandler, D.; Andersen, H. C., Role of repulsive forces in determining the equilibrium structure of simple liquids, *J. Phys. Chem.* **1971**, *54*, 5237–5247

15. Pratt, L. R.; Chandler, D., Theory of hydrophobic effect, *J. Chem. Phys.* **1977**, *67*, 3683–3704

16. Lee, J. K.; Barker, J. A.; Abraham, F. F., Theory and Monte Carlo simulation of physical clusters in the imperfect vapor, *J. Chem. Phys.* **1973**, *58*, 3166–3180

17. Mruzik, M. R.; Abraham, F. F.; Schreiber, D. E.; Pound, G. M., A Monte Carlo study of ion–water clusters, *J. Chem. Phys.* **1975**, *64*, 481–491

18. McGinty, D. J., Molecular dynamics studies of the properties of small clusters of argon atoms, *J. Chem. Phys.* **1973**, *58*, 4733–4742

19. Mezei, M.; Swaminathan, S.; Beveridge, D. L., Ab initio calculation of the free energy of liquid water, *J. Am. Chem. Soc.* **1978**, *100*, 3255–3256

20. Mezei, M., Excess free energy of different water models computed by Monte Carlo methods, *Mol. Phys.* **1982**, *47*, 1307–1315

21. Patey, G. N.; Valleau, J. P., A Monte Carlo method for obtaining the interionic potential of mean force in ionic solution, *J. Chem. Phys.* **1975**, *63*, 2334–2339

22. Okazaki, S.; Nakanishi, K.; Touhara, H., Monte Carlo studies on the hydrophobic hydration in dilute aqueous solutions on nonpolar molecules, *J. Theor. Biol.* **1979**, *71*, 2421–2429

23. Pangali, C. S.; Rao, M.; Berne, B. J., A Monte Carlo simulation of the hydrophobic effect, *J. Chem. Phys.* **1979**, *71*, 2975–2981

24. Chipot, C.; Kollman, P. A.; Pearlman, D. A., Alternative approaches to potential of mean force calculations: free energy perturbation versus thermodynamic integration. Case study of some representative nonpolar interactions, *J. Comput. Chem.* **1996**, *17*, 1112–1131

25. Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R., Thermodynamics of cavity formation in water: a molecular dynamics study, *Faraday Symp. Chem. Soc.* **1982**, *17*, 55–67

26. Lee, C. Y.; Scott, H. L., The surface tension of water: a Monte Carlo calculation using an umbrella sampling algorithm, *J. Chem. Phys.* **1980**, *73*, 4591–4596

27. Quirke, N.; Jacucci, G., Energy difference functions in Monte Carlo simulations: application to the calculation of free energy of liquid nitrogen. II. The calculation of fluctuation in Monte Carlo averages, *Mol. Phys.* **1982**, *45*, 823–838

28. Shing, K. S.; Gubbins, K. E., The chemical potential in dense fluids and fluid mixtures via computer simulation, *Mol. Phys.* **1982**, *46*, 1109–1128

29. Warshel, A., Dynamics of reactions in polar solvents. Semiclassical trajectory studies of electron transfer and proton transfer reactions, *J. Phys. Chem.* **1982**, *86*, 2218–2224

30. Northrup, S. H.; Pear, M. R.; Lee, C. Y.; McCammon, J. A.; Karplus, M., Dynamical theory of activated processes in globular proteins, *Proc. Natl Acad. Sci. USA* **1982**, *79*, 4035–4039

31. Tembe, B. L.; McCammon, J. A., Ligand–receptor interactions, *Comput. Chem.* **1984**, *8*, 281–283

32. Jorgensen, W. L.; Ravimohan, C., Monte Carlo simulation of differences in free energies of hydration, *J. Chem. Phys.* **1985**, *83*, 3050–3054

33. Jorgensen, W. L.; Briggs, J. M.; Gao, J., A priori calculations of pK_a's for organic compounds in water. The pK_a of ethane, *J. Am. Chem. Soc.* **1987**, *109*, 6857–6858

34. Jorgensen, W. L.; Briggs, J. M., A priori pK_a calculations and the hydration of organic anions, *J. Am. Chem. Soc.* **1989**, *111*, 4190–4197

35. Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. S_N2 reaction profiles in the gas phase and aqueous solution. *J. Am. Chem. Soc.* **1984**, *106*, 3049–3050

36. Tobias, D. J.; Brooks III, C. L., Calculation of free energy surfaces using the methods of thermodynamic perturbation theory, *Chem. Phys. Lett.* **1987**, *142*, 472–476

37. Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A., Free energy calculations by computer simulation, *Science* **1987**, *236*, 564–568

38. Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A., Calculation of the relative change in binding free energy of a protein–inhibitor complex, *Science* **1987**, *235*, 574–576

39. Rao, B. G.; Singh, U. C.; Bash, P. A.; Kollman, P. A., Free energy perturbation calculations on binding and catalysis after mutating Asn 155 in subtilisin, *Nature* **1987**, *328*, 551–554

40. Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A., An approach to the application of free energy perturbation methods using molecular dynamics: applications to the transformations of methanol → ethane, oxonium → ammonium, glycine → alanine, and alanine → phenylalanine in aqueous solution and to $H_3O^+(H_2O)_3$ → $NH_4^+(H_2O)_3$ in the gas phase, *J. Am. Chem. Soc.* **1987**, *109*, 1607–1611

41. Jarzynski, C., Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **1997**, *78*, 2690–2693

42. Dang, L. X.; Pearlman, D. A.; Kollman, P. A., Why do A·T base pairs inhibit Z-DNA formation?, *Proc. Natl Acad. Sci. USA* **1990**, *87*, 4630–4634

43. Fleischman, S. H.; Brooks III, C. L., Thermodynamics of aqueous solvation: Solution properties of alchohols and alkanes, *J. Chem. Phys.* **1987**, *87*, 3029–3037

44. Lu, N.; Kofke, D. A.; Woolf, T. B., Staging is more important than perturbation method for computation of enthalpy and entropy changes in complex systems, *J. Phys. Chem. B* **2003**, *107*, 5598–5611

45. Peter, C.; Oostenbrink, C.; van Dorp, A.; van Gunsteren, W. F., Estimating entropies from molecular dynamics simulations, *J. Chem. Phys.* **2004**, *120*, 2652–2661

46. Straatsma, T. P.; Berendsen, H. J. C., Free energy of ionic hydration: analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulations, *J. Chem. Phys.* **1988**, *89*, 5876–5886

47. Wang, C. X.; Liu, H. Y.; Shi, Y. Y.; Huang, F. H., Calculations of relative free energy surfaces in configuration space using an integration method, *Chem. Phys. Lett.* **1991**, *179*, 475–478

48. Jorgensen, W. L., in *Computer simulation of biomolecular systems: Theoretical and experimental applications*, Van Gunsteren, W. F.; Weiner, P. K., Eds. Escom: The Netherlands, 1989, p. 60

49. Jorgensen, W. L. Free energy calculations, a breakthrough for modeling organic chemistry in solution. *Acc. Chem. Res*. **1989**, *22*, 184–189

50. Beveridge, D. L.; DiCapua, F. M., Free energy via molecular simulation: applications to chemical and biomolecular systems, *Annu. Rev. Biophys. Biophys.* **1989**, *18*, 431–492

51. Beveridge, D. L.; DiCapua, F. M., Free energy via molecular simulation: a primer, in *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Van Gunsteren, W. F.; Weiner, P. K., Eds. Escom: The Netherlands, 1989, pp. 1–26

52. Kollman, P. A., Free energy calculations: applications to chemical and biochemical phenomena, *Chem. Rev.* **1993**, *93*, 2395–2417

53. Ferrenberg, A. M.; Swendsen, R. H., Optimized Monte Carlo data analysis, *Phys. Rev. Lett.* **1989**, *63*, 1195–1198

54. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted histogram analysis method for free energy calculations on biomolecules. I. The method, *J. Comput. Chem.* **1992**, *13*, 1011–1021

55. Bartels, C.; Karplus, M., Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations, *J. Comput. Chem.* **1997**, *18*, 1450–1462

56. Mezei, M., Adaptive umbrella sampling: self-consistent determination of the non-Boltzmann bias, *J. Comput. Phys.* **1987**, *68*, 237–248

57. Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F., Avoiding singularities and neumerical instabilities in free energy calculations based on molecular simulations, *Chem. Phys. Lett.* **1994**, *222*, 529–539

58. Hummer, G.; Pratt, L.; Garcia, A. E., Free energy of ionic hydration, *J. Phys. Chem.* **1996**, *100*, 1206–1215

59. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., Equation of state calculations by fast computing machines, *J. Chem. Phys.* **1953**, *21*, 1087–1092

60. Vorontsov-Velyaminov, P. N.; Elyashevich, A. M.; Morgenshtern, L. A.; Chasovskikh, V. P., Investigation of phase transitions in argon and coulomb gas by the Monte Carlo method using an isothermally isobaric ensemble, *High Temp. USSR* **1970**, *8*, 261–268

61. Adams, D.J., Grand canonical ensemble Monte Carlo for a Lennard-Jones fluid, *Mol. Phys.* **1975**, *29*, 307–311

62. Frenkel, D.; Ladd, A. J. C., New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres, *J. Chem. Phys.* **1984**, *81*, 3188–3193

63. Panagiotopoulos, A. Z., Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.* **1987**, *61*, 813–826

64. Wilding, N. B., Critical-point and coexistence-curve properties of the Lennard-Jones fluid: a finite-size scaling study, *Phys. Rev. E* **1995**, *52*, 602–611

65. Kofke, D. A., Gibbs–Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation, *Mol. Phys.* **1993**, *78*, 1331–1336

66. Pearlman, D. A.; Kollman, P. A., The overlooked bond-stretching contribution in free energy perturbation calculations, *J. Chem. Phys.* **1991**, *94*, 4532–4545

67. Boresch, S.; Karplus, M., The Jacobian factor in free energy simulations, *J. Comp. Chem.* **1996**, *105*, 5145–5154

68. Fixman, M., Classical statistical mechanics of constraints: A theorem and application to polymers, *Proc. Natl Acad. Sci. USA* **1974**, *71*, 3050–3053

69. Gō, N.; Scheraga, H. A. S., On the use of classical statistical mechanics in the treatment of polymer chain conformation, *Macromolecules* **1976**, *9*, 535–542

70. van Gunsteren, W. F. Methods for calculation of free energies and binding constants: successes and problems, in *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Van Gunsteren, W. F.; Weiner, P. K., Eds. Escom: The Netherlands, 1989, pp. 27–59

71. Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R., Constrained reaction coordinate dynamics for the simulation of rare events, *Chem. Phys. Lett.* **1989**, *156*, 472–477

72. den Otter, W. K.; Briels, W. J., The calculation of free-energy differences by constrained molecular dynamics simulations, *J. Chem. Phys.* **1998**, *109*, 4139–4146

73. den Otter, W. K.; Briels, W. J., Free energy from molecular dynamics with multiple constraints, *Mol. Phys.* **2000**, *98*, 773–781

74. Darve, E.; Pohorille, A., Calculating free energies using average force, *J. Chem. Phys.* **2001**, *115*, 9169–9183

75. Hénin, J.; Chipot, C., Overcoming free energy barriers using unconstrained molecular dynamics simulations, *J. Chem. Phys.* **2004**, *121*, 2904–2914

76. Chipot, C.; Hénin, J., Exploring the free energy landscape of a short peptide using an average force, *J. Chem. Phys.* **2005**, *123*, 244906

77. Laio, A.; Parrinello, M., Escaping free energy minima, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 12562–12565

78. Berg, B. A.; Neuhaus, T., Multicanonical ensemble: a new approach to simulate first-order phase transitions, *Phys. Rev. Lett.* **1992**, *68*, 9–12

79. Wang, F.; Landau, D. P., An efficient, multiple range random walk algorithm to calculate the density of states, *Phys. Rev. Lett.* **2001**, *86*, 2050–2053

80. Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z., Generalization of the Wang–Landau method for off-lattice simulations, *Phys. Rev. E* **2002**, *90*, 056703

81. Yan, Q.; de Pablo, J. J., Fast calculation of the density of states of a fluid by Monte Carlo simulations, *Phys. Rev. Lett.* **2003**, *90*, 035701

82. Smith, G. R.; Bruce, A. D., A study of the multi-canonical Monte Carlo method, *J. Phys. A* **1995**, *28*, 6623–6643

83. Smith, G. R.; Bruce, A. D., Multicanonical Monte Carlo study of solid–solid phase coexistence in a model colloid, *Phys. Rev. E* **1996**, *53*, 6530

84. Valleau, J. P. The Coulombic phase transition: density-scaling Monte Carlo. *J. Chem. Phys.* **1991**, *95*, 584–589

85. Valleau, J. P. Temperature-and-density-scaling Monte-Carlo: methodology and the canonical thermodynamics of Lennard-Jonesium. *Mol. Sim.* **2005**, *31*, 223–253

86. Kong, X.; Brooks III, C. L., λ-dynamics: a new approach to free energy calculations, *J. Chem. Phys.* **1996**, *105*, 2414–2423

87. Bitetti-Putzer, R.; Yang, W.; Karplus, M., Generalized ensembles serve to improve the convergence of free energy simulations, *Chem. Phys. Lett.* **2003**, *377*, 633–641

88. Frantz, D.D.; Freeman, D.L.; Doll, J.D., Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking: applications to atomic clusters, *J. Chem. Phys.* **1990**, *93*, 2769–2784

89. Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N., New approach to Monte Carlo calculation of the free energy: method of expanded ensembles, *J. Chem. Phys.* **1992**, *96*, 1776–1783

90. Marinari, E.; Parisi, G., Simulated tempering: a new Monte Carlo scheme, *Europhys. Lett.* **1992**, *19*, 451–458

91. Hansmann, U. H. E., Parallel tempering algorithm for conformational studies of biological molecules, *Chem. Phys. Lett.* **1997**, *281*, 140–150

92. Roitberg, A.; Elber, R., Modeling side chains in peptides and proteins: application of the locally enhanced sampling technique and the simulated annealing methods to find minimum energy conformations, *J. Chem. Phys.* **1991**, *95*, 9277–9287

93. Verkhivker, G.; Elber, R.; Nowak, W., Locally enhanced sampling in free energy calculations: application of mean field approximation to accurate calculation of free energy differences, *J. Chem. Phys.* **1992**, *97*, 7838–7841

94. Hummer, G.; Garde, S.; García, A.; Pohorille, A.; Pratt, L., An information theory model of hydrophobic interactions, *Proc. Natl Acad. Sci. USA* **1996**, *93*, 8951–8955

95. Pohorille, A.; Pratt, L. R., Cavities in molecular liquids and the theory of hydrophobic solubilities, *J. Am. Chem. Soc.* **1990**, *112*, 5066–5074

96. Pratt, L. R.; Pohorille, A., Theory of hydrophobicity: Transient cavities in molecular liquids, *Proc. Natl Acad. Sci. USA* **1992**, *89*, 2995–2999

97. Pratt, L. R.; LaViolette, R. A.; Gomez, M. A.; Gentile, M. E., Quasi-chemical theory for the statistical thermodynamics of the hard-sphere fluid, *J. Phys. Chem. B* **2001**, *105*, 11662–11668

98. Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S., Absolute hydration free energies ofions, ion–water clusters and quasichemical theory, *J. Chem. Phys.* **2003**, *119*, 2702–2708

99. Jarzynski, C., Equilibrium free-energy differences from nonequilibrium measurements: a master-equation approach, *Phys. Rev. E* **1997**, *56*, 5018–5035

100. Crooks, G. E., Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, *Phys. Rev. E* **1999**, *60*, 2721–2726

101. Ritort, F.; Bustamante, C.; Tinoco Jr., I., A two-state kinetic model for the unfolding of single molecules by mechanical force, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 13544–13548

102. Hummer, G.; Szabo, A., Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3658–3661

103. Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K., Steered molecular dynamics, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Deuflhard, P.; Hermans, J.; Leimkuhler, B.; Mark, A. E.; Skeel, R.; Reich, S., Eds., vol. 4, *Lecture Notes in Computational Science and Engineering*. Springer: Berlin, Heidelberg, New York, 1998, pp. 39–65

104. Jensen, M. Ø.; Park, S.; Tajkhorshid, E.; Schulten, K., Energetics of glycerol conduction through aquaglyceroporin GlpF, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 6731–6736

105. Ytreberg, F. M.; Zuckerman, D. M., Single-ensemble nonequilibrium path-sampling estimates of free energy differences, *J. Chem. Phys.* **2004**, *120*, 10876–10879

106. Chipot, C.; Millot, C.; Maigret, B.; Kollman, P. A., Molecular dynamics free energy perturbation calculations. Influence of nonbonded parameters on the free energy of hydration of charged and neutral species, *J. Phys. Chem.* **1994**, *98*, 11362–11372

107. Zuckerman, D.M.; Woolf, T.B., Theory of a systematic computational error in free energy differences, *Phys. Rev. Lett.* **2002**, *89*

108. Lu, N.; Kofke, D. A.; Woolf, T. B., Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods, *J. Comput. Chem.* **2003**, *25*, 28–39

109. Chipot, C., Free energy calculations in biological systems. How useful are they in practice? in *New Algorithms for Macromolecular Simulation*, Leimkuhler, B.; Chipot, C.; Elber, R.; Laaksonen, A.; Mark, A. E.; Schlick, T.; Schütte, C.; Skeel, R., Eds., vol. 49. Springer: Berlin, Heidelberg, New York, 2005, pp. 183–209

110. Åqvist, J.; Medina, C.; Sammuelsson, J. E., A new method for predicting binding affinity in computer-aided drug design, *Protein Eng.* **1994**, *7*, 385–391

111. Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A., The statistical-thermodynamic basis for computation of binding affinities: a critical review, *Biophys. J.* **1997**, *72*, 1047–1069

112. Hermans, J.; Wang, L., Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme, *J. Am. Chem. Soc.* **1997**, *119*, 2707–2714

113. Duffy, E. M.; Jorgensen, W. L., Prediction of properties from simulations: free energies of solvation in hexadecane, octanol and water, *J. Am. Chem. Soc.* **2000**, *122*, 2878–2888

114. Jorgensen, W. L.; Ruiz-Caro, J.; Tirado-Rives, J.; Basavapathruni, A.; Anderson, K. S.; Hamilton, A. D. Computer-aided design of non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Bioorg. Med. Chem. Lett.* **2006**, *16,* 663–667

115. Pearlman, D. A.; Charifson, P. S., Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system, *J. Med. Chem.* **2001**, *44*, 3417–3423

116. Smith, P. E.; van Gunsteren, W. F., Predictions of free energy differences from a single simulation of the initial state, *J. Chem. Phys.* **1994**, *100*, 577–585

117. Oostenbrink, C.; van Gunsteren, W. F., Free energies of ligand binding for structurally diverse compounds, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6750–6754

118. Amadei, A.; Apol, M. E. F.; Berendsen, H. J. C., The quasi-Gaussian entropy theory: free energy calculations based on the potential energy distribution function, *J. Chem. Phys.* **1996**, *104*, 1560–1574

119. Simonson, T.; Archontis, G.; Karplus, M., Continuum treatment of long-range inter-actions in free energy calculations. Application to protein–ligand binding, *J. Phys. Chem. B* **1997**, *101*, 8349–8362

120. Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A., Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy, *Biophys. J.* **2004**, *86*, 67–74

121. Feynman, R. P., *Statistical Mechanics*, Benjamin/Cummings: London, 1972

122. Feynman, R. P.; Hibbs, A. R., *Quantum Mechanics and Path Integrals*, McGraw-Hill: New York, 1965

123. Kleinert, H., *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, World Scientific: Singapore, 1995

124. Beck, T. L., Quantum path integral extension of Widom's test particle method for chemical potentials with application to isotope effects on hydrogen solubilities in model solids, *J. Chem. Phys.* **1992**, *96*, 7175–7177

125. Beck, T. L.; Marchioro, T. L., The quantum potential distribution theorem, in *Path integrals from meV to MeV: Tutzing 1992* (1993), Grabert, H.; Inomata, A.; Schulman, L.; Weiss, U., Eds., World Scientific: Singapore, pp. 238–243

126. Wang, Q.; Johnson, J. K.; Broughton, J. Q., Thermodynamic properties and phase equilibrium of fluid hydrogen from path integral simulations, *Mol. Phys.* **1996**, *89*, 1105–1119

127. Wang, Q.; Johnson, J. K.; Broughton, J. Q., Path integral grand canonical Monte Carlo, *J. Chem. Phys.* **1997**, *107*, 5108–5117

128. Simonson, T.; Archontis, G.; Karplus, M., Free energy simulations come of age: protein–ligand recognition, *Acc. Chem. Res.* **2002**, *35*, 430–437

129. Chipot, C.; Pearlman, D. A., Free energy calculations. the long and winding gilded road, *Mol. Simul.* **2002**, *28*, 1–12

130. Berne, B. J.; Straub, J. E., Novel methods of sampling phase space in the simulation of biological systems, *Curr. Opin. Struct. Biol.* **1997**, *7*, 181–189

131. Rodinger, T.; Pomès, R., Enhancing the accuracy, the efficiency and the scope of free energy simulations, *Curr. Opin. Struct. Biol.* **2005**, *15*, 164–170

132. Hill, T. L., *An Introduction to Statistical Thermodynamics*, Dover: New York, 1986

133. McQuarrie, D. A., *Statistical Mechanics*, Harper and Row: New York, 1976

134. Chandler, D., *Introduction to Modern Statistical Mechanics*, Oxford University Press: Oxford, 1987

135. Frenkel, D.; Smit, B., *Understanding Molecular Simulations: From Algorithms to Applications*, Academic: San Diego, 1996

136. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon: Oxford, 1987

137. Pearson, E. M.; Halicioglu, T.; Tiller, W. A., Laplace-transform technique for deriving thermodynamic equations from the classical microcanonical ensemble, *Phys. Rev. A* **1988**, *32*, 3030–3039

138. Cagin, T.; Ray, J. R., Fundamental treatment of molecular-dynamics ensembles, *Phys. Rev. A* **1988**, *37*, 247–251

139. Ruelle, D., *Statistical Mechanics: Rigorous Results*, World Scientific: Singapore, 1999

140. Ray, J. R., Microcanonical ensemble Monte Carlo method, *Phys. Rev. A* **1991**, *44*, 4061–4064

141. Lustig, R., Microcanonical Monte Carlo simulation of thermodynamic properties, *J. Chem. Phys.* **1998**, *109*, 8816–8828

142. Tabor, M. *Chaos and Integrability in Nonlinear Dynamics: An Introduction.* Wiley: New York, 1989

143. Srinivasan, R., *Importance Sampling*, Springer: Berlin, Heidelberg, New York, 2002

2

Calculating Free Energy Differences Using Perturbation Theory

Christophe Chipot and Andrew Pohorille

2.1 Introduction

Perturbation theory is one of the oldest and most useful, general techniques in applied mathematics. Its initial applications to physics were in celestial mechanics, and its goal was to explain how the presence of bodies other than the sun perturbed the elliptical orbits of planets. Today, there is hardly a field of theoretical physics and chemistry in which perturbation theory is not used. Many beautiful, fundamental results have been obtained using this approach. Perturbation techniques are also used with great success in other fields of science, such as mathematics, engineering, and economics.

Although applications of perturbation theory vary widely, the main idea remains the same. One starts with an initial problem, called the unperturbed or reference problem. It is often required that this problem be sufficiently simple to be solved exactly. Then, the problem of interest, called the target problem, is represented in terms of a perturbation to the reference problem. The effect of the perturbation is expressed as an expansion in a series with respect to a small quantity, called the perturbation parameter. It is expected that the series converges quickly, and, therefore, can be truncated after the first few terms. It is further expected that these terms are markedly easier to evaluate than the exact solution.

This is precisely the approach that was followed by the pioneers of free energy perturbation (FEP) theory [1–3]. The Hamiltonian of the target system was represented as the sum of the reference Hamiltonian and the perturbation term. The free energy difference between the two systems was expressed exactly as the ensemble average of the appropriate function of the perturbation term over the reference system. Finally, this statistical average was represented as a series. The first two terms in this series were easy to evaluate and interpret. With the advent of digital computers it was, however, realized that, with sufficient care, one might be able to evaluate the free energy difference directly from the exact formula which was the starting point for the expansion. At present, most FEP calculations are based on this approach. Even though this is somewhat inconsistent with the original idea of perturbation theory, the name of the method remained unchanged. This is well justified – the presently used FEP methodology is still focused on the perturbation term in the Hamiltonian

averaged over the reference state, and the expansion of the free energy remains a helpful theoretical and practical research tool.

FEP is not only the oldest but also one of the more useful, general-purpose strategies for calculating free energy differences. In the early years of molecular-level computer simulations in chemistry and biology it was applied to small solutes dissolved in water [4]. Today, it is used for some of the most challenging applications, such as protein–ligand interactions and in silico protein engineering [5]. It can also be applied to examine the effect of force fields on the computed free energies. Finally, perturbation theory forms the conceptual framework for a number of approximate theories.

In this chapter, we shall focus on calculating Helmholtz free energies in the canonical ensemble. Extension to other ensembles is fairly straightforward. We also note that, for most systems, differences between free energies calculated in different ensembles vanish rather quickly as the system size increases. In Sects. 2.2 and 2.3, we will derive the general expression for free energy perturbation and discuss its meaning. This will be followed by the expansion of free energy in a perturbative series. Next, we will show how to deal with large perturbations that cannot be treated satisfactorily by direct application of the general formula. At that point, we will introduce the concept of the order parameter, which describes transformations between different thermodynamic states. The basic FEP methodology will be illustrated through two simple examples. Next, *alchemical transformations* used to estimate, for instance, relative binding affinities resulting from site-directed point mutations, will be discussed in detail as an important application of FEP theory. In particular, we will examine the differences between the so-called single-topology and the dual-topology paradigms used for in silico transformations between states.

Considering that FEP calculations require significant computational effort, we will discuss a number of techniques for improving their efficiency. Discussion in the present chapter of these important, practical aspects of FEP is, however, far from complete. The issues of efficiency and the closely related topic of error analysis will be considered again in Chap. 6, yet from a somewhat different perspective. This distinction, which may seem somewhat inconvenient to the reader, is motivated by the fact that the analyses developed in Chap. 6 apply not only to FEP but also to calculations of free energies from nonequilibrium simulations, which are discussed in Chap. 5. The last topic of this chapter will be the extension of the FEP formalism to calculations of energy and entropy differences, and the relevance of free energy contributions obtained by breaking down the potential energy function into terms that have different physical interpretations.

2.2 The Perturbation Formalism

Let us start by considering an N-particle reference system described by the Hamiltonian $\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)$, which is a function of $3N$ Cartesian coordinates, $\mathbf{x}$, and their conjugated momenta $\mathbf{p}_x$. We are interested in calculating the free energy difference between this system and the target system characterized by the Hamiltonian $\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x)$

$$\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x) = \mathscr{H}_0(\mathbf{x}, \mathbf{p}_x) + \Delta\mathscr{H}(\mathbf{x}, \mathbf{p}_x) \tag{2.1}$$

Let us assume, for instance, that we seek the free energy of solvation of a chemical species at infinite dilution. Then, $\Delta\mathscr{H}(\mathbf{x}, \mathbf{p}_x)$ consists of all terms in $\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x)$ that describe solute–solvent interactions. In another example, we might be interested in calculating the difference between the hydration free energies of sodium and argon described as Lennard-Jones particles with and without a charge, respectively. For these systems, $\Delta\mathscr{H}(\mathbf{x}, \mathbf{p}_x)$ contains contributions to solute–solvent potential energy due to the presence of the charge and the change in the Lennard-Jones parameters. Also present is a kinetic energy term associated with the difference in the mass of the solute.

The difference in the Helmholtz free energy between the target and the reference systems, ΔA, can be written in terms of the ratio of the corresponding partition functions, Q_1 and Q_0 [see (1.15)]

$$\Delta A = -\frac{1}{\beta} \ln \frac{Q_1}{Q_0} \tag{2.2}$$

where $\beta = (k_{\mathrm{B}}T)^{-1}$, and

$$Q = \frac{1}{h^{3N}N!} \int\int \exp\left[-\beta\mathscr{H}\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x \tag{2.3}$$

Substituting (2.3) to (2.2) and using (2.1), we obtain

$$\Delta A = -\frac{1}{\beta} \ln \frac{\int\int \exp\left[-\beta\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x)\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x}{\int\int \exp\left[-\beta\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x} \tag{2.4}$$

$$= -\frac{1}{\beta} \ln \frac{\int\int \exp\left[-\beta\Delta\mathscr{H}(\mathbf{x}, \mathbf{p}_x)\right] \exp\left[-\beta\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x}{\int\int \exp\left[-\beta\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x}$$

As has already been discussed in the context of (1.19), the probability density function of finding the reference system in a state defined by positions $\mathbf{x}$ and momenta $\mathbf{p}_x$ is

$$P_0(\mathbf{x}, \mathbf{p}_x) = \frac{\exp\left[-\beta\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)\right]}{\int\int \exp\left[-\beta\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x} \tag{2.5}$$

If this definition is used, (2.4) becomes

$$\Delta A = -\frac{1}{\beta} \ln \int\int \exp\left[-\beta\Delta\mathscr{H}(\mathbf{x}, \mathbf{p}_x)\right] \, P_0(\mathbf{x}, \mathbf{p}_x) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x \tag{2.6}$$

or, equivalently,

$$\boxed{\Delta A = -\frac{1}{\beta} \ln\langle\exp\left[-\beta\Delta\mathscr{H}(\mathbf{x}, \mathbf{p}_x)\right]\rangle_0.} \tag{2.7}$$

Here, $\langle \cdots \rangle_0$ denotes an ensemble average over configurations sampled from the reference state. This is the fundamental FEP formula, which is the basis for all further developments in this chapter. It states that ΔA can be estimated by sampling only equilibrium configurations of the reference state.

Note that integration over the kinetic term in the partition function, (2.3), can be carried out analytically. This term is identical for the solvated and the gas-phase molecule in the first example given at the beginning of this section. Thus, it cancels out in (2.2), and (2.7) becomes

$$\Delta A = -\frac{1}{\beta} \ln\langle \exp\left(-\beta\Delta U\right)\rangle_0 \qquad (2.8)$$

where ΔU is the difference in the potential energy between the target and the reference states. The integration implied by the statistical average is now carried out over particle coordinates only. This simplification is true for any two systems of particles with the same masses. If, however, the masses differ, as is the case in the second example, there is an additional term due to the change in the kinetic energy. If this term is neglected the left-hand side of (2.8) should be identified as the *excess* Helmholtz free energy of the solute over that in the ideal gas. This issue has already been discussed in Sects. 1.2 and 1.3. Since this is the quantity that can be obtained experimentally and, therefore, is of interest in most cases, we will further use (2.8) rather than (2.7), and do not consider any change of mass during transformations between states.

If we reverse the reference and the target systems, and repeat the same derivation, using the same convention for ΔA and ΔU as before, we obtain

$$\Delta A = \frac{1}{\beta} \ln\langle \exp\left(\beta\Delta U\right)\rangle_1 \qquad (2.9)$$

Although expressions (2.8) and (2.9) are formally equivalent, their convergence properties may be quite different. As will be discussed in detail in Chap. 6, this means that there is a preferred direction to carry out the required transformation between the two states.

Using a similar approach we can derive a formula for the statistical average of any mechanical property, $\mathscr{F}(\mathbf{x}, \mathbf{p}_x)$ in the target system in terms of statistical averages over conformations representative of the reference ensemble

$$\langle \mathscr{F}(\mathbf{x}, \mathbf{p}_x)\rangle_1 = \frac{\displaystyle\int\int \mathscr{F}(\mathbf{x}, \mathbf{p}_x) \exp\left[-\beta U_1\right]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x}{\displaystyle\int\int \exp\left[-\beta U_1\right]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x} \qquad (2.10)$$

$$= \frac{\displaystyle\int\int \mathscr{F}(\mathbf{x}, \mathbf{p}_x) \exp\left[-\beta\Delta U\right]\exp\left[-\beta U_0\right]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x}{\displaystyle\int\int \exp\left[-\beta\Delta U\right]\exp\left[-\beta U_0\right]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x}$$

After multiplying both the numerator and the denominator by Q_0, we obtain

$$\langle \mathscr{F}(\mathbf{x}, \mathbf{p}_x) \rangle_1 = \frac{\langle \mathscr{F}(\mathbf{x}, \mathbf{p}_x) \exp\left(-\beta \Delta U\right) \rangle_0}{\langle \exp\left(-\beta \Delta U\right) \rangle_0} \qquad (2.11)$$

Examples of properties that can be calculated using (2.11) are the average potential energy, forces, molecular dipole moment or torsional angles in a flexible molecule. Equation (2.11) is quite useful and, not too surprisingly, will reappear later in the book, for example, in Sect. 2.10 and in Chap. 11. It is worth noting, however, that in practice, calculating $\langle \mathscr{F}(\mathbf{x}, \mathbf{p}_x) \rangle_1$ is not easier than calculating ΔA, which is related to the denominator of (2.11). In particular, obtaining accurate estimates of the average potential energy in the target state by sampling from the reference state is at least as difficult as estimating the corresponding free energy difference. We will return to this point is Sect. 2.10.

2.3 Interpretation of the Free Energy Perturbation Equation

The formulas for free energy differences, (2.8) and (2.9), are formally exact for any perturbation. This does not mean, however, that they can always be successfully applied. To appreciate the practical limits of the perturbation formalism, we return to the expressions (2.6) and (2.8). Since ΔA is calculated as the average over a quantity that depends only on ΔU, this average can be taken over the probability distribution $P_0(\Delta U)$ instead of $P_0(\mathbf{x}, \mathbf{p}_x)$ [6]. Then, ΔA in (2.6) can be expressed as a one-dimensional integral over energy difference

$$\Delta A = -\frac{1}{\beta} \ln \int \exp\left(-\beta \Delta U\right) P_0(\Delta U) \, \mathrm{d}\Delta U \qquad (2.12)$$

If U_0 and U_1 were the functions of a sufficient number of identically distributed random variables, then ΔU would be Gaussian distributed, which is a consequence of the central limit theorem. In practice, the probability distribution $P_0(\Delta U)$ deviates somewhat from the ideal Gaussian case, but still has a 'Gaussian-like' shape. The integrand in (2.12), which is obtained by multiplying this probability distribution by the Boltzmann factor $\exp\left(-\beta \Delta U\right)$, is shifted to the left, as shown in Fig. 2.1. This indicates that the value of the integral in (2.12) depends on the low-energy tail of the distribution – see Fig. 2.1.

Even though $P_0(\Delta U)$ is only rarely an exact Gaussian, it is instructive to consider this case in more detail. If we substitute

$$P_0(\Delta U) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\Delta U - \langle \Delta U \rangle_0)^2}{2\sigma^2}\right] \qquad (2.13)$$

where

$$\sigma^2 = \langle \Delta U^2 \rangle_0 - \langle \Delta U \rangle_0^{\,2} \qquad (2.14)$$
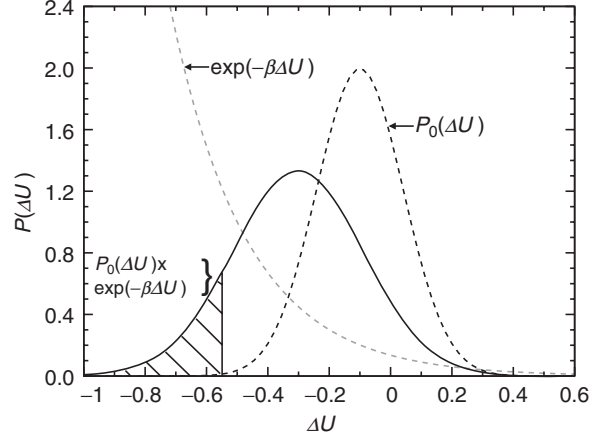
Fig. 2.1. $P_0(\Delta U)$, the Boltzmann factor $\exp\left(-\beta\Delta U\right)$ and their product, which is the integrand in (2.12). The low-ΔU tail of the integrand, marked with stripes is poorly sampled with $P_0(\Delta U)$ and, therefore, is known with low statistical accuracy. However, it provides an important contribution to the integral

to (2.12), we obtain

$$\exp(-\beta\Delta A) = \frac{C}{\sqrt{2\pi}\sigma} \int \exp\left[-\left(\Delta U - \langle\Delta U\rangle_0 - \beta\sigma^2\right)^2 / 2\sigma^2\right] d\Delta U \quad (2.15)$$

Here, C is independent of ΔU

$$C = \exp\left[-\beta\left(\langle\Delta U\rangle_0 - \frac{1}{2}\beta\sigma^2\right)\right] \quad (2.16)$$

Comparing (2.13) and (2.15), we note that $\exp\left(-\beta\Delta U\right)P_0(\Delta U)$ is a Gaussian, as is $P_0(\Delta U)$, but is not normalized and shifted toward low ΔU by $\beta\sigma^2$. This means that reasonably accurate evaluation of ΔA it via direct numerical integration is possible only if the probability distribution function in the low-ΔU region is sufficiently well known up to two standard deviations from the peak of the integrand or $\beta\sigma^2$ + 2 standard deviations from the peak of $P_0(\Delta U)$, located at $\langle\Delta U\rangle_0$. This statement is clearly only qualitative — the reader is referred to Chap. 6 for detailed error analysis in FEP methods. This simple example, nevertheless, clearly illustrates the limitations in the direct application of (2.12). If σ is small, e.g., equal to k_BT, 95% of the sampled values of ΔU are within 2σ of the peak of $\exp\left(-\beta\Delta U\right)P_0(\Delta U)$ at room temperature. However, if σ is large, for example equal to $4k_BT$, this percentage drops to 5%. Moreover, most of these samples correspond to ΔU larger than $\langle\Delta U\rangle_0 - \beta\sigma^2$ (the peak of the integrand). For this value of σ, ΔU smaller than the peak of the integrand will be sampled, on average, only 63 out of 10^6 times. Not surprisingly, estimates of ΔA will be highly inaccurate in this case, as illustrated in Fig. 2.1. Several techniques for dealing with this problem will be discussed later in this chapter and in Chap. 6.

If $P_0(\Delta U)$ is Gaussian, there is, of course, no reason to carry out a numerical integration, since the integral in (2.15) can be readily evaluated analytically. This yields

$$\Delta A = \langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2 \qquad (2.17)$$

As anticipated, the free energy difference in the Gaussian case is expressed in terms of the expectation value and the variance of the probability distribution. Both these quantities are relatively easy to estimate reliably from computer simulation. Formula (2.17) also has a broader significance that will become clear in Sect. 2.4. At this point, two observations ought to be made. The first term, which is simply equal to the average energy difference measured in the reference state, can be either positive or negative, whereas the second term, which depends on fluctuations of ΔU, is always negative. In addition, a small value of ΔA does not imply that this quantity is easy to estimate in computer simulations. In fact, if $\langle \Delta U \rangle_0$ and $-\beta\sigma^2 / 2$ were equal but large, an accurate estimate of ΔA would evidently be hard to achieve.

One consequence of the positivity of σ is that $\Delta A \leq \langle \Delta U \rangle_0$. If we repeat the same reasoning for the backwards transformation, in (2.9), we obtain $\Delta A \geq \langle \Delta U \rangle_1$. These inequalities, known as the Gibbs–Bogoliubov bounds on free energy, hold not only for Gaussian distributions, but for any arbitrary probability distribution function. To derive these bounds, we consider two spatial probability distribution functions, F and G, on a space defined by N particles. First, we show that

$$\int F \ln F \, \mathrm{d}\mathbf{x} - \int F \ln G \, \mathrm{d}\mathbf{x} \geq 0 \qquad (2.18)$$

To do so, we rewrite this expression using the fact that both F and G are normalized to 1, and, hence

$$\int F \mathrm{d}\mathbf{x} - \int G \, \mathrm{d}\mathbf{x} = 0 \qquad (2.19)$$

This expression can be added to or subtracted from any other expression without affecting its value. It follows that:

$$\int F \ln F \mathrm{d}\mathbf{x} - \int F \ln G \, \mathrm{d}\mathbf{x} = \int \left(F \ln \frac{F}{G} - F + G \right) \, \mathrm{d}\mathbf{x}$$
$$= \int G \left(\frac{F}{G} \ln \frac{F}{G} - \frac{F}{G} + 1 \right) \, \mathrm{d}\mathbf{x} \qquad (2.20)$$

The quantity in the parenthesis must be nonnegative because $x \ln x - x + 1$ is non-negative for any real, non-negative x. Since G is also non-negative, (2.18) is satisfied.

If we identify in (2.18) $P_0(\mathbf{x})$ and $P_1(\mathbf{x})$ as F and G, respectively, then, after some algebra, in which we use the expressions (2.4) and (2.5), we obtain

$$\langle \Delta U \rangle_0 - \Delta A \geq 0 \qquad (2.21)$$

or, after appropriate rearrangement,

$$\Delta A \leq \langle \Delta U \rangle_0 \tag{2.22}$$

If we reverse the assignment of F and G, we get

$$\Delta A \geq \langle \Delta U \rangle_1 \tag{2.23}$$

The Gibbs–Bogoliubov inequalities set bounds on ΔA of $\langle \Delta U \rangle_0$ and $\langle \Delta U \rangle_1$, which are easier a priori to estimate. These bounds are of considerable conceptual interest, but are rarely sufficiently tight to be helpful in practice. Equation (2.17) helps to explain why this is so. For distributions that are nearly Gaussian, the bounds are tight only if σ is small enough.

2.4 Cumulant Expansion of the Free Energy

In this section, we take an approach that is characteristic of conventional perturbation theories, which involves an expansion of a desired quantity in a series with respect to a small parameter. To see how this works, we start with (2.8). The problem of expressing $\ln\langle \exp(-tX) \rangle$ as a power series is well known in probability theory and statistics. Here, we will not provide the detailed derivation of this expression, which relies on the expansions of the exponential and logarithmic functions in Taylor series. Instead, the reader is referred to the seminal paper of Zwanzig [3], or one of many books on probability theory – see for instance [7]. The basic idea of the derivation consists of inserting

$$\left\langle e^{tX} \right\rangle = \exp\left(\sum_{n=1}^{\infty} \frac{\kappa_n t^n}{n!} \right) \tag{2.24}$$

into (2.8), which leads, to the cumulant expansion

$$\Delta A = -\frac{1}{\beta} \ln \langle \exp(-\beta \Delta U) \rangle_0 = \sum_{n=1}^{\infty} (-1)^{n-1} \kappa_n \frac{\beta^n}{n!} \tag{2.25}$$

in which κ_n is the nth cumulant of the probability distribution $P_0(\Delta U)$. Consecutive cumulants can be obtained from lower-order cumulants and raw moments, μ_n, of $P_0(\Delta U)$, which are defined as $\langle \Delta U^n \rangle_0$, using the recursion formula

$$\kappa_n = \mu_n - \sum_{m=1}^{n-1} \binom{n-1}{m-1} \kappa_m \, \mu_{n-m} \tag{2.26}$$

Cumulants have several interesting properties. All κ_n for $n > 1$ are shift-independent, i.e., they do not depend on the value of $\langle \Delta U \rangle_0$. Homogeneity, expressed by the relationship

$$\kappa_n(cX) = c^n \kappa_n(X) \tag{2.27}$$

in which c is a constant, indicates how to transform cumulants with a change of energy scale. Additivity for two independent variables ΔU_i and ΔU_j

$$\kappa_n(X + Y) = \kappa_n(X) + \kappa_n(Y) \tag{2.28}$$

provides the prescription for constructing the total cumulant expansion from the results obtained using the stratification method described in Sect. 2.5.

The first four terms called, respectively, the average (or expectation value), variance, skewness, and kurtosis, are equal to

$$\kappa_1 = \langle \Delta U \rangle_0 \tag{2.29}$$

$$\kappa_2 = \langle \Delta U^2 \rangle_0 - \langle \Delta U \rangle_0^2$$

$$\kappa_3 = \langle \Delta U^3 \rangle_0 - 3 \langle \Delta U^2 \rangle_0 \langle \Delta U \rangle_0 + 2 \langle \Delta U \rangle_0^3$$

$$\kappa_4 = \langle \Delta U^4 \rangle_0 - 4 \langle \Delta U^3 \rangle_0 \langle \Delta U \rangle_0 - 3 \langle \Delta U^2 \rangle_0^2 + 12 \langle \Delta U^2 \rangle_0 \langle \Delta U \rangle_1^2 - 6 \langle \Delta U \rangle_0^4$$

As may be seen, the formulas for higher-order cumulants become more complicated. More importantly, they are increasingly difficult to estimate accurately from simulations.

If the expansion is terminated after the second order, the free energy takes the form

$$\boxed{\Delta A = \langle \Delta U \rangle_0 - \frac{\beta}{2} \left(\langle \Delta U^2 \rangle_0 - \langle \Delta U \rangle_0^2 \right),} \tag{2.30}$$

which is identical to (2.17). This means that the second-order perturbation theory is accurate for Gaussian probability distribution functions. In fact, these are the only probability distributions that have this property. In other words, truncating the expansion for ΔA at the second order is equivalent to replacing $P_0(\Delta U)$ by a Gaussian distribution with the same variance. This is a fundamental result because it forms the basis for many approximate methods for estimating free energies. The Born and Onsager formulas for the free energy of an ion or a dipolar particle in a homogeneous liquid, respectively, are well-known examples of applying second-order perturbation theory. We will discuss this point in the following sections.

No probability density function exists that can be expanded into a finite number of cumulants larger than two. In other words, the cumulant expansion either has less than three terms, or it must be infinite, or it does not exist – i.e., it diverges. One limit at which the expansion is usually well behaved – i.e., converges quickly – is the high temperature. It is then clear from (2.25) that β becomes a true 'small parameter,' as required in conventional perturbation theories. Concerns about convergence imply that, in general, the cumulant expansion should be used beyond the second-order with great care. Including higher terms in (2.25) may not be more accurate than the second-order or direct free energy calculations. Doing so might, nevertheless, be advantageous because it yields a $P_0(\Delta U)$ that is smoother in the tails than the one

obtained directly from the data. Other, and probably better, methods for modeling $P_0(\Delta U)$ by smooth functions will be discussed in Sect. 2.9.3.

2.5 Two Simple Applications of Perturbation Theory

In this section, we discuss applications of the FEP formalism to two systems and examine the validity of the second-order perturbation approximation in these cases. Although both systems are very simple, they are prototypes for many other systems encountered in chemical and biological applications. Furthermore, the results obtained in these examples provide a connection between molecular-level simulations and approximate theories, especially those based on a dielectric continuum representation of the solvent.

2.5.1 Charging a Spherical Particle

In the first example, we consider the transformation in which an uncharged Lennard-Jones particle immersed in a large container of water acquires a charge q. The free energy change associated with charging is given by (2.8), in which the subscript 0 refers to the reference state of the solvated Lennard-Jones sphere and ΔU is the electrostatic energy of interaction between the charge q and all water molecules.

$$\Delta U = qV \tag{2.31}$$

Here, V is the electrostatic potential created by the solvent that acts on the charge in the center of the cavity. Recalling (2.30), the second-order perturbation theory yields

$$\Delta A = q\langle V \rangle_0 - \frac{\beta}{2}q^2[\langle (V - \langle V \rangle_0)^2 \rangle_0]. \tag{2.32}$$

If water is considered a homogeneous dipolar liquid, $\langle V \rangle_0 = 0$ and the expression for the free energy change further simplifies to

$$\Delta A = -\frac{\beta}{2} q^2 \left\langle V^2 \right\rangle_0 \tag{2.33}$$

This result implies that ΔA should be a quadratic function of the ionic charge. This is exactly what is predicted by the Born model, in which the ion is a spherical particle of radius a and the solvent is represented as a dielectric continuum characterized by a dielectric constant ε [1]

$$\Delta A = -\frac{\varepsilon - 1}{\varepsilon} \frac{q^2}{2a}. \tag{2.34}$$

It is instructive to compare these predictions with the results of computer simulations. This comparison, however, requires care. In practice, the computed values of ΔA exhibit considerable system-size dependence, i.e., they vary with the size of the simulation box. This is because charge–dipole interactions between the solute and

solvent molecules decay slowly as $1/R^2$ with the distance R. For typical simulation systems, they are not negligible even at the largest solute–solvent separations. System-size effects may, however, be greatly reduced by properly correcting for the self-interaction term, which is due to interactions between the charge and its images and the neutralizing background [8]. This is true for both Ewald lattice summation and generalized reaction field (GRF) treatments of finite-size effects [9]. In general, free energy calculations, in which the system is transformed such that its electrical charge changes, should include system-size corrections.

Monte Carlo simulations, in which a methane-like particle was progressively charged to $q = 1$ or $q = -1$ in intervals of 0.25 e led to the conclusion that the quadratic dependence of ΔA on q, predicted by the second-order perturbation theory, is essentially correct [8]. In agreement with experimental data [10] negative ions are, however, better hydrated than positive ions. This is reflected by the different slopes of the straight line in Fig. 2.2, and can be ascribed to the different arrangements of water molecules in the vicinity of the ion. The positively charged hydrogen atoms of water, which possess small van der Waals radii, can approach negative ions closer than the large, negative oxygen atoms can approach positive ions. This asymmetry leads to a net positive potential acting on the uncharged particle, and for this reason the lines in Fig. 2.2 do not intersect at $q = 0$.
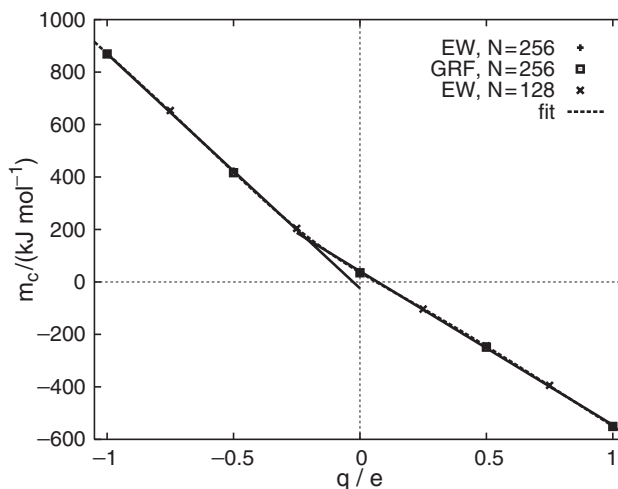
Fig. 2.2. Average electrostatic potential m_c at the position of the methane-like Lennard-Jones particle Me as a function of its charge q. m_c contains corrections for the finite system size. Results are shown from Monte Carlo simulations using Ewald summation with $N = 256$ (*plus*) and $N = 128$ (*cross*) as well as GRF calculations with $N = 256$ water molecules (*square*). Statistical errors are smaller than the size of the symbols. Also included are linear fits to the data with $q < 0$ and $q > 0$ (*solid lines*). The fit to the tanh-weighted model of two Gaussian distributions is shown with a *dashed line*. Reproduced with permission of the American Chemical Society

We note that accurate values of ΔA for larger values of q – e.g., $q = 1$ – cannot be obtained from single-step FEP calculations. Instead, a series of calculations, in which q is progressively changed from zero to its final value, need to be performed. This approach is discussed in Sect. 2.6.

2.5.2 Dipolar Solutes at an Aqueous Interface

In this second example, we examine simple systems near the water–hexane interface. Specifically, we calculate the difference in the free energy of hydrating a hard-sphere solute of radius a, considered as the reference state, and a model solute consisting of a point dipole p located at the center of a cavity [11]. We derive the formula for ΔA assuming that the solute is located at a fixed distance z from the interface, and subsequently we examine the dependence of the free energy on z. The geometry of the system is shown in Fig. 2.3.

For the model solutes, the difference in the potential energy is equal to the electrostatic energy of solvating a dipole

$$\Delta U = -\mathbf{p} \cdot \mathbf{E} = -pE \cos \theta \tag{2.35}$$

and ΔA, statistically averaged over dipolar orientations, is expressed as

$$\Delta A = -\frac{1}{\beta} \ln \langle \exp(-\beta \Delta U) \rangle_0 \tag{2.36}$$

$$= -\frac{1}{\beta} \ln \left[\frac{\int \mathrm{d}\mathbf{x} \exp(-\beta U_0) \int_0^{2\pi} \mathrm{d}\phi \int_0^\pi \exp(\beta pE \cos \theta) \sin \theta \, \mathrm{d}\theta}{4\pi \int \mathrm{d}\mathbf{x} \exp(-\beta U_0)} \right]$$

where $\mathbf{p}$ and $\mathbf{E}$ are, respectively, the vectors of the dipole moment of the solute and the electric field created by the solvent and acting on the dipole, $\mathbf{x}$ abbreviates the coordinates of all particles in the system, and ϕ, θ are angles in the cylindrical coordinate system that describe the orientation of the dipole.

After integrating over ϕ and θ, we obtain

$$\Delta A = -\frac{1}{\beta} \ln \left[\frac{\int \mathrm{d}\mathbf{x} \exp(-\beta U_0) 4\pi \sinh(\beta pE)/\beta pE}{4\pi \int \mathrm{d}\mathbf{x} \exp(-\beta U_0)} \right]$$

$$= -\frac{1}{\beta} \ln \left\langle \frac{\sinh(\beta pE)}{\beta pE} \right\rangle_0 \tag{2.37}$$

Next, we consider the second-order perturbation approximation. Since $\langle \Delta U \rangle_0$ averaged over dipolar orientations vanishes

$$\langle \Delta U \rangle_0 = \frac{1}{4\pi} \left\langle \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} pE \cos\theta \sin\theta \, \mathrm{d}\theta \right\rangle_0 = 0 \qquad (2.38)$$

the expression for the free energy difference simplifies to

$$\Delta A = -\frac{\beta}{2} \langle \Delta U^2 \rangle_0 = -\frac{\beta}{6} p^2 \langle E^2 \rangle_0 \qquad (2.39)$$

Equation (2.39) leads to the prediction that ΔA should be proportional to p^2. For a bulk solvent, this can be considered as a molecular equivalent of the well-known Onsager formula derived for the continuum dielectric model [12].

$$\Delta A = -\frac{p^2(\varepsilon - 1)}{a^3(2\varepsilon + 1)} \qquad (2.40)$$

Indeed, both expressions predict quadratic dependence of ΔA on the dipole moment of the solute. As in the previous example, it is of interest to test whether this prediction is correct. Such a test was carried out by calculating ΔA for a series of model solutes immersed in water at different distances from the water–hexane interface [11]. The solutes were constructed by scaling the atomic charges and, consequently, the dipole moment of a nearly spherical molecule, CH_3F, by a parameter λ, which varied between 0 and 1.2. The results at two positions – deep in the water phase and at the interface – are shown in Fig. 2.3. As can be seen from the linear dependence of ΔA on p^2, the accuracy of the second-order perturbation theory
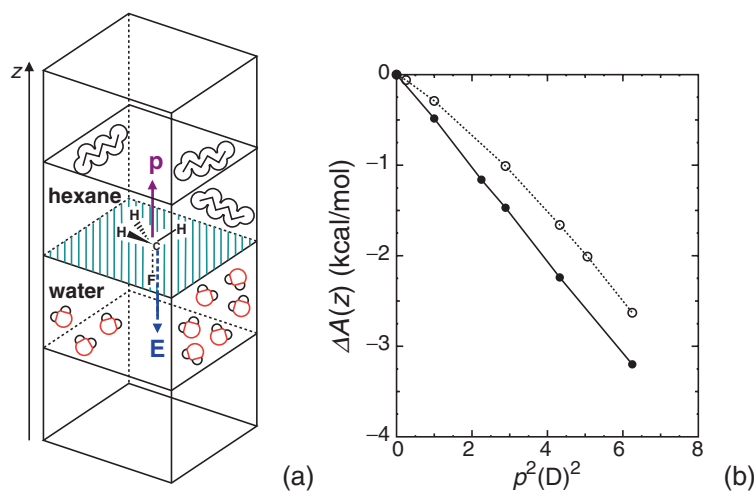
Fig. 2.3. Schematic representation of the water–hexane system (**a**). The z-coordinate is perpendicular to the water–hexane interface. Each medium is in equilibrium with its vapor phase. Periodic boundary conditions are applied. The electrostatic part of the free energy of dissolving CH_3F with scaled atomic charges as a function of the square of the molecular dipole moment, p^2, in water (*solid line*) and at the water–hexane interface (*dashed line*) at 310 K (**b**). Reproduced with permission of the American Chemical Society

in bulk water is excellent. In the anisotropic, interfacial environment this approximation is less satisfactory, and some deviations from linearity are observed. It should be pointed out, however, that long-range electrostatic effects were not taken into account in the simulations. Were they included, the agreement with the second-order theory might improve.

In both examples discussed in this section, the second-order approximation to ΔA turned out to be satisfactory. We, however, do not want to leave the reader with the impression that this is always true. If this were the case, it would imply that probability distributions of interest were always Gaussian. Statistical mechanics would then be a much simpler field. Since this is obviously not so, we have to develop techniques to deal with large and not necessarily Gaussian-distributed perturbations. This issue is addressed in the remainder of this chapter.

2.6 How to Deal with Large Perturbations

As has been already stressed, using (2.8) directly can be successful only if $P_0(\Delta U)$ is a narrow function of ΔU. This does not imply that the free energy difference between the reference and the target states must be small. For instance, although the hydration free energy of benzene – i.e., the free energy for transferring benzene from the gas phase to the aqueous phase – is only $-0.4\,\text{kcal mol}^{-1}$ at 298 K, this quantity cannot be successfully calculated by direct application of (2.8) in a simulation of a reasonable length. This is because low-energy configurations in the target ensemble, which do not suffer from the overlap between the solute and solvent molecules, are not sampled in simulations of the reference state. This point is discussed amply in Chap. 6.

This difficulty in applying FEP theory can be circumvented through a simple stratification strategy, also often called staging. It relies on constructing several intermediate states between the reference and the target state such that $P(\Delta U_{i,i+1})$ for two consecutive states i and $i + 1$ sampled at state i is sufficiently narrow for the direct evaluation of the corresponding free energy difference, $\Delta A_{i,i+1}$. Then, (2.8) can be used serially to yield ΔA. If we construct $N-2$ intermediate states then

$$\Delta A = \sum_{i=1}^{N-1} \Delta A_{i,i+1} = -\frac{1}{\beta} \sum_{i=1}^{N-1} \ln \left\langle \exp\left(-\beta \Delta U_{i,i+1}\right)\right\rangle_i. \tag{2.41}$$

Intermediate states do not have to be physically meaningful, i.e., they do not have to correspond to systems that actually exist. As an example, assume that we want to calculate the difference in hydration free energies of a Lennard-Jones particle and an ion with a positive charge q of 1e. For simplicity, we further assume that the Lennard-Jones parameters remain unchanged upon charging the particle. Since a direct calculation of the free energy difference is not likely to succeed in this case, we construct intermediate states in which the particle carries fractional charges q_i such that $q_i < q_j$ for $i < j$ and $0 < q_j < q$.

More generally, we can consider the Hamiltonian as a function of some parameter, λ. Without loss of generality, we can choose $0 \leq \lambda \leq 1$, such that $\lambda = 0$ and $\lambda = 1$ for the reference and target states, respectively. A simple choice for the dependence of the Hamiltonian on λ is a linear function

$$\mathscr{H}(\lambda_i) = \lambda_i \mathscr{H}_1 + (1 - \lambda_i)\mathscr{H}_0 = \mathscr{H}_0 + \lambda_i \Delta\mathscr{H}, \tag{2.42}$$

which justifies calling λ a coupling parameter. Here $\mathscr{H}_0$ and $\mathscr{H}_1$ denote, as before, the Hamiltonian of the reference and the target systems, respectively, but to simplify the notation we do not explicitly specify their dependence on $\mathbf{x}$ and $\mathbf{p}_x$. Similarly, $\Delta\mathscr{H}$ is the perturbation term in the target Hamiltonian, equal to $\mathscr{H}_1 - \mathscr{H}_0$. If we create $N-2$ intermediate states linking the reference and the target states such that $\lambda_1 = 0$ and $\lambda_N = 1$ then the change in the Hamiltonian, $\Delta\mathscr{H}_i$, between two consecutive states is given by

$$\Delta\mathscr{H}_i = \mathscr{H}(\lambda_{i+1}) - \mathscr{H}(\lambda_i) = (\lambda_{i+1} - \lambda_i)\Delta\mathscr{H} = \Delta\lambda_i \Delta\mathscr{H} \tag{2.43}$$

where $\Delta\lambda_i = \lambda_{i+1} - \lambda_i$, and the formula for the total free energy difference becomes

$$\Delta A = -\frac{1}{\beta} \sum_{i=1}^{N-1} \ln \langle \exp(-\beta\Delta\lambda_i \Delta\mathscr{H}) \rangle_{\lambda_i}. \tag{2.44}$$

If we recall the discussion about integrating out the kinetic term in the Hamiltonian in Sects. 1.2.1 and 2.2, then we can rewrite ΔA as

$$\Delta A = -\frac{1}{\beta} \sum_{i=1}^{N-1} \ln \langle \exp(-\beta\Delta\lambda_i \Delta U) \rangle_{\lambda_i}. \tag{2.45}$$

In the example of charging a neutral particle, $\lambda_i = q_i/q$ is the linear parameter. Choosing intermediate states separated by a constant $\Delta\lambda$ is, however, not a good choice for this problem because, as has been seen in Sect. 2.5, ΔA is a quadratic function of q. A better choice would be to decrease $\Delta\lambda$ quadratically. Alternatively, one could define $\mathscr{H}(\lambda)$ as a quadratic function of λ. Then, using a constant $\Delta\lambda$ would be appropriate.

So far, the issue of choosing N and $\Delta\lambda_i$ has remained open. Clearly, if each state is equally sampled, increasing N should improve accuracy at the expense of efficiency. There is, however, no general and practical method for striking the perfect balance between these two conflicting criteria, because that would require prior knowledge of the dependence of ΔA on λ. One method for optimizing both N and $\Delta\lambda_i$ is to start with short runs with a large N, and then select the number of intermediate states and the corresponding values of the coupling parameter on the basis of these runs, such that the estimated variances in $P(\Delta U_{i,i+1})$ are sufficiently small and approximately equal [13]. In practice, however, it might be simpler to make reasonable, if not optimal choices, remembering that it is always possible to add intermediate states if required. A deeper insight into this issue will come from the error analysis discussed in Chap. 6.

Stratification is not specific to FEP – it is a universal strategy that improves the efficiency of many other methods for calculating free energies. Not surprisingly, we will return to this strategy several times, in particular in Chaps. 3 and 6.

2.7 A Pictorial Representation of Free Energy Perturbation

In FEP calculations, configurations are sampled according to the probability, $P_0(U)$, of finding the reference system in a state corresponding to the potential energy U. Does this guarantee that the key quantity for calculating ΔA, i.e., $P_0(\Delta U)$, is estimated accurately? Unfortunately, this does not have to be the case. As has been discussed previously, FEP will only provide accurate estimates of free energy differences under the sine qua non condition that the target system is sufficiently similar to the reference system. This somewhat vague statement can be understood better by introducing the concept of *important regions* in phase space. These regions are volumes that encompass configurations of the system with highly probable energy values. More specifically, a configuration in the important region should have a potential energy the probability of which is higher than a predefined value. Configurations that belong to the important regions are expected to make significant contributions to the free energy and, therefore, should be adequately sampled.

Using the concept of important regions, it is possible to develop a pictorial representation of the relationship between the reference and the target systems, which has proven to be a useful tool to detect inaccuracies caused by incomplete sampling [14]. This is depicted in Fig. 2.4. If the important region of the target system fully overlaps with or, more precisely, is a subset of the important region of the reference system, as shown in Fig. 2.4b, $P_0(\Delta U)$ estimated from FEP calculations should be reliable – because good sampling of the important region in the reference system will also yield good sampling of the important region in the target system. Conversely, if the important regions of these two systems do not overlap (see Fig. 2.4a), the important region of the target state is not expected to be sufficiently sampled during a simulation of the reference system. Then, it is unlikely that satisfactory estimates of ΔA will be obtained.

In many instances, the important region of the reference system overlaps with only a part of the important region of the target system. This is shown in Fig. 2.4c. The poorly sampled remainder of the latter important region contributes to inaccuracies in the estimated free energy differences, which, in some circumstances, could be substantial. Note a special case of the situation discussed here, in which the important region of the reference system is a subset of the important region of the target system. Then only a part of the latter region will be adequately sampled. This deficiency, however, can be readily remedied by switching the reference and the target systems. If sampling is conducted from the target system, then the relationship between the important regions corresponds to that shown in Fig. 2.4b.

If the two important regions do not overlap, or overlap only partially, it is usually necessary to use the enhanced sampling techniques introduced in Sect. 1.4. This is schematically illustrated in Fig. 2.4d. One of these techniques, stratification, has
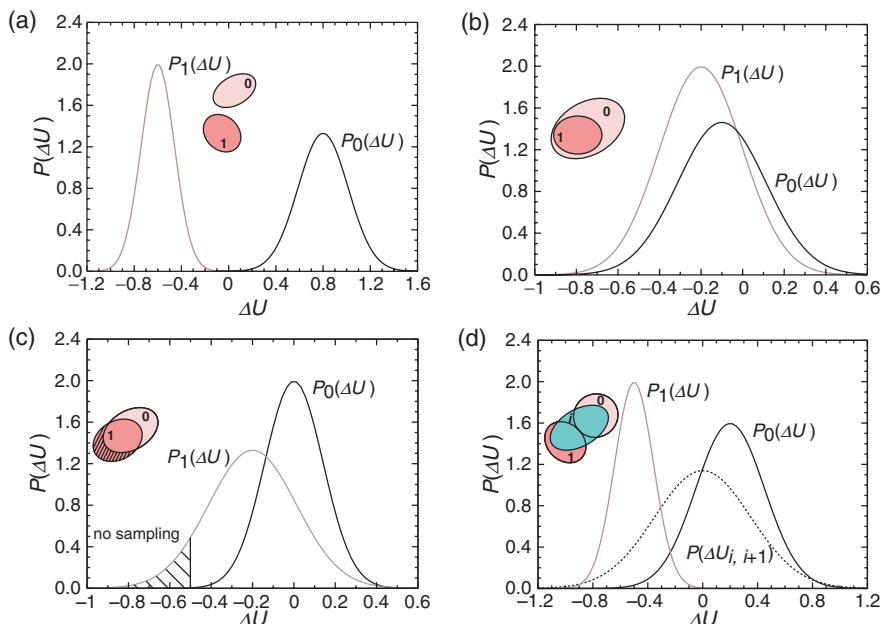
Fig. 2.4. Schematic representation of the different relationships between the important regions in phase space for the reference (0) and the target (1) systems, and their possible interpretation in terms of probability distributions – it should be clarified that because ΔU can be distributed in a number of different ways, there is no obvious one-to-one relation between $P_0(\Delta U)$, or $P_1(\Delta U)$, and the actual level of overlap of the ensembles [14]. (**a**) The two important regions do not overlap. (**b**) The important region of the target system is a subset of the important region of the reference system. (**c**) The important region of the reference system overlaps with only a part of the important region of the target state. Then enhanced sampling techniques of stratification or importance sampling that require the introduction of an intermediate ensemble should be employed (**d**)

been presented in Sect. 2.6, whereas the application of important sampling to FEP will be discussed briefly in Sect. 2.9.1. This discussion will be expanded considerably in Chap. 6. Anticipating these developments, we just mention that the optimal enhanced sampling strategy is largely determined by the relationship between the two important regions in phase space [14].

This pictorial representation is useful for understanding under what circumstances satisfactory estimates of $P_0(\Delta U)$ can be expected, and how to deal with situations when this is not the case. It should, however, also be appreciated that the reasoning behind this representation is only qualitative and may occasionally fail. For example, if the energy landscapes in the important regions of the reference and the target systems were markedly different, obtaining an accurate estimate of ΔA would be a challenge even if these regions overlapped perfectly. A similar difficulty would be encountered if a large fraction of the important region of the target system overlapped with a low-probability part of the important region of the reference

state. Conversely, it would be possible to obtain a reliable estimate of ΔA even if the two regions overlapped only partially, providing that the nonoverlapping parts corresponded to relatively low-probability, high-energy configurations.

2.8 'Alchemical Transformations'

In the earlier sections, we have developed the theoretical framework for the FEP approach. In this section, we outline some specific methodologies built upon this framework to calculate the free energy differences associated with the transformation of a chemical species into a different one. This computational process is often called *alchemical transformation* because, in a sense, this is a realization of the inaccessible dream of the proverbial alchemist – to transmute matter. Yet, unlike lead, which was supposed to turn into gold in the alchemist's furnace, the potential energy function is sufficiently malleable in the hands of the computational chemist that it can be gently altered to transform one chemical system into another, slightly modified one.

Over the years, the FEP methodology has been employed widely to some of the most computationally challenging problems in theoretical chemistry and biology. Applications of this approach include, among others, protein–ligand binding, host–guest chemistry, and solvation properties [5, 15–17]. The reader is referred to Chap. 13 for a brief review of these applications. Here, we focus much of our discussion on the free energy of binding between a protein and its potential ligands, which is a problem of great interest to computer-aided drug design and protein engineering. This focus should help to make the following material less abstract, but without limiting the generality of the methodologies discussed. It will be quite clear that they can be applied to other, related problems without change. Before we proceed further, however, we introduce the concept of an order parameter, which is essential not only for further developments is this section but for free energy calculations in general.

2.8.1 Order Parameters

The coupling parameter λ, discussed in Sect. 2.6, is an example of an *order parameter*. In different fields, this term has very specific and often somewhat different meanings, but here, we use it in the most generic sense. An order parameter indicates the degree of order in the system, or, even more generally, it is a variable chosen to describe changes in a system. In the context of free energy calculations, order parameters are collective variables that are used to describe transformations from the initial, reference system to the final, target one. An order parameter may, although does not necessarily have to, correspond to the path along which the transformation takes place in nature. If this were the case, it would be called the *reaction coordinate* or the *reaction path*.

Several examples of order parameters are shown in Fig. 2.5. Some of them, e.g., the torsional angles **a**, are dynamical variables, which means that they can be fully
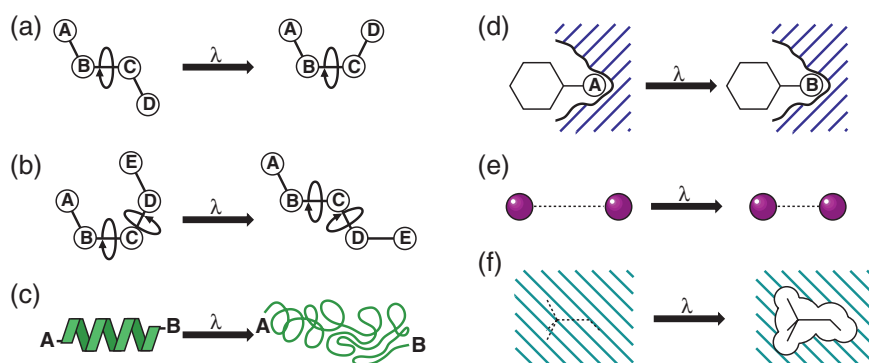
Fig. 2.5. Possible applications of a coupling parameter, λ, in free energy calculations. (**a**) and (**b**) correspond, respectively, to simple and coupled modifications of torsional degrees of freedom, involved in the study of conformational equilibria; (**c**) represents an intramolecular, end-to-end reaction coordinate that may be used, for instance, to model the folding of a short peptide; (**d**) symbolizes the alteration of selected nonbonded interactions to estimate relative free energies, in the spirit of site-directed mutagenesis experiments; (**e**) is a simple distance separating chemical species that can be employed in potential of mean force (PMF) calculations; and (**f**) corresponds to the annihilation of selected nonbonded interactions for the estimation of e.g., free energies of solvation. In the examples (**a**), (**b**), and (**e**), the coupling parameter, λ, is not independent of the Cartesian coordinates, $\{\mathbf{x}\}$. Appropriate metric tensor correction should be considered through a relevant transformation into generalized coordinates

represented as functions of Cartesian coordinates. Other order parameters, for instance, the charging parameter – as part of point mutation **d** – introduced at the end of Sect. 2.6, are not. This distinction is useful, because these two types of order parameters may require different treatments, as will become clear in Chap. 4. Fortunately, once these treatments are completed, it turns out that almost all theoretical developments in this chapter apply to both cases.

The concept of order parameter is central to free energy calculations. Yet, even a casual reader will easily observe that in most cases there is more than one way to define it. This immediately raises the question: how to make the best, or at least an appropriate, choice of an order parameter? Unfortunately, this remains a major, essentially still unresolved, problem in the field. The choice of order parameters may have a significant effect on the efficiency and accuracy of free energy calculations. Some order parameters may map a smooth path between the reference and target states whereas others may lead through a rough energy landscape. Then, estimating ΔA in the former case should be easier and should require fewer intermediate states than in the latter. This does not mean, however, that we are totally helpless. In many cases, there is a 'natural' choice of the order parameter, dictated by the physical problem at hand. Furthermore, it is possible to formulate several criteria that should guide our choices. A full discussion of this issue would, however, be premature at this point. In several subsequent chapters, we will consider order parameters from

different perspectives, and only in the last chapter will we summarize our under-standing of the problem.

2.8.2 Creation and Annihilation

Central to the application of FEP to alchemical transformations is the concept of a thermodynamic cycle. This is a series of reversible transformations of the system constructed such that, at the end, the initial state is restored and, therefore, the total change of free energy in the cycle is zero. The cycle may be hypothetical in the sense that individual transformations can be carried out only on a computer. Usually, we are interested in only one or two steps of the cycle. In practice, however, it might be simpler to calculate free energy changes associated with all other steps of the cycle and obtain the free energy of interest using this route. Considering that all transformations are reversible, the cycle can be run in both directions and the free energy of interest can be calculated from either forward or backward transformations (with proper attention to the signs of the associated free energies). The computational efficiencies of carrying out these transformations may, however, differ.

To illustrate how this approach works, we return to the first example given at the beginning of Sect. 2.2, in which we considered the transfer of a molecule between the aqueous solution and the gas phase. In the FEP framework this process can be described by turning on the perturbation term in the Hamiltonian, which, in this case, is responsible for solute–solvent interactions. This is represented by the upper horizontal arrow in the thermodynamic cycle shown in Fig. 2.6. The corresponding free energy change, viz. $\Delta A_{\mathrm{hydration}}$, can be obtained using the dual-topology app-roach, described later in this section. As we have already pointed out in Sect. 2.2, $\Delta A_{\mathrm{hydration}}$ calculated this way is the *excess* free energy over that in the gas phase. This quantity is of interest because it is directly related to the solubility, s, of the solute, which can be determined experimentally or calculated as

$$s = C \, \exp\left(-\beta \Delta A_{\mathrm{hydration}}\right) \tag{2.46}$$

where C is a constant that determines the units of s.

When solute–solvent interactions are turned on, the solute–solvent potential energy no longer vanishes. This is not the only change in the system. Usually, the solvent undergoes substantial reorganization, and conformational equilibria in flexible solutes may also be affected. For example, the *trans–gauche* equilibrium in 1,2-dichloroethane shifts towards the *gauche* state upon hydration. All these reorganization effects are correctly taken into account in the FEP formalism.

$\Delta A_{\mathrm{hydration}}$ can be also obtained from the reverse process, in which solute–solvent interactions are turned off. This corresponds to moving the solute from the aqueous solution to the gas phase. Then the calculated quantity is the negative of $\Delta A_{\mathrm{hydration}}$. If the same order parameter, λ, is used for the forward and the reverse transformations, the changes in the free energy with λ should be re-versible, and, consequently, the sum of the calculated free energies differences should be zero. This is shown in Fig. 2.7. Discrepancies between the forward and the reverse
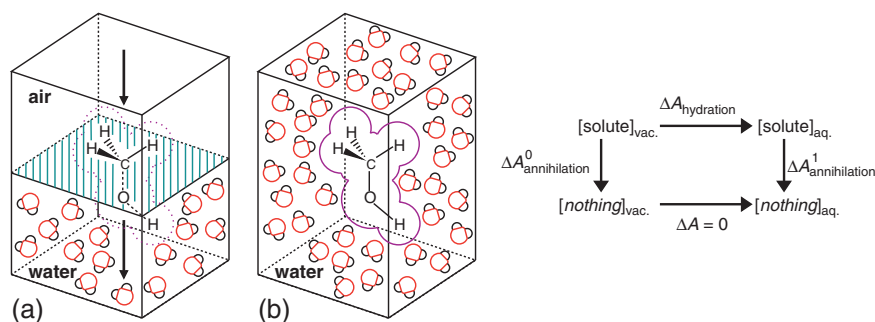
Fig. 2.6. The thermodynamic cycle for estimating the hydration free energy, $\Delta A_{\text{hydration}}$, of a small solute (the right side of the figure). One route is the direct evaluation of $\Delta A_{\text{hydration}}$ along the upper vertical arrow. The solute, originally placed in vacuum (**a**) is moved to the bulk water (**b**). Another route consists of annihilating, or creating, the solute both in vacuo and in the aqueous medium and corresponds to the vertical lines in the thermodynamic cycle. As suggested by the cycle, these two routes are formally equivalent, as: $\Delta A_{\text{hydration}} = \Delta A^0_{\text{annihilation}} - \Delta A^1_{\text{annihilation}}$
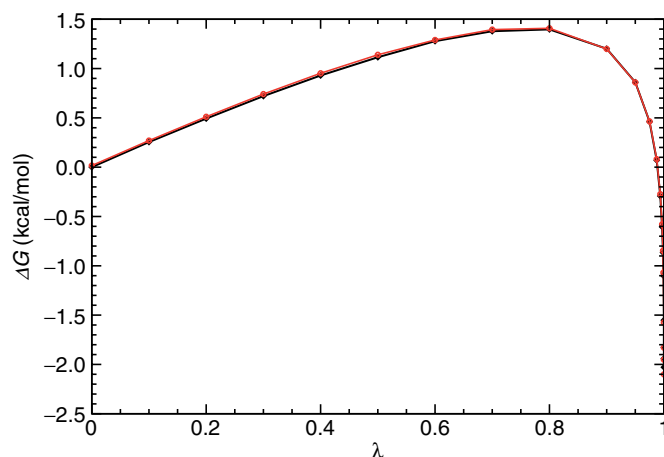
Fig. 2.7. Hydration free energy of argon. Using NAMD [18], two molecular dynamics simulation were carried out in the isothermal–isobaric ensemble, at 300 K and 1 atm, to annihilate (*solid curve*, λ: $0 \rightarrow 1$) and create (*dashed curve*, λ: $1 \rightarrow 0$) an argon atom in liquid water. Twenty one windows of uneven width, consisting of 40 ps of equilibration and 400 ps of data collection, were utilized for each transformation. This corresponds to a total of 9.24 ns of sampling for the creation and annihilation of argon. The equations of motion were integrated with a time step of 2 fs. The TIP3P model was chosen to describe water molecules [19]. Long-range electrostatic interactions were taken into account using the particle-mesh Ewald (PME) method. Van der Waals interactions were truncated smoothly beyond 10 Å. The calculated free energies in the forward (creation) and the backward (annihilation) transformations are, respectively, $+2.11$ kcal mol^{-1} and to -2.08 kcal mol^{-1}, which yields a negligible hysteresis. For comparison, the experimentally determined free energy of hydration at 298 K is 2.002 kcal mol^{-1}

transformations yield the hysteresis of the reaction, which constitutes a measure of the error in the free energy calculation. If the hysteresis is markedly larger than the estimated statistical errors, it is usually indicative of ergodicity issues during the transformations. Yet, even if hysteresis is negligible and statistical errors are small, this does not imply that the calculated free energy difference is accurate, because it may be burdened with systematic errors, due, for instance, to unsuitable potential energy functions. These points are fully developed in Chap. 6.

An alternative approach to calculating the free energy of solvation is to carry out simulations corresponding to the two vertical arrows in the thermodynamic cycle in Fig. 2.6. The transformation to 'nothing' should not be taken literally – this means that the perturbed Hamiltonian contains not only terms responsible for solute–solvent interactions – viz. for the right vertical arrow – but also all the terms that involve intramolecular interactions in the solute. If they vanish, the solvent is reduced to a collection of noninteracting atoms. In this sense, it 'disappears' or is 'annihilated' from both the solution and the gas phase. For this reason, the corresponding computational scheme is called *double annihilation*. Calculations of the corresponding free energy differences, $\Delta A^0_{\text{annihilation}}$ and $\Delta A^1_{\text{annihilation}}$, are amenable to the single-topology approach, which will be discussed shortly. Since the total free energy change in the cycle must be zero it follows that $\Delta A_{\text{hydration}} = \Delta A^0_{\text{annihilation}} - \Delta A^1_{\text{annihilation}}$.

As before, we can perform reverse simulations. Instead of annihilating the solute, we can 'create' it by turning on the perturbation part of the Hamiltonian. The resulting free energy differences are connected through the relation: $\Delta A^1_{\text{creation}} - \Delta A^0_{\text{creation}} = \Delta A^0_{\text{annihilation}} - \Delta A^1_{\text{annihilation}}$. Comparison of this creation scheme with the transformation described by the horizonal arrow reveals two important differences. First, the vertical transformations require two sets of simulations instead of one, although one of them involves only solute in the gas phase and, is, therefore, much less computationally intensive. Second, the two methods differ in their description of the solute in the reference state. In both cases the solute does not interact with the solvent. For the vertical transformations, however, all interactions between atoms forming the solute vanish, whereas in the horizontal transformation, the molecule remains intact.

In closing, we note that FEP may not be the most efficient approach for calculating ΔA in the examples given in this section. The free energy of solvation can be obtained efficiently by considering a system in which a water lamella coexists with its vapor phase, and then using methods described in Chaps. 3 and 4 to compute the free energy change associated with translating the solvent along the normal to the interface formed by the liquid and the vapor phases [20, 21]. This path is shown in Fig. 2.6a. The free energy of hydrating argon can be determined accurately using the particle insertion method, described in Chap. 9 . For more-complicated problems that require the determination of binding free energies, FEP, however, still remains the method of choice.

2.8.3 Free Energies of Binding

The free energy of binding of two molecules, $\Delta A_{\text{binding}}$, defined as the free energy difference between these molecules in the bound and the free, unbound states, can be determined experimentally through the measurement of binding constants using, for instance, BIAcore [22] or microcalorimetry techniques. If we focus on protein–ligand binding, the computationally equivalent procedure corresponds to calculating $\Delta A_{\text{binding}}$ along the top, horizontal transformation of Fig. 2.8. This, in principle, can be done via FEP, provided that we use an order parameter that measures the separation between the ligand and the binding center of the protein. It should be appreciated, however, that, in general, defining a relevant order parameter that describes protein–ligand association may be quite difficult, in particular when the ligand is buried deep in the binding pocket, and access to the latter involves large conformational changes in the protein. Furthermore, the choice of an order parameter that is a function of the Cartesian coordinates, $\mathbf{x}$, causes conceptual difficulties, because it cannot be treated as an independent variable. This problem can be handled through a transformation to generalized coordinates incorporating the appropriate mass tensor correction to $\Delta A_{\text{binding}}$. The theoretical formalism for this treatment will be developed in considerable detail in Chap. 4. Here, we only note that probability distribution methods or thermodynamic integration, described in Chaps. 3 and 4 are excellent alternatives to FEP if the direct route for calculating $\Delta A_{\text{binding}}$ is used.

As in Sect. 2.8.2, there is an alternative route to calculating $\Delta A_{\text{binding}}$. It requires obtaining $\Delta A^0_{\text{annihilation}}$ and $\Delta A^1_{\text{annihilation}}$ along the vertical legs of Fig. 2.8. It follows that $\Delta A_{\text{binding}} = \Delta A^0_{\text{annihilation}} - \Delta A^1_{\text{annihilation}}$. If we opt for this route, then FEP is an appropriate technique. This requires, however, some care. The ligand in the binding pocket is annihilated from a strongly constrained position, whereas the free, unbound ligand can rotate freely during annihilation. This means that the free energy of the lower horizontal transformation may not be zero unless
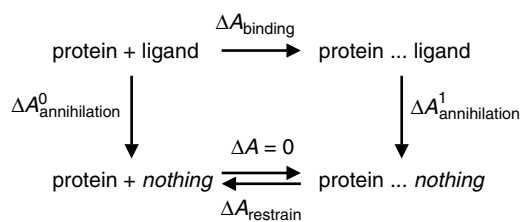
Fig. 2.8. The thermodynamic cycle used for the determination of protein–ligand binding free energy, $\Delta A_{\text{binding}}$. In general, FEP cannot be used for calculating $\Delta A_{\text{binding}}$ directly, following the upper horizontal transformation. Considering that the lower horizontal transformation corresponds to a zero free energy change, annihilation of the ligand in the reference, free state – i.e., the left, vertical transformation, and in the bound state – i.e., the right, vertical transformation, yields the binding free energy: $\Delta A_{\text{binding}} = \Delta A^0_{\text{annihilation}} - \Delta A^1_{\text{annihilation}}$. The contribution $\Delta A_{\text{restrain}}$ that appears in the reverse, lower horizontal transformation characterizes the loss of rotational and translational entropies due to restraining the position of the ligand
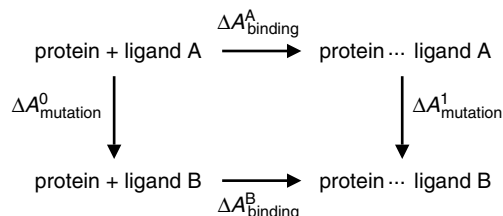
Fig. 2.9. The thermodynamic cycle used for the determination of protein–ligand relative binding free energies. Instead of carrying the horizontal transformations one can mutate the ligand in the free state – i.e., the left, vertical 'alchemical transformation', and in the bound state – i.e., the right, vertical 'alchemical transformation.' This yields the difference in the binding free energies: $\Delta A_{\mathrm{binding}}^{\mathrm{B}} - \Delta A_{\mathrm{binding}}^{\mathrm{A}} = \Delta A_{\mathrm{mutation}}^{1} - \Delta A_{\mathrm{mutation}}^{0}$

proper corrections for the loss of translational and rotational entropies are taken into account [23, 24]. For flexible solutes, corrections associated with the conformational degrees of freedom might also be required. Similar considerations apply to ligand creation.

In many instances we are not interested in determining just a single binding free energy, but rather the binding free energies of several different ligands. This is the case, for example, if we want to evaluate a host of potential inhibitors of an enzyme in the context of computer-aided drug design. Such evaluation can be, of course, handled by repeating the transformations of Fig. 2.8 for each ligand of interest. An alternative route, likely to be more efficient, is also available. It involves mutating ligand A into ligand B in both the bound and the free states. The corresponding thermodynamic cycle is shown in Fig. 2.9.

The basic FEP algorithm for ligand binding can be improved in several ways. One method is to use a nonphysical ligand that serves as the common reference state for a variety of ligands of interest [25]. This method, referred to as the one-step perturbation approach, appears to be quite successful even for complex and fairly diverse ligands [26].

2.8.4 The Single-Topology Paradigm

In this section, as well as the next, we shall discuss how in silico transformations can be carried out in practice. The stratification scheme, as described in Sect. 2.6, is almost always required. Two approaches have been devised for this purpose, in which the reference, target and all intermediate states are described by either a unique topology, or two separate topologies.

In the single-topology paradigm, a common topology is sought for the initial state and the final states of the 'alchemical transformation' [4, 27]. In practice, the most complex topology serves as the common denominator for both states, and the missing atoms are treated as vanishing particles, the nonbonded parameters of which are progressively set to zero as λ varies from 0 to 1. For instance, in the mutation of ethane

into methanol, the former serves as the common topology. As the carbon atom is transformed into oxygen, two hydrogen atoms of the methyl moiety are turned into noninteracting, *ghost* particles by annihilating their point charges and van der Waals parameters. A single-topology transformation is shown schematically in Fig. 2.10a. Modifications of the perturbation term in the Hamiltonian are represented as a linear combination of the relevant atomic and interatomic parameters as a function of λ

$$\begin{cases} q_i(\lambda) & = \lambda\, q_i^{(1)} + (1-\lambda)\, q_i^{(0)} \\[2mm] R_{ij}^*(\lambda) = \lambda\, R_{ij}^{*(1)} + (1-\lambda)\, R_{ij}^{*(0)} \\[2mm] \varepsilon_{ij}(\lambda) & = \lambda\, \varepsilon_{ij}^{(1)} + (1-\lambda)\, \varepsilon_{ij}^{(0)} \end{cases} \qquad (2.47)$$

Here, q_i stands for the net atomic charge borne by atom i. R_{ij}^* and ε_{ij} are the van der Waals parameters for the pair of atoms $\{i, j\}$ and the superscripts (1) and (0) refer, respectively, to the target and the reference states. Clearly, this is just one example of how modifications of the Hamiltonian can be handled. To optimize performance of FEP simulations, one might use several order parameters in
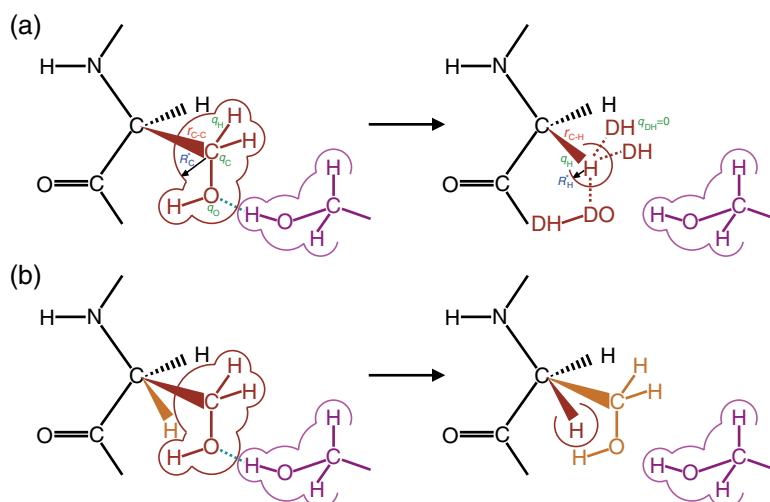
Fig. 2.10. Comparison of the single- (**a**) and the dual-topology (**b**) paradigms through the example of a serine-to-glycine point mutation. In the single-topology approach, all atoms of the side chain, except the β-carbon atom are annihilated by progressive cancelation of their charge and van der Waals parameters. The β-carbon atom is altered into an aliphatic hydrogen atom, bearing the appropriate net atomic charge. In the dual-topology paradigm, the side chain of serine and glycine always coexist, yet without 'seeing' each other through a list of excluded atoms. Interaction of the side chains with their environment is scaled progressively as λ varies between 0 and 1

lieu of a single λ parameter to transform the different nonbonded contributions of the potential energy function. It might also be preferable to use a different transformation of the parameters describing changes in the van der Waals energy. Additional terms in (2.47) might also be needed, should the potential energy function contain terms other than point-charge electrostatics and Lennard-Jones potentials.

It has been shown that, in the case of appearing or disappearing atoms, simultaneous modification of the electrostatic and the van der Waals terms in the potential energy function leads to numerical instabilities in the molecular dynamics trajectories, especially if the system contains vanishing atoms. When the van der Waals parameters of these atoms become quite small, for λ approaching 0 or 1, they can come extremely close to other particles in the system. Since the vanishing atoms still carry residual charges, the resulting nonbonded interactions often increase dramatically, which is incompatible with the basic idea of the perturbative approach. For this reason, a number of authors have opted for decoupling the mutation of the electrostatic and the van der Waals contributions, taking advantage of the fact that free energy is a state function, and, therefore, its value does not depend on the pathway chosen for its evaluation [28, 29]. This implies that the computational effort for estimating ΔA increases. Usually, this effort is smaller than it appears, because $P(\Delta U)$ for individual stages are narrower, and, consequently, easier to evaluate if electrostatic and van der Waals energies are modified individually rather then concomitantly.

An important aspect of the single-topology paradigm that has not been discussed so far is related to the modification of chemical bonds between the transformed atoms. The lengths and force constants of these bonds change during the transformation. In particular, if atoms disappear, their bonds with other atoms progressively shrink to zero length. In most 'alchemical transformations', the corresponding free energy contributions in the bound and in the free states are expected nearly to cancel out, providing that the affected chemical bonds are not strongly deformed by steric hindrances in the bound geometry. This explains why such bonded contributions were often neglected in free energy calculations [30, 31]. The same applies to contributions from bending planar angles. If needed, they can be calculated explicitly. In the simplest approach they can be approximated as the difference in the corresponding average energies. Assuming that the second-order perturbation theory is sufficiently accurate for these contributions, this is equivalent to taking an equal σ for the bound and the free states. A more accurate approach consists of carrying out FEP or thermodynamic integration (see Chap. 4) simulations to account for changes in the bond lengths and the valence angles concurrently with the other modifications of the Hamiltonian described in (2.47).

2.8.5 The Dual-Topology Paradigm

In sharp contrast with the single-topology paradigm, the topologies of the reference, 0, and the target, 1, states coexist in the dual-topology scheme throughout the

'alchemical transformation' [32–34]. This is shown in Fig. 2.10b. Using an exclusion list, atoms that are not common to 0 and 1 never interact during the simulation. Their intra- and intermolecular interaction with other atoms in the system are scaled by λ, which varies from 0 to 1. In the initial state, only topology 0 interacts with the rest of the system, whereas in the final state, only topology 1 does

$$U(\mathbf{x}; \lambda) = \lambda U_1(\mathbf{x}) + (1 - \lambda)U_0(\mathbf{x}), \tag{2.48}$$

where $U_0(\mathbf{x})$ and $U_1(\mathbf{x})$ are the potential energies characteristic of the reference and the target states. The same scaling of the Hamiltonian in (2.42) has already been introduced in Sect. 2.6.

This paradigm avoids several complications inherent to the single-topology approach. First, the problem of growing or shrinking chemical bonds is not present here. Second, decoupling the electrostatic and nonelectrostatic contributions during simulations is no longer needed. Unfortunately, the dual-topology approach also suffers from problems when λ approaches 0 or 1, which are often referred to as 'end-point catastrophes.' At these end points, interaction of the reference or the target topology with its environment is extremely weak, yet still nonzero, which in turn allows the surrounding atoms to clash against the appearing or vanishing chemical moieties. The resulting numerical instabilities cause large fluctuations in the estimated $\langle \Delta U \rangle$, which are attenuated only after extensive sampling. Additional difficulties arise if the mutated groups are flexible and more than one conformation needs to be taken into account in the reference and/or target state.

A number of strategies have been devised to circumvent such undesirable effects. The most trivial one consists in splitting the reaction pathway into windows of uneven widths, $\delta\lambda$, and using a large number of narrow windows when λ approaches 0 or 1. In essence, this is equivalent to adopting a nonlinear dependence of the interaction potential energy on the coupling parameter. It may be shown, however, that clashes between the appearing atoms and the rest of the system occur even for windows as narrow as $\delta\lambda \simeq 10^{-5}$. This difficulty may be overcome by modifying the parametrization of the van der Waals term in the potential energy function that governs the interaction of an appearing, or disappearing, atom, i, with an unaltered one, j [35]

$$U_{ij}^{\mathrm{vdW}}(r_{ij}; \lambda) = 4\,\epsilon_{ij}\,\lambda^n \left\{ \frac{1}{\left[\alpha_{\mathrm{vdW}}\,(1-\lambda)^2 + \left(\dfrac{r_{ij}}{\sigma_{ij}}\right)^6\right]^2} \right.$$
$$\left. - \frac{1}{\alpha_{\mathrm{vdW}}\,(1-\lambda)^2 + \left(\dfrac{r_{ij}}{\sigma_{ij}}\right)^6} \right\} \tag{2.49}$$

Here, α_{vdW} is a positive constant, and σ_{ij} and ϵ_{ij} are the usual Lennard-Jones parameters found in macromolecular force fields. The role played by the term $\alpha_{\mathrm{vdW}} (1 - \lambda)^2$ in the denominator is to eliminate the singularity of the van der Waals interaction. Introduction of this soft-core potential results in bounded derivatives of the potential energy function when λ tends towards 0.

2.8.6 Algorithm of an FEP Point-Mutation Calculation

In this section, we present a pseudocode for an FEP 'alchemical transformation' based on the dual-topology paradigm. The steps followed in this algorithm, specifically (c)–(f), may be implemented independently of the core of the program that generates an ensemble of configurations at a given λ state – either Monte Carlo or molecular dynamics. This is probably the simplest scheme, which may be improved in several ways, as will be discussed in Sect. 2.9.

(a) Build the topologies representative of state 0 and state 1, and establish an exclusion list to prevent atoms that are not common to 0 and 1 from interacting.

(b) Generate an ensemble of configurations that are representative of the reference state, λ.

(c) For each configuration, evaluate the potential energy using the reference-state Hamiltonian, $U(\mathbf{x}; \lambda = 0)$.

(d) Repeat the same calculations using the Hamiltonian of the target state.

(e) For each configuration, evaluate the potential energy difference using (2.48).

(f) Compute the ensemble average $\langle \exp\{-\beta[U(\mathbf{x}; \lambda + \Delta\lambda) - U(\mathbf{x}; \lambda)]\}\rangle_\lambda$, from which the free energy difference $\Delta A = A(\lambda + \delta\lambda) - A(\lambda)$ can be derived.

(g) Increment λ and go to stage (b).

2.9 Improving the Efficiency of FEP

From Sect. 2.8.6, it is clear that FEP calculations for many systems of practical interest are expected to be computationally very demanding. It is, thus, important to develop numerical techniques that allow us to apply the theory outlined so far in an efficient manner. If properly used, these techniques make calculations better in every sense – i.e., they improve both their accuracy and efficiency. It is, therefore, highly recommended that they be employed in practical applications of FEP. Chapter 6 is devoted entirely to this topic. Here, we only give the reader a preview of a few issues that will be covered in that chapter. In addition, we will discuss two other promising techniques that fall outside the conceptual framework developed in Chap. 6.

2.9.1 Combining Forward and Backward Transformations

At every stage of a stratified free energy calculation, one can easily perform perturbations to several states defined by different values of $\Delta\lambda$. This is because the only quantities that we need to evaluate are the differences in the potential energy between the states λ and $\lambda + \Delta\lambda$ determined for configurations sampled from the ensemble defined by λ. If we return to (2.42) or (2.48), we notice that for every value of $\Delta\lambda$, these differences can be obtained from the potential energy functions of the reference and the target states. If these values are stored in the course of the simulation, the appropriate potential energy differences can be conveniently calculated by means of post-simulation processing.

Computing ΔU for several different values of λ might be helpful in optimizing choices of intermediate states for stratification. The main power of this strategy, however, was realized in techniques that are aimed at improving the accuracy of free energy estimates through appropriate choices of positive and negative values of $\Delta\lambda$ or, equivalently, through combining forward and backward calculations. The simplest implementation of this idea is to calculate the free energy difference, $\Delta A_{i,i+1}$, between two consecutive stages i and $i+1$ in the forward and backward directions, starting from i or $i+1$ and using (2.8) and (2.9), respectively. The results of these two calculations may then be combined, for example by simple averaging. This procedure, however, has a serious drawback. In general, the accuracies of estimating $\Delta A_{i,i+1}$ from the forward and the backward simulations are not identical. In fact, it is common that they differ substantially, because the corresponding probability distributions, $P_i\left(\Delta U_{i,i+1}\right)$ and $P_{i+1}\left(\Delta U_{i+1,i}\right)$ have different widths. This means that combining free energy differences obtained from the forward and the backward calculations might turn out to be less accurate than the results of a calculation carried out along a single direction.

Another simple idea consists in performing a forward calculation from state i, corresponding to λ_i, to an additional intermediate at $\lambda_i + \Delta\lambda/2$, and a backward calculation from state $i+1$ – which corresponds to $\lambda_i + \Delta\lambda$ – to the same additional intermediate. The difference in the free energies obtained from these calculations is equal to $\Delta A_{i,i+1}$. Combining (2.8) and (2.9), we obtain

$$\Delta A_{i,i+1} = -\frac{1}{\beta} \ln \left[\frac{\langle \exp\left(-\beta\Delta U_{i,i+1}/2\right)\rangle_i}{\langle \exp\left(\ \beta\Delta U_{i,i+1}/2\right)\rangle_{i+1}} \right]. \tag{2.50}$$

In this equation, we used the fact that the potential energy difference between the states λ_i, or $\lambda_i + \Delta\lambda$, and state $\lambda_i + \Delta\lambda/2$ is equal to $\Delta U_{i,i+1}/2$, which is a consequence of the linear form of (2.42).

This approach is one of the oldest techniques for improving FEP calculations [36]. It is often called the simple overlap sampling (SOS) method and is usually markedly more accurate than simple averaging. It requires that one forward and one backward calculation be performed at every intermediate state. It is worth noting that no sampling is performed from the ensemble characterized by $\lambda_i + \Delta\lambda/2$, so that the number of stages is the same as in the pure forward, or backward calculation.

From what has been seen so far, it is obvious that the additional intermediate does not have to be located at $\lambda_i + \Delta\lambda/2$, but, instead, may be chosen at any value between λ_i and $\lambda_i + \Delta\lambda$. What we would like to do is to find the location of this intermediate that minimizes the statistical error of the calculated free energy difference, $\Delta A_{i,i+1}$. This problem was studied 30 years ago by Bennett [37]. As it turns out, it is equivalent to calculating $\Delta A_{i,i+1}$ from the formula

$$\exp(-\beta\Delta A_{i,i+1}) = \frac{\left\langle \{1 + \exp[\beta(\Delta U_{i,i+1} - C)]\}^{-1} \right\rangle_i}{\left\langle \{1 + \exp[-\beta(\Delta U_{i,i+1} - C)]\}^{-1} \right\rangle_{i+1}} \exp(-\beta C), \quad (2.51)$$

in which the constant, C, that determines the position of the additional intermediate is chosen such that

$$C = \Delta A_{i,i+1} + \frac{1}{\beta} \ln \frac{n_i}{n_{i+1}}. \quad (2.52)$$

Here, n_i and n_{i+1} are the sample sizes collected in the states i and $i + 1$. Equation (2.52) cannot be solved directly because it involves the unknown value of $\Delta A_{i,i+1}$. Instead, (2.51) and (2.52) may be solved iteratively during post-simulation processing.

A detailed presentation of the overlap sampling approach will be given in Sect. 6.6 of Chap. 6. In the present chapter, we merely note that applying this scheme, or any other similar technique that will be discussed extensively later on in the book, almost always improves the quality of the results. It is, therefore, highly recommended that they be routinely used in FEP calculations, perhaps in combination with other techniques.

2.9.2 Hamiltonian Hopping

It is not unusual that one encounters problems with quasi-nonergodicity along some segments of the transformation pathway from the reference to the target state. These problems can be solved, at least in part, by employing the Hamiltonian hopping technique. In essence, Hamiltonian hopping is just a special case of a general strategy, called parallel tempering. This method will be presented in detail in Chap. 8. Here, it also serves as an illustration of a recurring theme of this book – it is often advantageous to combine several strategies for free energy calculations in order to exploit the strength of each of them.

Hamiltonian hopping, as any other version of parallel tempering, is highly efficient if it is implemented on parallel computer architectures. In a stratified FEP calculation involving N states of the system, the simulations of the different λ states are carried out in parallel on separate processors. After a predefined number of steps, N_{sample}, $N/2$ swaps between two randomly chosen simulation cells are attempted [38]. This procedure is illustrated in Fig. 2.11. Acceptance of the proposed exchange between cells i and j is ruled by the following probability [39]:

$$\exp\left\{-\beta\left[\Delta\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x)^{i\to j} - \mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)^{i\to j}\right]\right\} \geq \mathrm{rand}[0;1], \quad (2.53)$$
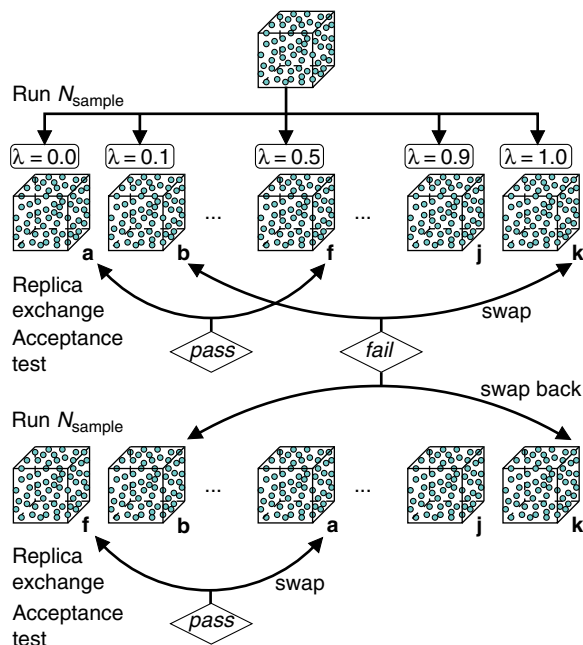
Fig. 2.11. Schematic representation of a parallel tempering simulation, in which the N states of a stratified FEP calculation are run concurrently. After a predefined number of steps, N_{sample}, the cells are swapped randomly across the different processors. A Metropolis-based acceptance criterion is used to determine which of the $N / 2$ exchanged λ-states should be accepted. Pairs of boxes that fail the test are swapped back. Then additional sampling is performed until the next exchange of the replicas

where $\Delta\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)^{i \to j} = \mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)^j - \mathscr{H}_1(\mathbf{x}, \mathbf{p}_x)^i$ denotes the change in the Hamiltonian representative of state $\mathbf{a}$ between simulation cells i and j. rand[0;1] represents a uniform random number generated from $0.0 \leq \mathrm{rand}[0;1] \leq 1.0$. If the random exchange for a given pair of λ states is rejected, the simulation cells are swapped back. N_{sample} steps are being performed again, until the next exchange.

Provided that the modeler has access to parallel architectures, this approach offers an advantage over sequential update of the interaction Hamiltonian as a function of the coupling parameter: for a given λ-state, swapping the simulation cells allows regions of phase space that might be separated by large barriers to be visited with appropriate statistical weights and over reasonable simulation times. This strategy may be valuable in quasi-nonergodic systems, in which satisfactory sampling of different conformational states is important for the accurate estimation of free energy differences.

2.9.3 Modeling Probability Distributions

Let us return to (2.12). As we have already discussed in Sect. 2.3, the probability distribution $P_0(\Delta U)$ is integrated in this equation with the Boltzmann weighting factor $\exp(-\beta\Delta U)$. This means that, especially for broad $P_0(\Delta U)$, the poorly sampled, negative-ΔU tail of the distribution provides the dominant contribution to the integral, whereas the contributions from the well-sampled region around the peak of $P_0(\Delta U)$ is small. Thus, the range of $P_0(\Delta U)$ which is known with a high accuracy is not useful for calculating ΔA, unless the perturbation is small and the corresponding probability distribution is sufficiently narrow.

It is only natural to consider ways that would allow us to use our knowledge of the whole distribution $P_0(\Delta U)$, rather than its low-ΔU tail only. The simplest strategy is to represent the probability distribution as an analytical function or a power-series expansion. This would necessarily involve adjustable parameters that could be determined primarily from our knowledge of the function in the well-sampled region. Once these parameters are known, we can evaluate the function over the whole domain of interest. In a way, this approach to modeling $P_0(\Delta U)$ constitutes an extrapolation strategy.

In general, this type of extrapolation is not very successful, because its reliability deteriorates, often rapidly, as we move away from the region in which the function is known with a good accuracy. In the particular case of $P_0(\Delta U)$, we might, however, be more successful, because this function is smooth and Gaussian-like. We shall exploit these features by considering three different representations of $P_0(\Delta U)$.

In fact, we have already used a modeling strategy when $P_0(\Delta U)$ was approximated as a Gaussian. This led to the second-order perturbation theory, which is only of limited accuracy. A simple extension of this approach is to represent $P_0(\Delta U)$ as a linear combination of n Gaussian functions, $p_i(\Delta U)$, with different mean values and variances [40]

$$P_0(\Delta U) = \sum_{i=1}^{n} c_i p_i(\Delta U) \tag{2.54}$$

where c_i is the weight of the ith Gaussian function, subject to the constraints $c_i \geq 0$, $\sum c_i = 1$. Then, using (2.12), (2.13), and (2.54), we obtain

$$\Delta A = -\frac{1}{\beta} \ln \sum_{i=1}^{n} c_i \exp(-\beta\langle\Delta U\rangle_i + \beta^2\sigma_i^2/2), \tag{2.55}$$

where $\langle\Delta U\rangle_i$ and σ_i are the mean and the variance of the ith Gaussian, respectively. In numerical tests it was shown that the free energy of hydrating a water molecule or an ion can be recovered accurately by taking $n = 6$ or 7, even for fairly large perturbations [40].

A different expansion relies on using Gram–Charlier polynomials, which are the products of Hermite polynomials and a Gaussian function [41]: These polynomials are particularly suitable for describing near-Gaussian functions. Even and odd terms of the expansion describe symmetric and asymmetric deformations of the Gaussian, respectively. To ensure that $P_0(\Delta U)$ remains positive for all values of ΔU, we take

$$P_0(\Delta U) = \left(\sum_{i=1}^{n} c_i \varphi_i(\Delta U) \right)^2, \qquad (2.56)$$

where c_i are coefficients of the expansion and ϕ_i are the normalized Gram–Charlier polynomials

$$\varphi_i(x) = \frac{1}{\sqrt{2^n \pi^{1/2} n!}} H_i(x) \exp\left(-x^2/2\right). \qquad (2.57)$$

In the last equation $H_i(x)$ is the ith Hermite polynomial. The reader may readily recognize that the functions ϕ_i look familiar. Indeed, these functions are identical to the wave functions for the different excitation levels of the quantum harmonic oscillator. Using the expansion (2.56), it is possible to express ΔA as a series, as has been done before for the cumulant expansion. To do so, one takes advantage of the linearization theorem for Hermite polynomials [42] and the fact that $\exp(-t^2 + 2tx)$ is the generating function for these polynomials. In practice, however, it is easier to carry out the integration in (2.12) numerically, using the representation of $P_0(\Delta U)$ given by expressions (2.56) and (2.57).

The expansion in (2.56) is complete and convergent. This means that any positive function normalized to unity can be represented in this form, and, for sufficiently large n, the absolute values of the coefficients in the expansion for $i > n$ will be smaller than any arbitrary small, predefined value, ϵ. This nice and formal property is, however, not particularly useful in practice because, by and large, only the first few coefficients in the expansion can be determined from simulations with sufficient accuracy. This means that (2.56), or any other expansion, is useful only if it converges quickly.

These considerations raise a question: how can we determine the optimal value of n and the coefficients $\{c_i\}$, $i \leq n$ in (2.54) and (2.56)? Clearly, if the expansion is truncated too early, some terms that contribute importantly to $P_0(\Delta U)$ will be lost. On the other hand, terms above some threshold carry no information, and, instead, only add statistical noise to the probability distribution. One solution to this problem is to use physical intuition [40]. Perhaps a better approach is that based on the *maximum likelihood* (ML) method, in which we determine the maximum number of terms supported by the provided information. For the expansion in (2.54), calculating the number of Gaussian functions, their mean values and variances using ML is a standard problem solved in many textbooks on Bayesian inference [43]. For the expansion in (2.56), the ML solution for n and $\{c_i\}$ also exists. Just like in the case of the multistate Gaussian model, this equation appears to improve the free energy estimates considerably when $P_0(\Delta U)$ is a broad function.

The two expansions discussed so far appear to be quite different. In the multistate Gaussian model, different functions are centered at different values of ΔU. In the Gram–Charlier expansion, all terms are centered at $\langle \Delta U \rangle_0$. The difference, however, is smaller that it appears. In fact, one can express a combination of Gaussian functions in the form of (2.56) taking advantage of the addition theorem for Hermite polynomials [44]. Similarly, another, previously proposed representation of $P_0(\Delta U)$ as a Γ function [45] can also be transformed into the more general form of (2.56).

The third model for $P_0(\Delta U)$ that we discuss here is sometimes called the 'universal' probability distribution function (UPDF) [46]. The origin of this name can be traced to the suggestion that the model describes well the statistical properties of global quantities in a broad class of finite-size, equilibrium or nonequilibrium systems characterized by strong correlations and self-similarity. It has been proposed that the UPDF applies even if a macroscopic system cannot be divided, on account of strong correlations, into statistically independent mesoscopic subsystems. For such systems, the underlying assumption of the central limit theorem is not satisfied, and, therefore, fluctuations are not expected to be Gaussian distributed. UPDF can be represented in the following form [46, 47]:

$$P(y) = K \exp\left\{a[b(y - s) - \exp\{b(y - s)\}]\right\}, \tag{2.58}$$

where a, b and the shift, s, are adjustable parameters, and K is a normalization constant. Here, $y = (\Delta U - \langle \Delta U \rangle_0)/\sigma$, so that the distribution has zero mean and a unit variance. UPDF was used with success to model $P_0(\Delta U)$ obtained from 'alchemical transformations' involving an anion, adenosine and a fatty acid [47].

Even though modeling the probability distributions has not been utilized nearly as extensively as some other techniques for increasing the efficiency of FEP calculations, it appears to be a highly promising area for further research. Initial applications of this technique lead to the conclusion that it can considerably improve estimates of ΔA, especially when the tails of $P_0(\Delta U)$ are poorly sampled. Furthermore, it can be readily combined with other techniques, such as forward and backward calculations [47]. The method also has drawbacks. The physical underpinnings for choosing a model distribution are, so far, not very strong, and estimates of the errors introduced by the procedure are presently not available.

2.10 Calculating Free Energy Contributions

The free energy difference associated with a given process is not the only thermodynamic quantity of interest. In many instances, one would also like to know the entropic and energetic, or enthalpic in the isobaric–isothermal ensemble, contributions to ΔA. This is because they may reveal something new about the nature of the process that is not necessarily apparent from knowledge of the free energy alone. Indeed, we often say that a process is enthalpy- or entropy-driven, or the barrier to a transformation is enthalpic or entropic. This information not only improves our understanding of the process, but it also provides clues about how to control it. For example, if ligand–enzyme interactions are primarily entropy-driven, then preorganization, which relies on making either the ligand or the active center more rigid, might be an effective strategy to enhance ligand–enzyme affinity.

Similarly, we often try to interpret changes in the free energy in terms of contributions to the potential energy function. For instance, one might want to know whether ΔA is primarily driven by electrostatic or van der Waals interactions. Alternatively, one might be interested in finding out what are the contributions to ΔA arising from

changes in the internal structure of the solute and the solute–solvent interactions, as well as in the reorganization of the solvent.

Addressing these issues from an FEP perspective is the main goal of this section. The main conclusions that we reach are, however, of a general nature and are independent of the method used for calculating the free energy.

2.10.1 Estimating Energies and Entropies

The three quantities of interest — the free energy difference, ΔA, the difference in potential energies, $\Delta U_{0\rightarrow1}$, and the entropy difference, $\Delta S_{0\rightarrow1}$, are connected through a basic thermodynamic relation

$$\Delta A = \Delta U_{0\rightarrow1} - T\Delta S_{0\rightarrow1}, \tag{2.59}$$

in which

$$\Delta U_{0\rightarrow1} = \langle U_1 \rangle_1 - \langle U_0 \rangle_0 \tag{2.60}$$

and

$$\Delta S_{0\rightarrow1} = \langle S_1 \rangle_1 - \langle S_0 \rangle_0. \tag{2.61}$$

Here, U_0 and U_1 are the potential energies of the reference and the target system, respectively. Note that $\Delta U_{0\rightarrow1}$ is not the same quantity as the previously used $\langle \Delta U \rangle_0$, because the latter is the average of energy differences between the target and the reference states taken over the reference ensemble.

It would appear at first glance that, once ΔA has been calculated, obtaining $\Delta U_{0\rightarrow1}$ and $\Delta S_{0\rightarrow1}$ with comparable accuracy should be a simple task. Unfortunately, this is not the case. The simplest approach would be to calculate $\langle U_0 \rangle_0$ and $\langle U_1 \rangle_1$ from simulations of the reference and the target state, respectively, and then to extract $\Delta S_{0\rightarrow1}$ from (2.59). This naive strategy is, however, not very successful. This is because the average total energies are usually large, approximately proportional to the number of particles in the system. Estimating a small quantity, $\Delta U_{0\rightarrow1}$, from a difference of two, independently measured, large numbers, subject to large fluctuations, is usually unreliable. It has been shown that the uncertainty in estimating $\Delta U_{0\rightarrow1}$ based on this approach can be one to two orders of magnitude larger than the uncertainty in the corresponding ΔA [48].

The difficulty in calculating a small number as a difference of two large ones, both known with limited accuracy, is not new and arises in many other fields. It often justifies using perturbation theory, which is aimed at estimating the quantity of interest directly. We will attempt to follow the same approach here. To do so, we combine (2.60) with (2.11) from Sect. 2.2, in which we substitute U_1 for $\mathscr{F}$. This yields

$$\Delta U_{0\rightarrow1} = \frac{\langle U_1 \exp\left(-\beta\Delta U\right) \rangle_0}{\langle \exp\left(-\beta\Delta U\right) \rangle_0} - \langle U_0 \rangle_0 \tag{2.62}$$

If this equation is used, ΔA and $\Delta U_{0\rightarrow1}$ can be calculated from a single simulation, or, as is usually the case, from a single series of stratified simulations. This is why it is sometimes called a 'single-state perturbation' method. Unfortunately, the problem

discussed in the context of (2.60) persists in this approach, i.e., $\Delta U_{0\to 1}$ is again obtained as a difference between two large numbers. This time, however, both numbers are obtained from the same simulation, which improves the accuracy because of partial error cancelation.

Another approach to calculating $\Delta S_{0\to 1}$ or $\Delta U_{0\to 1}$ relies on the classical thermodynamic relationships

$$\Delta S_{0\to 1} = -\left(\frac{\partial \Delta A}{\partial T}\right)_{N,V},\tag{2.63}$$

$$\Delta U_{0\to 1} = \left(\frac{\partial \beta \Delta A}{\partial \beta}\right)_{N,V}.\tag{2.64}$$

These relations may be utilized to calculate $\Delta S_{0\to 1}$ or $\Delta U_{0\to 1}$ in a finite-difference approximation

$$\Delta S_{0\to 1}(T) = -\frac{\Delta A(T+\Delta T) - \Delta A(T-\Delta T)}{2\Delta T}\tag{2.65}$$

$$\Delta U_{0\to 1} = \frac{\beta^+ \Delta A(\beta^+) - \beta^- \Delta A(\beta^-)}{2\Delta\beta},\tag{2.66}$$

where $\beta^+ = \beta + \Delta\beta$ and $\beta^- = \beta - \Delta\beta$. To estimate the energetic and entropic contributions employing this approach, it is required that ΔA be determined not only at the temperature T, but also at $T + \Delta T$ and $T - \Delta T$. As a result, at least three different series of simulations are needed. In a concrete application, ΔT should be properly chosen. If it is too large, then deviations from the linear dependence of ΔA on the temperature, implicitly assumed in the finite-difference method, become large. If it is too small, then statistical errors in evaluating the numerator overwhelm the calculation. It has been reported that, at least in some applications, 30–50 K represents a reasonable choice for ΔT [48]. The estimates can be further improved, at an additional cost, by using more points in the finite-difference formulas.

The finite-difference method can be combined with the perturbation technique that was previously used to derive the basic formulas in Sect. 2.2. This yields another single-state perturbation formula [49, 50]. Starting from (2.66), we get

$$\beta^+ \Delta A(\beta^+) = -\ln\langle\exp(-\beta^+\Delta U)\rangle_{0,\beta^+}$$

$$\equiv -\ln\frac{\displaystyle\int\int \exp(-\beta^+\Delta U)\exp(-\beta^+ U_0)\ \mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{p}_x}{\displaystyle\int\int \exp(-\beta^+ U_0)\ \mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{p}_x}$$

$$= -\ln\frac{\displaystyle\int\int \exp(-\beta^+\Delta U - \Delta\beta U_0 - \beta U_0)\ \mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{p}_x}{\displaystyle\int\int \exp(-\Delta\beta U_0 - \beta U_0)\ \mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{p}_x}$$

$$= -\ln\frac{\langle\exp(-\Delta\beta U_0 - \beta^+\Delta U)\rangle_0}{\langle\exp(-\Delta\beta U_0)\rangle_0}.\tag{2.67}$$

Similarly,

$$\beta^- \Delta A \left(\beta^- \right) = -\ln \frac{\langle \exp \left(\Delta \beta U_0 - \beta^- \Delta U \right) \rangle_0}{\langle \exp \left(\Delta \beta U_0 \right) \rangle_0}, \tag{2.68}$$

where $\langle \ldots \rangle_{0,\beta^+}$, $\langle \ldots \rangle_{0,\beta^-}$ and $\langle \ldots \rangle_0$ denote ensemble averages over the reference system at β^+, β^-, and β, respectively.

After inserting (2.67) and (2.68) to (2.66) we obtain

$$\Delta U_{0 \to 1} = \frac{1}{2\Delta\beta} \ln \left[\frac{\langle \exp \left(\Delta\beta U_0 - \beta^- \Delta U \right) \rangle_0 \langle \exp \left(-\Delta\beta U_0 \right) \rangle_0}{\langle \exp \left(-\Delta\beta U_0 - \beta^+ \Delta U \right) \rangle_0 \langle \exp \left(\Delta\beta U_0 \right) \rangle_0} \right]. \tag{2.69}$$

This is the desired formula, which requires only statistical averages over the reference system at temperature T. If, instead, we start from (2.65) and perform identical steps, we obtain a similar, single-state perturbation formula for $\Delta S_{0 \to 1}$. As it turns out, that formula, however, is more cumbersome to use than (2.69) but does not seem to offer any benefits in terms of accuracy.

One way to improve this approach it to carry out simultaneous perturbations for several, different values of ΔT or, equivalently, $\Delta\beta$, and then estimate the appropriate derivative through averaging or graphical interpolation. This can be done with only small additional computational effort.

At this point, the reader might ask: which of the three methods is the most accurate? Unfortunately, there is no clear-cut theoretical guidelines to answer this question, and empirical evidence is inconclusive. This has been discussed by Lu [50], where the reader can also find many references to earlier studies on extracting entropies and enthalpies from free energy calculations. In general, it appears that even in the simple case of the 'zero-sum' ethane $\to$ ethane alchemical transformation, the accuracy of determining $\Delta U_{0 \to 1}$ and $\Delta S_{0 \to 1}$ is inferior to the accuracy of ΔA. This is illustrated in Fig. 2.12. The rule of thumb appears to be that stratification is more effective for increasing the accuracy of the computed free energy components than the actual choice of the computational method. Additional considerations about the system of interest may also come into play. For instance, for systems close to phase transitions, such as some membrane systems, taking temperature derivatives might not be appropriate. An approach based on (2.62) would then be the method of choice.

2.10.2 How Relevant are Free Energy Contributions?

The total potential energy of the system can easily be divided into physically meaningful terms. Depending upon the problem of interest, one might wish to consider it, for instance, as a sum of electrostatic and van der Waals contributions, or as a sum of terms representing interactions within and between different subsystems. The average change in each component of the total potential energy upon the transformation of the system can be estimated using an expression analogous to (2.62). Unfortunately, a similar division is not possible for entropy, and, consequently, for free energy.

To understand better the difficulties connected with the partitioning of free energy into contributions, let us break down the potential energy into a sum of two terms,
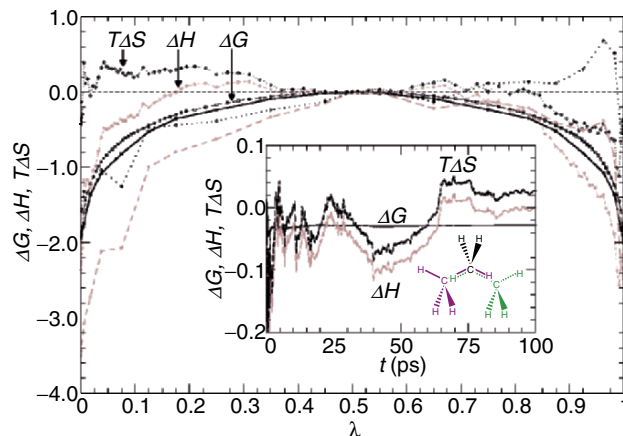
Fig. 2.12. Enthalpy, entropy, and free energy differences for the ethane → ethane 'zero-sum' alchemical transformation in water. The molecular dynamics simulations are similar to those described in Fig. (2.7). 120 windows (*thin lines*) and 32 windows (*thick lines*) of uneven widths were utilized to switch between the alternate topologies, with, respectively, 20 and 100 ps of equilibration and 100 and 500 ps of data collection, making a total of 14.4 and 19.2 ns. The enthalpy (*dashed lines*) and entropy (*dotted lines*) difference amount to, respectively, −0.1 and +1.1 kcal mol^{-1}, and −0.5 and +4.1 cal mol^{-1} K^{-1}. For comparison purposes, the free energy difference is equal to, respectively, +0.02 and −0.07 kcal mol^{-}1, significantly closer to the target value. *Inset*: Convergence of the different thermodynamic quantities

U_a and U_b. Assume further that the second-order perturbation theory applies. This means that $P_0(\Delta U)$ can be represented as a bivariate Gaussian. Then, ΔA, from (2.30), is given by

$$\Delta A = \langle \Delta U_a \rangle_0 + \langle \Delta U_b \rangle_0 - \frac{\beta}{2} \left(\sigma_a^2 + \sigma_b^2 + 2\rho\sigma_a\sigma_b \right), \qquad (2.70)$$

where

$$\sigma_a = \langle \Delta U_a^2 \rangle_0 - \langle \Delta U_a \rangle_0^2 \qquad (2.71)$$

is the variance of the probability distribution $P_0(\Delta U_a)$ and

$$\rho = \langle \left(\Delta U_a - \langle \Delta U_a \rangle_0 \right) \left(\Delta U_b - \langle \Delta U_b \rangle_0 \right) \rangle_0 / \sigma_b \sigma_b \qquad (2.72)$$

is the correlation coefficient for fluctuations in ΔU_a and ΔU_a.

From (2.70), it follows that the free energy cannot be divided simply into two terms, associated with the interactions of type a and type b. There are also coupling terms, which would vanish only if fluctuations in ΔU_a and ΔU_b were uncorrelated. One might expect that such a decoupling could be accomplished by carrying out the transformations that involve interactions of type a and type b separately. In Sect. 2.8.4, we have already discussed such a case for electrostatic and van der Waals interactions in the context of single-topology 'alchemical transformations.' Even then, however, correlations between these two types of interactions are not

eliminated. Let us consider a simple example of transforming a hydrated argon atom into a sodium ion. Even if charging is carried out separately from the modifications to the Lennard-Jones parameters, the van der Waals energy of the system changes in a way that is correlated with charging. In particular, as the charge borne by the solute increases, negatively charged oxygen atoms of water can approach it closer, which in turn causes not only the electrostatic, but also the van der Waals energies to change.

In addition, separation of the free energy into contributions depends on the path taken to transform the reference system into the target one. In other words, in contrast to ΔA, the contributions from the interactions of type a, type b and their coupling change if the transformation is performed along a different pathway. This is due to the fact that free energy is a state function of the system, but its contributions are not.

From this discussion, it would appear that decomposition of free energy into components may not be very helpful. This is, however, not necessarily so, especially if a physically meaningful path can be identified. For instance, dividing the free energy of dissolving a solute into contributions for creating a cavity in the solvent that is sufficiently large to accommodate the solute, and subsequently 'turning on' solute–solvent interactions have proven to be highly informative. Also, it is possible to obtain valuable results from free energy decomposition along a dynamical variable that closely approximates the reaction coordinate [51]. The formalism associated with this decomposition is based on thermodynamic integration and, therefore, will be discussed in Chap. 4. Another possible approach to make the partition of free energy useful is to find paths that minimize the correlation term in (2.70) [52].

2.11 Summary

Thermodynamic perturbation theory represents a powerful tool for evaluating free energy differences in complex molecular assemblies. Like any method, however, FEP has limitations of its own, and particular care should be taken not only when carrying out this type of statistical simulations, but also when interpreting their results. We summarize in a number of guidelines the important concepts and features of FEP calculations developed in this chapter:

(a) Formally, FEP is *exact* for any perturbation. In practice, however, even for moderately large perturbations, the method suffers from convergence issues. It is, therefore, recommended to use a stratification strategy by breaking the reaction pathway into a series of intermediate states through the introduction of an order, or 'coupling' parameter. The choice of the number of intermediate states in a staged FEP calculation should not be dictated by the corresponding change in free energy, but rather by the similarity between the reference and the target ensembles.

(b) Although the general FEP theory applies equally to both forward and backward simulations, the efficiencies of these two types of simulations may differ considerably. A properly converged FEP calculation for a $0 \rightarrow 1$ transformation does

not necessarily imply that the reverse, $1 \rightarrow 0$, transformation converges equally efficiently to the correct free energy difference.

(c) The free energy difference between the reference and the target states can be represented as a cumulant expansion. Retaining only the first two terms of this expansion is equivalent to assuming that $P_0(\Delta U)$ is a Gaussian. Second-order perturbation theory is a very useful tool for analyzing free energy calculations and developing approximate theories. Beyond the second order, however, the cumulant expansion diverges, and should, therefore, be used with extreme care.

(d) Since free energy is a state function, free energy differences are independent of the pathway chosen for their evaluation. Consequently, 'alchemical transformations,' during which a chemical species is mutated into an alternate one, may be carried out using either a single- or dual-topology paradigm by scaling the non-bonded parameters or the potential energy functions with the order parameter.

(e) Several techniques are available for improving the efficiency and accuracy of free energy calculations. These techniques require only very modest additional computational effort. Carrying out forward and backward simulations in an appropriate way is one of the more powerful schemes. It is strongly advised that these techniques be used in practice.

(f) The FEP methodology may be extended to the computation of potential energy, enthalpy and entropy differences. Yet, compared to free energy differences, these quantities are more difficult to estimate with good accuracy, because they inherently depend upon *all* molecular interactions in the system, and not only on those that are perturbed during the transformation.

(g) Particular attention should be paid to the interpretation of free energy components obtained by perturbing individual contributions of the potential energy function. These free energy components reflect the pathway defined for their determination, which is not unique.

References

1. Born, M., Volumen und Hydratationswarme der Ionen, *Z. Phys.* **1920**, *1*, 45–48
2. Kirkwood, J. G., Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **1935**, *3*, 300–313
3. Zwanzig, R. W., High-temperature equation of state by a perturbation method. I. Nonpolar gases, *J. Chem. Phys.* **1954**, *22*, 1420–1426
4. Jorgensen, W. L.; Ravimohan, C., Monte Carlo simulation of differences in free energies of hydration, *J. Chem. Phys.* **1985**, *83*, 3050–3054
5. Simonson, T.; Archontis, G.; Karplus, M., Free energy simulations come of age: protein–ligand recognition, *Acc. Chem. Res.* **2002**, *35*, 430–437

6. Shing, K. S.; Gubbins, K. E., The chemical potential in dense fluids and fluid mixtures via computer simulation, *Mol. Phys.* **1982**, *46*, 1109–1128

7. Kenney, J. F.; Keeping, E. S., *Mathematics of Statistics*, [2nd edition], Van Nostrand: Princeton, NJ, 1951

8. Hummer, G.; Pratt, L.; Garcia, A. E., Free energy of ionic hydration, *J. Phys. Chem.* **1996**, *100*, 1206–1215

9. Hünenberger, P. H.; McCammon, J. A., Ewald artifacts in computer simulations of ionic solvation and ion–ion interactions: a continuum electrostatics study, *J. Chem. Phys.* **1999**, *110*, 1856–1872

10. Markus, Y., *J. Chem. Soc. Faraday Trans.* **1991**, *87*, 2995–2999

11. Pohorille, A.; Chipot, C.; New, M.; Wilson, M. A. Molecular modeling of protocellular functions, in *Pacific Symposium on Biocomputing '96*, Hunter, L.; Klein, T. E., Eds. World Scientific: Singapore, 1996, pp. 550–569

12. Onsager, L., Electric moments of molecules in liquids, *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493

13. Pearlman, D. A.; Kollman, P. A., A new method for carrying out free energy perturbation calculations: dynamically modified windows, *J. Chem. Phys.* **1989**, *90*, 2460–2470

14. Lu, N.; Kofke, D. A., Accuracy of free-energy perturbation calculations in molecular simulation. I. Modeling, *J. Chem. Phys.* **2001**, *114*, 7303–7312

15. Kollman, P. A., Free energy calculations: Applications to chemical and biochemical phenomena, *Chem. Rev.* **1993**, *93*, 2395–2417

16. King, P. M. Free energy via molecular simulation: A primer. in *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J., Eds., vol. 2. ESCOM: Leiden, 1993, pp. 267–314

17. Kollman, P. A., Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules, *Acc. Chem. Res.* **1996**, *29*, 461–469

18. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K., NAMD2: Greater scalability for parallel molecular dynamics, *J. Comput. Phys.* **1999**, *151*, 283–312

19. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* **1983**, *79*, 926–935

20. Pohorille, A.; Benjamin, I., Structure and energetics of model amphiphilic molecules at the water liquid–vapor interface. A molecular dynamic study, *J. Phys. Chem.* **1993**, *97*, 2664–2670

21. Chipot, C.; Wilson, M. A.; Pohorille, A., Interactions of anesthetics with the water–hexane interface. A molecular dynamics study, *J. Phys. Chem. B* **1997**, *101*, 782–791

22. Silin, V.; Plant, A., Biotechnological applications of surface plasmon resonance, *Trends Biotechnol.* **1997**, *15*, 353–359

23. Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A., The statistical–thermodynamic basis for computation of binding affinities: a critical review, *Biophys. J.* **1997**, *72*, 1047–1069

24. Hermans, J.; Wang, L., Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme, *J. Am. Chem. Soc.* **1997**, *119*, 2707–2714

25. Liu, S. Y.; Mark, A. E.; van Gunsteren, W. F., Estimating the relative free energy of different molecular states with respect to a single reference state, *J. Phys. Chem.* **1996**, *100*, 9485–9494

26. Oostenbrink, C.; van Gunsteren, W. F., Free energies of ligand binding for structurally diverse compounds, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6750–6754

27. Pearlman, D. A., A comparison of alternative approaches to free energy calculations, *J. Phys. Chem.* **1994**, *98*, 1487–1493

28. Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A., Free energy calculations by computer simulation, *Science* **1987**, *236*, 564–568

29. Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A., Calculation of the relative change in binding free energy of a protein–inhibitor complex, *Science* **1987**, *235*, 574–576

30. Pearlman, D. A.; Kollman, P. A., The overlooked bond-stretching contribution in free energy perturbation calculations, *J. Chem. Phys.* **1991**, *94*, 4532–4545

31. Wang, L.; Hermans, J., Change of bond length in free-energy simulations: algorithmic improvements, but when is it necessary?, *J. Chem. Phys.* **1994**, *100*, 9129–9139

32. Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M., Hidden thermodynamics of mutant proteins: a molecular dynamics analysis, *Science* **1989**, *244*, 1069–1072

33. Boresch, S.; Karplus, M., The role of bonded terms in free energy simulations: I. Theoretical analysis, *J. Phys. Chem. A* **1999**, *103*, 103–118

34. Boresch, S.; Karplus, M., The role of bonded terms in free energy simulations: II. Calculation of their influence on free energy differences of solvation, *J. Phys. Chem. A* **1999**, *103*, 119–136

35. Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F., Avoiding singularities and neumerical instabilities in free energy calculations based on molecular simulations, *Chem. Phys. Lett.* **1994**, *222*, 529–539

36. Lee, C. Y.; Scott, H. L., The surface tension of water: A Monte Carlo calculation using an umbrella sampling algorithm, *J. Chem. Phys.* **1980**, *73*, 4591–4596

37. Bennett, C. H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comp. Phys.* **1976**, *22*, 245–268

38. Hansmann, U. H. E., Parallel tempering algorithm for conformational studies of biological molecules, *Chem. Phys. Lett.* **1997**, *281*, 140–150

39. Fukunishi, O. Watanabe; Takada, S., On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction, *J. Chem. Phys.* **2002**, *116*, 9058–9067

40. Hummer, G.; Pratt, L.; Garcia, A. E., Multistate Gaussian model for electrostatic solvation free energies, *J. Am. Chem. Soc.* **1997**, *119*, 8523–8527

41. Szegö, G., *Orthogonal Polynomials*, [4th edition], American Mathematical Society: Providence, 1975

42. Andrews, G. E.; Askey, R.; Roy, R., *Special Functions*, Cambridge University Press: Cambridge, 1999

43. Sivia, D. S., *Data Analysis. A Bayesian Tutorial*, Clarendon: Oxford, 1996

44. Sansone, G., *Orthogonal Functions*, Dover: New York, 1991

45. Amadei, A.; Apol, M. E. F.; Berendsen, H. J. C., The quasi-Gaussian entropy theory: free energy calculations based on the potential energy distribution function, *J. Chem. Phys.* **1996**, *104*, 1560–1574

46. Bramwell, S. T.; Christensen, K.; Fortin, J. Y.; Holdsworth, P. C. W.; Jensen, H. J.; Lise, S.; López, J. M.; Nicodemi, M.; Pinton, J. F.; Sellitto, M., Universal fluctuations in correlated systems, *Phys. Rev. Lett.* **2000**, *84*, 3744–3747

47. Nanda, H.; Lu, N.; Kofke, D. A., Using non-Gaussian density functional fits to improve relative free energy calculations, *J. Chem. Phys.* **2005**, *122*, 134110:1–8

48. Pearlman, D. A.; Rao, B. G. Free energy calculations: Methods and applications. in *Encyclopedia of computational chemistry*, Schleyer, P. v. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer III, H. F.; Schreiner, P. R., Eds., vol. 2. Wiley: Chichester, 1998, pp. 1036–1061

49. Fleischman, S. H.; Brooks III, C. L., Thermodynamics of aqueous solvation: solution properties of alchohols and alkanes, *J. Chem. Phys.* **1987**, *87*, 3029–3037

50. Lu, N.; Kofke, D. A.; Woolf, T. B., Staging is more important than perturbation method for computation of enthalpy and entropy changes in complex systems, *J. Phys. Chem. B* **2003**, *107*, 5598–5611

51. Hénin, J.; Pohorille, A.; Chipot, C., Insights into the recognition and association of transmembrane α-helices. The free energy of α-helix dimerization in glycophorin A., *J. Am. Chem. Soc.* **2005**, *127*, 8478–8484

52. Smith, P. E.; van Gunsteren, W. F., Predictions of free energy differences from a single simulation of the initial state, *J. Chem. Phys.* **1994**, *100*, 577–585

3

Methods Based on Probability Distributions and Histograms

M. Scott Shell, Athanassios Panagiotopoulos, and Andrew Pohorille

3.1 Introduction

One of the most powerful tools molecular simulation affords is that of measuring distribution functions and sampling probabilities. That is, we can easily measure the frequencies with which various macroscopic states of a system are visited at a given set of conditions – e.g., composition, temperature, density. We may, for example, be interested in the distribution of densities sampled by a liquid at fixed pressure or that of the end-to-end distance explored by a long polymer chain. Such investigations are concerned with *fluctuations* in the thermodynamic ensemble of interest, and are fundamentally connected with the underlying statistical–mechanical properties of a system.

In order to reconstruct probability distributions, we employ histograms in our simulation. These are simply arrays (or matrices), the indices of which correspond to variations in some parameter of interest, such as the number of particles, energy, distances, etc. For cases in which this variable is continuous, like the energy, we must be careful to discretize the histogram to sufficient resolution. During the course of the simulation, we then treat each *bin* of the histogram as an indicator of *counts*, or the number of visits to that state. Counts are added to the appropriate bins at either each step in the simulation or after a predetermined block of steps of specified length. Typically one requires a simulation with a length that is at least several correlation times of the system at hand in order to accrue data with good accuracy.

The utility of histograms stems from their rigorous connection to statistical mechanics, and the ability to extract from them fundamental thermodynamic potentials which can be used to predict properties of the system at conditions other than those of the original simulation. In this chapter, we provide an overview of the theory of histograms, their implementation in simulations, and free energy algorithms which make use of them. We begin by reviewing the statistical foundation of histograms and how they might be 'reweighted' in order to obtain information at multiple state conditions. We subsequently review a class of algorithms called flat-histogram methods, which utilize histograms to obtain entropies and free energies of model systems directly. Finally, we briefly comment on the use of histograms in extended ensemble

and reaction coordinate formulations. Several examples of the methods described here – as applied to phase equilibria – may be found in Chap. 10.

3.2 Histogram Reweighting

Early in the history of development of simulation methods it was realized that a single calculation can, in principle, be used to obtain information on the properties of a system for a range of state conditions [1–3]. However, the practical application of this concept was severely limited by the performance of computers available at the time. Many groups have now confirmed the usefulness of this concept, first in the context of simulations of spin systems [4–6] and later for continuous-space fluids [7–11]. In the following sections, we give a pedagogical review of histogram reweighting methods as applied to one-component systems.

Since many of the histogram-based methods we will outline in the following sections make use of the density of states, the reader may wish to review the material in Chap. 1 regarding its correct treatment in continuum systems. That chapter provides a brief discussion of the subtleties involved, although the reader who is already comfortable with the physical significance behind the density of states may wish to proceed directly.

3.2.1 Free Energies from Histograms

The connection between histogram measurements in simulation at a given temperature and the statistical mechanics of a system is given by the macrostate probability distribution. That is, when the partition function is expressed as a sum over macroscopic states of a system, the individual terms in the sum are proportional to the probability with which those states will be visited in the associated ensemble. A histogram taken from a simulation measuring the frequency with which macrostates are sampled therefore reflects the same probabilities. For example, the canonical partition function illustrates that the probability of observing a potential energy at a given temperature obeys:

$$\wp(U; N, V, T) = \frac{\mathrm{e}^{-\beta U} \Omega(N, V, U)}{Z(N, V, T)}, \tag{3.1}$$

where Z is the configurational integral introduced in (1.10) and normalization constant. A normalized histogram taken from an equilibrated canonical simulation provides an estimate of $\wp(U; N, V, T)$. Hereafter we will drop the dependence on N and V for simplicity.

Note that the expression in (3.1) is a continuous probability distribution in that $\wp(U; T)\mathrm{d}U$ gives the probability of macrostates with energy $U \pm \mathrm{d}U/2$. In an NVT simulation, we measure this distribution to a finite precision by employing a nonzero bin width ΔU. Letting $f(U)$ be the number of times an energy within the range $[U, U + \Delta U]$ is visited in the simulation, the normalized observed energy distribution is then

$$\boxed{\widetilde{\wp}(U;T) = \frac{f(U)}{\Delta U \sum_{U'} f(U')}} \tag{3.2}$$

where the tilde indicates a simulation estimate. For ergodic systems and in the joint limit of $\Delta U \to 0$ and $\sum f(U) \to \infty$, we will find that $\widetilde{\wp}(U) \to \wp(U)$. We clearly cannot take this limit rigorously in a simulation, but a reasonable bin width can usually be selected by a very short test run, while the total number of bin entries required (as dictated by the simulation duration) can depend more substantially on details of the molecular interactions and state point. Given a well-formed measurement $\widetilde{\wp}$, however, one can substitute this into (3.1) and rearrange to obtain an estimate of the density of states:

$$\widetilde{\Omega}(U) = \widetilde{\wp}(U;T_0) \mathrm{e}^{\beta_0 U} Z(T_0). \tag{3.3}$$

Notice that this equation allows us to calculate $\Omega(U)$ from a probability distribution measured from a simulation at temperature T_0. We do not know the value of $Z(T_0)$, but it is a constant independent of U. Furthermore, since Ω has no dependence on T, measurement of $\wp$ at any temperature should in principle permit its complete determination. In practice, however, the potential energies in a canonical simulation are sharply distributed around their average, away from which the statistical quality of $\widetilde{\wp}$ and hence $\widetilde{\Omega}$ in (3.3) becomes extremely poor.

Proceeding conceptually for a moment without these logistical difficulties, once we have determined the density of states we can calculate thermodynamic properties at any temperature of interest. The average potential energy is

$$\langle U \rangle (T) = \frac{\int U \mathrm{e}^{-\beta U} \widetilde{\Omega}(U) \mathrm{d}U}{\int \mathrm{e}^{-\beta U} \widetilde{\Omega}(U) \mathrm{d}U}, \tag{3.4}$$

or, substituting (3.3),

$$\langle U \rangle (T) = \frac{\int U \widetilde{\wp}(U;T_0) \mathrm{e}^{-(\beta-\beta_0)U} \mathrm{d}U}{\int \widetilde{\wp}(U;T_0) \mathrm{e}^{-(\beta-\beta_0)U} \mathrm{d}U}. \tag{3.5}$$

It is clear here that any multiplicative factor in the density of states, such as $Z(T_0)$ from (3.3), will not affect $\langle U \rangle$. Notice that in the case of $T = T_0$, this expression simply returns the original average energy at T_0, which we would have obtained through integrating the probability distribution. Beyond the energy, we might also wish to determine the configurational heat capacity, $kT^2 C_V(T) = \langle U^2 \rangle - \langle U \rangle^2$. When we allow additional fluctuating quantities in the simulation such as volume or number of particles, this approach also allows calculation of thermodynamic properties which depend on density derivatives of the entropy. The corresponding reweighting equations can be derived in exactly the same manner, replacing (3.3) with the NPT (fluctuating volume and energy) or grand-canonical (fluctuating particle number and energy) ensemble expressions for the probability distribution. This general technique, in which measured distributions at one set of state conditions are projected

onto another, is termed 'histogram reweighting' because the macrostate probabilities are adjusted along changes in the microstate weight factors.

A more general version of the canonical reweighting scheme in (3.5), in which the value of any order parameter ξ is reweighted to different temperatures, is given by:

$$\langle \xi \rangle (T) = \frac{\int \tilde{\xi}(U) \widetilde{\wp}(U; T_0) \mathrm{e}^{-(\beta - \beta_0)U} \mathrm{d}U}{\int \widetilde{\wp}(U; T_0) \mathrm{e}^{-(\beta - \beta_0)U} \mathrm{d}U}, \tag{3.6}$$

where $\tilde{\xi}(U)$ is the average value of ξ for a particular potential energy, a calculation which needs to occur alongside histogram collection during the simulation course. Compare this formula with equation (2.11). Formally, the expression for $\xi(U)$ is given by the microcanonical expression:

$$\xi(U') = \frac{\int_{V^N} \xi(\mathbf{q}) \delta[U(\mathbf{q}) - U'] \mathrm{d}\mathbf{q}}{\int_{V^N} \delta[U(\mathbf{q}) - U'] \mathrm{d}\mathbf{q}}, \tag{3.7}$$

where $\xi(\mathbf{q})$ is the value of ξ for a particular configuration $\mathbf{q}$. Such order parameters have already been defined in Sect. 2.8.1 and are discussed in more detail at the end of this chapter in Sect. 3.5.

As alluded to previously, numerical issues actually create a more complex situation than that just been described. For starters, the density of states is almost never calculated directly, as it typically spans many orders of magnitudes. This, in turn, would quickly overwhelm standard double-precision calculations in personal computers. This is easily remedied by working instead with the dimensionless entropy: $\mathscr{S} = \ln \Omega$, which for the purposes of this chapter will inherit all of the same notation used for the density of states in Chap. 1 – subscripts "tot," "ex," etc.

The more important issue concerns the statistics in the tails of the measured probability distributions. We clearly cannot get a good estimate for the density of states from data at energies which are rarely visited at our run temperature. The solution is to use *multiple* run temperatures to generate the estimate. The exact procedure will be presented more comprehensively later, but anticipating that discussion, let us first consider the situation when two separate Ω predictions are made from data collected at two temperatures T_1 and T_2. If the temperatures are close enough, there is a range of energies that are visited with sufficient frequency during *both* simulations to provide estimates of Ω from each of them. We know that the two density of states predictions must be the same, and so dividing (3.3) for T_1 by the same for T_2, taking the logarithm, and using $\beta A = -\ln Z$, we obtain

$$\beta_1 A(T_1) - \beta_2 A(T_2) = \ln \left[\frac{\widetilde{\wp}(U; T_1)}{\widetilde{\wp}(U; T_2)} \right] + (\beta_1 - \beta_2)U \tag{3.8}$$

which provides an estimate of the free energy difference between the system at temperatures 1 and 2. Notice that the right-hand side of this equation is a function of U.

In principle, any value of U should return the same free energy difference; however, we can choose the value of U so as to minimize the statistical error in the prediction. Assuming the histograms from both runs have the same total number of entries and that $\mathrm{var}(\widetilde{\wp}) \propto \widetilde{\wp}$ (i.e., the variance in a histogram bin is proportional to its number of entries), we find that the squared relative error in the argument of the logarithm in (3.8) is proportional to $f_1(U)^{-1} + f_2(U)^{-1}$, where f_1 and f_2 are the numbers of counts of energy U in the histograms corresponding to T_1 and T_2, respectively. Therefore, the optimal U in (3.8) will minimize this expression. These kinds of arguments involving propagation of statistical error were used by Ferrenberg and Swendsen to create a comprehensive methodology for extracting accurate free energy estimates from multiple simulation runs [12], described next.

3.2.2 Ferrenberg–Swendsen Reweighting and WHAM

In the earlier sections, we saw that the density of states could be reconstructed to a multiplicative constant from probability distributions measured during canonical simulations. This enables, in principle, the application of reweighting techniques to predict ensemble averages at conditions other than those of the original simulation. However, the statistical errors at the tails of the measured distribution from a single run limit the practical application of this approach. We hinted that the way around this problem was to use data from multiple runs and temperatures to reconstruct the density of states. The optimal procedure for incorporating multiple run results in this way, which has since become standard practice, was initially developed in 1989 by Ferrenberg and Swendsen [12] and later generalized by Kumar et al. [13]. The approach is often called Ferrenberg–Swendsen reweighting, multiple histogram reweighting, or the weighted histogram analysis method (WHAM). The basic idea proposed by these authors is that the contribution of each run to a reweighting estimate should be weighed based on the magnitude of errors in their histograms. That is, runs that have greater overlap with the reweighting conditions should contribute more to the estimation of property averages.

A brief sketch of the derivation of the WHAM equations follows; we note that a detailed explanation is available in the book by Frenkel and Smit [14]. Consider the canonical reweighing (3.5). Our goal will be to combine the histograms $\widetilde{\wp}_i(U)$ from several runs at different temperatures T_i to predict the distribution of potential energies at a new temperature T. Individually, each run enables us to reweight its histogram to obtain the distribution at T

$$
\begin{aligned}
\widetilde{\wp}_{\mathrm{r},i}(U;T) &= \frac{\widetilde{\wp}(U;T_i)\mathrm{e}^{-(\beta-\beta_i)U}}{\displaystyle\int \widetilde{\wp}(U;T_0)\mathrm{e}^{-(\beta-\beta_i)U}\mathrm{d}U} \\
&= \frac{Z_i}{Z}\widetilde{\wp}(U;T_i)\mathrm{e}^{-(\beta-\beta_i)U}
\end{aligned}
\tag{3.9}
$$

with $Z_i = Z(T_i)$ and $Z = Z(T)$. The subscript "r" simply indicates that $\widetilde{\wp}$ has been reweighted from a measured $\widetilde{\wp}$. The ratio of the partition functions in the last line of

this equation imparts a multiplicative constant to the probability distribution that is related to the free energy difference between the run temperature T_i and the target temperature T. Practically speaking, we could simply normalize the probabilities to determine this constant, but for this discussion it will be instructive to leave the partition functions explicit. Recall that each run provides an estimate of $\wp(U;T)$ according to (3.9). Ferrenberg and Swendsen proposed to combine all of these estimates linearly using a weighting factor w that depends both on the energy and the run number

$$\widetilde{\wp}_{\mathrm{r}}(U;T) = \sum_i w_i(U)\widetilde{\wp}_{\mathrm{r},i}(U;T)$$

$$= \sum_i w_i(U)\frac{Z_i}{Z}\widetilde{\wp}(U;T_i)\mathrm{e}^{-(\beta-\beta_i)U}. \tag{3.10}$$

The weighting function $w_i(U)$ is normalized such that the sum total contribution to $\widetilde{\wp}_{\mathrm{r}}$ from each run at a given U is equal to one: $\sum_i w_i(U) = 1$. Within this constraint, we determine the optimal functional form of w by minimizing the statistical error in the predicted $\widetilde{\wp}_{\mathrm{r}}$. Note that the variance in histogram measurements changes as the number of counts in a bin, $f(U)$, divided by the total number of counts, $f_{\mathrm{tot}} = \sum_k f(U_k)$, squared. This translates to a variance in the measured distributions $\widetilde{\wp}(U;T_i)$ that is dependent both on the energy and the run number. Using standard error propagation rules, we can predict the expected variance in the composite $\widetilde{\wp}_{\mathrm{r}}$ at the target temperature. Hence, we minimize this error by optimizing the weight function w, using Lagrange multipliers to ensure normalization of w. Here, we simply present the final result when the optimal w is substituted

$$\widetilde{\wp}_{\mathrm{r}}^*(U;T) = \frac{\sum_i f_i(U)\exp(-\beta U)}{\sum_i f_{\mathrm{tot},i}\exp(\beta_i\widetilde{A}_i - \beta_i U)}, \tag{3.11a}$$

$$\exp(-\beta_i\widetilde{A}_i) = \sum_U \widetilde{\wp}_{\mathrm{r}}^*(U;T_i), \tag{3.11b}$$

$$\widetilde{\wp}_{\mathrm{r}}(U;T) = \frac{\widetilde{\wp}_{\mathrm{r}}^*(U;T)}{\sum_U \widetilde{\wp}_{\mathrm{r}}^*(U;T)}. \tag{3.11c}$$

Here, $\widetilde{\wp}_{\mathrm{r}}^*(U;T)$ is the un-normalized probability distribution, $\widetilde{A}_i$ gives the free energy for run i, $f_i(U)$ is the number of counts of energy U for run i, and $f_{\mathrm{tot},i}$ is the total number of counts in run i. The values for $\widetilde{A}_i$ are solutions to the set of (3.11a) and (3.11b), and are usually solved by iterating between the two; initial values are often taken to be $\widetilde{A}_i = 0$ and iteration between (3.11a) and (3.11b) proceeds until the values no longer change significantly. Actually, it is only possible to determine relative values for the free energies using these equations, so typically one run is chosen to be the reference state for which $\widetilde{A}$ is always set to zero. When the free energies

have converged, the final probability distribution is given through normalization via (3.11c). This distribution can then be used directly to calculate various moments of the potential energy.

We can make several observations about the WHAM equations. First, the optimal weights depend on the target reweighting temperature through (3.11a) only. This means that, once the free energies $\widetilde{A}_i$ have been determined, they do not need to be recalculated upon reapplication of (3.11b) at different temperatures, meaning the iteration procedure only needs to happen once. Second, the optimal weights entail the free energies of each run. This has a practical benefit, as it allows us to extract free energy estimates using WHAM, but in addition it implies a deep connection between free energies and optimal ensemble overlap. In fact, there is a close connection between WHAM and Bennett's method [4] – which yields the optimal free energy estimate between two ensembles – as discussed in Chaps. 5 and 6 and in [14].

The WHAM equations can be generalized to many simulation settings. In particular, it is fairly straightforward to adapt them to other ensembles. Chapter 10, for instance, demonstrates the use of the reweighting equations in the grand canonical ensemble, where the calculated probability distribution is a function of N in addition to U. The general derivation of the WHAM equations in [13] allows runs to differ not only in temperature, but in the potential energy function as well, and it permits reweighting of general order parameters as (3.6). Here, one assumes that the energy function of each run can be expressed as a set of values λ_j that are coefficients to component energy functions V_j in a master potential energy expression $U = \sum_j \lambda_j V_j$. Thus, for one of the individual runs, the energy function is expressed as: $U_i = \sum_j \lambda_{i,j} V_j$. For this case, the WHAM equations become

$$\widetilde{\wp}_\mathrm{r}^*(\mathbf{V}, \xi; T) = \frac{\sum_i f_i(\mathbf{V}, \xi) \exp\left(-\beta \sum_j \lambda_j V_j\right)}{\sum_i f_{\mathrm{tot},i} \exp\left(\beta_i \widetilde{A}_i - \beta_i \sum_j \lambda_{i,j} V_j\right)}, \qquad (3.12a)$$

$$\exp(-\beta_i \widetilde{A}_i) = \sum_U \widetilde{\wp}_\mathrm{r}^*(\mathbf{V}, \xi; T), \qquad (3.12b)$$

$$\widetilde{\wp}_\mathrm{r}(\mathbf{V}, \xi; T) = \frac{\widetilde{\wp}_\mathrm{r}^*(\mathbf{V}, \xi; T)}{\sum_U \widetilde{\wp}_\mathrm{r}^*(\mathbf{V}, \xi; T)}. \qquad (3.12c)$$

Here, $\mathbf{V}$ is vector notation for the set of all component energies V_j, and $\lambda_{i,j}$ gives the coefficient of V_j in the ith run. The λ_j, without subscript i, indicate the values of λ in the target ensemble. The histograms collected in the runs are multidimensional in that they are tabulated as functions of the component energies as well as the order parameter ξ. Similarly, the final result of the WHAM calculation is a multidimensional probability distribution in V_j and ξ.

3.3 Basic Stratification and Importance Sampling

In this section, we describe briefly the most basic applications of two fundamental enhanced sampling techniques – stratification and importance sampling – to probability density and histogram-based methods for free energy calculations. We cover such popular approaches as the 'windows' method, non-Boltzmann sampling, and umbrella sampling, all which are routinely combined with histogram reweighting and WHAM. Their popularity is well deserved – they are not only among the oldest, but also among the most successful methods for improving the efficiency of free energy calculations. For this reason, these methods have been described numerous times before, and are probably well known to readers who have even rudimentary knowledge of the field.

The emphasis of this section is slightly different from that in the rest of this chapter. Instead of focusing on the density of states, we will consider probability density functions and we will not specifically discuss the methods in the context of Monte Carlo simulations. Also, the applications that we have in mind here are slightly different. We are particularly interested in problems in which we want to know not only the free energy difference between the initial and final states of the system, but also the free energy change along an order parameter. This allows us to estimate the relative stabilities of different local free energy minima and the magnitude of the barriers that separate them. Typical problems of this sort are conformational transitions associated with rotation around flexible bonds, for example in macromolecules. In some instances, the free energy between the initial and final states is known and the real focus of a calculation is on the shape of the free energy in between. This is the case, for example, in assisted or protein-mediated ion transport across membranes. In the absence of transmembrane electrical field and ionic gradients, the free energy difference between the ion on both sides of the membrane is zero. The interest is in determining the free energy profile associated with the transport, as it reveals valuable information about the mechanism and kinetics of this process.

Even though we take a specific perspective on the topics addressed in this section, we will make an effort not to obscure two important points. First, the methods described here are highly versatile and can be used for many different problems. Second, there are deep conceptual connections between the material covered in this section and the methods described in the remainder of this chapter and in many other chapters.

3.3.1 Stratification

Assume that we are interested in how the free energy of a system changes as a function of an order parameter, ξ, which changes between ξ_0 and ξ_1. A direct approach to this problem is to carry out MD or MC simulations long enough to obtain a sufficiently accurate estimate of the probability density function, $\mathscr{P}(\xi)$, of finding the system in a state corresponding to ξ. Then, it follows from (1.22) of

Chap. 1 that the free energy difference, $\Delta A(\xi)$, between the states described by ξ and ξ_0 is

$$\Delta A\left(\xi\right) = A\left(\xi\right) - A\left(\xi_0\right) = -\beta^{-1} \ln \frac{\wp\left(\xi\right)}{\wp\left(\xi_0\right)}. \tag{3.13}$$

In practice, the continuous function $\wp(\xi)$ is represented as a histogram consisting of M bins. If all bins have equal size $\Delta\xi = (\xi_1 - \xi_0)/M$ then

$$\wp\left[\xi_0 + (i - 0.5)\,\xi\right] = \frac{f_i}{\displaystyle\sum_j f_j} \tag{3.14}$$

where f_i is the number of sampled configurations for which the order parameter takes a value between $\xi_0 + (i - 1)\,\xi$ and $\xi_0 + i\xi$.

Combining (3.13) and (3.14) leads to a formula for histogram-based estimates of $\wp\left(\xi\right)$

$$\Delta A\left(\xi_0 + (i - 0.5)\,\xi\right) = -\beta^{-1} \ln \frac{f_i}{f_1}. \tag{3.15}$$

In practice, this simple formula will hardly ever work, especially if the free energy changes appreciably with ξ. Consider, for example, two states of the systems, ξ_i and ξ_j such that $\Delta A(\xi_i) - \Delta A(\xi_j) = 5k_\mathrm{B}T$. Then, on average, the former state is sampled only seven times for every 1,000 configurations sampled from the latter state. Such nonuniform sampling is undesirable, as it leads to a considerable loss of statistical accuracy. For the free energy profile shown in Fig. 3.1, transitions between
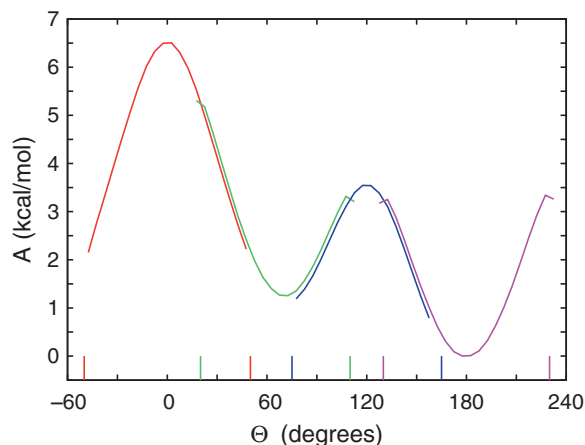
Fig. 3.1. Free energy of isomerization of butane as a function of the C–C–C–C torsional angle, Θ. Because of high energy barriers, transitions between stable *trans* and *gauche* rotamers are rare, which makes calculation of the free energy in a single simulation highly inefficient. Instead, the calculation was performed in four overlapping windows, whose edges are marked on the x-axis. In each window, the probability density functions and the free energies were determined as functions of Θ. They were subsequently shifted so that they matched in the overlapping regions, yielding the free energy profile in the full range of Θ

stable regions are rare because they require traversing a free energy barrier. This impedes equilibration along ξ, which in turn yields poor estimates of the free energy in the whole range of the order parameter.

The problem described earlier is perfectly suited for applying stratification, outlined in Sect. 1.4. The full range of ξ is divided into strata, or 'windows', for which separate simulations are performed. The system can be restrained to remain mostly within a window by adding to the potential energy function an extra term that depends only on ξ and is equal to zero inside the window, but increases, for example harmonically, as the system moves beyond the edge of the window. In Monte Carlo simulations, it is also possible to reject outright those moves which take the system outside of the current window, which corresponds to an added energy that is infinite outside of the window bounds. Clearly, $\wp(\xi)$ within each window changes less than in the whole range of ξ, which leads to a more uniform sampling of ξ and improved efficiency of free-energy calculations.

For each window, $\wp(\xi)$ is estimated by using the exact analog of (3.14). However, reconstruction of the full probability distribution directly is not possible because the total normalization constant is not known. Instead, we exploit the fact that $\wp(\xi)$ (or, equivalently, the free energy) is a continuous function of ξ. If consecutive windows overlap one can build the complete probability distribution by matching $\wp(\xi)$ in the overlapping regions, as illustrated in Fig. 3.1. How to do this in a systematic fashion will be discussed later in this section.

From the discussion so far, it might appear that stratification is advantageous only if the free energy changes as a function of ξ. This is, however, not so. Stratification improves efficiency even if the free energy is constant and the motion along ξ is strictly diffusive. If the full range of the order parameter is divided into L windows of equal size, the computer time needed to acquire the desired statistics in each window, τ_{w}, is proportional to the characteristic time of diffusion within a window

$$\tau_{\mathrm{w}} \propto \frac{\left[\left(\xi_1 - \xi_0\right)/L\right]^2}{D_\xi}. \tag{3.16}$$

Then, if we neglect overlaps between consecutive windows the total computer time, τ, is

$$\tau = L\tau_{\mathrm{w}} \propto \frac{\left(\xi_1 - \xi_0\right)^2}{LD_\xi}. \tag{3.17}$$

This means that τ decreases linearly with the number of windows, at least for large windows. For small windows, other factors can influence τ. First, the statistics accumulated in small windows are more correlated than the statistics obtained from larger windows. Second, for any widow size, τ_{w} must be longer than the time needed to equilibrate the system along the degrees of freedom orthogonal to ξ. For this reason, at some point reduction of window size no longer reduces τ. Finally, for some systems, motions along ξ and along orthogonal degrees of freedom are highly correlated. If windows become too small, motion in the ξ direction is so restricted that equilibration along other degrees of freedom becomes severely impeded, causing quasi-nonergodicity. We will return to these issues in Chaps. 4 and 14.

3.3.2 Importance Sampling

As outlined in Sect. 1.4, a powerful strategy to improve the efficiency of free energy calculations is based on modifying the underlying sampling probability such that important regions in phase space are visited more frequently. Not surprisingly, this method is called importance sampling. To see how it works we continue to consider an example of a system that is transformed along an order parameter ξ. The conventional Boltzmann distribution for this system, $\wp(\xi)$, is given by:

$$\wp(\xi) = \frac{\int \exp\left[-\beta U(\Gamma)\right] \delta\left(\xi - \hat{\xi}\right) d\Gamma}{\int \exp\left[-\beta U(\Gamma)\right] d\Gamma},$$
(3.18)

where $d\Gamma$ denotes integration over all degrees of freedom in the system and $\hat{\xi}$ can be either the value of ξ for the coordinates Γ or one of the coordinates themselves, depending on the nature of the order parameter. Contrary to the Boltzmann scheme, we wish to sample from a modified distribution, $\wp'(\xi)$, defined as

$$\wp'(\xi) = \frac{\int \omega(\hat{\xi}) \exp\left[-\beta U(\Gamma)\right] \delta\left(\xi - \hat{\xi}\right) d\Gamma}{\int \omega(\hat{\xi}) \exp\left[-\beta U(\Gamma)\right] d\Gamma}$$

$$= \omega(\xi) \frac{\int \exp\left[-\beta U(\Gamma)\right] \delta\left(\xi - \hat{\xi}\right) d\Gamma}{\int \omega(\hat{\xi}) \exp\left[-\beta U(\Gamma)\right] d\Gamma},$$
(3.19)

where $\omega(\xi)$ is a positive function of ξ. In the second line, the $\omega(\hat{\xi})$ term in the numerator can be taken outside of the integral and switched to $\omega(\xi)$ due to the presence of the delta function. Because $\wp'(\xi)$ differs from the correct Boltzmann distribution for the system, the methods based on sampling $\wp'(\xi)$ are often called non-Boltzmann sampling techniques.

If we combine (3.13), (3.18), and (3.19), we get

$$\Delta A(\xi) = -\beta^{-1} \ln \frac{\omega(\xi_0)\, \wp'(\xi)}{\omega(\xi)\, \wp'(\xi_0)}$$

$$= -\beta^{-1} \left[\ln \frac{\wp'(\xi)}{\wp'(\xi_0)} + \ln \omega(\xi_0) - \ln \omega(\xi) \right].$$
(3.20)

Consider the free energy difference, $\Delta A'(\xi)$, obtained from sampling the system with the non-Boltzmann probability density $\wp'(\xi)$, given by:

$$\Delta A'(\xi) = -\beta^{-1} \ln \frac{\wp'(\xi)}{\wp'(\xi_0)}.$$
(3.21)

Substituting this expression into (3.20), we obtain the important result

$$\Delta A\left(\xi\right) = \Delta A'\left(\xi\right) + \beta^{-1}\left[\ln\omega\left(\xi\right) - \ln\omega\left(\xi_0\right)\right]. \tag{3.22}$$

This formula provides a prescription for recovering the free energy of interest from a simulation carried out using non-Boltzmann sampling. Note that the left-hand side corresponds to the usual Boltzmann-weighted free energy difference, while the right-hand side contains the sampling probabilities ($\wp'$) and measurements ($\Delta A'$) from a non-Boltzmann simulation. A slightly different and more frequent casting of this formula can be obtained by rewriting ω as

$$\omega\left(\xi\right) = \exp\left[-\beta V(\xi)\right]. \tag{3.23}$$

Since ω is always positive this can be done without loss of generality. We also define

$$U'(\xi) = U(\xi) + V(\xi). \tag{3.24}$$

Then, after substituting (3.23) and (3.24) in (3.19), we obtain

$$\wp'\left(\xi\right) = \frac{\displaystyle\int \exp\left[-\beta U'\left(\Gamma\right)\right]\delta\left(\xi - \hat{\xi}\right)d\Gamma}{\displaystyle\int \exp\left[-\beta\,U'\left(\Gamma\right)\right]d\Gamma}, \tag{3.25}$$

and $\Delta A\left(\xi\right)$ can be expressed as

$$\boxed{\Delta A\left(\xi\right) = \Delta A'\left(\xi\right) + \left[V\left(\xi\right) - V\left(\xi_0\right)\right].} \tag{3.26}$$

The last two equations mean that $\wp'(\xi)$ can be interpreted as the Boltzmann distribution for a system in which the original potential energy function has been modified by a 'biasing' potential $V(\xi)$. They also provide clues how to choose ω or, equivalently, V. The desired choice is such that $V(\xi)$ 'counterbalances' (is the negative of) $\Delta A(\xi)$. This would make $\Delta A'(\xi)$, and consequently $\wp'(\xi)$, in a biased simulation as uniform as possible. Then, in most cases, calculations would approach optimal efficiency. Of course, perfect uniformity is difficult to achieve because this implies that the free energy profile along ξ is accurately anticipated before the simulation. In many instances, however, a reasonable guess can be made based on our understanding of the problem at hand. If this guess is correct, the benefits from increased efficiency could be substantial, as shown in Fig. 3.2. However, if the guessed $V(\xi)$ does not provide a good estimate of the negative of ΔA, the gains from applying a biasing potential will be negligible and can even lead to a loss of efficiency. Such a situation is shown in the bottom panel of Fig. 3.2. This occurs most often for new and challenging problems for which we lack good intuition about the shape of $\Delta A(\xi)$.

Despite this deficiency, importance sampling is a very powerful and versatile technique. It can be used with different types of order parameters, including those which describe an actual Hamiltonian coordinate of the system and those which
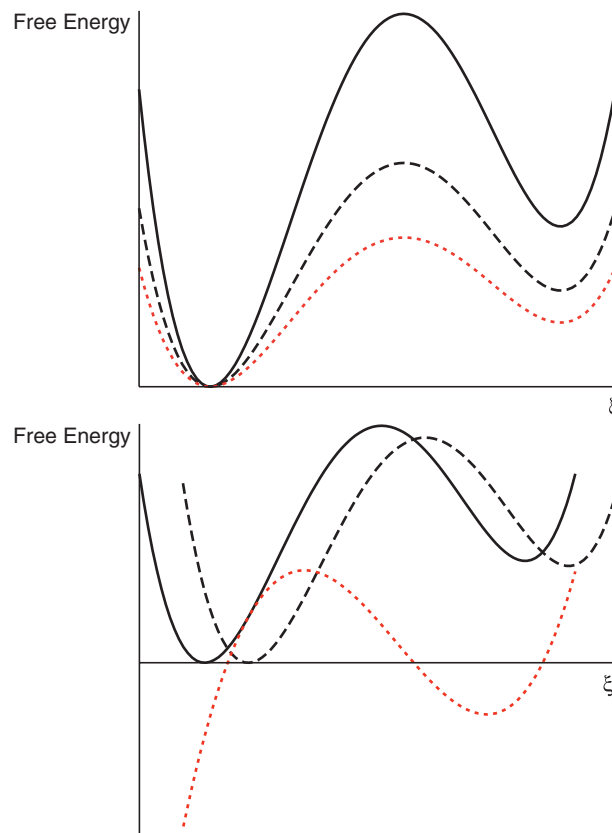
Fig. 3.2. The unbiased free energy $\Delta A(\xi)$ (*solid line*), the biasing potential $V(\xi)$ (*dashed line*), and the free energy from the importance sampling (biased) simulation $\Delta A'(\xi)$ (*dotted line*) as functions of the order parameter ξ. $\Delta A'(\xi)$ was obtained by subtracting $V(\xi)$ from $\Delta A(\xi)$. In the top panel, the biasing potential was well chosen because $\Delta A'(\xi)$ is more uniform than $\Delta A(\xi)$. In the bottom panel, this is not the case and determining the free energy from the importance sampling simulation is not expected to be more efficient than from the unbiased simulation. Note that $V(\xi)$ in this case has exactly the same shape as $\Delta A(\xi)$, but a slightly wrong guess was made regarding its position along ξ. If $V(\xi)$ and $\Delta A(\xi)$ were aligned the resulting free energy profile would be flat

are calculated parameters. It is equally easy to apply in both molecular dynamics and Monte Carlo simulations, and it can be seamlessly combined with stratification simply by applying separate biasing potentials to different windows. Finally, it is compatible with all general methods for free energy calculations. Not surprisingly, it is used in some form in almost every chapter of this book. In fact, we could have introduced it in the previous chapter, in that one of its first and most influential applications was in the context of free energy perturbation (FEP) [15].

The connection between importance sampling and FEP can be easily established by noting that free energy difference between the reference and the target system in (2.8) can be rewritten as

$$\langle \exp\left(-\beta \Delta U\right) \rangle_0 = \frac{\int \exp\left(-\beta \Delta U\right) \exp\left(-\beta U_0\right) \mathrm{d}\mathbf{x}}{\int \exp\left(-\beta U_0\right) \mathrm{d}\mathbf{x}} \qquad (3.27)$$

$$= \frac{\int \exp\left(-\beta \Delta U\right) w^{-1} w \exp\left(-\beta U_0\right) \mathrm{d}\mathbf{x}}{\int w^{-1} w \exp\left(-\beta U_0\right) \mathrm{d}\mathbf{x}}$$

$$= \frac{\left\langle w^{-1} \exp\left(-\beta \Delta U\right)\right\rangle_w}{\left\langle w^{-1}\right\rangle_w},$$

where $\omega = \omega(\mathbf{x})$ is a weighting function and $< \cdots >_\omega$ denotes the ensemble average over the distribution

$$\wp'\left(\mathbf{x}\right) = \frac{w \exp\left(-\beta U_0\right)}{\int w \exp\left(-\beta U_0\right) \mathrm{d}\mathbf{x}}. \qquad (3.28)$$

If we substitute

$$w\left(\mathbf{x}\right) = \exp\left[-\beta V\left(\mathbf{x}\right)\right] \qquad (3.29)$$

into (3.27), we obtain

$$\langle \exp\left(-\beta \Delta U\right) \rangle_0 = \frac{\langle \exp\left[-\beta \Delta U + \beta V\right] \rangle_w}{\langle \exp\left(\beta V\right) \rangle_w}. \qquad (3.30)$$

The same expression for free energy is obtained if one calculates ΔA between the reference and the target system by sampling from an 'intermediate' state, for which the potential energy is $U_0 + V$. This approach was discussed in Sect. 2.9.1.

As we have stressed in the previous chapter, reliable calculations of ΔA require that the range of ΔU around the peak of the integrand in (2.12) of Chap. 2 is well sampled. This peak coincides with the peak of the probability distribution function, $\wp_1(\Delta U)$ obtained through sampling ΔU from the target state. This means that $\wp'$ needs to be sufficiently wide in ΔU to extend to both $\wp_0(\Delta U)$ and $\wp_1(\Delta U)$. For this reason it has been called an 'umbrella distribution,' and the corresponding sampling technique is referred to as 'umbrella sampling' (US) [15]. The name has become so popular that it is commonly used for many different implementations of importance sampling and their combination with stratification.

3.3.3 Importance Sampling and Stratification with WHAM

The material presented earlier in this section raises several interesting practical questions. How can one optimally reconstruct the total probability distribution from

histograms recorded in overlapping windows? Is it possible to combine statistics acquired in importance sampling simulations with different biasing potentials (weighting functions)? Is there a systematic way to improve the choice of the biasing potential, so that it yields uniform sampling along the order parameter? As it turns out, the weighted histograms analysis method (WHAM) [16] described in Sect. 3.2.2 provides answers to all of these questions. As discussed there, WHAM provides a systematic approach for reconstructing the free energy profile, $\Delta A(\xi)$, from histograms tracking the probability distribution function for finding a system at different locations along ξ, acquired from a series of simulations that employed different biasing potentials. The relevant equations are the general case described by (3.12a)–(3.12c).

As we mentioned in the previous subsection, efficient sampling along the order parameter ξ is typically achieved if the biasing potential, $V(\xi)$ is chosen such that it is equal to $-\Delta A(\xi)$ [see (3.26)]. This implies that the free energy profile, $\Delta A'(\xi)$ obtained from the biased simulation is flat and ξ is sampled uniformly in its full range. Another criterion for efficiency might be that statistical errors in all bins of the histogrammed $\wp(\xi)$ are equal. These two criteria are closely related and become identical if the diffusion coefficient along ξ is constant. In either case, it is clearly possible to use the WHAM equations to refine the choice of V in an adaptive fashion [17]. One simply makes an initial guess of V and performs a number of MD or MC steps sufficiently large to obtain a reasonably accurate histogram representing $\wp'(\xi)$. These results are used to reconstruct $\Delta A(\xi)$, which is applied as an improved guess for V in the next simulation. The process continues iteratively until satisfactory convergence of the free energy profile is reached. In each iteration, the histograms obtained from all of the previous runs are used to solve the WHAM equations for the latest estimate of $\Delta A(\xi)$. If applied carefully, iterative WHAM is a very useful strategy. This is, however, not the only adaptive method for free energy calculations. Later in this chapter, we will learn about several strategies for adaptive biasing in so-called flat-histogram simulations. Furthermore, Chap. 4 presents another, powerful method of this type.

We close this section with some comments about errors arising from the discretization of the probability density in histogram-based free energy calculations. The bin width in histograms must be sufficiently small that the probability density does not change significantly within a single bin. Otherwise a systematic error arises because the probability density averaged in a bin is higher than the probability density at its center, yielding free energy estimates that are systematically too low [18]. On the other hand, as the bin size decreases, statistical fluctuations in each bin increase. Thus, the optimal choice for bin size reflects a balance between systematic and statistical errors. Since systematic error depends on the shape of the free energy profile, which is known only after the calculations have been completed, the optimal choice can be done only in post-simulation analysis. The existence of both statistical and systematic errors is common to different methods for estimating free energy and can be traced to rapid changes in the underlying probability distributions. This issue will be discussed at length in Chap. 6 in the context of FEP and nonequilibrium work methods for free energy calculation.

3.4 Flat-Histogram Methods

We turn now to a powerful class of algorithms called 'flat-histogram' methods that rely fundamentally on the ideas of importance sampling mentioned in Sect. 3.3.3. A flat-histogram simulation simply aspires to what its name suggests: equal visits to states defined by one or more macroscopic parameters, such as the potential energy or number of particles. Generally speaking, the MC sampling scheme which one must employ to recover such a flat histogram is unknown at the start of the simulation. This is because, as we will see later, the configurational probabilities entail thermodynamic partition functions, which are not known a priori. Therefore, the goal of these methods is to determine, over the course of the simulation, the appropriate state weights which will produce the flat histogram. These weights, in turn, provide information about the free energy or entropy of the system.

Why would one want to use flat-histogram sampling in a simulation, rather than that of conventional ensembles? There does not seem to be any immediate advantage, since there is no connection to systems in nature – where, for example, would we find a liquid, the potential energy of which fluctuates uniformly between its value at freezing and vaporization? Indeed, the flat-histogram technique is not a direct simulation of any physical state of the system; rather, it is a way of exploring numerous states in a single run. When we employ this approach, the nature of our results differs from that of conventional ensembles in that run averages have little significance. Instead of quantities such as the average potential energy or virial, we use information from the calculated state weights and the observed distribution of visited states (which we desire to be very nearly flat) to determine *post*-simulation the expected averages in conventional ensembles. Equivalently, we are approximating the density of states or a partition function in generating the correct flat-histogram sampling scheme. This information is then used to 'reweight' state probabilities for the calculation of ensemble averages, as in the histogram reweighting method discussed in Sect. 3.2.

Flat-histogram sampling is a very flexible and powerful addition to our library of Monte Carlo techniques. It allows us to sample a range of macrostate space explicitly. Take the case of potential energy, for example. In the canonical ensemble, the range of potential energies explored by the system becomes increasingly narrow as the system size grows, scaling as $N^{1/2}$ (compared to the average energy as N) [19]. Thus if one wanted to perform an overlapping histogram method in potential energy, with canonical results obtained over a range of temperatures, the number of histograms required would increase significantly as the system size grew. This feature would be prohibitive for very large simulations. Alternatively, one could use the flat histogram approach to *force* the system to explore a predetermined range of potential energy. It would be possible to perform a single simulation corresponding to the entire temperature range of interest, or the potential energy range could be divided arbitrarily into various overlapping 'windows' if multiple processors were available. Though in the latter case, the mathematical details of patching the data from the multiple runs would differ from the Ferrenberg–Swendsen approach, the underlying conceptual task would be identical.

Flat-histogram sampling is also often useful in successfully surmounting ergodic bottlenecks. The foremost example of this is its application to first-order phase transitions, which was one of the earliest applications of the approach [6]. Consider a liquid in an NPT simulation at its liquid–vapor coexistence pressure, but well below the critical temperature. Although we would expect the system volume to alternate between its values in the gas and liquid states, in reality this is an extremely slow if not negligible process in simulation because the intermediate densities have an extremely low probability of being visited. In applying flat-histogram sampling to the volume moves, we can weight these intermediate states with higher probability, thereby encouraging the system to travel more frequently between liquid and vapor states. Flat-histogram sampling can also be implemented in potential energy space, which may facilitate the attainment of equilibrium simulation data by allowing greater mobility in the system's crossing of energy barriers. This is in contrast with the conventional Boltzmann weights at low temperature, for which high energy barriers are rarely crossed.

3.4.1 Theoretical Basis

Historically, there have been two approaches to flat-histogram simulation. In the 'weights' scheme, one begins with a conventional MC ensemble and adds a weighting factor to the state probabilities, which later forces an equiprobable distribution in one or more macroscopic parameters [6, 20]. In the direct partition function approach, one samples explicitly according to an initially unknown partition function (possibly the density of states), which is systematically determined over the course of the simulation and which produces a flat histogram [21, 22]. The two methods yield identical sampling although the implementation often differs and in specific cases one might be conceptually more straightforward than the other.

We begin with the weights formalism in general form. Starting from a conventional ensemble, we introduce a weighting factor into the microscopic sampling probabilities which contains the weights η [23]. We construct this factor such that the original sampling distribution is recovered when the weights are zero

$$\boxed{\wp(\mathbf{q}) \propto \rho_0(\mathbf{q}) \exp[-\eta(\mathbf{X})],}$$ (3.31)

where ρ_0 is the configuration space density in a conventional unweighted ensemble, and $\mathbf{X} = \{X_1, X_2, \cdots\}$ contains all the parameters for which we wish to generate a flat distribution, such as the potential energy, number of particles, or volume. We have introduced the exponential involving the weights $\eta(\mathbf{X})$, which we will later tune to obtain this flat distribution. Integrating (3.31) over all configurations with specific values of $\mathbf{X}$ to find the macroscopic distribution, we obtain

$$\wp(\mathbf{X}) \propto \Xi_0(\mathbf{X}) \exp[-\eta(\mathbf{X})],$$ (3.32)

where $\Xi_0(\mathbf{X})$ is a generic partition function, which depends on the parameters $\mathbf{X}$ as well as the original ensemble density ρ_0. When the weights are zero, we find

that this partition function is proportional to the observed macroscopic distribution: $\wp_0(\mathbf{X}) \propto \Xi_0(\mathbf{X})$, where again the subscript zero indicates the unweighted case. In order to sample according to a flat distribution, therefore, we must choose the weights as

$$\boxed{\eta(\mathbf{X}) \propto \ln \Xi_0(\mathbf{X}) \propto \ln \wp_0(\mathbf{X}).} \tag{3.33}$$

In other words, we simply introduce a factor in the microstate probabilities which is inversely proportional to the conventional macroscopic distribution. As a result, this factor cancels the integrated macroscopic probabilities and leaves the distribution constant – exactly the flat-histogram scenario of interest.

Let us illustrate this procedure with the grand-canonical ensemble, and take the scenario in which we desire to achieve a uniform distribution in particle number N at a given temperature. In the weights formalism, we introduce the weighting factor $\eta(N)$ into the microstate probabilities from (3.31) so that

$$\wp(\mathbf{q}, N) \propto \frac{1}{\Lambda^{3N} N!} \exp\left[-\beta U + \beta \mu N - \eta(N)\right] \tag{3.34a}$$

$$\wp(N) \propto Q(N, V, T) \exp\left[\beta \mu N - \eta(N)\right], \tag{3.34b}$$

where μ is the chemical potential. Note that we have made the dependence of $\wp$ on V, T, μ, and η implicit. In order to achieve a uniform distribution in N, we require the second of these expressions to be constant. According to (3.33), this implies that our weights should be

$$\begin{aligned} \eta(N) &\propto \beta \mu N + \ln Q(N, V, T) \\ &= \beta \mu N - \beta A. \end{aligned} \tag{3.35}$$

This last expression contains the Helmholtz free energy and is reminiscent of the familiar thermodynamic identity $\beta \mu N - \beta A = \beta PV$. The converged flat-histogram weights in this scheme, therefore, seem to contain information related to the pressure (although not necessarily the true pressure due to the fixed μ). This will be a recurring theme: by calculating the correct weights to obtain uniform sampling, we obtain thermodynamic information about the system. If we insert the ideal weights back into (3.34a), we see that a flat histogram in N in the grand-canonical ensemble yields microscopic sampling according to the inverse canonical partition function weighted by the Boltzmann factor

$$\wp(\mathbf{q}, N) \propto \frac{1}{\Lambda^{3N} N!} \frac{\exp\left[-\beta U\right]}{Q(N, V, T)}. \tag{3.36}$$

In general, any flat-histogram technique corresponds to sampling microstates according to an inverse partition function. This is because, when the microstate sampling probability is integrated to determine the overall macrostate distribution, any partition function must be canceled out to obtain a uniform scheme. In the weights approach, we never directly reference this partition function: we implement (3.34a) as our MC sampling scheme and adjust η until a flat histogram is obtained. The

acceptance probabilities for particle insertions and deletions are straightforwardly extended from those in the pure grand-canonical ensemble

$$\mathscr{P}_{\mathrm{acc}}(N \to N+1) = \min\left[1, \frac{V}{\Lambda^3(N+1)}\exp\left(-\beta\Delta U + \beta\mu - \Delta\eta\right)\right] \quad (3.37\mathrm{a})$$

$$\mathscr{P}_{\mathrm{acc}}(N \to N-1) = \min\left[1, \frac{\Lambda^3 N}{V}\exp\left(-\beta\Delta U - \beta\mu - \Delta\eta\right)\right], \quad (3.37\mathrm{b})$$

where $\Delta\eta$ is the change in the weight factor for the move.

In contrast to the weights formalism, the partition function approach directly employs the ideal flat-histogram expression in (3.36). Its goal is not to determine η but $Q(N,V,T)$ directly, or more precisely in this case, the N dependence of Q. Due to numerical reasons, we usually work instead with the associated thermodynamic potential which is the logarithm of the partition function of interest; in this case it is $\ln Q = -\beta A \equiv \mathscr{F}$, where we have used script $\mathscr{F}$ as an abbreviation. Thus our sampling scheme becomes

$$\wp(\mathbf{q}, N) \propto \frac{1}{\Lambda^{3N}N!}\exp\left[-\beta U - \mathscr{F}(N)\right], \quad (3.38\mathrm{a})$$

$$\wp(N) \propto Q(N,V,T)\exp\left[-\mathscr{F}(N)\right], \quad (3.38\mathrm{b})$$

where $\mathscr{F}$ is now what we will tune in order to obtain a flat histogram, and which will converge upon the true Helmholtz free energy expression $(-\beta A)$. Again, we find that determining the correct flat-histogram sampling scheme provides thermodynamic information. The corresponding acceptance criteria differ somewhat from (3.37a) and (3.37b)

$$\mathscr{P}_{\mathrm{acc}}(N \to N+1) = \min\left[1, \frac{V}{\Lambda^3(N+1)}\exp\left(-\beta\Delta U - \Delta\mathscr{F}\right)\right], \quad (3.39\mathrm{a})$$

$$\mathscr{P}_{\mathrm{acc}}(N \to N-1) = \min\left[1, \frac{\Lambda^3 N}{V}\exp\left(-\beta\Delta U - \Delta\mathscr{F}\right)\right]. \quad (3.39\mathrm{b})$$

To reach a flat histogram in either approach, eventually we must determine an unknown thermodynamic function of N, whether it be η or $\mathscr{F}$. Both functions play into the microstate sampling scheme and hence the acceptance probability in our simulation. Usually we can only determine these functions to an arbitrary constant since additive shifts in them have no effect on the resulting probability distribution (the reason is identical to that discussed in Sect. 3.2.1). We, therefore, usually set their minimum value to be zero and shift accordingly. Typically we also start off each simulation by setting η or $\mathscr{F}$ to be zero for *all* values of N. This has a distinct effect in the two approaches, as the initial ensembles yield different distributions. During the course of the simulation, we then have a nice feedback mechanism for systematic adjustment: if we obtain a flat histogram, we have converged to the true functions of interest; otherwise we must make changes to our working approximation.

If we wish to generate a uniform distribution in *all* of the macrostates that fluctuate during the simulation (in this case both N and U), the same arguments necessitate the following microstate sampling scheme:

$$\wp(\mathbf{q}, N) \propto \frac{1}{N!}\Omega(N, V, U)^{-1}, \tag{3.40}$$

where the relevant quantity is the density of states. This expression is actually quite general, as it applies to all situations in which we wish to generate equiprobable distributions in all fluctuating macroscopic parameters of interest – a specific flat-histogram scenario we will term a density-of-states simulation (Fig. 3.3). In fact, it is possible to define density-of-states expressions for and run flat-histogram simulations using *arbitrary* macroscopic parameters, beyond the familiar N, V, U. In such cases, the sampling scheme in (3.40) is simply modified by introducing additional dependent variables into Ω. We elaborate on this in Sect. 3.5, but for the present discussion we simply emphasize that any density-of-states simulation requires us to determine the dependence of Ω on the fluctuating quantities of interest. Again in practice, we usually determine its logarithm instead, $\mathscr{S} = \ln \Omega$, where $\mathscr{S}$ is the dimensionless entropy. The acceptance criterion for single-particle displacements, additions and deletions, and volume scaling moves for these simulations can all be written as [24]

$$\mathscr{P}_{\mathrm{acc}}(o \rightarrow n) = \exp\left[-\Delta\mathscr{S}_{\mathrm{ex}}\right], \tag{3.41}$$

where o and n are the original and new states, and $\mathscr{S}_{\mathrm{ex}} = \ln \Omega_{\mathrm{ex}} = \mathscr{S} - \ln N! - N \ln V$ is the dimensionless excess entropy, as defined in Chap. 1.

Ultimately from these simulations, we would like to recover thermodynamic data appropriate to natural ensembles. This is readily accomplished by histogram-reweighting techniques, in which we convert a measured probability distribution
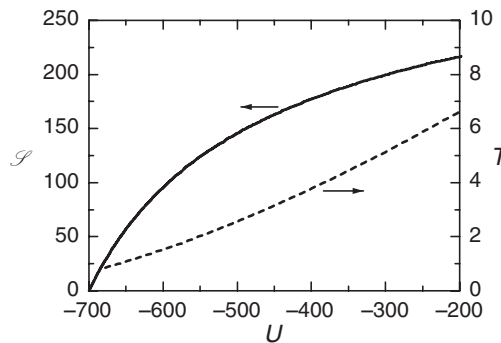
Fig. 3.3. Typical results from a density-of-states simulation in which one generates the entropy for a liquid at fixed N and V (i.e., fixed density) (adapted from [29]). The dimensionless entropy $\mathscr{S} = \ln \Omega$ is shown as a function of potential energy U for the 110-particle Lennard-Jones fluid at $\rho = 0.88$. Given an input temperature, the entropy function can be reweighted to obtain canonical probabilities. The most probable potential energy U^* for a given temperature is related to the slope of this curve, $\mathrm{d}\mathscr{S}/\mathrm{d}U(U^*) = 1/k_{\mathrm{B}}T$, and this temperature–energy relationship is shown by the *dotted line*. Energy and temperature are expressed in Lennard-Jones units

from one set of state conditions to another. In the weights approach, for example, we can combine the weighted and unweighted ($\eta = 0$) instances of (3.34b) to give

$$\wp(N; \mu) \propto \widetilde{\wp}(N; \mu_0, \eta) \exp\left[\beta(\mu - \mu_0)N + \eta(N)\right], \tag{3.42}$$

where $\widetilde{\wp}$ is the measured probability distribution using the weights and μ_0 is the original chemical potential at which the weights were run. This expression allows us to determine the N distribution for any chemical potential at β, as the proportionality is fixed by the normalization condition. In practice, we will be limited by chemical potentials corresponding to the range of N explored in the flat-histogram simulation. There is an underlying generality here: for each fluctuating macroscopic parameter for which we enforce a flat histogram, we can reweight simulation results in the conjugate thermodynamic field.[1] So if we had enforced a flat distribution of both particle number and energy, for example, it would be straightforward to determine $\wp(N, U; \mu, T)$ for arbitrary chemical potential and temperature. For the partition function approach, we instead divide the unweighted grand canonical probability by (3.38b)

$$\wp(N; \mu) \propto \widetilde{\wp}(N; \mathscr{F}) \exp\left[\beta\mu N + \mathscr{F}(N)\right], \tag{3.43}$$

where $\widetilde{\wp}(N; \mathscr{F})$ is the distribution measured from the ensemble using $\mathscr{F}$ in (3.38a). In general for either the weights or partition function approach, the reweighting procedure is easily determined by dividing the macrostate probability scheme of the desired ensemble by that of the simulated flat-histogram one. The flat-histogram sampling and reweighting procedures for common ensembles are summarized in Table 3.1.

The reader may be concerned by the appearance of $\widetilde{\wp}$ in (3.42) and (3.43). After all, the very definition of a flat-histogram simulation says that this expression should be a constant, independent of N, so why do we need to measure it? The answer is simple: we do not, provided our weights are determined to sufficient statistical accuracy. Occasionally, however, extremely precise determination of η can become computationally demanding, and rather than perform very long simulations which guarantee a completely flat histogram, it is more efficient to run for moderate length with a rough estimate of η which provides a 'flat enough' histogram – one in which all macrostates are visited. In this latter case, measurement of $\widetilde{\wp}$ for (3.42) compensates for the statistical uncertainty in the weights. Still, recent algorithms [21, 25–28] have emerged which permit both efficient and accurate determination of flat-histogram sampling schemes, from which the calculated weights or partition function is used directly. If we apply these methods using the partition function

---

[1] Of course we need not implement a flat-histogram scheme to perform histogram reweighting. Practically speaking, however, we are always limited in reweighting to the finite range of macrostate space explored by the original simulation. Flat-histogram sampling often greatly increases this range relative to single runs in conventional ensembles, and therefore significantly increases our reweighting ability in the associated macrostates.

Table 3.1. Common flat-histogram ensembles and their reweighing procedures

Variable(s)	Microstate probabilities	Reweighting probabilities
U	$\wp(\mathbf{q}) \propto \mathrm{e}^{-\mathscr{S}(U)}$	$\wp(U;T) \propto \widetilde{\wp}(U)\mathrm{e}^{\mathscr{S}(U)-\beta U}$
U, N	$\wp(\mathbf{q}, N) \propto \frac{1}{N!}\mathrm{e}^{-\mathscr{S}(N,U)}$	$\wp(N, U;\mu, T) \propto \widetilde{\wp}(N, U)\Lambda^{-3N}$ $\times \mathrm{e}^{\mathscr{S}(N,U)-\beta U+\beta\mu N}$
U, V	$\wp(\mathbf{q}, V) \propto \mathrm{e}^{-\mathscr{S}(V,U)}$	$\wp(V, U;P, T) \propto \widetilde{\wp}(V, U)$ $\times \mathrm{e}^{\mathscr{S}(N,V)-\beta U-\beta PV}$
N	$\wp(\mathbf{q}, N;T_0) \propto \frac{1}{N!\Lambda^{3N}}$ $\times \mathrm{e}^{-\beta_0 U-\mathscr{F}(N)}$	$\wp(U;\mu, T_0) \propto \widetilde{\wp}(N)\mathrm{e}^{\mathscr{F}(N)+\beta_0\mu N}$
V	$\wp(\mathbf{q}, V;T_0) \propto \mathrm{e}^{-\beta_0 U-\mathscr{F}(V)}$	$\wp(V;P, T_0) \propto \widetilde{\wp}(V)\mathrm{e}^{\mathscr{F}(V)-\beta_0 PV}$

The first column indicates the flat-histogram variables, the second the prescribed microstate sampling scheme, and the third the appropriate reweighing probabilities. The script variables $\mathscr{S}$ and $\mathscr{F}$ are the weights to be determined, which converge on $\ln \Omega$ and $\ln Q$, respectively, in the flat-histogram limit. $\widetilde{\wp}$ is the measured distribution from the flat-histogram simulation, frequently dropped if the weights are calculated to high accuracy

formalism, we can directly determine quantities such as the free energy ($\mathscr{F}$) or entropy ($\mathscr{S}$) along various macrostate coordinates.[2]

We have hitherto overlooked two important points. One is that we certainly cannot perform a flat-histogram simulation over the *entire* range of many macroscopic parameters, including N, V, U which are all unbounded to the right. For practical reasons, we also desire to sample a range of macrostate space of our choosing, for example a particular potential energy range corresponding to the temperatures we wish to study. Therefore, these simulations always entail macrostate bounds, and moves are rejected which would take the system beyond them. Algorithmically, these rejection moves are treated as any other rejection. The second important point is that reaching 'equilibrium' in a run is no longer indicated by the convergence of simulation averages, but rather by the accurate determination of state weights. An important measure of the dynamics in these MC simulations, analogous to the structural relaxation time in conventional ensembles, is the so-called tunneling time, generically defined as the time it takes the system to sample all of its macrostates [6, 29]. For example, this might be the number of steps required to move between the minimum and maximum energy boundaries during a flat-histogram simulation in potential energy.

Not surprisingly, the essential component of flat-histogram algorithms is the determination of the weights, η, or the thermodynamic potential, e.g., $\mathscr{F}$ or $\mathscr{S}$. There exist a number of techniques for accomplishing this task. The remainder of this section is dedicated to reviewing a small but instructive subset of these methods, the multicanonical, Wang–Landau, and transition-matrix approaches. We subsequently discuss their common and sometimes subtle implementation issues, which become of practical importance in any simulation.

---

[2] More accurately, we can determine free energy and entropy *differences*, since their absolute value remains unspecified.

3.4.2 The Multicanonical Method

The multicanonical algorithm of Berg and Neuhaus [6] is a weights approach to flat-histogram simulation, building closely upon the ideas of reweighting in Sect. 3.2.1. In this method, near-ideal flat-histogram weights η are first iteratively determined through several small preproduction runs, and then a longer simulation interval using these results is performed to collect histogram data for reweighting [20]. We again take the grand canonical ensemble as our example, with weights in the particle number N. The multicanonical approach specifies the state probabilities as:

$$\wp(\mathbf{q}, N; \eta_i) = \frac{1}{\Lambda^{3N} N!} \frac{\exp\left[-\beta U + \beta\mu N - \eta_i(N)\right]}{\Xi(\mu, V, T; \eta_i)}, \tag{3.44a}$$

$$\wp(N; \eta_i) = \frac{Q(N, V, T) \exp\left[\beta\mu N - \eta_i(N)\right]}{\Xi(\mu, V, T; \eta_i)}, \tag{3.44b}$$

where the difference from (3.34a) and (3.34b) is the subscript on η and the specification of the denominators. For two sets of weights η_1 and η_2 at the same μ, V, and T, we have two instances of (3.44b) which can be combined to yield:

$$\boxed{\ln \wp(N; \eta_1) - \ln \wp(N; \eta_2) = \eta_2(N) - \eta_1(N) + \ln \Xi(\eta_2) - \ln \Xi(\eta_1).} \tag{3.45}$$

This expression forms the basis of the weight update scheme in our multicanonical simulation. For the set of weights η_1, we can perform a short simulation and measure the $\wp(N; \eta_1)$ distribution. Then, by setting $\wp(N; \eta_2)$ in (3.45) to a uniform distribution, we are able to estimate a new set of weights η_2 corresponding to a flat histogram. When we run with this new set of weights, however, we still may not find a completely flat histogram owing to numerical uncertainty in our original measurements at the tails of $\wp(N; \eta_1)$. Therefore, we may have to iterate this process several times until the statistics of the measured $\wp(N)$ become good enough over the complete N range to permit accurate estimation of all the weight values. In more general terms, therefore, we can write this iterative process as

$$\eta_{i+1}(N) = \ln f(N; \eta_i) + \eta_i(N) + k, \tag{3.46}$$

where i refers to the ith simulation run, f gives the raw histogram counts of N, and k is a constant which we simply adjust such that the minimum value of η_{i+1} is zero. Here we have omitted all terms from (3.45) which are additive and independent of N, as they have no effect on the probability scheme. An immediate problem with (3.46), however, is that weight estimates become ill-defined for values of N at which $f = 0$. Using a Bayesian-statistical framework, Smith and Bruce have shown that an alternate update scheme is justified and avoids this shortcoming [25]

$$\eta_{i+1}(N) = \ln\left[f(N; \eta_i) + 1\right] + \eta_i(N) + k. \tag{3.47}$$

The procedure implied by (3.46) is simple: we start off with any weights of our choice, although often the most convenient choice is simply $\eta_1(N) = 0$. We perform

a short simulation with these values, measuring $\wp(N; \eta_1)$, and use this information to predict the next set of weights. We continue this process until η is reasonably converged and until we achieve, more or less, an equal sampling in particle number. At this point, we fix our final weights η_f and perform a much longer simulation in order to gather accurate statistics for $\widetilde{\wp}(N; \eta_f)$. The resulting data set η_f and $\widetilde{\wp}(N; \eta_f)$ can then be used in conjunction with (3.42) to generate conventional averages at the desired chemical potential. A typical evolution of the weights for a grand-canonical simulation of the Lennard-Jones system is shown in Fig. 3.4.

Beyond our grand canonical example, the multicanonical procedure for determining η is quite general and straightforward. In most cases, the weights update scheme is nearly identical to (3.46) and (3.47), with changes only in the dependent variables of η and f. This is intuitive, as a flat f will leave the weights unchanged from the previous iteration, as desired. A weakness of the multicanonical approach, however, is that statistical information from previous runs is discarded as one iterates the weights. That is, only data from run i is directly utilized for the $i + 1$ weights, not the cumulative data from runs $1, 2, \ldots, i$. A remedy developed by Smith and Bruce [25] is to incorporate uncertainties in the weight update scheme

$$\frac{\eta_{i+1}(N)}{\sigma_{i+1}^2(N)} = \frac{\eta_i(N)}{\sigma_i^2(N)} + \frac{\hat{\eta}_{i+1}(N)}{\hat{\sigma}_{i+1}^2(N)} \tag{3.48}$$
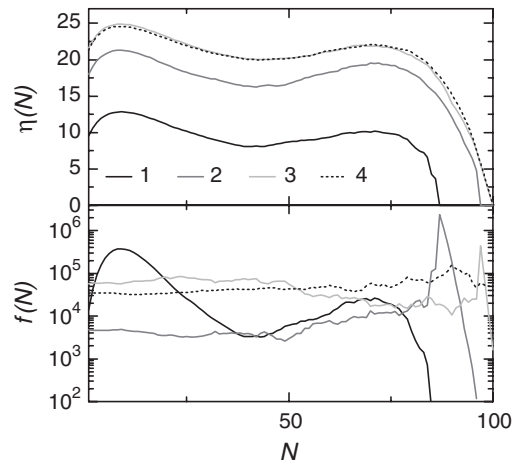
Fig. 3.4. Evolution of the weights $\eta_i(N)$ and the histograms $f_i(N)$ in a grand-canonical implementation of the multicanonical method for the Lennard-Jones fluid at $V = 125$. The temperature is $T = 1.2$ and the initial chemical potential is $\mu = -3.7$. The weights are updated after each 10-million-step interval, and the numbers indicate the iteration number. The second peak in the weights at large particle numbers indicates that the initial chemical potential is close to its value at coexistence

with:

$$\frac{1}{\sigma_{i+1}^2(N)} = \frac{1}{\sigma_i^2(N)} + \frac{1}{\hat{\sigma}_{i+1}^2(N)} \tag{3.49a}$$

$$\hat{\eta}_{i+1}(N) = \langle \ln f(N; \eta_i) + \eta_i(N) + k \rangle \tag{3.49b}$$

$$\hat{\sigma}_{i+1}^2(N) = \mathrm{var}\left[\ln f(N; \eta_i) + \eta_i(N) + k\right], \tag{3.49c}$$

where $\hat{\eta}$ and $\hat{\sigma}$ are the average weights and their uncertainties predicted from iteration i. Unlike previously, this procedure requires one to determine the uncertainty in each $\eta_{i+1}(N)$. Smith and Bruce suggested subdividing the total histogram $f(N; \eta_i)$ into M subhistograms, each producing an estimate of η_{i+1} from which means and variance are extracted. They had moderate success with this particular approach, which required special attention in the initial iterations and local averaging of $\hat{\sigma}$ for good convergence.

3.4.3 Wang–Landau Sampling

The Wang–Landau (WL) method is a flat-histogram technique of the partition function variety, and is designed to calculate thermodynamic potentials directly to a high level of accuracy [21, 22]. Originally designed for discrete lattice systems, the WL approach was successfully adapted to continuum fluids by Yan et al. [30] and Shell et al. [24]. Wang and Landau originally introduced their algorithm for the calculation of the density of states. For that reason, we will take as our example the determination of $\Omega(U)$ for a fixed-density fluid. However, the method is equally applicable to the calculation of other partition functions as discussed in Sect. 3.4.1.

Unlike the multicanonical method, the WL algorithm does not consist of separate weight-determining and production periods, but entails a series of stages over which the density of states is successively approximated with increasing precision. Perhaps the defining feature of the WL approach is the continuous modification of its sampling scheme at each step, which would be analogous to multicanonical updating of η with each Monte Carlo move. We begin with the microstate probabilities in our example

$$\wp(\mathbf{q}) = \exp\left[-\mathscr{S}(U)\right], \tag{3.50}$$

the single-particle displacement acceptance criterion of which is

$$\mathscr{P}_{\mathrm{acc}}(U_{\mathrm{o}} \to U_{\mathrm{n}}) = \min\left\{1, \exp\left[\mathscr{S}(U_{\mathrm{o}}) - \mathscr{S}(U_{\mathrm{n}})\right]\right\}, \tag{3.51}$$

where the "ex" subscript from (3.41) can be dropped owing to constant N and V. Here we have written the sampling scheme in terms of the dimensionless entropy rather than the density of states; for convenience and its connection with the actual implementation, we will deal with the former in this example, although we note that the original formulation differs in this respect. The initial estimate for $\mathscr{S}$ is set to zero for all U, and the simulation proceeds according to (3.51). The crucial element is that after each MC move, the value of the entropy at the ending energy (or in general, the

value of the logarithm of the partition function at the current macroscopic state) is updated according to

$$\boxed{\mathscr{S}(U_{\text{end}}) \leftarrow \mathscr{S}(U_{\text{end}}) + g,}$$

(3.52)

where g is a parameter greater than zero, termed the modification factor. Equation (3.52) is applied regardless of whether the ending state is the original configuration after a rejection or a new configuration after an acceptance. In this way, the estimate for the partition function is dynamically modified, with g controlling the magnitude of those modifications. Another way to think about it is that we are 'building up' an estimate for $\mathscr{S}$ as we go. The particular form of the update scheme in (3.52) has an important property: when all energies are visited on an equal basis, on average the entire $\mathscr{S}(U)$ curve will simply shift by an increasing amount with time. We immediately realize that this does not affect the entropy calculation, since we always adjust $\mathscr{S}(U)$ so that its minimum value is zero (recall that additive shifts in the entropy have no effect on the microstate probabilities). Therefore, the convergence of the entropy and the attainment of a flat histogram are directly linked in this update scheme. Initially the system will sample only energies of the highest entropy, but as $\mathscr{S}(U)$ is built up, those states will be sampled less frequently according to (3.50) until ultimately a flat histogram is reached.

The modification factor plays a central role in a WL simulation and has several effects. First, its presence violates microscopic detailed balance because it continuously alters the state probabilities, and hence acceptance criterion. Only for $g = 0$ do we obtain a true Markov sampling of our system. Furthermore, we obviously cannot resolve entropy differences which are smaller than g, yet we need the modification factor to be large enough to build up the entropy estimate in a reasonable amount of simulation time. Wang and Landau's resolution of these problems was to impose a schedule on g, in which it starts at a modest value on the order of one and decreases in stages until a value very near to zero (typically in the range 10^{-5}–10^{-8}). In this manner, detailed balance is satisfied asymptotically toward the end of the simulation.

Wang and Landau proposed several heuristics for automating the modification factor schedule. A stage is deemed 'complete' when the histogram for that stage is deemed flat enough, which they suggested was the point at which no histogram bin was less than 80% of the average bin value. Yan and de Pablo [27] showed that tuning this percentage higher has the effect of increasing the statistical efficiency of method, but at the expense of greater simulation time. Other heuristics have also been suggested; for example, Shell et al. suggest that a stage is complete when each histogram bin has been visited a minimum number of times [24]. Regardless, once a stage is finished the modification factor must be decreased. Wang and Landau suggested

$$g \leftarrow \frac{1}{2} g,$$

(3.53)

which has proved adequate in numerous studies (some discussion of this choice is provided in [22]).

The WL simulation terminates when the last stage takes g below some cutoff value; at that point, the output is the calculated $\mathscr{S}$. Contrary to the multicanonical

scheme, the flat-histogram probability distribution is not reweighted, but instead the converged entropy is used directly as the true entropy. Canonical averages such as the energy, for example, follow (3.4)

$$\langle U \rangle (T) = \frac{\displaystyle\int U \mathrm{e}^{-\beta U + \mathscr{S}(U)} \mathrm{d}U}{\displaystyle\int \mathrm{e}^{-\beta U + \mathscr{S}(U)} \mathrm{d}U}. \tag{3.54}$$

The same approach applies to alternate partition functions; the WL output is used directly in the macroscopic probability scheme. For example, in the previous grand canonical scenario, the WL simulation would yield $\mathscr{F}$, which would be substituted directly into the macrostate probabilities as $Q = \exp(\mathscr{F})$ for subsequent results generation.

A nice feature of the WL algorithm is that it is readily amenable to labor division by multiple processors. As suggested by the original authors, the energy range of interest can be divided into multiple overlapping windows, and an individual simulation can be performed on each subrange. Such stratification has been described in greater detail in Sect. 3.4.2. The procedure is illustrated graphically in Fig. 3.5. Multiple windows are especially useful when attempting to explore large energy ranges, as otherwise the long tunneling time of the complete range might be prohibitive. At the end of the subrange simulations, the data are patched together by shifting the entropy functions to obtain agreement in their overlap. Assuming the statistical error in $\mathscr{S}$ is uniform, one minimizes the following variance [24]
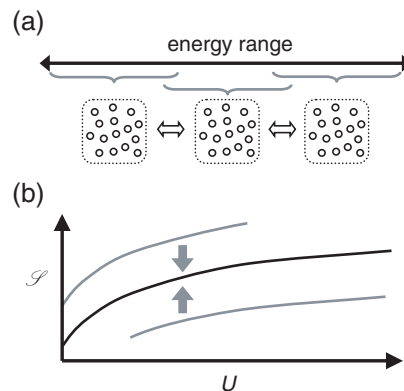
Fig. 3.5. (**a**) Schematic of a parallel processing setup for a flat-histogram simulation in energy. Each processor is assigned a 'window' of energy, and moves which take the system outside of the window are rejected. To restore ergodicity, periodic configurations swaps are performed between adjacent windows when the respective configurations are both within the overlapping region. (**b**) Illustration of the postsimulation patching procedure for the generation of an overall entropy curve from adjacent windows. The two curves are shifted to obtain overlap in the region of common energy; entropy values at the overlapping energies are then averaged

$$\sigma_{\text{tot}}^2 = \sum_{i<j} \sum_k \left[\mathscr{S}_i(k) + C_i - \mathscr{S}_j(k) - C_j \right]^2, \tag{3.55}$$

where i and j are indices of the subrange results, the constants C are the values by which the entropies are to be shifted, and k is an index for all overlapping points between subranges i and j. In our example, k runs over the overlapping energies between two subranges. One C value must be specified to obtain a solution to this equation; once the remainder are determined, the overall entropy is pieced together from each shifted section with values at areas of overlap averaged. For two adjacent $\mathscr{S}(U)$ curves, minimization of σ_{tot}^2 is given by:

$$C_j - C_i = \langle \mathscr{S}_i(U) - \mathscr{S}_j(U) \rangle_{\text{overlap}}, \tag{3.56}$$

where the average is performed over the overlap region. The equivalent patching procedure for multiple macrostate variables simply involves the introduction of additional sums in (3.55) over the appropriate parameters, although its minimization may be more complex than (3.56).

An important caveat associated with multiple-range patching is the possible restriction of system ergodicity. It is possible that a small window would not permit the system sufficient flexibility in exploring ergodically all of the states in that window. This might occur, for example, if the system is trapped in a local potential energy minimum, the escape energy of which resides outside of the subrange upper energy bound. To avoid restricting the system's ability to explore its phase space, one can link each subrange together in a parallel computing fashion with occasional configuration swaps between adjacent windows. This is reminiscent of the replica exchange or parallel tempering methods [31–33] discussed in Chap. 8. Configurations in the overlapping region of adjacent energy windows are periodically swapped with an acceptance probability of

$$\mathscr{P}_{\text{acc}}(1 \leftrightarrow 2) = \min \left\{ 1, \exp \left[\mathscr{S}_1(U_1) + \mathscr{S}_2(U_2) - \mathscr{S}_1(U_2) - \mathscr{S}_2(U_1) \right] \right\}. \tag{3.57}$$

However, since $\mathscr{S}_1$ and $\mathscr{S}_2$ asymptote to the same function, one might approximate $\mathscr{S}_1(U) = \mathscr{S}_2(U)$ in (3.57) so that the acceptance probability is a constant.[3] The procedure allows trial swaps to be accepted with 100% probability. This general parallel processing scheme, in which the macrostate range is divided into windows and configuration swaps are permitted, is not limited to density-of-states simulations or the WL algorithm in particular. Alternate partition functions can be calculated in this way, such as $\mathscr{F}$ from previous discussions, and the parallel implementation is also feasible for the multicanonical approach [34] and transition-matrix calculations [35].

The principle benefits of the WL method are its wide applicability, ease of implementation, and rapid initial sampling of the macrostate space of interest. It

---

[3] This simplification of the acceptance criterion must be justified a posteriori by the observation that $\mathscr{S}_1$ and $\mathscr{S}_2$ are statistically identical in their overlap region. Otherwise, strict detailed balance is not reached toward the end of the WL simulation.

can be used in both continuum and lattice systems, and a number of methodological variants exist, including 'multibondic' [36], N-fold way [37], quantum [38], and reverse MC [39] extensions. It has since been recognized, however, that the approach can reach a limiting statistical accuracy, beyond which additional stages of the modification factor do not diminish the errors in the calculated density of states [27]. This appears particularly problematic in the simulation of fluids, when ultimately the changes in the modification factor occur too quickly relative to the ability of the calculations to adjust to them. Not surprisingly, increasing the stringency of the flat-histogram requirement – say, from 80% to 90% flatness – postpones this statistical deadlock, but at the expense of increased computational time [27]. The problem is illustrated in Fig. 3.6, which plots the typical evolution of the statistical error in $\mathscr{S}$ for a density of states simulation of the Lennard-Jones fluid.

Several authors have proposed improvements to the precision of WL calculations through the incorporation of additional 'estimators' of the entropy. For random walks in potential energy, Yan and de Pablo [27] have suggested the incorporation of temperature estimators which contain information about the slope of the entropy curve, $\beta = \partial\mathscr{S}/\partial U$, and hence can be integrated to produce an estimate of $\mathscr{S}$. During the WL simulation, they record instantaneous temperatures from each configuration based on an expression due to Butler and co-workers [40]: $\beta = \langle \nabla^2 U/|\nabla U^2| \rangle$, where the derivatives are with respect to the configurational coordinates. They tabulate the average temperature as a function of potential energy and, at the end of at each modification factor stage, generate an entirely new entropy curve by integrating this data. This procedure generates entropy estimates of higher statistical quality
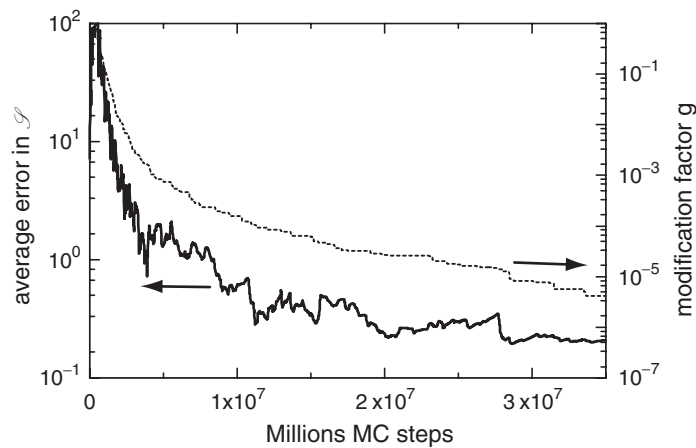
Fig. 3.6. Evolution of a WL simulation for the Lennard-Jones fluid at $\rho = 0.88$ and $N = 110$. The calculated quantity of interest is the dimensionless entropy, $\mathscr{S}$, as a function of potential energy. The average statistical error is determined from the standard deviation of $\mathscr{S}$ from 10 independent runs. The modification factor curve (the *dotted line*) has also been averaged over these runs, and consequently appears smoother than would normally be the case

than the original method, but requires computational overhead in the calculation of forces and their derivatives.

Building on these ideas, Shell and co-workers proposed the incorporation of transition-matrix estimates of $\mathscr{S}$ [28]. They record the move proposal statistics during the simulation, which are used to calculate transition probabilities from one macrostate to another. Then, rather than integrating a temperature–energy curve at the end of each stage, they determine a new entropy estimate from this history of move proposals (see Sect. 3.4.4 for formal details). As with the configurational temperature approach, the calculations for $\mathscr{S}$ are much improved with the additional estimator; however, no energy derivatives are required and the transition-matrix formalism remains general to arbitrary macroscopic variables.

Finally, Zhou, and Bhatt have performed a mathematical analysis of the WL method, which explores the relationship between the modification factor schedule, flat-histogram requirement, and rate of calculation of $\mathscr{S}$ [41]. In particular, they suggests that an optimal strategy for making changes to the modification factor is to make the flatness requirement increasingly stringent with each stage.

In summary, Wang and Landau's algorithm appears to be a robust and straightforward tool for generating rapid flat-histogram sampling and reasonably accurate estimates of the entropy or free energy. Highly precise calculations, though, may require the incorporation of additional estimators.

3.4.4 Transition-Matrix Estimators

Whereas the histograms we have discussed deal with the raw number of visits to each state, a transition matrix concerns itself with the number of moves between pairs of states. Transition-matrix calculations are bookkeeping schemes for transition probabilities between macrostates, and typically generate excellent estimates of thermodynamic functions with high statistical accuracy. This latter observation was originally made by Smith and Bruce [25, 42], and later explored by Oliveira [43, 44]; Wang and Swendsen [26, 45]; and Fitzgerald, Picard, and Silver [46, 47] as the basis of several simulation methods. We present here a general outline of their use in simulation.

We begin with the microstate probability $t(i \rightarrow j)$ of making a move from configuration i to j, each characterized by a volume, number of particles, and set of coordinates $\mathbf{q}$. This probability and its reverse satisfy the detailed balance condition:

$$\wp(i)t(i \rightarrow j) = \wp(j)t(j \rightarrow i). \tag{3.58}$$

Now consider the macrostates I and J to which the microstates belong. These macrostates may be characterized by a combination of N, V, U or any other well-defined order parameter. If we sum all instances of (3.58) for the pairs of microstates $i \in I$ and $j \in J$, we can write

$$\sum_{i \in I} \sum_{j \in J} \wp(i)t(i \rightarrow j) = \sum_{i \in I} \sum_{j \in J} \wp(j)t(j \rightarrow i). \tag{3.59}$$

Note that we have employed discrete notation, although rigorously this sum is an integral in configuration space. The choice is for clarity in the following derivations, and our conclusions will be unaffected by the notation. With minor manipulation, (3.59) gives this important result

$$\boxed{\wp(I)T(I \to J) = \wp(J)T(J \to I),}$$ (3.60)

where we have made the definitions

$$T(I \to J) = \frac{\displaystyle\sum_{i \in I} \sum_{j \in J} \wp(i)t(i \to j)}{\displaystyle\sum_{i \in I} \wp(i)},$$ (3.61a)

$$\wp(I) = \sum_{i \in I} \wp(i).$$ (3.61b)

The physical significance of $T(I \to J)$ in this definition is crucial to its practical use: it gives the probability a move will be made from *macrostate* I to J, and thus is the macroscopic analog of $t(i \to j)$. Essentially, $T(I \to J)$ is the ensemble average of all of the microstate transition probabilities for moving from a member of macrostate I to any member of macrostate J. The function $\wp(I)$, on the other hand, gives the total probability of macrostate I in the ensemble, and is related to a partition function (see, for example, Sect. 3.2). Equation (3.60) is therefore rather remarkable, implying that the macrostate transition probabilities also satisfy detailed balance. That is, the equilibrium flux from macrostate I to J equals that of J to I. Even more importantly, it provides a key thermodynamic estimator. Given numerical values for $T(I \to J)$ and $T(J \to I)$, we can determine the ratio $\wp(I)/\wp(J)$, which is related to a free energy difference.

Fortunately for us, measurement of the macroscopic transition probabilities is straightforward. We could accomplish this, for example, by counting the number of times moves are made between every I and J macrostate in our simulation. The estimate for $T(I \to J)$ would then be the number of times a move from I to J occurred, divided by the total number of attempted moves from I. The latter is simply given by the sum of counts for transitions from I to any state. A more precise procedure that retains more information than simple counts is to record the acceptance probabilities themselves, regardless of the actual acceptance of the moves [46, 47]. In this case, one adds a fractional probability to the running tallies, rather than a count (the number one). This data is stored in a matrix, which we will notate $C(I, J)$ and which initially contains all zeros. With each move, we then update C as

$$C(I, J) \leftarrow C(I, J) + \mathscr{P}_{\text{acc}}$$
$$C(I, I) \leftarrow C(I, I) + [1 - \mathscr{P}_{\text{acc}}].$$ (3.62)

The estimates for the transition probabilities are then given by

$$\widetilde{T}(I \to J) = C(I, J) / \sum_{K} C(I, K).$$ (3.63)

Note that (3.63) would also be the appropriate expression if C recorded counts. By recording the actual acceptance probabilities, however, we are in effect maintaining more information than the binary accept or reject data. This approach is also more flexible; if the acceptance probabilities change during the simulation, due to changes in temperature or weighting function, for example, the C matrix does not need to be re-zeroed.[4] Instead, one simply calculates what the acceptance probability would have been in the original sampling scheme – with the original temperature or weights – for use in (3.62). In other words, all of the $\mathscr{P}_{\mathrm{acc}}$ data used to update C are derived from or adjusted to the original ensemble. The general expression for updates to the C matrix is therefore

$$
\begin{aligned}
C(I, J) &\leftarrow C(I, J) + (\mathscr{P}_{\mathrm{acc}})_0 \\
C(I, I) &\leftarrow C(I, I) + [1 - (\mathscr{P}_{\mathrm{acc}})_0],
\end{aligned}
\tag{3.64}
$$

where the subscript "0" indicates that the acceptance probabilities are calculated in the reference "0" ensemble. This feature also permits a straightforward implementation in a parallel processing scenario like the one described in Sect. 3.4.3; each processor simply updates a section of the overall transition matrix [35].

Of practical importance, we must consider the situation in which the proposed move J lies outside our macrostate boundaries. For such moves, we only need to update the diagonal term as

$$
C(I, I) \leftarrow C(I, I) + 1.
\tag{3.65}
$$

Actually, any time a move is proposed to a region outside the bounds of our transition matrix, whether or not it actually takes the system outside of some constraint, the equivalent transition-matrix update is to add the extrinsic $\mathscr{P}_{\mathrm{acc}}$ to the diagonal C term instead. Since the diagonal terms enter in the final calculations only to normalize the transition probabilities [i.e., they appear only in the denominator of (3.63)], adding such data to the diagonal maintains the correct normalization without affecting the results.

Let us illustrate the transition-matrix procedure with the grand-canonical example [48]. Initially, we will consider the unweighted case (pure grand-canonical ensemble). Here, I and J denote particle numbers and $C(I, J)$ records only the statistics of particle addition and deletion moves. This means that $C(I, J)$ is a tridiagonal matrix. Substituting the macrostate probabilities of (3.34b) into (3.60) and taking the logarithm, we find that the transition probabilities provide an estimate of the Helmholtz free energy

$$
\beta A(N + 1, V, T) - \beta A(N, V, T) - \beta\mu = \ln T(N + 1 \to N) - \ln T(N \to N + 1)
\tag{3.66}
$$

---

[4] If the sampling scheme is changed, $C(I, J)$ can continue to be updated with an adjusted $\mathscr{P}_{\mathrm{acc}}$ provided that the distribution of states within the macrostates does not change. This would not be the case, for example, if we were only monitoring transition probabilities between particle numbers and the temperature changed, as it would redistribute the microstate probabilities within each value of N.

or more compactly,

$$\mathscr{F}(N) - \mathscr{F}(N+1) - \beta\mu = \ln T(N+1 \to N) - \ln T(N \to N+1). \quad (3.67)$$

Implicitly, the $T(I \to J)$ are a function of the chemical potential and temperature at which they were measured. After our simulation, we can determine the complete Helmholtz function (to an additive constant) by considering the transition probabilities between each neighboring pair of particle numbers. Of course, our results will be practically limited to the range of N explored by the simulation, beyond which the statistics of $T(I \to J)$ will be too poor. In order to broaden this range, we could change the chemical potential. By continuing to add data to the C matrix as if the chemical potential had not changed, i.e., by calculating $\mathscr{P}_{\text{acc}}$ for the original μ, the C matrix would not need to be re-zeroed between the change.

A more systematic approach for sampling a wider range of particle numbers is to incorporate flat-histogram sampling. The remarkable feature of the transition-matrix approach is that it yields exactly the information we need to do this. Recall that, in order to achieve a flat distribution in N in the grand canonical ensemble, we need an estimate for the macrostate probabilities $\wp \propto Q(N) \exp(\beta\mu N)$. From (3.66), this is precisely the information we obtain from the transition probabilities. In general, the relevant $T(I \to J)$ will always provide us with the necessary flat-histogram information because they relate directly to the macroscopic probabilities as in (3.60). The usual flat-histogram procedure in this case is to update periodically the weights η or potential $\mathscr{F}$ from the transition-matrix data, in the same sense as multicanonical iterations [25, 47–49]. Detailed balance is violated at the step after an update, and so each iteration should be sufficiently long to guarantee ergodic sampling. Regardless of the iteration procedure, however, the transition matrix is always updated with the acceptance probabilities of the original ensemble, for which the weights are zero (see Fig. 3.7).

As we have seen one of the most attractive features of this formalism is that it permits the continual addition of data to a transition matrix – when state probabilities change, we can continue calculating acceptance probabilities for the transition matrix according to the original ensemble weights. If the original weights do not matter because they will be systematically adjusted (as in a flat-histogram setting, for example), we can choose a simple transition-matrix scheme which is based exclusively on move proposal probabilities. In this case, we let the original microstate probabilities be such that the acceptance of moves is always 100%, which for single-particle displacements corresponds to $\wp(\mathbf{q}) = \text{const}$. Even though the actual acceptance of moves is tuned to less than 100% by adjusting a weight factor, the transition matrix is then always updated in the following manner: entries in the $C(I, J)$ matrix are incremented by a numerical value of 1 whenever a move is *proposed* between macrostates I and J. This procedure is often identical to an initial infinite-temperature ensemble, and is sometimes referred to as such [26]. The proposal approach is readily implemented in a density-of-states algorithm. The desired sampling scheme is $\exp(-\mathscr{S})$, and by incrementing $C(I, J)$ by one every time a move is *proposed* from I to J, an estimate for entropy can be extracted [28]
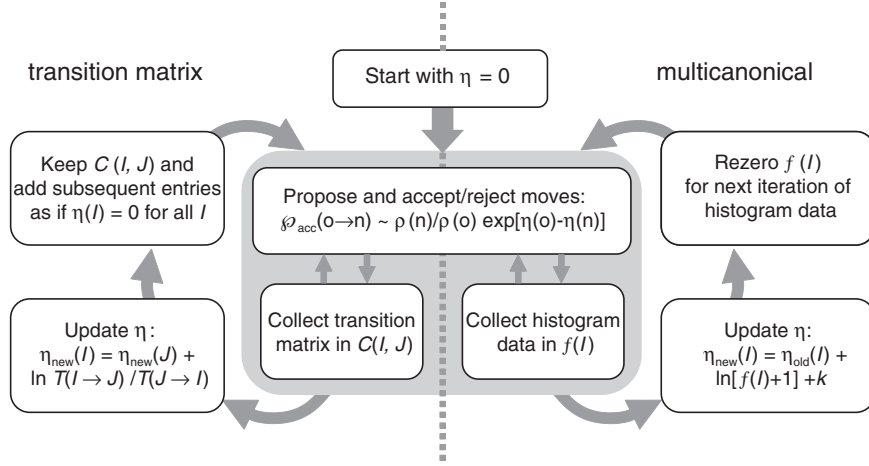
Fig. 3.7. Schematic comparison of the multicanonical and iterative transition-matrix methods. The *light gray box* indicates a single iteration of several thousand or more individual Monte Carlo steps. The acceptance criterion includes the configuration-space density of the original ensemble, ρ, and is presented for symmetric moves such as single-particle displacements

$$\mathscr{S}(I) - \mathscr{S}(J) = \ln T_{\mathrm{prop}}(J \to I) - \ln T_{\mathrm{prop}}(I \to J), \qquad (3.68)$$

where T_{prop} is the proposal probability. This expression is applicable to symmetric moves, such as single-particle displacements in continuum systems and single-spin flips in lattice models. A typical proposal transition matrix in potential energy for single-particle displacements is shown in Fig. 3.8.

An iteration process similar to the multicanonical approach can also be applied to density-of-states transition-matrix calculations; the function $\mathscr{S}$ is updated from the proposal probabilities over several iterations. In the density-of-states formulation, there is no need to 'adjust' the additions to C to be congruent with the initial ensemble, since the move proposal probabilities are naturally ensemble independent. In practice, however, the actual convergence of this iterative procedure can become quite slow when one is trying to sample large ranges of energy. The problem is that, for each iteration, a large number of MC steps are required to reach previously unexplored states, which means that numerous iterations are necessary in order to generate an initial estimate of $\mathscr{S}$ over the macrostate range of interest. To sidestep this difficulty, Shell et al. [28] suggested the addition of dynamic WL modification of $\mathscr{S}$ in the early stages of the simulation, with the transition-matrix calculations taking over once an initial estimate is determined.

In some cases, the data in a single transition matrix provides more thermodynamic estimates than there are unknowns. This routinely occurs when one examines transition probabilities among potential energy levels, since each energy is usually connected to a handful of neighbors rather than two. The simplest approach is to consider only immediate neighbors, and to discard the information provided by transitions between energy levels which are not consecutive. However, Wang
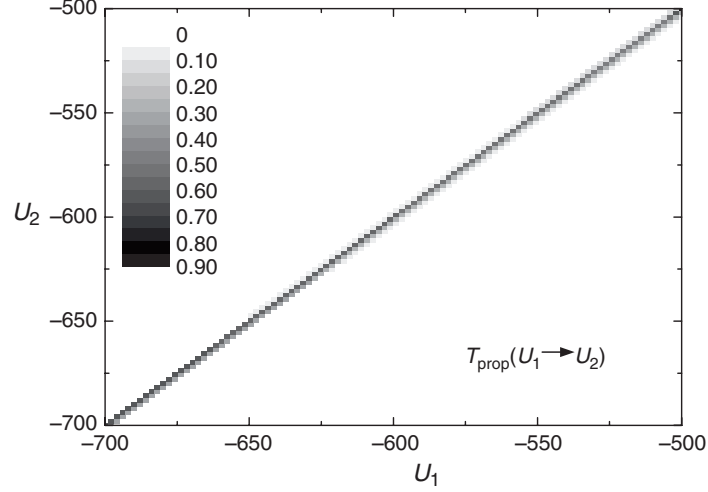
Fig. 3.8. Transition matrix of move proposal probabilities for the Lennard-Jones fluid at $\rho = 0.88$ and $N = 110$. The energy range of -700 to -500 in Lennard-Jones units has been discretized into 100 bins. Due to the adjustment of the random displacement moves to achieve 50% acceptance, the transition probabilities are highly banded. The tuned moves change the potential energy by only a small amount, and as a result, each energy level is effectively only connected to a few neighbors

and Swendsen suggested incorporating the complete set of data by minimizing the variance among the predicted values of $\ln \wp(I)$ [26], where the variance is defined by

$$\sigma_{\text{tot}}^2 = \sum_{I,J} \frac{[\ln \widetilde{\wp}(I) - \ln \widetilde{\wp}(J) + \ln \widetilde{T}(I \to J) - \ln \widetilde{T}(J \to I)]^2}{\sigma_{IJ}^2}, \qquad (3.69)$$

where σ_{IJ}^2 is the variance of the single prediction for $[\ln \widetilde{\wp}(I) - \ln \widetilde{\wp}(J)]$. The individual variances might be estimated using the number of observed transitions; in the density-of-states version, for example, one might assume $\text{var}\,[C(I,J)] \propto C(I,J)$, which yields

$$\sigma_{IJ}^2 = C(I,J)^{-1} + f(I)^{-1} + C(J,I)^{-1} + f(J)^{-1}, \qquad (3.70)$$

where f contains the histogram counts. Once a form for σ_{IJ} has been selected, one value of $\ln \widetilde{\wp}$ must be fixed and the remainder are determined from minimization of (3.69). Thermodynamic estimates then follow from the ensemble-specific relevance of $\ln \widetilde{\wp}$.

Transition-matrix estimators are typically more accurate than their histogram counterparts [25, 26, 46], and they offer greater flexibility in accumulating simulation data from multiple state conditions. This statistical improvement over histograms is likely due to the 'local' nature of transition probabilities, which are more readily equilibrated than global measures such as histograms [25]. Fenwick and Escobedo

[50, 51] have shown that transition-matrix calculations are closely related to Bennet's acceptance ratio method [4], in which minimization of the statistical error is attempted in order to predict free energy differences. In general, these estimators are an excellent tool for the measurement of free energies and entropies in simulation, and can usually be appended to existing simulation methodologies with little overhead.

3.4.5 Implementation Issues

Flat-histogram methods are a significant departure from simulation in conventional ensembles, and this warrants additional discussion about the logistics of their implementation. We have already seen that it is necessary to determine the logarithm of partition functions rather than the partition functions themselves, owing to finite numerical precision. In the case of continuous macroscopic variables, such as the potential energy, we must also discretize our calculations for η, $\mathscr{S}$, or $\mathscr{F}$. This requires some consideration; smaller bin widths yield better approximations, but necessitate longer simulation times in order to sample each bin accurately. Treating the discretized function as an interpolation can sometimes lower the number of bins necessary, particularly in the WL algorithm [24]. However, such interpolation is often not compatible with transition-matrix calculations unless nontrivial accommodations are made to (3.61a) [28]. Fortunately with transition probabilities, the number of thermodynamic estimates scales as the square of the number of bins, which can help offset the statistical loss at high bin number if using an error-minimization scheme such as (3.69). Practically speaking, though, past work has suggested that a reasonable bin width for energy can be selected from simple consideration of the pair potential; in the Lennard-Jones case, values of 1 to 2ϵ have fared well [24, 27, 28, 30].

In conventional MC simulations, it is not uncommon for the range of move attempts to be varied to achieve a desired acceptance rate. For example, in single-particle displacement moves, $\delta r_{\max}$ is routinely tuned to obtain around 50% acceptance. During this dynamic adjusting, detailed balance is not rigorously satisfied; however, the distribution of sampled states in a conventional simulation is sharply peaked and fluctuations in $\delta r_{\max}$ are small. In a flat-histogram simulation, however, these fluctuations are large. It is imperative, therefore, to explicitly maintain constant $\delta r_{\max}$ after finding a suitable initial value.

In transition-matrix simulations, the memory demands can often grow quite large since the size of the C matrix grows as the square of the number of bins n in the parameter of interest. In the case where two parameters are studied, the matrix is four-dimensional and quickly becomes unwieldy. One way to minimize the memory requirements is to use a *neighbor* format for C, which takes advantage of the fact that most entries are concentrated along its diagonal elements. Rather than tabulate data in $C(1 \ldots n, 1 \ldots n)$, where n is the number of bins, one records $C'(1 \ldots n, -\Delta n \ldots 0 \ldots \Delta n)$, where the second index is measured relative to the first and Δn is the preset number of neighbors covered by the matrix, typically in the range 1–10. Many moves can be exactly accommodated by $\Delta n = 1$, including

single-particle additions and deletions when C' measures transitions between N values, and volume scaling moves of fixed ΔV for transitions between V values. For moves which take the system beyond Δn, one simply updates the diagonal element $C'(I, 0)$ instead [46].

3.5 Order Parameters, Reaction Coordinates, and Extended Ensembles

We can generalize the density of states to include arbitrary macroscopic parameters, which allows us to perform extended ensemble simulations in which we wish to generate flat histograms in particular variables. Such simulations are useful for studying the behavior of the system in specific states or along particular paths. We may, for example, wish to use some geometric order parameter to bias our system toward a particular configuration or to sample the complete reaction coordinate for chemical events or protein folding. Here, we will let ξ denote the variables associated with these additional parameters, and assume they are continuous functions of the configurational coordinates only. We will consider ξ simply as additional properties of the system, rather than as new variables which form a new set of generalized coordinates, but note that a more in-depth discussion of such issues can be found in Chap. 4. Let the functions $\hat{\xi}_i(\mathbf{q})$ return the value of ξ_i for a given configuration. Then, the extended density of states simply incurs additional delta functions in the configurational integral

$$\Omega_\xi(N, V, E, \boldsymbol{\xi}) = \int_{V^N} \delta[U(\mathbf{q}) - E] \left(\prod_i \delta[\hat{\xi}_i(\mathbf{q}) - \xi_i] \right) d\mathbf{q}, \qquad (3.71)$$

where $\boldsymbol{\xi}$ is vector notation for the collection of ξ_i. Notice that, if any of the additional coordinates are dimensional, the density of states will acquire the inverse units on account of the delta function. This is normal, as it is the quantity $\Omega_\xi dE d\boldsymbol{\xi}$ which is dimensionless and proportional to the number of microstates with E and $\boldsymbol{\xi}$. We can recover the usual density of states by integrating the extended one

$$\Omega(N, V, E) = \int \Omega_\xi(N, V, E, \boldsymbol{\xi}) d\boldsymbol{\xi}, \qquad (3.72)$$

where the integral is multidimensional for more than one coordinate. Similarly, we can define an extended canonical partition function by integrating Ω_ξ over the Boltzmann factor as

$$Q_\xi(N, V, T, \boldsymbol{\xi}) = \Lambda^{-3N} \int e^{-\beta E} \Omega_\xi(N, V, E, \boldsymbol{\xi}) dE, \qquad (3.73)$$

where likewise:

$$Q(N, V, T) = \int Q_\xi(N, V, T, \boldsymbol{\xi}) d\boldsymbol{\xi}. \qquad (3.74)$$

Notice that Q_ξ also assumes the inverse units of any dimensional coordinates.

Associated with both the extended density of states and canonical partition function are the *coordinate* entropy S_ξ and free energy A_ξ, respectively. Often A_ξ is termed a generalized potential of mean force. Much of the same discussion regarding flat-histogram methods and free-energy estimators in the last section also applies to these extended thermodynamic functions. We simply need to introduce additional independent variables in our calculations of η, $\mathscr{S}$, or $\mathscr{F}$, and realize that, in order to obtain conventional thermodynamic results, we must integrate over these coordinates as in (3.72) and (3.74). Let us say we wish to examine the free energy along some reaction pathway ξ at a given temperature. The extended canonical partition function in (3.74) is our starting point, as it tell us that the probability of observing a specific point in the pathway is proportional to $Q_\xi(N, V, T, \xi)$. In order to generate a flat histogram along this coordinate, we must bias our microstate probabilities by the inverse factor, as

$$\wp(\mathbf{q}) \propto \exp\left[-\beta U - \mathscr{F}_\xi(\xi)\right], \tag{3.75}$$

which leads to

$$\wp(\xi) \propto Q_\xi(N, V, T, \xi) \exp\left[-\mathscr{F}_\xi(\xi)\right], \tag{3.76}$$

where $\mathscr{F}_\xi \equiv \ln Q_\xi = -\beta A_\xi$ for the ideal flat histogram. We can then use any of the previous techniques for determining $\mathscr{F}_\xi$. We could, for example, use WL modification to generate an estimate as Calvo has done [52], in which case the relevant histograms for the dynamic updating and the changing of the modification factor would be along the coordinate ξ. Alternatively, we could employ a transition-probability estimator

$$\mathscr{F}_\xi(\xi_1) - \mathscr{F}_\xi(\xi_2) = \ln T(\xi_2 \to \xi_1) - \ln T(\xi_1 \to \xi_2), \tag{3.77}$$

which is a direct extension of (3.60). Once we have performed a flat-histogram simulation and converged $\mathscr{F}_\xi$, normal thermodynamic quantities are recovered using (3.74). The energy is given by $\langle U \rangle (T) = -\partial \ln Q/\partial \beta$, and combining this expression with (3.74) and (3.76) yields

$$\langle U \rangle (T) = \frac{\displaystyle\int \langle U \rangle (T; \xi) \widetilde{\wp}(\xi) e^{\mathscr{F}_\xi(\xi)} \mathrm{d}\xi}{\displaystyle\int \widetilde{\wp}(\xi) e^{\mathscr{F}_\xi(\xi)} \mathrm{d}\xi}, \tag{3.78}$$

where $\langle U \rangle (T; \xi)$ is the average potential energy at a given coordinate value, which is easily measured in the simulation. Furthermore, if $\mathscr{F}_\xi$ is determined to sufficient precision, we can drop the measured $\widetilde{\wp}(\xi)$ from this expression.

It is also possible to build the order parameter directly into the potential energy function. This would enable one to determine the free energy difference between a given potential and a reference potential. In a sense, this is complementary to the previous discussion of reaction coordinates; rather than further subdividing configuration space by various order parameters, we are considering multiple configuration

spaces, each defined by a different potential. We again take ξ as our additional coordinate, only this time we let $U_\xi(\mathbf{q})$ be the energy of a configuration evaluated using the ξ version of the potential energy function. A simple choice might be $U_\xi(\mathbf{q}) = (1-\xi)U_1(\mathbf{q}) + \xi U_2(\mathbf{q})$, which is a linear interpolation between the two functions U_1 and U_2. The microscopic sampling is identical to (3.75), and so the logistics of the simulation follow the same considerations outlined earlier. Once $\mathscr{F}_\xi(\xi)$ has been calculated, it provides an estimate of the free energy between the different potentials at a given T.

One advantage offered by these expanded-ensemble techniques is greater flexibility in studying metastable phenomena. Although a formal treatment of metastability is beyond the scope of this chapter (the interested reader is referred to [53]), it is useful to mention the practical simulation aspects of it. Generally speaking, a metastable state is one that persists for an observable amount of time, but which is not the true thermodynamically stable state at the conditions of interest. Classic examples include superheated or supercooled liquids. Microscopically, the partition function corresponding to a metastable state must exclude those configurations which do not correspond to it. We write this as

$$Q'(N, V, T) = \frac{1}{\Lambda^{3N} N!} \int_{\Gamma_0} e^{-\beta U} d\mathbf{q}, \qquad (3.79)$$

where Γ_0 is the metastable state's subset of configuration space. Equation (3.79) shows that the metastable free energy will always be greater than the associated stable one, because the integral must be less for the restricted space (the exponential is always positive) [19]. The essential difficulty in metastability is the determination of precise and nonheuristic methods for identifying which configurations belong and which do not. In fact, it is this element which creates problems when one wants to rigorously define metastable properties such as spinodals [53]. The precise microscopic definition of metastable states can be a deep and often ill-defined issue. For this discussion, we will skip over this subtle difficulty. We assume that we have found a specific microscopic function $\hat{\xi}(\mathbf{q})$ which is able to make the distinction between metastable and stable configurations, for instance by assigning $\xi > \xi_0$ to the metastable state. Then, a flat-histogram simulation in the extended parameter ξ will cast our results in a form relevant to studying the metastable averages. For example, the average energy of a metastable state could be calculated as

$$\langle U \rangle (T) = \frac{\displaystyle\int_{\xi > \xi_0} \langle U \rangle (T; \xi) \widetilde{\wp}(\xi) e^{\mathscr{F}_\xi(\xi)} d\xi}{\displaystyle\int_{\xi > \xi_0} \widetilde{\wp}(\xi) e^{\mathscr{F}_\xi(\xi)} d\xi}, \qquad (3.80)$$

where the only differences with (3.78) are the limits on the integral. In other words, we are limiting the contribution of specific configurations to the partition function through selecting classes of configurations as defined by the order parameter. The order parameter does not need to be exotic; for example, averages restricted to

specific densities are frequently encountered when generating single-phase properties at points of liquid–vapor phase coexistence.

References

1. McDonald, I. R.; Singer, K., Machine calculation of thermodynamic properties of a simple fluid at supercritical temperatures, *J. Chem. Phys.* 1967, *47*, 4766–4772
2. Wood, W. W., Monte Carlo calculations for hard disks in the isothermal–isobaric ensemble, *J. Chem. Phys.* 1968, *48*, 415–434
3. Card, D. N.; Valleau, J. P., Monte Carlo study of the thermodynamics of electrolyte solutions, *J. Chem. Phys.* 1970, *52*, 6232–6240
4. Bennett, C. H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comp. Phys.* 1976, *22*, 245–268
5. Ferrenberg, A. M.; Swendsen, R. H., New Monte Carlo technique for studying phase transitions, *Phys. Rev. Lett.* 1988, *61*, 2635–2638
6. Berg, B. A.; Neuhaus, T., Multicanonical ensemble: a new approach to simulate first-order phase transitions, *Phys. Rev. Lett.* 1992, *68*, 9–12
7. Wilding, N. B., Critical-point and coexistence-curve properties of the Lennard-Jones fluid: a finite-size scaling study, *Phys. Rev. E* 1995, *52*, 602–611
8. Panagiotopoulos, A. Z.; Wong, V.; Floriano, M. A., Phase equilibria of lattice polymers from histogram reweighting Monte Carlo simulations, *Macromolecules* 1998, *31*, 912–918
9. Wilding, N. B., Critical end point behavior in a binary fluid mixture, *Phys. Rev. E* 1997, *55*, 6624–6631
10. Wilding, N. B.; Schmid, F.; Nielaba, P., Liquid–vapor phase behavior of a symmetrical binary fluid mixture, *Phys. Rev. E* 1998, *58*, 2201–2212
11. Potoff, J. J.; Panagiotopoulos, A. Z., Critical point and phase behavior of the pure fluid and a Lennard-Jones mixture, *J. Chem. Phys.* 1998, *109*, 10914–10920
12. Ferrenberg, A. M.; Swendsen, R. H., Optimized Monte Carlo data analysis, *Phys. Rev. Lett.* 1989, *63*, 1195–1198
13. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted histogram analysis method for free-energy calculations on biomolecules, *J. Comput. Chem.* 1992, *13*, 1011–1021
14. Frenkel, D.; Smit, B., *Understanding Molecular Simulation*, [2nd edition]
15. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free energy estimation: umbrella sampling, *J. Comput. Phys.* 1977, *23*, 187–199
16. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., Multidimensional free-energy calculations using the weighted histogram analysis method, *J. Comput. Chem.* 1995, *16*, 1339–1350
17. Bartels, C.; Karplus, M., Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations, *J. Comput. Chem.* 1997, *18*, 1450–1462
18. Kobrak, M. N., Systematic and statistical error in histogram-based free energy calculations, *J. Comput. Chem.* 2003, *24*, 1437–1446
19. Hill, T. L., *An Introduction to Statistical Thermodynamics*, Dover: New York, 1986
20. Lee, J., New Monte Carlo algorithm: entropic sampling, *Phys. Rev. Lett.* 1993, *71*, 211–214

21. Wang, F.; Landau, D. P., Efficient, multiple-range random walk algorithm to calculate the density of states, *Phys. Rev. Lett.* 2001, *86*, 2050–2053
22. Wang, F.; Landau, D. P., Determining the density of states for classical statistical models: a random walk algorithm to produce a flat histogram, *Phys. Rev. E* 2001, *64*, 056101
23. Torrie, G. M.; Valleau, J. P., Monte Carlo free energy estimates using non-Boltzmann sampling: application to the sub-critical Lennard-Jones fluid, *Chem. Phys. Lett.* 1974, *28*, 578–581
24. Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z., Generalization of the Wang–Landau method for off-lattice simulations, *Phys. Rev. E* 2002, *90*, 056703
25. Smith, G. R.; Bruce, A. D., A study of the multicanonical Monte Carlo method, *J. Phys. A* 1995, *28*, 6623–6643
26. Wang, J. S.; Swendsen, R. H., Transition matrix Monte Carlo method, *J. Stat. Phys.* 2001, *106*, 245–285
27. Yan, Q.; de Pablo, J. J., Fast calculation of the density of states of a fluid by Monte Carlo simulations, *Phys. Rev. Lett.* 2003, *90*, 035701
28. Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z., An improved Monte Carlo method for direct calculation of the density of states, *J. Chem. Phys.* 2003, *119*, 9406–9411
29. Berg, B. A.; Celik, T., New approach to spin-glass simulations, *Phys. Rev. Lett.* 1992, *69*, 2292–2295
30. Yan, Q.; Faller, R.; de Pablo, J. J., Density-of-states Monte Carlo method for simulation of fluids, *J. Chem. Phys.* 2002, *116*, 8745–8749
31. Lyubartsev, A.; Martsinovski, A.; Shevkunov, S., New approach to Monte Carlo calculation of the free energy: method of expanded ensembles, *J. Chem. Phys.* 1992, *96*, 1776–1783
32. Marinari, E.; Parisi, G., Simulated tempering: a new Monte Carlo scheme, *Europhys. Lett.* 1992, *19*, 451–458
33. Geyer, C. J.; Thompson, E. A., Annealing markov chain Monte Carlo with applications to ancestral inference, *J. Am. Stat. Soc.* 1995, *90*, 909–920
34. Faller, R.; Yan, Q. L.; de Pablo, J. J., Multicanonical parallel tempering, *J. Chem. Phys.* 2002, *116*, 5419–5423
35. Errington, J. R., Prewetting transitions for a model argon on solid carbon dioxide system, *Langmuir* 2004, *20*, 3798–3804
36. Yamaguchi, C.; Kawashima, N., Combination of improved multibondic method and the Wang–Landau method, *Phys. Rev. E.* 2002, *65*, 056710
37. Schulz, B. J.; Binder, K.; Muller, M., Flat histogram method of Wang–Landau and *n*-fold way, *Int. J. Mod. Phys. C* 2002, *13*, 477–494
38. Troyer, M.; Wessel, S.; Alet, F., Flat histogram methods for quantum systems: algorithms to overcome tunneling problems and calculate the free energy, *Phys. Rev. Lett.* 2003, *12*, 120201
39. Almarza, N. G.; Lomba, E., Determination of the interaction potential from the pair distribution function: an inverse Monte Carlo technique, *Phys. Rev. E* 2003, *68*, 011202
40. Butler, B. D.; Ayton, G.; Jepps, O. G.; Evans, D. J., Configurational temperature: verification of Monte Carlo simulations, *J. Chem. Phys.* 1998, *109*, 6519–6522
41. Zhou, C.; Bhatt, R. N., Understanding and improving the Wang–Landau algorithm, *Phys. Rev. E* 2005, *72*, 025701
42. Smith, G. R.; Bruce, A. D., Multicanonical Monte Carlo study of solid–solid phase coexistence in a model colloid, *Phys. Rev. E* 1996, *53*, 6530–6543
43. de Oliveira, P. M. C.; Penna, T. J. P.; Herrmann, H. J., Broad histogram method, *Braz. J. Phys.* 1996, *26*, 677–683

44. de Oliveira, P. M. C., Broad histogram relation is exact, *Eur. Phys. J. B* 1998, *6*, 111–115

45. Wang, J. S.; Tay, T. K.; Swendsen, R. H., Transition matrix Monte Carlo reweighting and dynamics, *Phys. Rev. Lett.* 1999, *82*, 476–479

46. Fitzgerald, M.; Picard, R. R.; Silver, R. N., Monte Carlo transition dynamics and variance reduction, *J. Stat. Phys.* 1999, *98*, 321–345

47. Fitzgerald, M.; Picard, R. R.; Silver, R. N., Canonical transition probabilities for adaptive metropolis simulation, *Eur. Phys. Lett.* 1999, *46*, 282–287

48. Errington, J.R., Direct calculation of liquid–vapor phase equilibria from transition matrix Monte Carlo simulation, *J. Chem. Phys.* 2003, *118*, 9915–9925

49. Errington, J. R., Evaluating surface tension using grand-canonical transition-matrix Monte Carlo simulation and finite-size scaling, *Phys. Rev. E* 2003, *67*, 012102

50. Fenwick, M. K.; Escobedo, F. A., Expanded ensemble and replica exchange methods for simulation of protein-like systems, *J. Chem. Phys.* 2003, *119*, 11998–12010

51. Fenwick, M. K.; Escobedo, F. A., On the use of Bennett's acceptance ratio method in multi-canonical-type simulations, *J. Chem. Phys.* 2004, *120*, 3066–3074

52. Calvo, F., Sampling along reaction coordinates with the Wang–Landau method, *Mol. Phys.* 2002, *100*, 3421–3427

53. Debenedetti, P. G., *Metastable Liquids: Concepts and Principles*, Princeton University Press: Princeton, NJ, 1996

4

Thermodynamic Integration Using Constrained and Unconstrained Dynamics

Eric Darve

4.1 Introduction

In Chaps. 2 and 3 we discussed two general approaches for calculating free energies: free energy perturbation and probability density (histogram) methods. In this chapter we introduce another, general approach, which relies on calculating and subsequently integrating the derivatives of the free energy with respect to an order parameter (or several order parameters) along a transformation path. Not surprisingly, this class of methods is called thermodynamic integration (TI).

The order parameter can be defined in two different ways. It can be either a function of atomic coordinates or just a parameter in the Hamiltonian. Examples of both types of order parameters are given in Sect. 2.8.1 in Chap. 2 and illustrated in Fig. 2.5. This distinction is theoretically important. In the first case, the order parameter is, in effect, a generalized coordinate, the evolution of which can be described by Newton's equations of motion. For example, in an association reaction between two molecules, we may choose as order parameter the distance between the two molecules. Ideally, we often would like to consider a reaction coordinate which measures the progress of a reaction. However, in many cases this coordinate is difficult to define, usually because it cannot be defined analytically and its numerical calculation is time consuming. This reaction coordinate is therefore often approximated by simpler order parameters.

In contrast, no equations of motion exist naturally for a parameter in the Hamiltonian. It is, however, possible to extend the formalism to include the dynamics of such a parameter. This approach goes back to Kirkwood [1].

The case of order parameters is more difficult to consider. However, since it is applicable to parameterized Hamiltonians as well, we will discuss TI in this broader setting.

Figure 4.1 represents the typical free energy profile or potential of mean force (PMF) along a reaction coordinate. The x-axis is the reaction coordinate, which could be the distance between two molecules, a torsion angle along the backbone of a protein, or the relative orientation of an α-helix with respect to a membrane.
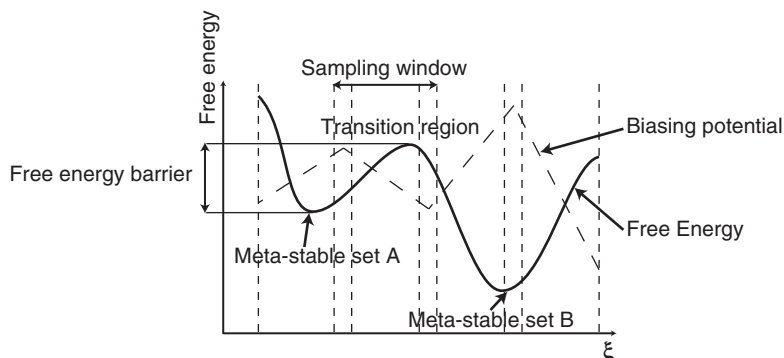
Fig. 4.1. Stratification consists of splitting the interval of interest into subintervals, thereby reducing the free energy barriers inside each window. The umbrella sampling method can bias the sampling and attempt to make it more uniform

The y-axis is the free energy. In general the free energy is related to the probability density function $P(\xi)$ of the reaction coordinate through

$$A(\xi) = -k_B T \ln P(\xi). \tag{4.1}$$

This equation means that when there is a free energy difference of a few $k_B T$ the probability $P(\xi)$ is reduced considerably, that is, those conformations with large $A(\xi)$ are sampled very rarely. This is a very important observation in terms of numerical efficiency. At the transition region for example, the free energy is maximum and typically very few sample points are obtained during the course of molecular dynamics simulation. In turn this results in very large statistical errors. Those errors can only be reduced by increasing the simulation time, sometimes beyond what is practically feasible.

A basic but powerful method to improve the efficiency of such computations is to split the interval of computation along the reaction coordinate ξ into subintervals, an approach termed stratification. In addition, in each window a biasing potential can be used in order to improve the sampling further: this is the umbrella sampling method [2–8], see also Sect. 2.2.1. In this way, the energy barriers inside each window can be made smaller. If this biasing potential is a function $U_b(\mathbf{x})$, the new Hamiltonian (or energy function) for the system is $\mathscr{H}(\mathbf{x}) + U_b(\mathbf{x})$. In general the biasing potential needs to be guessed beforehand or can be gradually improved using an iterative refinement process. This process consists of first running a short simulation to estimate the free energy, $A^{(0)}(\mathbf{x})$ and then using this estimate to bias the system using $U_b^{(0)}(\mathbf{x}) = -A^{(0)}(\mathbf{x})$. With this first bias, we improve the sampling and obtain a more accurate approximation of the free energy $A^{(1)}(\mathbf{x})$. The biasing potential $U_b^{(i)}(\mathbf{x})$ can be gradually improved in this fashion. Within each window, we obtain a relatively uniform sampling. This leads in general to small statistical errors.

It is clear, however, that in complex situations it is not possible to make an educated guess of the biasing potential. In particular the position of the transition regions

or the height of the barrier may be quite difficult to guess. There might be several intermediate states between A and B (see Fig. 4.1) corresponding to the local minima of the free energy. In addition, given the dependence of $P(\xi)$ on $A(\xi)$ [see (4.1)] even relatively small errors in the biasing potential can lead to regions of space which are not sampled adequately (for example, if one maximum in the free energy is not correctly estimated, see Fig. 3.2).

Many other approaches have been designed such as slow growth. In this approach, an external operator is exerting a force on the system such that ξ varies infinitely slowly from A to B. In the limit of infinitely slow speed, a classical result of statistical mechanics is that the work needed to change ξ is equal to the free energy difference ΔA. In practice, this works well if $\dot{\xi}$ is so small that the system stays close to equilibrium. An extension of this technique has been proposed by Jarzynski and coworkers [9–17] who provided an exact equation for the free energy difference for finite switching times. If $\dot{\xi}$ is not negligible, significant nonequilibrium effects will be present. These lead in general to a heating of the system and larger energies than usual. Despite this deviation from equilibrium statistics, the equation given by Jarzynski is exact. However, it has been observed by several authors [18] that, even though the equation is exact mathematically, this method leads to relatively large statistical errors which require very long simulation times to be reduced. For further discussion see the next chapter.

In this chapter, we focus on another class of techniques called TI. For a sufficiently smooth function, the free energy difference can always be written as the following integral $A(\xi_1) - A(\xi_0) = \int_{\xi_0}^{\xi_1} \mathrm{d}A/\mathrm{d}\xi \, \mathrm{d}\xi$. The key observation is that it is possible to calculate $\mathrm{d}A/\mathrm{d}\xi$ by recognizing that it is in fact equal to the following statistical average:

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \frac{\partial \mathscr{H}}{\partial \xi} \right\rangle_\xi . \tag{4.2}$$

The subscript ξ indicates that the average is computed for a fixed value of ξ, i.e., at a given point on our free energy plot (Fig. 4.1). Such an average corresponds very naturally to a 'generalized' force acting on the reaction coordinate ξ. Of course in general ξ is not a particle and therefore no 'real' mechanical force is acting on it. But if ξ is indeed a particle coordinate then this expression reduces to the mechanical force acting on this coordinate, i.e., $-\partial U/\partial \xi$, where U is the potential energy. Therefore this equation generalizes the notion of force to arbitrary variables which are functions of the atomic positions.

4.2 Methods for Constrained and Unconstrained Simulations

Several techniques are available to calculate $\langle \partial \mathscr{H}/\partial \xi \rangle$. Ciccotti and coworkers [17, 19–27] have developed a technique, called blue-moon ensemble method or the method of constraints, in which a simulation is performed with ξ fixed at some value. This can be realized by applying an external force, the constraint force, which prevents ξ from changing. From the statistics of this constraint force it is possible to

recover the derivative of the free energy $\mathrm{d}A/\mathrm{d}\xi$. In fact, the constraint force is of the form $\lambda\nabla\xi$, that is, it is pointing in the direction of the gradient of ξ. We will see later that $\langle\partial\mathcal{H}/\partial\xi\rangle \sim \langle\lambda\rangle$ (the actual equation is more complicated, as we will show). This means that the constraint force is a direct measurement of the derivative of the free energy.

An advantage of this technique is that it allows for getting as many sample points as needed at each location ξ along the interval of interest. In particular, it is possible to obtain very good statistics even in transition regions which are rarely visited otherwise. This leads in general to an efficient calculation and small statistical errors. Nevertheless, despite its many successes, this method has some difficulties. First, the system needs to be prepared such that ξ has the desired value (at which $\mathrm{d}A/\mathrm{d}\xi$ needs to be computed) and an equilibration run needs to be performed at this value of ξ. Second, it is not always obvious to determine how many quadrature points are needed to calculate the integral $\int_{\xi_0}^{\xi_1} \langle\partial\mathcal{H}/\partial\xi\rangle \, \mathrm{d}\xi$.

Finally, it may be difficult to sample all the relevant conformations of the system with ξ fixed. This problem is more subtle, but potentially more serious, as illustrated by Fig. 4.2. Several distinct pathways may exist between A and B. It is usually relatively easy for the molecule to enter one pathway or the other while the system is close to A or B. However, in the middle of the pathway, it may be very difficult to switch to another pathway. This means that, if we start a simulation with ξ fixed inside one of the pathway, it is very unlikely that the system will ever cross to explore conformations associated with another pathway. Even if it does, this procedure will likely lead to large statistical errors as the rate-limiting process becomes the transition rate between pathways inside the set $\xi = \mathrm{constant}$.

These difficulties can be circumvented by using the adaptive biasing force (ABF) method of Darve, Pohorille, and coworkers [18, 28, 29], which is based on unconstrained molecular dynamics simulations. This is a very efficient approach which begins by establishing a simple formula to calculate $\mathrm{d}A/\mathrm{d}\xi$ from regular molecular dynamics in which ξ is not constrained. This derivative represents the mean force acting on ξ. Therefore if we remove this force from the system we obtain
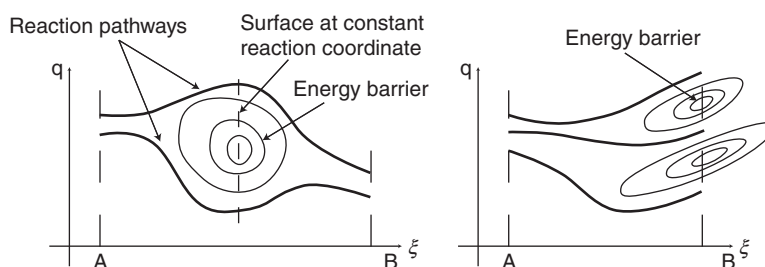
Fig. 4.2. Free energy computation using constraint forces. It may be difficult to sample the surface $\xi(\mathbf{x}) = \xi$ using a constrained simulation because of the presence of energy barriers separating different reaction pathways. *Left*: a barrier is shown in the middle of the pathway from reactant A to product B. *Right*: two barriers are shown at B

uniform sampling along ξ. This is done by adding an external, biasing force which is opposite to the one acting on ξ. In fact this biasing force is equal to $\mathrm{d}A/\mathrm{d}\xi\,\nabla\xi$. What is the resulting motion of the system? There is no average force acting on ξ. However, a nonzero fluctuating force with zero mean remains. Therefore, the dynamics of ξ is similar to a diffusive system or a random walk in 1D. Clearly the uniform sampling along ξ markedly reduces statistical error and yields an excellent convergence to the exact free energy profile.

Importantly, in contrast to constrained simulations, the system is allowed to evolve freely and in particular to explore the various pathways connecting A and B, see Fig. 4.2. This is one of the reasons why ABF can converge much faster than the method of constraints.

ABF shares some similarities with the technique of Laio et al. [30–34], in which potential energy terms in the form of Gaussian functions are added to the system in order to escape from energy minima and accelerate the sampling of the system. However, this approach is not based on an analytical expression for the derivative of the free energy but rather on importance sampling.

Organization of the Chapter

In this chapter, we focus on the method of constraints and on ABF. Generalized coordinates are first described and some background material is provided to introduce the different free energy techniques properly. The central formula for practical calculations of the derivative of the free energy is given. Then the method of constraints and ABF are presented. A newly derived formula, which is simpler to implement in a molecular dynamics code, is given. A discussion of some alternative approaches (steered force molecular dynamics [35–37] and metadynamics [30–34]) is provided. Numerical examples illustrate some of the applications of these techniques. We finish with a discussion of parameterized Hamiltonian functions in the context of alchemical transformations.

4.3 Generalized Coordinates and Lagrangian Formulation

In this section, we discuss some of the equations used to calculate the derivative of the free energy. In a different form, those results will be used in ABF to both calculate $\mathrm{d}A/\mathrm{d}\xi$ and bias the system in an adaptive manner.

4.3.1 Generalized Coordinates

To calculate $\mathrm{d}A/\mathrm{d}\xi$, we need to evaluate partial derivatives, such as $\partial\mathscr{H}/\partial\xi$, which measures the rate of change in energy with the order parameter. To do so we need to define generalized coordinates of the form $(\xi, q_1, \cdots, q_{N-1})$. Classical examples are spherical coordinates (r, θ, ϕ), cylindrical coordinates (r, θ, z) or polar coordinates in 2D. Those coordinates are necessary to form a full set that determines

the positions of all atoms. For example, given (r, θ, ϕ), we can calculate the Cartesian coordinates (x, y, z). In the context of protein modeling, order parameters are often defined as internal coordinates such as torsion angle, angle between three bonds or more generally groups of atoms, hydration number, gyration radius (distance between end points of a protein or polymer), etc.

Once an order parameter, ξ, is specified we can define the free energy A as a function of ξ through the relation $A(\xi) = -k_{\mathrm{B}}T \ln P(\xi)$. In the canonical ensemble, the probability density function $P(\mathbf{x}, \mathbf{p}_x)$ is equal to $\exp(-\mathscr{H}(\mathbf{x}, \mathbf{p}_x)/k_{\mathrm{B}}T)/Q$, where Q is the partition function:

$$Q = \int \exp(-\mathscr{H}(\mathbf{x}, \mathbf{p}_x)/k_{\mathrm{B}}T) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x.$$

The variables $\mathbf{x}$ and $\mathbf{p}$ are the positions and momenta of all the particles. With those definitions, it is possible to define $A(\xi)$ in terms of an integral in the phase space

$$A(\xi) = -k_{\mathrm{B}}T \, \ln \int \frac{\mathrm{e}^{-\beta\mathscr{H}}}{Q} \, \delta(\xi - \xi(\mathbf{x})) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x \qquad (4.3)$$

The Dirac delta function $\delta(\xi - \xi(\mathbf{x}))$ means that we are effectively integrating over all coordinates $\mathbf{x}$ such that $\xi(\mathbf{x}) = \xi$. In the rest of this chapter, since we will be interested in free energy differences only, we will omit the factor Q.

The delta function is not convenient to handle mathematically. However, if we define a set of generalized coordinates of the form $(\xi, q_1, \cdots, q_{N-1})$ and their associated momenta $(p^\xi, p_1^q, \cdots, p_{N-1}^q)$ then this integration simplifies to:

$$A(\xi) = -k_{\mathrm{B}}T \, \ln \int \mathrm{e}^{-\beta\mathscr{H}} \, \mathrm{d}_1^q \cdots \mathrm{d}_{N-1}^q \mathrm{d}p^\xi \cdots \mathrm{d}p_{N-1}^q. \qquad (4.4)$$

How are the generalized momenta defined? The easiest way to do so is to formulate our equations using the Lagrangian $\mathscr{L}$. Let us assume that a configuration of our mechanical system can be represented by a set of coordinates q. Then

$$\mathscr{L}(q, \dot{q}) = K(q, \dot{q}) - U(q, \dot{q}),$$

where K is the kinetic energy and U is the potential energy.

As an example, consider a simple system defined by two Cartesian coordinates x and y

$$\mathscr{L}(x, y, \dot{x}, \dot{y}) = \frac{1}{2} \, m \, \dot{x}^2 + \frac{1}{2} \, m \, \dot{y}^2 - U(x, y).$$

We may express the same Lagrangian using, for example, polar coordinates r and θ. The equation for $\mathscr{L}$ will then read

$$\mathscr{L}(r, \theta, \dot{r}, \dot{\theta}) = \frac{1}{2} \, m \left((\dot{r})^2 + (r\dot{\theta})^2\right) - U(r, \theta). \qquad (4.5)$$

In general, if we have $\mathbf{q} = \mathbf{q}(\mathbf{x})$ we can write

$$\dot{q}_i = \sum_j \frac{\partial q_i}{\partial x_j} \dot{x}_j \stackrel{\text{def}}{=} [J^{-1}(\mathbf{q})\dot{\mathbf{x}}]_i,$$

where $J(\mathbf{q})$ is the Jacobian matrix, $[J(\mathbf{q})]_{ij} = \partial x_i/\partial q_j$. The kinetic energy K can be expressed as

$$\sum_i \frac{1}{2}\, m_i \dot{x}_i^2 = \frac{1}{2}\, \dot{\mathbf{x}}^t \cdot M \cdot \dot{\mathbf{x}} \tag{4.6}$$

$$= \frac{1}{2}\, \dot{\mathbf{q}}^t \cdot [J^t(\mathbf{q})MJ(\mathbf{q})] \cdot \dot{\mathbf{q}}, \tag{4.7}$$

where M is a diagonal matrix with $M_{ii} = m_i$ and $\dot{\mathbf{x}}^t$ is the transpose of vector $\dot{\mathbf{x}}$. We denote by Z and M^{G} the following matrices:

$$Z \stackrel{\mathrm{def}}{=} J^{-1}M^{-1}(J^{-1})^t \tag{4.8}$$

$$M^{\mathrm{G}} \stackrel{\mathrm{def}}{=} Z^{-1} = J^t M J, \tag{4.9}$$

where Z_ξ is the first element in the matrix Z

$$Z_\xi = \sum_i \frac{1}{m_i} \left(\frac{\partial \xi}{\partial x_i} \right)^2. \tag{4.10}$$

For our 2D example we obtain for J and Z

$$J = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix},$$

$$Z = \frac{1}{m} \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{1}{r^2} \end{pmatrix}.$$

The kinetic energy is given by

$$K(\dot{r}, \dot{\theta}) = \frac{1}{2}\, (\dot{r}\ \dot{\theta})\ Z^{-1}\ \begin{pmatrix} \dot{r} \\ \dot{\theta} \end{pmatrix} = \frac{1}{2}\, m \left((\dot{r})^2 + (r\dot{\theta})^2 \right).$$

This is identical to (4.5).

For the general case of coordinates q, the generalized momenta are defined by

$$p_i = \frac{\partial \mathscr{L}}{\partial \dot{q}_i}.$$

If we use (4.7) and (4.8) to calculate p_i we obtain:

$$p_i = \frac{\partial}{\partial \dot{q}_i} \left(\frac{1}{2}\, \dot{\mathbf{q}}^t \cdot Z^{-1} \cdot \dot{\mathbf{q}} \right),$$

$$= \left[Z^{-1}\dot{\mathbf{q}} \right]_i.$$

The Hamiltonian in generalized coordinates is then defined as:

$$\mathscr{H}(\mathbf{q}, \mathbf{p}) = K(\mathbf{q}, \mathbf{p}) + U(\mathbf{q}, \mathbf{p}), \tag{4.11}$$

$$= \frac{1}{2}\, \mathbf{p}^t\, Z\, \mathbf{p} + U(\mathbf{q}, \mathbf{p}). \tag{4.12}$$

The equations of motion are given by

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial \mathscr{H}}{\partial p_i} = [Z\mathbf{p}]_i,$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \mathscr{H}}{\partial q_i} = -\frac{1}{2}\mathbf{p}^t \frac{\partial Z}{\partial q_i}\mathbf{p} - \frac{\partial U}{\partial q_i}.$$

Let us return to our 2D example to see how those equations are applied. The momenta are defined by

$$p_r = m\dot{r},$$

$$p_\theta = mr^2\dot{\theta},$$

where p_θ is the angular momentum of the particle. The equations of motion for the momenta are:

$$\frac{\mathrm{d}p_r}{\mathrm{d}t} = \frac{1}{m}\frac{1}{r^3}p_\theta^2 - \frac{\partial U}{\partial r} = mr\dot{\theta}^2 - \frac{\partial U}{\partial r},$$

$$\frac{\mathrm{d}p_\theta}{\mathrm{d}t} = -\frac{\partial U}{\partial \theta}.$$

It can be verified that $-\partial U/\partial \theta$ is the torque acting on the particle, and the second equation is the equation for the angular momentum.

Using the formalism we just developed for generalized coordinates we can now derive an expression for the derivative of the free energy. Let us differentiate (4.4) with respect to ξ

$$\boxed{\frac{\mathrm{d}A}{\mathrm{d}\xi} = \frac{\int \frac{\partial \mathscr{H}}{\partial \xi} \mathrm{e}^{-\beta\mathscr{H}} \, \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1} \mathrm{d}p^\xi \cdots \mathrm{d}p_{N-1}^q}{\int \mathrm{e}^{-\beta\mathscr{H}} \, \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1} \mathrm{d}p^\xi \cdots \mathrm{d}p_{N-1}^q} = \left\langle \frac{\partial \mathscr{H}}{\partial \xi} \right\rangle_\xi.} \qquad (4.13)$$

The angular brackets $\langle \cdots \rangle_\xi$ correspond to an average computed with ξ fixed. One way to calculate this average is to record the value of $\partial \mathscr{H}/\partial \xi$ every time $\xi(\mathbf{x})$ is equal to ξ during the trajectory, and then average the recorded values.

Equation (4.13) has a very important interpretation as the PMF exerted on ξ. Let us assume that ξ is one of the atomic coordinate, say x_1. Then

$$-\frac{\mathrm{d}A}{\mathrm{d}x_1} = -\left\langle \frac{\partial \mathscr{H}}{\partial x_1} \right\rangle_{x_1} = -\left\langle \frac{\partial U}{\partial x_1} \right\rangle_{x_1} = \langle F_1 \rangle_{x_1},$$

where F_1 is the force on the atomic coordinate x_1. $-\mathrm{d}A/dx_1$ is the force acting on x_1 averaged over all other variables and $A(x_1)$ can be thought of as the mean potential or PMF for x_1.

Let us denote $P_0(x_1) = \exp(-\beta A)/Z$, the probability density function of x_1. Assume now that one simulates a system at temperature T with a single coordinate

x_1 that satisfies the equation of motion $m_1 d^2 x_1/\mathrm{d}t^2 = -\mathrm{d}A/dx_1$. Then the probability density for x_1 at equilibrium will indeed be equal to $P_0(x_1)$. Thus the thermodynamic information about the variable x_1, considered by itself, is completely defined by the function $A(x_1)$.

More generally, ξ can be any function of atomic positions. Therefore, $\mathrm{d}A/\mathrm{d}\xi$ is not necessarily a force. Its interpretation, however, remains the same, but in a more general context: $-\mathrm{d}A/\mathrm{d}\xi$ is the mean force exerted on the generalized 'particle' ξ.

Computing the partial derivative of $\mathscr{H}$ with respect to ξ is in general a tedious process. A complete set of generalized coordinates needs to be defined. The correct expression for the kinetic energy in terms of generalized momenta need to be obtained. It involves a dense matrix which needs to be differentiated with respect to ξ. A great simplification can be obtained by showing that the derivative of the free energy can be written as

$$\boxed{\left\langle \frac{\partial \mathscr{H}}{\partial \xi} \right\rangle_\xi = \left\langle \frac{\partial U}{\partial \xi} - k_\mathrm{B}T \frac{\partial \ln|J|}{\partial \xi} \right\rangle_\xi .} \tag{4.14}$$

The term $|J|$ is the determinant of the Jacobian matrix upon changing from Cartesian to generalized coordinates. It measures the change in volume element between $\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_x$ and $\mathrm{d}\xi\,\mathrm{d}\mathbf{q}\,\mathrm{d}p^\xi\,\mathrm{d}\mathbf{p}_q$. For example in polar coordinate $|J| = r$ and therefore $\mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta$. The derivative of A is therefore the sum of two contributions: the mechanical forces acting along ξ ($\partial U/\partial \xi$), and the change of volume element. The term $-1/\beta\,\partial \ln|J|/\partial \xi$ is effectively an entropic contribution.

Let us assume that $\xi = r$ and we are using polar coordinates (r, θ). In this case, $|J| = r$. The formula therefore reads

$$\left\langle \frac{\partial \mathscr{H}}{\partial r} \right\rangle_r = \left\langle \frac{\partial U}{\partial r} - \frac{1}{\beta r} \right\rangle_r .$$

As we explained earlier, the difficulty in this formulation is that generalized coordinates appear explicitly in the form of the Jacobian $|J|$, which may be difficult to calculate in many cases. It is therefore desirable to find an expression in which the explicit knowledge of the generalized coordinates is not required. This is done in Sect. 4.4. We end this section with a proof of (4.14).

Proof. We start from the following equation for $A(\xi)$ expressed in terms of the position $\mathbf{x}$ of all the atoms only (no momenta):

$$A(\xi) = -k_\mathrm{B}T\,\ln \int \mathrm{e}^{-\beta U}\,\delta(\xi - \xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}.$$

We introduce generalized coordinates

$$A(\xi) = -k_\mathrm{B}T\,\ln \int \mathrm{e}^{-\beta U}|J|\,\mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}.$$

The Jacobian J appears because we have changed the variables of integration. We now differentiate with respect to ξ:

$$\frac{\mathrm{d}A(\xi)}{\mathrm{d}\xi} = \frac{\int \frac{\partial}{\partial \xi}\left(-k_{\mathrm{B}}T\,\mathrm{e}^{-\beta U}\,|J|\right)\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}{\int \mathrm{e}^{-\beta U}|J|\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}$$

$$= \frac{\int \left(\frac{\partial U}{\partial \xi}|J| - k_{\mathrm{B}}T\frac{\partial|J|}{\partial\xi}\right)\mathrm{e}^{-\beta U}\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}{\int \mathrm{e}^{-\beta U}\,\delta(\xi - \xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}}.$$

We can change the variables back to Cartesian coordinates and we obtain

$$\frac{\mathrm{d}A(\xi)}{\mathrm{d}\xi} = \frac{\int \left(\frac{\partial U}{\partial \xi} - k_{\mathrm{B}}T\frac{1}{|J|}\frac{\partial|J|}{\partial\xi}\right)\mathrm{e}^{-\beta U}\,\delta(\xi - \xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}}{\int \mathrm{e}^{-\beta U}\,\delta(\xi - \xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}}$$

$$= \frac{\int \left(\frac{\partial U}{\partial \xi} - k_{\mathrm{B}}T\frac{\partial \ln|J|}{\partial\xi}\right)\mathrm{e}^{-\beta U}\,\delta(\xi - \xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}}{\int \mathrm{e}^{-\beta U}\,\delta(\xi - \xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}}$$

$$= \left\langle \frac{\partial U}{\partial \xi} - k_{\mathrm{B}}T\frac{\partial \ln|J|}{\partial\xi}\right\rangle_\xi.$$

We have proved (4.14).

4.4 The Derivative of the Free Energy

As we pointed out earlier, calculating the derivative of the free energy appears to require a full set of generalized coordinates. However, this may seem quite surprising. Assume that we want to define the PMF as a function of the distance between two molecules. This force is clearly independent of the particular choice of generalized coordinates made to calculate it. In fact, we are now going to prove that an equation can be derived which does not require an explicit definition of generalized coordinates other than ξ.

For an arbitrary vector field $\mathbf{w}$ such that $\mathbf{w} \cdot \nabla\xi \neq 0$

$$\boxed{\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla\xi} - k_{\mathrm{B}}T\nabla \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla\xi}\right\rangle_\xi.} \tag{4.15}$$

This equation was also derived by den Otter [25] and Ciccotti et al. [38, 39]. It is not obvious at first sight that this formula should hold for an arbitrary vector field. We will come back to this point later on. Different convenient choices for $\mathbf{w}$ are possible

$$w_i = \frac{\partial \xi}{\partial x_i} \quad \text{or} \quad w_i = \frac{1}{m_i}\frac{\partial \xi}{\partial x_i}.$$

The second choice will turn out to be the more natural one as it leads to a direct interpretation in terms of the rate of change of the momentum p^ξ.

We prove this equation in the following proof. In Sect. 4.4.1, we discuss this fundamental equation further.

Proof of (4.15). As usual in TI, we try to compute the derivative of A rather than A itself. By differentiating directly equation (4.3) rather than introducing generalized coordinates we obtain:

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = -k_\mathrm{B}T\, \frac{\displaystyle\int \mathrm{e}^{-\beta U}\, \delta'(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}}{\displaystyle\int \mathrm{e}^{-\beta U}\, \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}}. \tag{4.16}$$

In order to simplify this expression and get rid of the δ' we are going to do an integration by parts. In the previous expression the derivative is with respect to ξ. To obtain derivatives with respect to $\mathbf{x}$ we use the chain rule of differentiation:

$$\frac{\mathrm{d}\,\delta(f(\mathbf{x}))}{\mathrm{d}\mathbf{x}} = f'(\mathbf{x})\, \delta'(f(\mathbf{x})).$$

Applying this equation to our case, we obtain:

$$\frac{\partial\,\delta(\xi - \xi(\mathbf{x}))}{\partial x_i} = -\frac{\partial \xi}{\partial x_i}\, \delta'(\xi - \xi(\mathbf{x})). \tag{4.17}$$

Let us assume now that we have a vector field $\mathbf{w}$ such that $\mathbf{w}\cdot\nabla\xi \neq 0$. Equation (4.17) leads to

$$(\mathbf{w} \cdot \nabla)\, \delta(\xi - \xi(\mathbf{x})) = -(\mathbf{w} \cdot \nabla\xi)\, \delta'(\xi - \xi(\mathbf{x})),$$

from which, after dividing by $-w \cdot \nabla\xi$

$$\delta'(\xi - \xi(\mathbf{x})) = -\left(\frac{\mathbf{w}}{\mathbf{w} \cdot \nabla\xi}\right) \cdot \nabla\, \delta(\xi - \xi(\mathbf{x})).$$

Inserting this result in (4.16), we obtain for the numerator

$$\int \mathrm{e}^{-\beta U}\, \delta'(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x} = -\int \mathrm{e}^{-\beta U}\, \left(\frac{\mathbf{w}}{\mathbf{w} \cdot \nabla\xi}\right) \cdot \nabla\, \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}.$$

We can now integrate by parts

$$\int \mathrm{e}^{-\beta U}\, \delta'(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x} = -\int \nabla\left(\mathrm{e}^{-\beta U}\, \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla\xi}\right) \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}.$$

The divergence can be computed using the product rule to obtain

$$
\nabla \left(e^{-\beta U} \, \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} \right) = e^{-\beta U} \left(-\beta \frac{\nabla U \cdot \mathbf{w}}{\mathbf{w} \cdot \nabla \xi} + \nabla \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} \right).
$$

For example, if $\mathbf{w} = \nabla \xi$ we obtain the following expression:

$$
e^{-\beta U} \left(-\beta \frac{\nabla U \cdot \nabla \xi}{|\nabla \xi|^2} + \nabla \cdot \frac{\nabla \xi}{|\nabla \xi|^2} \right).
$$

We have a new expression for $\mathrm{d}A/\mathrm{d}\xi$ which is valid for any choice of $\mathbf{w}$ satisfying the condition mentioned above

$$
\frac{\mathrm{d}A}{\mathrm{d}\xi} = -k_{\mathrm{B}} T \, \frac{\displaystyle\int e^{-\beta U} \left(-\beta \frac{\nabla U \cdot \mathbf{w}}{\mathbf{w} \cdot \nabla \xi} + \nabla \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} \right) \delta(\xi - \xi(\mathbf{x})) \, \mathrm{d}\mathbf{x}}{\displaystyle\int e^{-\beta U} \, \delta(\xi - \xi(\mathbf{x})) \, \mathrm{d}\mathbf{x}}.
$$

Using the notation $\langle \, \rangle$ we get (4.15)

$$
\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} - k_{\mathrm{B}} T \nabla \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} \right\rangle_{\xi}.
$$

4.4.1 Discussion of (4.15)

To give an example with a specific choice of $\mathbf{w}$, consider $\mathbf{w} = \nabla \xi$. Then

$$
\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \frac{1}{|\nabla \xi|^2} \left(\nabla U \cdot \nabla \xi - k_{\mathrm{B}} T \left(\nabla^2 \xi - 2 \frac{\nabla \xi \cdot \mathbf{H}(\xi) \cdot \nabla \xi}{|\nabla \xi|^2} \right) \right) \right\rangle_{\xi}, \qquad (4.18)
$$

where $\mathbf{H}(\xi)$ is the Hessian matrix of ξ:

$$
\mathbf{H}(\xi)_{ij} = \frac{\partial^2 \xi}{\partial x_i \partial x_j},
$$

and $\nabla^2 \xi$ is the Laplacian of ξ

$$
\nabla^2 \xi = \sum_i \frac{\partial^2 \xi}{\partial x_i^2}.
$$

As previously observed, the choice of $\mathbf{w}$ is to some extent arbitrary. This can be understood when one realizes that there is a direct connection between the choice of generalized coordinates in (4.14) and the choice of $\mathbf{w}$. Just as the choice of $(q_1, \ldots, q_{N-1})$ is arbitrary, so is the choice of $\mathbf{w}$.

The connection can be made by defining

$$
w_i = \frac{\partial x_i}{\partial \xi}.
$$

With this choice, the condition $\mathbf{w} \cdot \nabla \xi \neq 0$ is trivially satisfied since

$$\mathbf{w} \cdot \nabla \xi = \sum_i \frac{\partial x_i}{\partial \xi} \frac{\partial \xi}{\partial x_i} = 1.$$

We are now going to show that (4.15) is in fact identical to (4.14), which involved the Jacobian J.

Let us consider first $\nabla U \cdot (\mathbf{w}/(\mathbf{w} \cdot \nabla \xi))$ in (4.15)

$$\nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} = \sum_i \frac{\partial U}{\partial x_i} \frac{\partial x_i}{\partial \xi} = \frac{\partial U}{\partial \xi}.$$

This is the same term as in (4.14). The second term is equal to

$$\begin{aligned}
\frac{\partial \ln |J|}{\partial \xi} &= \mathrm{Tr}\left(J^{-1} \frac{\partial J}{\partial \xi} \right) \\
&= \sum_{ij} \frac{\partial q_i}{\partial x_j} \frac{\partial}{\partial \xi} \left[\frac{\partial x_j}{\partial q_i} \right] \\
&= \sum_{ij} \frac{\partial q_i}{\partial x_j} \frac{\partial}{\partial q_i} \left[\frac{\partial x_j}{\partial \xi} \right] \\
&= \sum_j \frac{\partial}{\partial x_j} \left(\frac{\partial x_j}{\partial \xi} \right) = \nabla \mathbf{w}.
\end{aligned}$$

Thus we have proved that

$$\boxed{\frac{\partial U}{\partial \xi} - k_{\mathrm{B}} T \frac{\partial \ln |J|}{\partial \xi} = \nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} - k_{\mathrm{B}} T \nabla \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi}}$$

for $w_i = \partial x_i / \partial \xi$.

The practical importance of this result is that choosing $\mathbf{w}$ can be done quite easily without explicitly defining a set of generalized coordinates, e.g., $w_i = \partial \xi / \partial x_i$ or $w_i = 1/m_i \, \partial \xi / \partial x_i$.

Let us consider an example with a specific choice of reaction coordinate. If we choose, $\xi = r$ and $\mathbf{w} = \nabla \xi$ we recover

$$\nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} - k_{\mathrm{B}} T \nabla \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} = \frac{\partial U}{\partial r} - \frac{1}{\beta r}.$$

4.5 The Potential of Mean Constraint Force

Having found a more convenient expression to calculate $\mathrm{d}A/\mathrm{d}\xi$ what remains to be developed is an algorithm which overcomes the sampling difficulty associated with high energy barriers. There are two approaches to this problem. The first one, called

the method of constraints [19, 27], consists of imposing a constraint force on the system such that ξ is constant. We will show that the average of this constraint force is related to $dA/d\xi$ through a simple equation. The procedure is then as follows. We discretize the interval of interest $[\xi_0, \xi_1]$ into a number of quadrature points. At each point, we run a constraint simulation from which $dA/d\xi$ can be computed. A quadrature can then be used to estimate $\Delta A = \int_{\xi_0}^{\xi_1} dA/d\xi \, d\xi$.

The second approach, ABF, consists of estimating the average force at different ξ and removing it from the system. This leads to a diffusive-like motion along ξ and a much better convergence of the calculation. We will describe ABF in Sect. 4.6.

4.5.1 Constrained Simulation

In the method of constraints, a force of the form $\lambda \nabla \xi$ is applied at each step such that ξ remains constant throughout the simulation. To see how λ can be related to $dA/d\xi$, we recall that the free energy A is the effective or average potential acting on ξ. From physical intuition, it should be true that $-dA/d\xi$ is the average of the force acting on ξ. In a constraint simulation, this force is equal to $-\lambda$. Therefore we can expect to have $dA/d\xi \sim \langle \lambda \rangle$. We now make this statement more rigorous.

In constrained simulations, the Hamiltonian $\mathscr{H}$ is supplemented by a Lagrange multiplier

$$\mathscr{H}^\lambda \stackrel{\text{def}}{=} \mathscr{H} + \lambda(\xi - \xi(\mathbf{x})) = K + U + \lambda(\xi - \xi(\mathbf{x})). \tag{4.19}$$

The additional term $\lambda(\xi - \xi(\mathbf{x}))$ is needed to enforce the constraint $\xi(\mathbf{x}) = $ constant. This corresponds to an additional force equal to $-\nabla(\lambda(\xi - \xi(\mathbf{x})))$. λ is chosen such that $\xi(\mathbf{x}) = \xi$. Therefore the force can be expressed more simply as $\lambda \nabla \xi$. Then, the interpretation is quite natural. In order to enforce the constraint we are applying a force parallel to $\nabla \xi$ which opposes the mechanical force acting on ξ.

The condition $\xi(\mathbf{x}) = $ constant implies in particular that $\dot{\xi} = 0$. With this condition it is possible to derive an expression for λ as a function of positions and velocities:

$$\begin{aligned}
\ddot{\xi} &= \frac{d}{dt}\dot{\xi} \\
&= \frac{d}{dt}(\nabla \xi \cdot \dot{\mathbf{x}}) \\
&= \nabla \xi \cdot \ddot{\mathbf{x}} + \frac{d\nabla \xi}{dt} \cdot \dot{\mathbf{x}} \\
&= \sum_i \frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \left(-\frac{\partial U}{\partial x_i} + \lambda \frac{\partial \xi}{\partial x_i} \right) + \dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}},
\end{aligned}$$

where we assumed that $\xi(\mathbf{x}) = \xi$. Solving for $\ddot{\xi} = 0$, we obtain

$$\lambda = Z_\xi^{-1} \left(\sum_i \frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \frac{\partial U}{\partial x_i} - \dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}} \right). \tag{4.20}$$

See (4.10) for the definition of Z_ξ. Therefore λ is in general a function of $\mathbf{x}$ and $\dot{\mathbf{x}}$. With this expression for λ, we see that if we start a simulation with $\xi = \xi(\mathbf{x})$ and $\dot{\xi} = 0$ then $\xi = \xi(\mathbf{x})$ at all times. Note that different methods, such as SHAKE or RATTLE [40], are employed numerically to avoid a 'drift' of ξ. Those methods are not based on a direct application of (4.20).

4.5.2 The Fixman Potential

Before we derive the appropriate expressions to calculate $\mathrm{d}A/\mathrm{d}\xi$ from constrained simulations, we note an important difference between sampling in constrained and unconstrained simulations. There are two ways to gather statistics at $\xi(\mathbf{x}) = \xi$. In unconstrained simulations, the positions are sampled according to $\exp -\beta U$ while the momenta are sampled according to $\exp -\beta K$. If a constraint force is applied to keep ξ fixed the positions are sampled according to $\delta(\xi(\mathbf{x}) - \mathbf{x}) \exp -\beta U$. The momenta, however, are sampled according to a more complex statistical ensemble. Recall that

$$p_\xi = (Z^{-1})_{\xi\xi}\dot{\xi} + \sum_{i \leq N-1} (Z^{-1})_{\xi i}\frac{\mathrm{d}q_i}{\mathrm{d}t}.$$

In constrained simulations, $\dot{\xi} = 0$ so that p_ξ is *not* an independent variable but rather a function of $\mathbf{q}$ and $\mathbf{p}_q$. Let us discuss the implications of this fact. Consider an arbitrary function $f(\mathbf{x})$ and the following average:

$$\frac{\displaystyle\int f(\mathbf{x})\mathrm{e}^{-\beta\mathscr{H}}\,\delta(\xi(\mathbf{x}) - \xi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_x}{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}}\,\delta(\xi(\mathbf{x}) - \xi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_x} \tag{4.21}$$

The Hamiltonian function of a system which is constrained with $\xi(\mathbf{x}) = \xi$ and $\dot{\xi} = 0$ is given by

$$H_\xi = \frac{1}{2}\,(\mathbf{p}_q)^t\,(M_q^{\mathrm{G}})^{-1}\,\mathbf{p}_q + U(\mathbf{q}), \tag{4.22}$$

where the matrix M_q^{G} is a submatrix of M^{G} associated with $\mathbf{q}$

$$\left[M_q^{\mathrm{G}}\right]_{ij} = \sum_k m_k \frac{\partial x_k}{\partial q_i}\frac{\partial x_k}{\partial q_j}. \tag{4.23}$$

What is the difference between sampling according to $\mathscr{H}$ (unconstrained) and to $\mathscr{H}_\xi$ (constrained)? If $\mathscr{H}$ is used p_ξ is sampled according to the correct distribution whereas, when $\mathscr{H}_\xi$ is used, p_ξ is a function of $\mathbf{q}$ and $\mathbf{p}$. However, averages obtained using $\mathscr{H}$ can easily be connected to averages using $\mathscr{H}_\xi$. Indeed for $\mathscr{H}$, we can analytically integrate the contribution of p_ξ to the integral

$$\int \mathrm{e}^{-\beta/2\,Z_\xi p_\xi^2}\,\mathrm{d}p_\xi \propto Z_\xi^{-1/2}.$$

Therefore, we have the following equality for the average of f in (4.21):

$$\frac{\int f(\mathbf{x})\mathrm{e}^{-\beta\mathscr{H}}\,\delta(\xi(\mathbf{x})-\xi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_x}{\int \mathrm{e}^{-\beta\mathscr{H}}\,\delta(\xi(\mathbf{x})-\xi)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_x} = \frac{\int f(\mathbf{x})Z_\xi^{-1/2}\mathrm{e}^{-\beta\mathscr{H}_\xi}\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}_q}{\int Z_\xi^{-1/2}\mathrm{e}^{-\beta\mathscr{H}_\xi}\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}_q}.$$

This equation is often written as

$$\langle f\rangle_\xi = \frac{\left\langle f\,Z_\xi^{-1/2}\right\rangle_{\xi,\dot\xi}}{\left\langle Z_\xi^{-1/2}\right\rangle_{\xi,\dot\xi}}, \tag{4.24}$$

where $\langle\ldots\rangle_{\xi,\dot\xi}$ denotes an average with ξ constant and $\dot\xi = 0$. Constrained and unconstrained simulations can therefore be easily connected through the additional weight factor $Z_\xi^{-1/2}$. This factor can be rewritten as

$$Z_\xi^{-1/2}\mathrm{e}^{-\beta\mathscr{H}_\xi} = \mathrm{e}^{-\beta(\mathscr{H}_\xi + \frac{1}{2\beta}\ln Z_\xi)}.$$

This means that the average of f $\langle f\rangle_\xi$ can be computed using a constrained simulation for a *modified* Hamiltonian

$$\boxed{\mathscr{H}_\xi^{\mathrm{F}}(\mathbf{q},\mathbf{p}_q) = \mathscr{H}_\xi(\mathbf{q},\mathbf{p}_q) + \frac{1}{2\beta}\ln Z_\xi(\mathbf{q}).} \tag{4.25}$$

The second term in this potential is the so-called Fixman potential. With this potential we simply have

$$\langle f\rangle_\xi = \langle f\rangle_{\xi,\dot\xi}^{\mathrm{F}},$$

where the F denotes an average at $\xi(\mathbf{x}) = \xi$, $\dot\xi = 0$ and for the Hamiltonian $\mathscr{H}_\xi^{\mathrm{F}}$.

4.5.3 The Potential of Mean Constraint Force

We are now going to establish a connection between the free energy $A(\xi)$ and constrained simulations. Specifically, there is a very simple relation between $A(\xi)$ and $\mathscr{H}_\xi^{\mathrm{F}}$.

$A(\xi)$ is defined in terms of a partition function

$$A(\xi) = -k_\mathrm{B}T\ln\int \mathrm{e}^{-\beta\mathscr{H}}\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}\mathrm{d}p^\xi\cdots\mathrm{d}p^q_{N-1}.$$

We have seen that the Hamiltonian $\mathscr{H}_\xi^{\mathrm{F}}$ is a function of $\mathbf{q}$ and $\mathbf{p}_q$ only. Let us integrate over p^ξ to get rid of this variable. The Hamiltonian $\mathscr{H}$ can be written as

$$\mathscr{H} = \frac{1}{2}\left((p^\xi)^t Z_\xi p^\xi + 2(p^\xi)Z_{\xi q}\mathbf{p}_q + (\mathbf{p}_q)^t Z_q\mathbf{p}_q\right) + U(\xi,\mathbf{q}).$$

The integral over p^ξ can be calculated analytically and is equal to

$$\int e^{-\frac{\beta}{2}\left((p^\xi)^t Z_\xi p^\xi + 2(p^\xi)Z_{\xi q}\mathbf{P}_q\right)}\mathrm{d}p^\xi \quad = \sqrt{\frac{\pi}{2\beta}}(Z_\xi)^{-1/2}e^{\frac{1}{2}(\mathbf{P}_q)^t Z_{q\xi}(Z_\xi)^{-1}Z_{\xi q}\mathbf{P}_q}.$$

With some additional algebra we get

$$A(\xi) = -k_\mathrm{B}T \ln \int e^{-\beta\mathscr{H}_\xi^\mathrm{F}}\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}_q - k_\mathrm{B}T \ln\sqrt{\frac{\pi}{2\beta}}$$

(we used the fact that the inverse of M_q^G is equal to $Z_q - Z_{q\xi}(Z_\xi)^{-1}Z_{\xi q}$).

A calculation similar to the one performed to obtain (4.13) finally gives

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \frac{\partial \mathscr{H}_\xi^\mathrm{F}}{\partial \xi} \right\rangle_{\xi,\dot\xi}^\mathrm{F}. \tag{4.26}$$

This equation is a very important result because it shows that constrained simulations can be used to calculate $\mathrm{d}A/\mathrm{d}\xi$. A possible algorithm would consist in running a constrained simulation with the Hamiltonian $\mathscr{H}_\xi^\mathrm{F}$ [which contains the extra Fixman potential $1/(2\beta)\ln Z_\xi(\mathbf{q})$], calculate the rate of change of $\mathscr{H}_\xi^\mathrm{F}$ with ξ at each step. Finally by averaging this rate of change the derivative of A can be computed.

In general, computing $\partial\mathscr{H}_\xi^\mathrm{F}/\partial\xi$ proves to be cumbersome. This derivative can be interpreted in terms of a mean force acting on ξ, which can be simply related to the external force $\lambda^\mathrm{F}\nabla\xi$ being exerted on the system to hold ξ constant. In fact, we show in Appendix A that $\mathrm{d}A/\mathrm{d}\xi$ is equal to

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \lambda^\mathrm{F} \right\rangle_{\xi,\dot\xi}^\mathrm{F}. \tag{4.27}$$

However, several authors took a slightly different route [17, 23, 25–27, 29] and derived an expression which does not explicitly introduce the Fixman potential.

If the Fixman potential is not added to the Hamiltonian, the system is sampled differently and therefore correction terms must be added to (4.27). First, the weight $Z_\xi^{-1/2}$ in (4.24) must be reintroduced. Second, the Lagrange multiplier is different since the Fixman potential $1/(2\beta)\ln Z_\xi$ is not used. Considering (4.20) for λ, it is possible to show that the Lagrange multiplier λ^F simply needs to be replaced by

$$\lambda + \frac{1}{2\beta Z_\xi}\sum_i \frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\frac{\partial\ln Z_\xi}{\partial x_i}.$$

The final expression therefore reads

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \frac{\left\langle Z_\xi^{-1/2}\left(\lambda + \frac{1}{2\beta Z_\xi}\sum_i \frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\frac{\partial\ln Z_\xi}{\partial x_i}\right)\right\rangle_{\xi,\dot\xi}}{\left\langle Z_\xi^{-1/2}\right\rangle_{\xi,\dot\xi}}. \tag{4.28}$$

Examples. Let us consider $\xi = r$ in 2D. In this case:

$$\nabla \xi = u_r, \quad Z_\xi = \frac{1}{m}$$

where $u_r = \mathbf{r}/r$. Thus

$$\frac{\mathrm{d}A}{\mathrm{d}r} = \langle \lambda \rangle_{r,\dot{r}}.$$

For $\xi = r^2$:

$$\nabla \xi = 2r\, u_r, \quad Z_\xi = \frac{4r^2}{m}, \quad \nabla \ln Z_\xi = \frac{2}{r} u_r.$$

The expression for $\mathrm{d}A/\mathrm{d}\xi$ then reads

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \langle \lambda \rangle_{\xi,\dot{\xi}} + \frac{1}{2\beta\xi}.$$

Let us consider another example, in which $\xi = \cos(\theta)$, $0 \le \theta \le \pi$, in 3D, where θ is the angle with the z-axis. Then

$$\nabla \xi = -\sin\theta \frac{u_\theta}{r}, \quad Z_\xi = \frac{\sin^2\theta}{mr^2}, \quad \nabla \ln Z_\xi = \frac{2\cot\theta}{r} u_\theta - \frac{2}{r} u_r,$$

where $\mathbf{u}_\theta$ is a unit vector defined by

$$\mathbf{u}_\theta = \mathbf{u}_r \times \frac{\mathbf{u}_r \times \mathbf{u}_z}{|\mathbf{u}_r \times \mathbf{u}_z|}.$$

The vector $\mathbf{u}_z$ is a unit vector pointing in the z-direction. We finally get

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \frac{\langle r\lambda \rangle_{\xi,\dot{\xi}}}{\langle r \rangle_{\xi,\dot{\xi}}} - \frac{\xi}{\beta(1-\xi^2)}. \tag{4.29}$$

Compared with (4.15), we see that this new expression for constrained simulations (4.28) is somewhat similar but a striking difference is that λ is a function of the velocity $\dot{\mathbf{x}}$ whereas (4.15) involves only an average in configurational space. Those two results are linked and we show in Appendix B that (4.28) is actually a special case of (4.15).

4.5.4 A More Concise Expression

The formula that we gave for calculating the PMF in constrained simulations (4.28) has the drawback of requiring knowledge of the second derivative of ξ with respect to $\mathbf{x}$. From a practical standpoint it would be convenient to have an expression involving first derivatives of ξ only. This can be done by introducing the constrained Hamiltonian $\mathscr{H}_\xi$ and carrying out the following expansion:

$$A(\xi) = -k_{\mathrm{B}}T \ \ln \int e^{-\beta\mathscr{H}} \ \delta(\xi - \xi(\mathbf{x})) \ \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x \tag{4.30}$$

$$= -k_{\mathrm{B}}T \ \ln \int e^{-\beta\mathscr{H}_\xi} \ \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_q \quad - k_{\mathrm{B}}T \ \ln \frac{\displaystyle\int e^{-\beta\mathscr{H}} \ \delta(\xi - \xi(\mathbf{x})) \ \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x}{\displaystyle\int e^{-\beta\mathscr{H}_\xi} \ \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_q}, \tag{4.31}$$

where $\mathscr{H}_\xi$ is the Hamiltonian for the constrained system and is a function of $(\mathbf{q}, \mathbf{p}_q)$ only [see (4.22) for its definition].

The first term can be computed in a straightforward way using the Lagrange multiplier λ. A derivation similar to the one given in Sect. 4.5.3 [see (4.27)] leads to

$$\frac{\mathrm{d}}{\mathrm{d}\xi} \left(-k_{\mathrm{B}}T \ \ln \int e^{-\beta\mathscr{H}_\xi} \ \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_q \right) = \langle \lambda \rangle_{\xi,\dot{\xi}}.$$

The Fixman potential is not needed here since we are sampling with Hamiltonian $\mathscr{H}_\xi$.

The second term is (4.31) can be expressed in terms of Z_ξ

$$\frac{\displaystyle\int e^{-\beta\mathscr{H}} \ \delta(\xi - \xi(\mathbf{x})) \ \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{p}_x}{\displaystyle\int e^{-\beta\mathscr{H}_\xi} \ \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_q} = \sqrt{\frac{\pi}{2\beta}} \frac{\displaystyle\int Z_\xi^{-1/2} e^{-\beta\mathscr{H}_\xi} \ \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_q}{\displaystyle\int e^{-\beta\mathscr{H}_\xi} \ \mathrm{d}\mathbf{q} \, \mathrm{d}\mathbf{p}_q}$$

$$= \sqrt{\frac{\pi}{2\beta}} \left\langle Z_\xi^{-1/2} \right\rangle_{\xi,\dot{\xi}}.$$

If we are interested (as is almost always the case) in computing free energy differences, the factor $\sqrt{\pi/2\beta}$ can be ignored. The factor $Z_\xi^{-1/2}$ represents a contribution to the free energy, in addition to the 'mechanical' forces given by λ, due to entropic effects.

The new expression for the derivative of the free energy is

$$\boxed{A(\xi) = \int \langle \lambda \rangle_{\xi,\dot{\xi}} \ \mathrm{d}\xi - k_{\mathrm{B}}T \ \ln \left\langle Z_\xi^{-1/2} \right\rangle_{\xi,\dot{\xi}}.} \tag{4.32}$$

This expression involves only the computation of the gradient of ξ. A comparison with (4.28) which requires the same constrained simulation but a different equation for $\mathrm{d}A/\mathrm{d}\xi$ underscores the simplicity of this new expression. A similar expression was derived by Schlitter et al. [41–44].

Let us see how (4.32) works in the examples from Sect. 4.5.3. If we take $\xi = r$ (the first example) Z_ξ is a constant and therefore can be omitted from (4.32). We directly get $A(\xi) = \int \langle \lambda \rangle_{\xi,\dot{\xi}} \ \mathrm{d}\xi$.

For $\xi = r^2$ (the second example), $Z_\xi = 4r^2/m$, we obtain

$$A(\xi) = \int \langle \lambda \rangle_{\xi,\dot{\xi}} \ \mathrm{d}\xi + \frac{\ln 4\xi/m}{2\beta}.$$

However, if we add a constant to $A(\xi)$ we get

$$A(\xi) = \int \langle \lambda \rangle_{\xi,\dot{\xi}} \, \mathrm{d}\xi + \frac{\ln \xi}{2\beta}.$$

This is the same result as before (Sect. 4.5.3).

The last example is $\xi = \cos(\theta)$. In this case $Z_\xi = \sin^2 \theta / mr^2$ and we get

$$A(\xi) = \int \langle \lambda \rangle_{\xi,\dot{\xi}} \, \mathrm{d}\xi + \frac{1}{2\beta} \ln(1 - \xi^2) - k_\mathrm{B} T \ln \langle r \rangle_{\xi,\dot{\xi}}.$$

This equation is different from (4.29). The derivative of $\ln(1 - \xi^2)/2\beta$ gives an identical term $-\xi/\beta(1-\xi^2)$, but the weight r appears as a separate term $-k_\mathrm{B} T \ln \langle r \rangle$ instead of $\langle r\lambda \rangle / \langle r \rangle$.

4.6 The Adaptive Biasing Force Method

We now describe a different approach which is simpler than the method of constraints and also very efficient. It does not require running a constrained simulation and can be performed entirely with a single molecular dynamics run.

One reason for the inefficiencies of constraint methods is that they may prevent an efficient sampling of the set $\xi(\mathbf{x}) = \xi$. This is illustrated by Fig. 4.2. It is common that many pathways separated by high energy barriers exist to go from A to B. In constrained simulation, the system can get trapped in one of the pathways. In the most serious cases, this leads to quasi-nonergodic effect where only a part of the set $\xi(\mathbf{x}) = \xi$ is effectively explored. In less serious cases, the convergence is quite slow.

An approach that does not suffer from such problems is the ABF method. This method is based on computing the mean force on ξ and then removing this force in order to improve sampling. This leads to uniform sampling along ξ. The dynamics of ξ corresponds to a random walk with zero mean force. Only the fluctuating part of the instantaneous force on ξ remains. This method is quite simple to implement and leads to a very small statistical error and excellent convergence.

In contrast to umbrella sampling, ABF removes the need to guess a priori the biasing potential or to refine it iteratively. Instead the biasing force is estimated locally from the sampled conformations of the system and continuously updated as the simulation progresses. This is therefore an adaptive algorithm. It uses all the statistics obtained so far to improve the sampling of the system. This is why it is superior to umbrella sampling, which requires global information (the probability density function across the whole range of ξ) in order to estimate the correct biasing potential properly. In ABF, the bias is applied as soon as enough samples have been accumulated in a given bin to estimate the mean force.

Let us first provide an expression to compute the derivative of the free energy for unconstrained simulations. Then, we will discuss the calculation of the biasing force and the algorithmic implementation of the method.

4.6.1 The Derivative of the Free Energy

Darve et al. [18, 28, 29] derived the following formula for the derivative of the free energy

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = -\left\langle m_\xi \ddot{\xi} - \frac{2m_\xi^2}{\beta} \sum_{ij} \frac{1}{m_i m_j} \frac{\partial \xi}{\partial x_i} \frac{\partial \xi}{\partial x_j} \frac{\partial^2 \xi}{\partial x_i \partial x_j} \right\rangle,$$

where

$$m_\xi = Z_\xi^{-1}.$$

This expression involves first- and second-order derivatives of ξ with respect to $\mathbf{x}$. While this equation has been successfully used for many computations, a more convenient expression can be derived which involves derivatives with respect to time, which are easier to calculate using molecular dynamics. No second derivatives are required, which significantly simplifies the implementation. This equation is

$$\boxed{\frac{\mathrm{d}A}{\mathrm{d}\xi} = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t} \left(m_\xi \frac{\mathrm{d}\xi}{\mathrm{d}t} \right) \right\rangle_\xi,} \tag{4.33}$$

where the averages are over ξ collected during unconstrained simulations.

For a single reaction coordinate $Z_\xi = \sum_r \frac{1}{m_r} (\partial \xi / \partial x_r)^2$. In the case where ξ is actually a Cartesian coordinate, for example $\xi = x_1$, this equation reduces to the mean force on x_1

$$\frac{\mathrm{d}A}{\mathrm{d}x_1} = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t} \left(m_1 \frac{dx_1}{\mathrm{d}t} \right) \right\rangle_{x_1} = \left\langle \frac{\partial U}{\partial x_1} \right\rangle_{x_1},$$

where we used Newton's equations of motion. Equation (4.33) is an extension to a general coordinate ξ.

This formula can be generalized to several reaction coordinates with little difficulty. In that case, m_ξ is a matrix defined as the inverse of matrix Z_ξ

$$[Z_\xi]_{kl} = \sum_r \frac{1}{m_r} \frac{\partial \xi_k}{\partial x_r} \frac{\partial \xi_l}{\partial x_r}.$$

Darve and Pohorille [28], for example, derived an equation that involved the derivative of m_ξ with respect to Cartesian coordinates. This is not required for our present equation (4.33).

Again we consider examples from the earlier section. For $\xi = r$, (4.33) becomes

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t} \left(m \frac{\mathrm{d}\xi}{\mathrm{d}t} \right) \right\rangle_\xi.$$

For $\xi = r^2$, we obtain

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{m}{4\xi} \frac{\mathrm{d}\xi}{\mathrm{d}t} \right) \right\rangle_\xi.$$

Finally for $\xi = \cos \theta$, we obtain

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{mr^2}{1-\xi^2}\frac{\mathrm{d}\xi}{\mathrm{d}t}\right)\right\rangle_\xi.$$

We now prove (4.33) and show that it is a consequence of the fundamental (4.15) (formulation with configurational derivatives) with the choice

$$w_i = \frac{1}{m_i}\frac{\partial \xi}{\partial x_i}.$$

The derivative $\mathrm{d}\xi/\mathrm{d}t$ can be written $\mathrm{d}\xi/\mathrm{d}t = \nabla \xi \cdot \dot{x}$. With the product rule for derivatives, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m_\xi\frac{\mathrm{d}\xi}{\mathrm{d}t}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(m_\xi\nabla\xi\right)\cdot\dot{\mathbf{x}} + \left(m_\xi\nabla\xi\right)\ddot{\mathbf{x}} \tag{4.34}$$

$$= (\dot{\mathbf{x}})^t \cdot \nabla\left(m_\xi\nabla\xi\right)\cdot\dot{\mathbf{x}} + \left(m_\xi\nabla\xi\right)\ddot{\mathbf{x}}. \tag{4.35}$$

In the last line, we used again $\mathrm{d}/\mathrm{d}t = (\dot{\mathbf{x}})^t \cdot \nabla$. The average over momenta with ξ fixed can be computed analytically. If we use $m_i\ddot{x}_i = -\partial U/\partial x_i$, we obtain

$$-\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(m_\xi\frac{\mathrm{d}\xi}{\mathrm{d}\mathbf{t}}\right)\right\rangle_\xi = -\left\langle (\dot{\mathbf{x}})^t \cdot \nabla\left(m_\xi\nabla\xi\right)\cdot\dot{\mathbf{x}} + \left(m_\xi\nabla\xi\right)\ddot{\mathbf{x}}\right\rangle_\xi$$

$$= -\left\langle k_\mathrm{B}T\,\mathrm{Tr}\left[M^{-1}\nabla\left(m_\xi\nabla\xi\right)\right] - m_\xi\sum_i\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\frac{\partial U}{\partial x_i}\right\rangle_\xi$$

$$= \left\langle\sum_i\left[m_\xi\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\frac{\partial U}{\partial x_i} - k_\mathrm{B}T\frac{\partial}{\partial x_i}\left(m_\xi\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\right)\right]\right\rangle_\xi,$$

which is identical to the general (4.15) with the choice $w_i = 1/m_i\,\partial\xi/\partial x_i$.

4.6.2 Numerical Calculation of the Time Derivatives

Equation (4.33) requires the computation of time derivatives. In molecular dynamics discretization errors due to the finite time step $\mathrm{d}t$ are of order $O(\mathrm{d}t^2)$. Therefore we would like to estimate the time derivative in (4.33) with the same accuracy.

If we calculate the time derivative at half time step $t + \Delta t/2$, we can approximate the instantaneous force

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m_\xi\frac{\mathrm{d}\xi}{\mathrm{d}t}\right) = \frac{(m_\xi\dot{\xi})(t+\mathrm{d}t) - (m_\xi\dot{\xi})(t)}{\mathrm{d}t} + O(\mathrm{d}t^2),$$

where $m_\xi\dot{\xi}$ can be computed using

$$m_\xi\dot{\xi} = m_\xi\nabla\xi\cdot\mathbf{v}, \quad \text{where } \mathbf{v} = \dot{\mathbf{x}}.$$

The first term $m_\xi\nabla\xi$ is a function of $\mathbf{x}$ only. Let us assume that we are using the velocity Verlet time integrator, which is the most common. In that case, $\mathbf{x}$ is computed with local accuracy $O(\mathrm{d}t^4)$ and global accuracy $O(\mathrm{d}t^2)$, and the velocity $\mathbf{v}$ at half-steps is computed with accuracy $O(\mathrm{d}t^2)$ if the following approximation is used:

$$\mathbf{v}(t + dt/2) \stackrel{\text{def}}{=} \frac{\mathbf{x}(t + dt) - \mathbf{x}(t)}{dt} + O(dt^2).$$

However, this approximation is not sufficient since it leads to an error of $O(dt)$ for $d(m_\xi \dot{\xi})/dt$.

Before introducing a more accurate approximation for $\mathbf{v}$, we recall the basic velocity Verlet algorithm

$$\mathbf{v}(t + dt/2) \stackrel{\text{def}}{=} \mathbf{v}(t - dt/2) + dt\, \mathbf{a}(t), \tag{4.36}$$

$$\mathbf{x}(t + dt) = \mathbf{x}(t) + dt\, \mathbf{v}(t + dt/2), \tag{4.37}$$

where $\mathbf{a}(t) = d^2\mathbf{x}/dt^2 = -M^{-1}\nabla U$.

It can be shown that the following expression provides an estimate of $\mathbf{v}(t)$ with accuracy $O(dt^4)$:

$$\mathbf{v}(t) = \frac{\mathbf{v}(t + dt/2) + \mathbf{v}(t - dt/2)}{2} - \frac{dt}{12}(\mathbf{a}(t + dt) - \mathbf{a}(t - dt)) + O(dt^4), \tag{4.38}$$

where $\mathbf{a}$ is the acceleration and the half-step velocities are given by the velocity Verlet algorithm.

Using this approximation we can now calculate the time derivative at $t + \Delta t/2$, as

$$\frac{d}{dt}\left(m_\xi \frac{d\xi}{dt}\right) = \frac{(m_\xi \dot{\xi})(t + dt) - (m_\xi \dot{\xi})(t)}{dt} + O(dt^3) \tag{4.39}$$

$$= \frac{m_\xi(t + dt)\nabla\xi(t + dt) \cdot \mathbf{v}(t + dt) - m_\xi(t)\nabla\xi(t) \cdot \mathbf{v}(t)}{dt} + O(dt^3) \tag{4.40}$$

$$= \frac{1}{2}\left(\frac{\mathbf{p}_\xi^+(t + dt) - \mathbf{p}_\xi^+(t)}{dt} + \frac{\mathbf{p}_\xi^-(t + dt) - \mathbf{p}_\xi^-(t)}{dt}\right), \tag{4.41}$$

where

$$\mathbf{p}_\xi^+(t) \stackrel{\text{def}}{=} m_\xi(t)\nabla\xi(t) \cdot \left[\mathbf{v}(t + dt/2) - \frac{dt}{6}\mathbf{a}(t + dt)\right], \tag{4.42}$$

$$\mathbf{p}_\xi^-(t) \stackrel{\text{def}}{=} m_\xi(t)\nabla\xi(t) \cdot \left[\mathbf{v}(t - dt/2) + \frac{dt}{6}\mathbf{a}(t - dt)\right]. \tag{4.43}$$

In this expression m_ξ, $\nabla\xi$ and $\mathbf{a}$ are functions of $\mathbf{x}$ and can computed at t and $t + \Delta t$. The velocity at the half-steps is directly provided by the velocity Verlet algorithm. To summarize, to calculate $d(m_\xi \dot{\xi})/dt$ at time $t + \Delta t/2$ we need to collect data from time steps $t - \Delta t$, t, $t + \Delta t$, and $t + 2\Delta t$. Recall again that these equations are used to calculate the mean force along ξ, not to advance the system in time.

4.6.3 Adaptive Biasing Force: Implementation and Accuracy

We previously showed that the free energy can be computed using (4.33). This is done by binning the force and computing the average force along the ξ interval.

However, a molecular dynamics simulation without a biasing force will be very inefficient if significant energy barriers along ξ are present in the system. Therefore, an external biasing force needs to be added to improve the sampling efficiency. This can be done by applying the ABF algorithm [18, 28, 29], which yields a uniform sampling along ξ, and by doing so leads to a significant reduction of statistical errors due to an increased sampling of transition regions.

Assume that we bin the interval of interest for ξ and that we have collected $n^k(N)$ samples in bin k after N steps in a molecular dynamics simulation. We use those samples to compute a *running* average of the force acting along ξ:

$$F_\xi^k(N) = \frac{1}{n^k(N)} \sum_{l=1}^{n^k(N)} \frac{\mathrm{d}}{\mathrm{d}t}\left(m_\xi \frac{\mathrm{d}\xi}{\mathrm{d}t}\right)(\mathbf{x}_l^k), \qquad (4.44)$$

where $\mathbf{x}_l^k$ corresponds to sample l in bin k. We will explain in the next section how the time derivatives can be computed. The external force applied to the system is chosen equal to $-F_\xi^k$.

In general, when very few samples are available the force $F_\xi^k(N)$ will not be an accurate approximation of $\mathrm{d}A/\mathrm{d}\xi$. Large variations in $F_\xi^k(N)$ may lead to nonequilibrium effects and systematic bias of the calculation. Mathematically, this can be expressed by introducing a perturbation $\Delta\mathscr{H}(\mathbf{q}, \mathbf{p}, N)$, which is a function of the number of steps N. At $N = 1$ if we average over all possible initial configurations, abbreviated by subscript 0, we obtain

$$\langle F_\xi^k(1)\rangle_0 = \frac{\int \frac{\mathrm{d}}{\mathrm{d}t}(m_\xi \dot{\xi})\,\mathrm{e}^{-\beta\mathscr{H}}\delta(\xi - \xi(\mathbf{q}))\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}{\int \mathrm{e}^{-\beta\mathscr{H}}\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}$$

so that we exactly have $\mathrm{d}A/\mathrm{d}\xi = -\langle F_\xi^k(1)\rangle_0$. Similarly, after a very long simulation time we have (again averaging over initial configurations)

$$\langle F_\xi^k(\infty)\rangle_0 = \frac{\int \frac{\mathrm{d}}{\mathrm{d}t}(m_\xi \dot{\xi})\,\mathrm{e}^{-\beta(\mathscr{H}-A(\xi))}\delta(\xi - \xi(\mathbf{q}))\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}{\int \mathrm{e}^{-\beta(\mathscr{H}-A(\xi))}\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}.$$

Again, we exactly have $\mathrm{d}A/\mathrm{d}\xi = -\langle F_\xi^k(\infty)\rangle_0$.

For intermediate values of N, a perturbation $\Delta\mathscr{H}(\mathbf{q}, \mathbf{p}, N)$ can be defined such that the weight in the average is $\exp(-\beta(\mathscr{H} + \Delta\mathscr{H}))$. In this case a systematic bias may be introduced and $\mathrm{d}A/\mathrm{d}\xi \neq -\langle F_\xi^k(N)\rangle_0$. In addition incorrect estimates of $\mathrm{d}A/\mathrm{d}\xi$ can lead to short-lived free energy barriers in the initial steps of the simulation.

In order to control those initial nonequilibrium effects, a ramp function can be added which reduces the variations from one step to the next of the external force applied in a given bin. The external force applied to the system can be chosen equal to

$$-R(n^k(N))F_\xi^k(N), \tag{4.45}$$

$$R(n) = \begin{cases} n/N_0 & \text{if } n \leq N_0 \\ 1 & \text{if } n > N_0. \end{cases} \tag{4.46}$$

It has been found in numerical tests that N_0 should be chosen relatively small, $N_0 = 100$ for example. In general it is important to have a good balance between rapidly improving the sampling of the system, which requires a small N_0, and avoiding large nonequilibrium effects, which requires a large N_0. However, the effects of nonequilibrium perturbations occur only at the beginning of the simulation and disappear rapidly as the number of sample points in the bin increases. For N samples, this initial systematic error decays as $1/N$. The statistical error, however, decays much more slowly, as $1/\sqrt{N}$. This means that initial sampling errors are negligible compared to statistical errors and therefore will not negatively affect the accuracy of the computation. Numerical tests have shown that choosing a small N_0 is preferred as it rapidly provides good sampling along the reaction coordinate ξ.

Other algorithms for rapid estimation of the average forces in individual bins can be used. In particular, when no or only a few samples have been gathered in a bin k, but good statistics in the surrounding bins are available, extrapolation schemes can provide an estimate of the biasing force in bin k. This accelerates sampling of conformations corresponding to transition states with large free energies.

4.6.4 The ABF Algorithm

When biasing the system, an external force $\lambda \nabla \xi$ is applied to improve the sampling along ξ. Since this force is added, the calculation of the derivative of $-m_\xi \dot{\xi}$ must be modified. Consider (4.35). The first term does not require any correction since $\mathbf{x}$ and $\dot{\mathbf{x}}$ are sampled according to the correct distribution. However $\ddot{\mathbf{x}}$ includes the ABF force whose contribution needs to be removed to compute the free energy derivative. The correction is equal to

$$(m_\xi \nabla \xi) M^{-1} (\lambda \nabla \xi) = \lambda.$$

Since we approximate $-\mathrm{d}(m_\xi \dot{\xi})/\mathrm{d}t$ at $t + \mathrm{d}t/2$ we need to add the following correction:

$$\frac{1}{2}(\lambda(t) + \lambda(t + \mathrm{d}t)).$$

The full ABF algorithm consists of two parts, which are summarized below (Algorithms 1 and 2). Any molecular dynamics code can easily be updated to include these modifications.

4.6.5 Additional Discussion of ABF

We now discuss a number of practical issues to illustrate the strengths and weaknesses of ABF.

Algorithm 1 Velocity Verlet loop with ABF

"←" is the assignment operator.

Loop over time steps $i = 1, \ldots, n$:

    a1 ← M^{-1} F(r0)

    /* Force computation at time t. M is a diagonal matrix with atom masses. F is the
       vector of mechanical forces: F(r) $= -\nabla U$. */

    a1 ← ABF(a1, r0, v)

    /* Call to ABF subroutine, which computes the derivative of the free energy using
       (4.33) and adds the adaptive biasing force. It takes as input the current
       acceleration a1 and returns the new acceleration with ABF added. */

    v ← v + dt a1

    /* The velocity vector is advanced from $t - \mathrm{d}t/2$ to $t + \mathrm{d}t/2$ */

    r0 ← r0 + dt v

    /* The position vector is advanced from t to $t + \mathrm{d}t$ */

End of loop

Probability distribution function of forces. In each bin, ABF attempts to compute $\mathrm{d}A/\mathrm{d}\xi$ by averaging $-\mathrm{d}/\mathrm{d}t(m_\xi \dot{\xi})$. If we ignore the correlation between samples, the statistical error can be approximated by $\sigma^k/\sqrt{n^k(N)}$ in bin k where σ^k is the standard deviation of the force. In most cases, the force has a Gaussian distribution. The efficiency of the calculation will depend on its standard deviation, which can be large compared to the mean. An example is shown in Fig. 4.3 where we considered polyalanine and define ξ as the distance between the α carbon in the first and last residue. Figure 4.4 shows the distribution of forces around $\xi = 19$ Å. The probability distribution is broad and the standard deviation (=13) is large compared to the mean (=−1). In that case, one needs at least 14,000 samples to reduce the error, when estimating the derivative, down to 10%. The distribution can also be seen to be nearly Gaussian.

Complete or partial removal of the mean force. In ABF, the biasing force is used only to improve the sampling. The biasing force can be modified without affecting the estimation of the derivative of $A(\xi)$. It is possible to remove free energy barriers completely by applying the full bias, or simply scale down the barriers by applying a fraction of the biasing force only. When using partial scaling, minima and maxima of the free energy remain at the same location but are reduced.

Optimal sampling. As was pointed out earlier, the error in the derivative of A is proportional to $\sigma^k/\sqrt{n^k(N)}$. The optimal sampling is therefore obtained when $\sigma^k/\sqrt{n^k(N)}$ is constant as a function of k. In regions where σ^k is large, additional sample points should be added to compensate. This is often a small effect but in some special cases is worth considering. In order to obtain the optimal sampling, the potential energy should be corrected as

$$U^{\mathrm{opt}} = U - 2k_{\mathrm{B}}T \ln(\sigma(\xi)/\sigma_0),$$

Algorithm 2 ABF routine

In this routine, the derivative of the free energy is calculated and the biasing force is added to the system.

"·" is the dot product.

a1 = function ABF(i,dt,a1, r0,v)

/* i: step number in the Verlet loop
   dt: time step
   a1: acceleration at time t
   r0: position at time t
   v: velocity at time $t - \mathrm{d}t/2$ */

  save n, F, pxi0, pxi1, a0, ZD0
  /* Those variables need to be saved between function calls. */

  k $\leftarrow$ bin corresponding to $\xi(r0)$
  /* Binning is used to calculate statistics for the average force. */

  Rk $\leftarrow$ ramp $R(n^k(N))$ in bin k
  /* Evaluates the ramp function $R(n^k(N))$ [see (4.46)]. $n^k(N)$ is the number of sample
     points in bin k after N molecular dynamics steps. */

  la $\leftarrow$ Rk F(k) / n(k)
  /* Calculates the biasing force to apply. n(k) $= n^k(N)$. F(k) is the sum of all the force
     samples in bin k. */

  a1 $\leftarrow$ a1 + la $M^{-1}\nabla\xi$(r0)
  /* Applies the adaptive biasing force. a1 is the new acceleration. This is an
     output variable. M is a diagonal matrix with atom masses. */

  pxip $\leftarrow$ ZD0 $\cdot$ (v/2 $-$ (dt/12) a1)
  /* This is equal to $p_\xi^+(t - \mathrm{d}t)/(2\mathrm{d}t)$ [see (4.42)]. ZD0 is equal to $m_\xi\nabla\xi/\mathrm{d}t$ at time
     $t - \Delta t$. a1 is the acceleration at time t. */

  ZD0 $\leftarrow$ 1/dt $(m_\xi$(r0) $\nabla\xi$(r0))
  /* Computes the new value of $m_\xi\nabla\xi/\mathrm{d}t$ at time t. */

  pxim $\leftarrow$ ZD0 $\cdot$ (v/2 + (dt/12) a0)
  /* Using the new ZD0, we compute $p_\xi^-(t)/(2\mathrm{d}t)$ [see (4.43)]. a0 contains the
     acceleration at the previous step $(t - \Delta t)$. */

  a0 $\leftarrow$ a1
  /* Saves the acceleration a1 in preparation for the next step. */

  pxi0 $\leftarrow$ pxi0 + pxip
  /* This steps completes the calculation of $d(m_\xi\dot\xi)/\mathrm{d}t$ at time $t - 3\mathrm{d}t/2$. */
  If $i \geq 4$:
  /* We need i to be at least 4 before we can save the first value of F_ξ^k. */
     k0$\leftarrow$bin index corresponding for ξ at step $t - 3\mathrm{d}t/2$.
     /* Calculates the bin index correspomding for ξ at step $t - 3\mathrm{d}t/2$. */
     n(k0) $\leftarrow$ n(k0) + 1
     /* Increments the counter $n^{k_0}(N)$ by 1. */
  F(k0) $\leftarrow$ F(k0) $-$ pxi0

/* Adds new sample $-d(m_\xi \dot{\xi})/\mathrm{d}t$ to bin k_0. */
    Endif
    pxi0 $\leftarrow$ $-$ pxip + pxim + pxi1 $-$ la/2
    /* pxi0 is now equal to $(-p_\xi^+(t - \mathrm{d}t) + p_\xi^-(t) - p_\xi^-(t - \mathrm{d}t))/(2\mathrm{d}t)$
        $-(\lambda(t - \mathrm{d}t) + \lambda(t))/2$. The only term missing to calculate $d(m_\xi \dot{\xi})/\mathrm{d}t$ at time
        $t - \mathrm{d}t/2$ is $p_\xi^+(t)/(2\mathrm{d}t)$. This is done at the next time step. */

    pxi1 $\leftarrow$ $-$ pxim $-$ la/2
    /* pxi1 is equal to $-p_\xi^-(t)/(2\mathrm{d}t) - \lambda(t)/2$. This value is used at the next step. */
End function ABF

At $t = 0$, the variables n, F, pxi0, pxi1, a0, ZD0 must be initialized to 0.
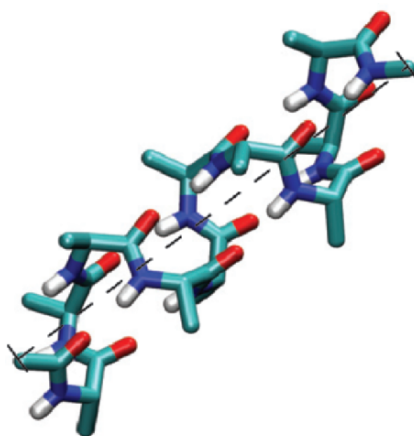
Fig. 4.3. Polyalanine. The order parameter is the distance between the first and last α carbon
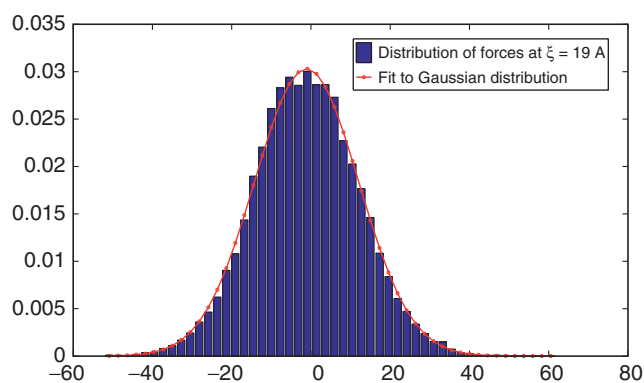
Fig. 4.4. Probability density function of the force. The mean is -1.1 and the standard deviation is 13.2. A fit with a Gaussian distribution with identical mean and variance is shown

where σ_0 is a constant with the same dimension as $\sigma(\xi)$. A practical approach is to run an ABF simulation for some time, compute an estimate of $\sigma(\xi)$ and then resume the simulation with ABF and the biasing potential $-2k_\mathrm{B}T\ln(\sigma(\xi)/\sigma_0)$.

Partition of forces. Going back to (4.33), we can derive a different equation

$$
\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}\xi} &= -\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(m_\xi \frac{\mathrm{d}\xi}{\mathrm{d}t}\right)\right\rangle \\
&= \left\langle \nabla U \cdot (m_\xi M^{-1}\nabla\xi) - \mathbf{v}\cdot\nabla(m_\xi\nabla\xi)\cdot\mathbf{v}\right\rangle
\end{aligned}
\tag{4.47}
$$

using the product rule and the chain rule for differentiation. The first term is related to the energy potential U and can easily be decomposed into several terms based on a partition of U. For example, let us say we write U as

$$
U = U_{\text{chemical bonds}} + U_{\text{electrostatic}} + U_{\text{Lennard-Jones}}
$$

The free energy can be decomposed into several contributions coming from each term in U.

The last term $\mathbf{v}\cdot\nabla(m_\xi\nabla\xi)\cdot\mathbf{v}$ corresponds to an entropic contribution. Even if U is constant, A might be nonconstant if ξ is a nonlinear function of $\mathbf{x}$. This is a purely entropic effect.

Free energy with several order parameters. The previous formulas can be extended to the case of multiple reaction coordinates $\xi_1,\ldots,\xi_p$. For example, the derivative of the free energy becomes a gradient and

$$
\nabla_\xi A(\xi_1,\ldots,\xi_p) = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(M_\xi \frac{\mathrm{d}\xi}{\mathrm{d}t}\right)\right\rangle,
$$

where ξ is a vector. The matrix M_ξ is defined by its inverse

$$
\left[M_\xi^{-1}\right]_{jk} = \sum \frac{1}{m_i}\frac{\partial\xi_j}{\partial x_i}\frac{\partial\xi_k}{\partial x_i}.
$$

This generalization is therefore quite straightforward. One difficulty is that in higher dimensions, reconstructing $A(\xi)$ from its derivatives is not straightforward. For example, let us choose a reference point ξ_0 and decide that $A^{\mathrm{ABF}}(\xi_0) = 0$. Say we define $A^{\mathrm{ABF}}(\xi)$ by

$$
A^{\mathrm{ABF}}(\xi_1) = \int_{C_{\xi_0,\xi_1}} \mathbf{D}^{\mathrm{ABF}}\cdot\mathrm{dl},
$$

where $\mathbf{D}^{\mathrm{ABF}}$ is the approximation of $\nabla_\xi A$ produced by the ABF procedure. C_{ξ_0,ξ_1} is a path joining ξ_0 and ξ_1. For an arbitrary closed loop C we always have

$$
\int_C \nabla_\xi A\cdot\mathrm{dl} = 0.
$$

However, this is not satisfied in general by the vector field $\mathbf{D}^{\mathrm{ABF}}$, which has some statistical error. As a consequence, the definition of $A^{\mathrm{ABF}}(\xi_1)$ above, in fact, depends on the path C_{ξ_0,ξ_1}, which is not desirable.
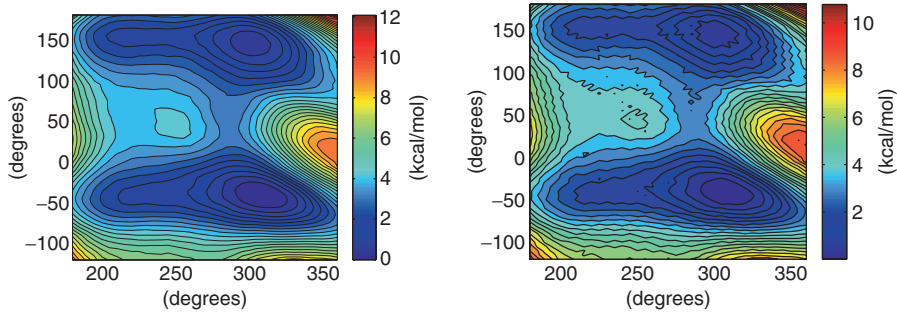
Fig. 4.5. Free energy profile of alanine dipeptide as a function of Φ and Ψ. The ABF method for second-order parameters was used in this calculation. The figure on the left shows the reconstruction using four control points per data point (as shown in Fig. 4.6). The figure on the right shows a reconstruction using only one control point per data point. This results in a more oscillatory solution
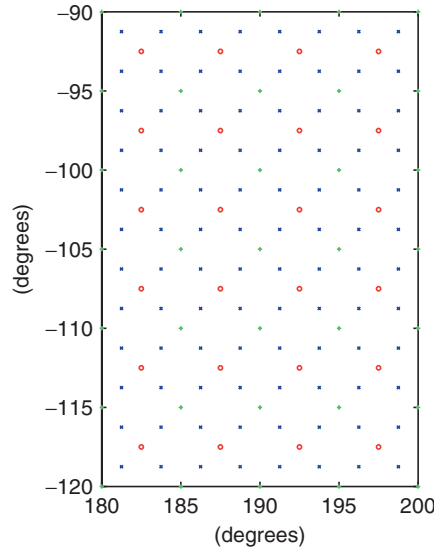
Fig. 4.6. Control points and Q_1 nodes used to reconstruct A. The Q_1 basis function is a bilinear function equal to 1 at a Q_1 node and zero at the surrounding Q_1 nodes. Four control points per cell were used. The derivatives of A were evaluated at each control point using a linear interpolation based on the neighboring data points

To circumvent this, the function A can be approximated using spline functions

$$A^{\mathrm{ABF}}(\xi) = \sum_l \alpha_l B_l(\xi).$$

The coefficients α_l can be computed by minimizing

$$\sum_k \left| \nabla_\xi A^{\mathrm{ABF}}(\xi_k) - D^{\mathrm{ABF}}(\xi_k) \right|^2 = \sum_k \left| \left(\sum_l \alpha_l \nabla_\xi B_l(\xi_k) \right) - D^{\mathrm{ABF}}(\xi_k) \right|^2,$$

where ξ_k are the sample points at which $\mathbf{D}^{\text{ABF}}$ was computed. This minimization problem for α_l has a unique solution if in addition we enforce that $A^{\text{ABF}}(\xi_0) = 0$.

As an example, this approach was applied to the calculation of the PMF for alanine dipeptide as a function of the two torsion angles Φ and Ψ. The resulting free energy surface is shown in Fig. 4.5. Bilinear Q_1 elements were used to approximate the free energy. Control points were chosen such that there are four of them around each data point. This was done in order to increase the smoothness and quality of the reconstructed free energy. The position of the Q_1 nodes and control points is shown in Fig. 4.6.

4.7 Discussion of Other Techniques

Several other techniques are available to compute the PMF. In this section we discuss the steered force molecular dynamics (SMD) [35–37] and the metadynamics of Laio and Parrinello [30–34].

In SMD an external force is applied in an attempt to accelerate a chemical process such as unfolding of a protein or dissociation of two molecules. In that respect, it is similar to AFM or optical-tweezers experiments. The applied force is given by:

$$F = k(vt - x),$$

where x is the displacement of the pulled atom, v is a pulling velocity and k is a stiffness constant. Variants of SMD include simulations, in which the applied force is constants or where the rate of change of x is constant. Several proteins, such as titin [35, 36], cadherin [37], V-CAM [37], fibronectin [37], cytochrome C6 [37], immunoglobulin binding protein [37] and synaptotagmin I [37], have been studied using SMD.

Metadynamics defines coarse-grained variables which are assumed to be slow coordinates of the system. Those coordinates are similar to the order parameters considered earlier in this chapter. The coarse variables are evolved independently following a steepest-descent equation. In the case of a single variable, Laio and Parrinello [34] use

$$\xi_{n+1} = \xi_n - h\,\Delta\xi\,\text{sign}(dA_n^{\text{b}}/d\xi), \tag{4.48}$$

where $\text{sign}(x) = 1$ if $x \geq 0$, -1 otherwise; h is a fixed stepping parameter and $\Delta\xi$ is the estimated size of the free energy well at the current point. If A_n^{b} were the free energy, this equation would simply move ξ towards the nearest free energy minimum. In order to guarantee an efficient exploration, A_n^{b} is a history-dependent potential defined by

$$A_n^{\text{b}}(\xi) = A(\xi) + W \sum_{k \leq n} \exp\left(-\frac{|\xi - \xi_k|^2}{2(h\Delta\xi)^2}\right).$$

The extra term with Gaussian functions helps push the variable ξ away from regions which have already been visited; see Fig. 4.7. This enhances the efficiency of the
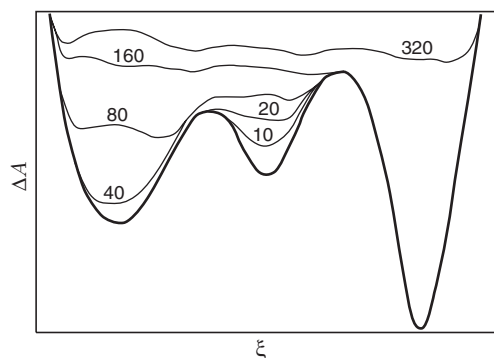
Fig. 4.7. The metadynamics technique of Laio and Parrinello progressively fills up free energy wells by adding Gaussian terms to the energy function. This figure shows the effective free energy profile experienced by the coarse-grained variable as the simulation progresses. The length of the simulation is indicated by the numbers above the *curves*. The free energy profile is shown using a thick *solid line*

simulation and leads to a uniform sampling along the coarse-grained variable. In this respect the method is similar to ABF.

The steepest descent (4.48) requires calculation of $dA/d\xi$. Laio and Parrinello chose to estimate this term by performing short constrained molecular dynamics simulations and computing the potential of mean constraint force (see Carter et al. [45]).

In the limit of a large number of sample points and for sufficiently small h

$$-W \sum_{k \leq n} \exp\left(-\frac{|\xi - \xi_k|^2}{2(h\Delta\xi)^2}\right) \ \rightarrow \ A(\xi).$$

A variant of this method [32] was also developed which uses an extended Lagrangian formulation of the form

$$\mathscr{L} = \mathscr{L}_0 + \frac{1}{2}M\dot{\xi}^2 - \frac{1}{2}k(\xi(\mathbf{x}) - \xi)^2 - A_t^{\mathrm{b}}(\xi),$$

where ξ is now an extra degree of freedom and A_t^{b} is a history-dependent potential formed by superposing Gaussian functions in a fashion similar to that described above.

Several problems have been studied with this technique, including the dissociation of NaCl [34], alanine dipeptide [34], C_4H_6 [32], and phase transitions in silicon [33].

4.8 Examples of Application of ABF

4.8.1 Two Simple Systems

The first example involves calculating the potential of mean force for the rotation of the C–C bond in 1,2-dichloroethane (DCE) dissolved in water. In the second

example, the potential of mean force for the transfer of fluoromethane (FMet) across a water–hexane interface is obtained. Details of the simulations can be found in [28, 29, 46, 47].

For DCE in water, the Cl–C–C–Cl torsional angle ξ was taken as the reaction coordinate. For the transfer of FMet across the water–hexane interface, ξ was defined as the z component of the distance between the centers of mass of the solute and the hexane lamella.

In Fig. 4.8 we show the free energy profile for the rotation of DCE around the Cl–C–C–Cl torsional angle. This profile, obtained using ABF, is in excellent agreement with the previously calculated reference curve [28]. The free energies for the *gauche* and *trans* conformations in water are nearly the same. In contrast, in the gas phase the *trans* conformation is favored by 1.1 kcal mol^{-1} [46, 47]. This means that, compared to the *trans* rotamer, the *gauche* conformation is stabilized in the aqueous environment. This can be explained by favorable interactions between the permanent dipole of DCE in the *gauche* state and the surrounding water. These interactions are absent if DCE is in the *trans* conformation because, by symmetry, this state has no dipole moment.

The free energy profile for the transfer of FMet across a water–hexane interface obtained using ABF is shown in Fig. 4.9. The free energy exhibits a minimum at the interface, which is approximately 2 kcal mol^{-1} deep (compared to the free energy in bulk water). The existence of this minimum is due to the lower density at the interface between weekly interacting liquids, such as water and oil, compared to the densities in the bulk solvents. As a result, the probability of finding a cavity sufficiently large to accommodate the solute increases and the corresponding free energy cost of inserting a small, nonpolar or weakly polar solute decreases [48]. Similar free energy profiles
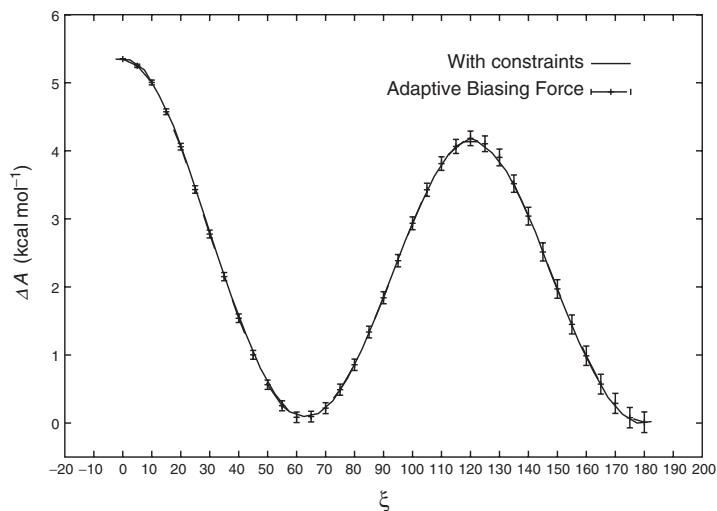
Fig. 4.8. Free energy computation using ABF and a constrained simulation. Reprinted with permission from Darve et al. 2001 [28]. Copyright 2001, American Institute of Physics
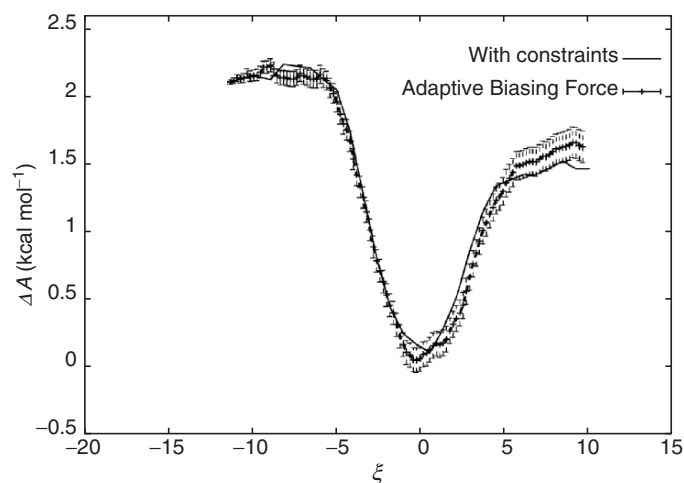
Fig. 4.9. Free energy computation using ABF and a constrained simulation. Reprinted with permission from Gomez et al. 2004 [18]. Copyright 2004, American Institute of Physics.

were found for a wide range of other solutes [49–51]. The free energy difference between FMet in water and hexane is approximately equal to $0.5\,\mathrm{kcal\,mol^{-1}}$, which corresponds well to the measured partition coefficient between these two liquids [48].

Despite its apparent simplicity, this calculation is relatively difficult to perform. Gomez et al. [18] demonstrated that ABF performs much better than both slow-growth and fast-growth implementations of the nonequilibrium method of Jarzynski and Crooks (see the next chapter).

4.8.2 Deca-L-alanine

ABF was probed through the reversible unfolding of a short peptide, deca-L-alanine, in vacuo [52] (see Fig. 4.10). The reaction coordinate, ξ, is the distance separating the first and the last C_α carbon atom of the peptide chain. ξ was varied between 12 and $32\,\text{Å}$, thereby allowing the peptide to sample the full range of conformations between the native α-helical structure and the extended structures. The force acting along ξ was accrued in bins $0.1\,\text{Å}$ wide.

Starting from the α-helical state of the peptide, the complete free energy profile is obtained from a molecular dynamics trajectory 5 ns long, with a reasonably uniform sampling distribution over the entire range of ξ values (see Fig. 4.11). This profile possesses a unique minimum around $14\,\text{Å}$, corresponding to the native helical segment. As the peptide chain stretches out, the intramolecular $i \rightarrow i + 4$ hydrogen bonds responsible for the scaffold of the α-helix are successively broken, leading to a progressive increase of the free energy.

It is interesting to note that the free energy profiles obtained with ABF are essentially identical to those derived by Park et al. [53] from a reversible, 200-ns steered molecular dynamics simulation and from a set of 100 shorter runs, pulling very
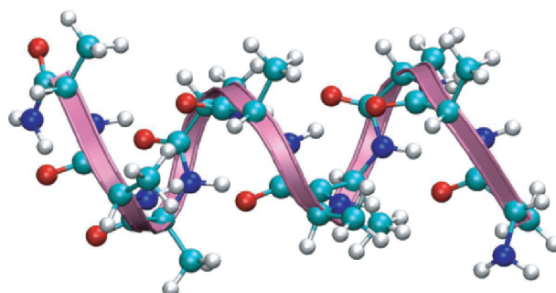
Fig. 4.10. Deca-L-alanine in its folded configuration
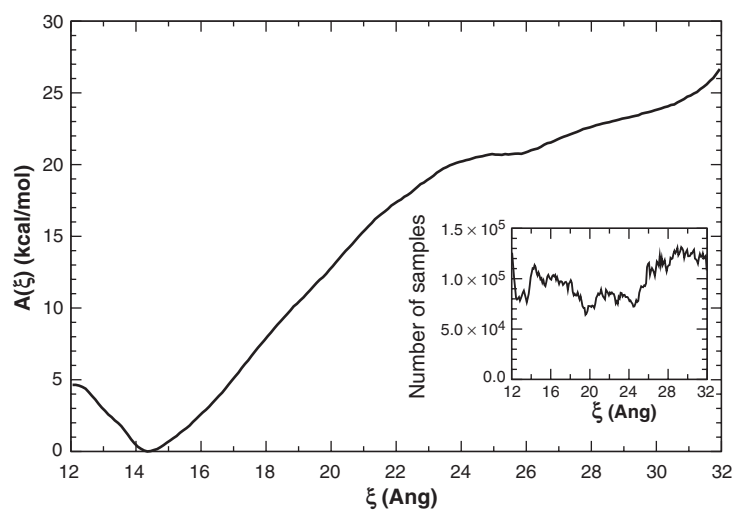
Fig. 4.11. Free energy profile for Deca-L-alanine calculated using the ABF. The *inset* shows the number of samples as a function of ξ

slowly the terminal end of deca-L-alanine and applying the Jarzynski identity to infer equilibrium free energy differences along ξ. They were however obtained at a much lower computational cost.

4.9 Glycophorin A

Hénin et al. [54] used ABF to model the association of Glycophorin A inside a membrane mimetic. GpA was modeled using two *trans*-membrane helical segments. The key interactions between the two segments are shown in Fig. 4.12.
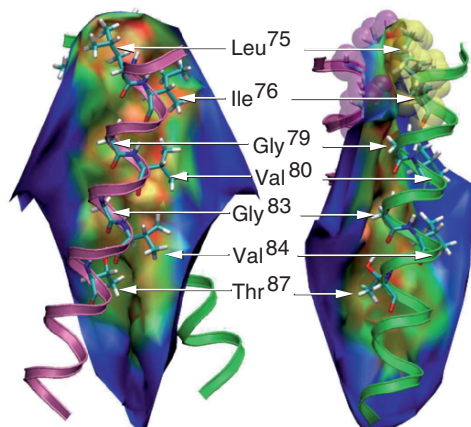
Fig. 4.12. Glycophorin A: key interactions between the two helical segments. The color scale reflects decreasing interhelical distances. Reprinted in part with permission from Hénin et al. 2005 [54]. Copyright 2005 American Chemical Society.
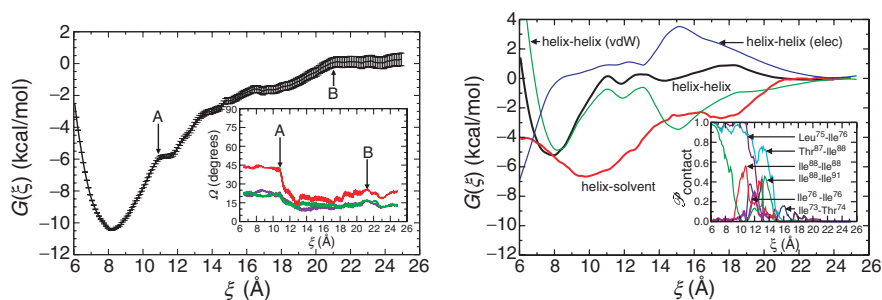
Fig. 4.13. Free energy profile for glycophorin A as a function of helix–helix distance in Å. The figure on the *right* shows the individual contributions of helix–helix van der Waals and electrostatic forces, and helix–solvent forces. Reprinted in part with permission from Hénin et al. 2005 [54]. Copyright 2005 American Chemical Society.

An important difficulty of this simulation is the fact that, once the two helices are close to each other, it is very difficult to slide one with respect to the other or change their relative orientation because of steric clashes. This means that a simulation where the helix–helix distance is constrained would not be able to sample phase space efficiently but rather would remain near the local energy minimum where it started. In contrast, using ABF, the helices have the opportunity to move closer and farther apart so that their relative position can vary more freely.

Figure 4.13 shows the free energy profile as a function of the helix–helix distance. Equation (4.47) allows the computation of the contributions to the profile by the different intermolecular potentials. The helix–helix and helix–solvent interactions were considered. The helix–helix van der Waals potential shows a significant minimum

at short distances around 8 Å. The helix–helix electrostatic interactions has a strong attractive component at short distances. The helix–solvent interaction is weaker and attractive in the range up to 21 Å.

4.10 Alchemical Transformations

So far we have discussed various techniques for computing the PMF. The other type of free energy calculation commonly performed is alchemical transformation where two different systems are compared. Such calculations have many applications such as: Lennard-Jones fluid with and without dipoles for each particles, comparison of ethanol (CH_3CH_2OH) and ethane thiol (CH_3CH_2SH), replacing one amino acid by another in a protein, changing the formula for a compound in drug discovery, etc.

For those applications, each system is modeled using different Hamiltonian functions $\mathscr{H}_0$ and $\mathscr{H}_1$. The free energy difference is defined as

$$\Delta A = -k_\mathrm{B} T \ln \frac{Q_1}{Q_0} = -k_\mathrm{B} T \ln \frac{\displaystyle\int \exp(-\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x)/k_\mathrm{B} T)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x}{\displaystyle\int \exp(-\mathscr{H}_0(\mathbf{x}, \mathbf{p}_x)/k_\mathrm{B} T)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x}.$$

Several techniques exist to compute ΔA. Following our earlier discussion for the PMF we will discuss TI. In this approach a parameterized Hamiltonian $\mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)$ is defined such that when $\lambda = 0$, $\mathscr{H}_\lambda = \mathscr{H}_0$ and when $\lambda = 1$, $\mathscr{H}_\lambda = \mathscr{H}_1$. $\mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)$ interpolates smoothly between the two Hamiltonian functions. The free energy A becomes itself a function of λ and we have

$$\frac{\mathrm{d}A}{\mathrm{d}\lambda} = \frac{\displaystyle\int \frac{\partial \mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)}{\partial \lambda} \exp(-\mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)/k_\mathrm{B} T)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x}{\displaystyle\int \exp(-\mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)/k_\mathrm{B} T)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x} = \left\langle \frac{\partial \mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)}{\partial \lambda} \right\rangle_\lambda.$$

$$(4.49)$$

If $\mathscr{H}_\lambda$ is a linear interpolation between Hamiltonian $\mathscr{H}_0$ and $\mathscr{H}_1$, then this equation reduces to

$$\frac{\mathrm{d}A}{\mathrm{d}\lambda} = \langle \mathscr{H}_1(\mathbf{x}, \mathbf{p}_x) - \mathscr{H}_0(\mathbf{x}, \mathbf{p}_x) \rangle_\lambda.$$

Calculations are usually performed by considering a set of quadrature points λ_k between 0 and 1 and associated weights ω_k, e.g., Gaussian points and weights. At each λ_k, an MD simulation is performed and the average of $\partial \mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)/\partial \lambda$ is computed. Finally the free energy is computed using

$$\Delta A \sim \sum_k \omega_k \left\langle \frac{\partial \mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)}{\partial \lambda} \right\rangle_{\lambda_k}.$$

Example: Ion in Water
Consider the case of an ion in water. The force field for the ion may be described using a Lennard-Jones and electrostatic potential with charge q. Assume that we

want to compute the free energy difference between a charged and uncharged parti-
cle. In that case we can define a charge $q_\lambda = \lambda q$. The Hamiltonian function can be
defined as

$$\mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x) = \mathscr{H}_0(\mathbf{x}, \mathbf{p}_x) + \lambda q V(\mathbf{x}_{\text{ion}}),$$

where $V(\mathbf{x}_{\text{ion}})$ is the electrostatic potential due to water molecules at the ion location.
In this case

$$\frac{\partial \mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)}{\partial \lambda} = q V(\mathbf{x}_{\text{ion}}).$$

The derivative of the free energy is simply computed using

$$\frac{\mathrm{d}A}{\mathrm{d}\lambda} = \langle q V(\mathbf{x}_{\text{ion}}) \rangle_\lambda.$$

Note that, even though $q V(\mathbf{x}_{\text{ion}})$ is not a function of λ, $A(\lambda)$ is since we are com-
puting the average using the Hamiltonian $\mathscr{H}_\lambda(\mathbf{x}, \mathbf{p}_x)$. As λ changes the structure of
the water around the ion changes and, as a result, the mean value of $V(\mathbf{x}_{\text{ion}})$ also
changes.

In this case, $A(\lambda)$ is a concave function which can be very useful to check the
accuracy of the calculation. This is proved by computing the second derivative of
$A(\lambda)$

$$\frac{d^2 A}{\mathrm{d}\lambda^2} = -\frac{1}{k_\mathrm{B}T} \left[\left\langle (\Delta U)^2 \right\rangle_\lambda - \langle \Delta U \rangle_\lambda^2 \right] \leq 0, \quad \Delta U = U_1 - U_0,$$

where U_0 and U_1 are the potential energies of system 0 and 1. This property, in
general, holds for any parametrization such that $\mathscr{H}_\lambda$ is a linear function of the
parameter λ, that is: $\mathscr{H}_\lambda = (1 - \lambda)\mathscr{H}_0 + \lambda \mathscr{H}_1$.

Entropy

In the previous examples, we considered a parameterized Hamiltonian function $\mathscr{H}_\lambda$
and derived equations to compute $A(\lambda)$. Let us now consider the dependence of A
with temperature. Based on the definition of A, we have

$$\frac{\partial A}{\partial T} = \frac{1}{T} (A - \langle \mathscr{H} \rangle).$$

If we recall the classical equation of thermodynamics relating the entropy S to the
free energy: $A = \langle \mathscr{H} \rangle - TS$, we get

$$S = -\frac{\partial A}{\partial T}.$$

Computing the absolute free energy A is in general a difficult task. For the same
reasons the absolute entropy cannot be computed except in special cases. How-
ever, entropy differences can be computed. Indeed, using finite differences, we can
approximate

$$\Delta S = -\frac{\Delta A(T + \Delta T) - \Delta A(T - \Delta T)}{2 \Delta T}.$$

The free energy ΔA can be computed using TI and (4.49) for example. This is a somewhat indirect way of calculating ΔS and practically difficult in many cases since it requires computing ΔA with great accuracy.

Following the same procedure which was applied to the calculation of ΔA, TI allows direct computation of the difference in entropy:

$$\frac{\partial S_\lambda}{\partial \lambda} = \frac{1}{T}\left(\frac{\partial \langle \mathscr{H}_\lambda \rangle}{\partial \lambda} - \frac{\partial A_\lambda}{\partial \lambda}\right)$$

$$= \frac{-\beta}{T}\left(\left\langle \mathscr{H}_\lambda \frac{\partial \mathscr{H}_\lambda}{\partial \lambda}\right\rangle - \langle \mathscr{H}_\lambda \rangle \left\langle \frac{\partial \mathscr{H}_\lambda}{\partial \lambda}\right\rangle\right).$$

After integration, the final expression for ΔS is

$$\Delta S = \frac{-\beta}{T}\int_0^1 \left(\left\langle \mathscr{H}_\lambda \frac{\partial \mathscr{H}_\lambda}{\partial \lambda}\right\rangle - \langle \mathscr{H}_\lambda \rangle \left\langle \frac{\partial \mathscr{H}_\lambda}{\partial \lambda}\right\rangle\right)\,\mathrm{d}\lambda.$$

Other methods for calculating entropy are discussed in Sect. 2.10.

4.10.1 Parametrization of $\mathscr{H}_\lambda$

For simple calculations such as the one given above, a linear parametrization of the Hamiltonian is sufficient. However in most cases, one is interested in 'growing' new atoms or 'removing' atoms. This is the case if for example one is transforming a glycine residue into an alanine. See Fig. 4.14.

In that case, one might run into problems when trying to calculate $dA/\mathrm{d}\lambda$ at $\lambda = 0$. For that value, CH_3 does not 'feel' any other atoms since the Hamiltonian functions does not contain any terms related to CH_3. However, $\partial \mathscr{H}/\partial \lambda$ will be extremely large each time a water molecule overlaps with CH_3. Those very large values are van der Waals energy terms which scale like $(\sigma/r)^{12} - (\sigma/r)^6$ and become large for small values of r. Consequently, doing such a calculation might be very challenging.

As we have discussed in Sect. 2.8.5 a convenient approach to remedy this is to use soft-core potentials [55], in which the Lennard-Jones potential is replaced by

$$4(1-\lambda)\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right] \Rightarrow 4(1-\lambda)\epsilon\left[\frac{\sigma^{12}}{[\alpha\lambda^2\sigma^6 + r^6]^2} - \frac{\sigma^6}{\alpha\lambda^2\sigma^6 + r^6}\right].$$
$$(4.50)$$
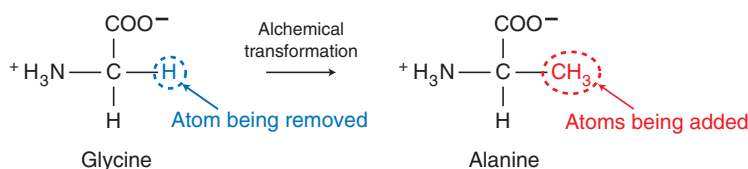
Fig. 4.14. A glycine residue is changed into an alanine. The H atom disappears while CH_3 appears

When $\lambda = 0$ one recovers the Lennard-Jones potential. When $\lambda = 1$, the atom is annihilated smoothly and the singularity disappears progressively. The parameter α can be chosen to increase the smoothness of the free energy. A small α results in a near singularity around $\lambda = 1$ while a large α results in a near singularity around $\lambda = 0$. See article by Beutler et al. [55] for an algorithm to calculate an appropriate value of α.

4.10.2 Thermodynamic Cycle

It is often the case that alchemical transformations are used to compare the binding affinity of two ligands $\mathscr{L}_1$ and $\mathscr{L}_2$ to a receptor molecule R. For example $\mathscr{L}_1$ and $\mathscr{L}_2$ might be two putative inhibitors of an enzyme R. If ΔA_1 (respectively, ΔA_2) is the free energy of binding $\mathscr{L}_1$ (respectively, $\mathscr{L}_2$) to R, we can define the relative binding affinity by $\Delta\Delta A = \Delta A_2 - \Delta A_1$.

The thermodynamic cycle shown in Fig. 4.15 can be used to greatly simplify this type of calculation. For large molecules, computing the binding energy directly, by modeling the association of say $\mathscr{L}_1$ and R, is a difficult task since large rearrangements in R are involved. However, if one is interested in $\Delta\Delta A$ only, then a shortcut can be taken by observing that $\Delta A_1 + \Delta A_4 = \Delta A_3 + \Delta A_2$. Therefore instead of computing $\Delta A_2 - \Delta A_1$ one can compute $\Delta A_4 - \Delta A_3$. ΔA_3 and ΔA_4 correspond to the mutation of $\mathscr{L}_1$ to $\mathscr{L}_2$ either in water (ΔA_3) or bound to R (ΔA_4). Even though experimentally only ΔA_1 and ΔA_2 are accessible, numerically it is easier to calculate ΔA_3 and ΔA_4. See the discussion in Sect. 2.8.3.

4.10.3 λ Dynamics

It is possible to treat the parameter λ in the alchemical transformation as a dynamic variable using an extended ensemble [56]. For simplicity of implementation, it has been proposed to use two variables λ_0 and λ_1 such that $\lambda_0^2 + \lambda_1^2 = 1$. The Hamiltonian function is then defined as [57, 58]

$$\mathscr{H}^e(\mathbf{x}, \mathbf{p}, \lambda_0, \lambda_1, p_0, p_1) = K(\mathbf{p}) + \frac{1}{2m_\lambda}(p_0^2 + p_1^2) + \lambda_0^2 U_0(\mathbf{x}) + \lambda_1^2 U_1(\mathbf{x}), \quad (4.51)$$
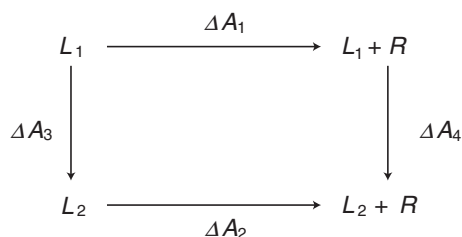
Fig. 4.15. Thermodynamic cycle for the binding of $\mathscr{L}_1$ and $\mathscr{L}_2$ to a receptor molecule R. Calculating ΔA_3 and ΔA_4 is often easier than ΔA_1 and ΔA_2

where K is the kinetic energy, U_0 and U_1 the potential energy for the two systems, and p_0 and p_1 are momentum variables associated with λ_0 and λ_1. With this approach we do not need to enforce $0 < \lambda < 1$. However, we do need to enforce $\lambda_0^2 + \lambda_1^2 = 1$. Bitetti-Putzer et al. [59] has argued that this approach leads to an overall improved sampling compared to a simulation with λ fixed.

Here we propose a different parametrization which removes the need for constraints. The following Hamiltonian $\mathscr{H}^e$ is defined:

$$\mathscr{H}^e(\mathbf{x}, \mathbf{p}, \theta, p_\theta) = K(\mathbf{p}) + \frac{p_\theta^2}{2m_\theta} + \cos^2(\theta)\, U_0(\mathbf{x}) + \sin^2(\theta)\, U_1(\mathbf{x}).$$

The equations of motion for θ are

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{p_\theta}{m_\theta}, \quad \frac{\mathrm{d}p_\theta}{\mathrm{d}t} = -\sin(2\theta)\Delta U, \quad \Delta U = U_1 - U_0.$$

The sin function results from the use of $\cos^2(\theta)$ as the weight function. The free energy difference can now be calculated as

$$\Delta A = \int_0^{\pi/2} \left\langle \frac{\partial \mathscr{H}^e}{\partial \theta} \right\rangle_\theta \mathrm{d}\theta = \int_0^{\pi/2} \sin(2\theta) \left\langle \Delta U \right\rangle_\theta \mathrm{d}\theta = \int_0^1 \left\langle \Delta U \right\rangle_\lambda \mathrm{d}\lambda$$

with $\lambda = \sin^2\theta$.

The λ dynamic approach can be combined with other techniques used for calculating the PMF. One of them is the non-Boltzmann TI of Ota and Brunger [60, 61]. ABF can also be used in combination with λ dynamics to accelerate the sampling along λ.

It has been argued that using several λ parameters λ_0, λ_1, ..., λ_L one can perform several mutation studies *simultaneously* in so-called competitive binding experiments [57]. This approach allows the calculation of ΔA_{ij} for any two ligands and rapid ranking of compounds according to their binding affinity. In this case, the potential U is defined as

$$U(\mathbf{x}, \lambda_0, \ldots, \lambda_L) = \sum_{i=1}^{L} \lambda_i^2\, U_i(\mathbf{x}), \quad \sum_{i=1}^{L} \lambda_i^2 = 1.$$

Guo et al. [57] used this approach to calculate the binding affinities of different inhibitors of trypsin, shown in Fig. 4.16. They proposed to improve the sampling by adding the following biasing potential:

$$U(\mathbf{x}, \lambda_0, \cdots, \lambda_L) = \sum_{i=1}^{L} \lambda_i^2\, (U_i(\mathbf{x}) - F_i),$$

where F_i is iteratively modified so that all molecule types are visited with equal probability. The weighted histogram method [62] can be used to obtain the final estimate for the free energy differences ΔA_{ij}.
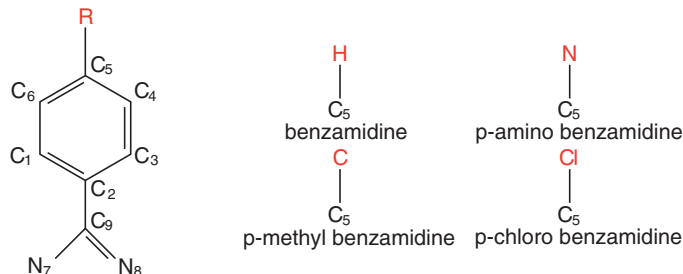
Fig. 4.16. The λ dynamics method for alchemical transformations was developed by Guo and Brooks [57] for rapid screening of binding affinities. In this approach the parameter λ is a dynamic variable. Techniques like ABF or metadynamics [34] can be used to accelerate this type of calculation. λ dynamics was used by Guo [57] to study the binding of benzamidine to trypsin. One simulation is sufficient to gather data on several benzamidine derivatives. Substitutions were made at the *para* position C_5 (H, NH_2, CH_3 and Cl). The hydrogen atoms are not shown for clarity

Again a different parametrization allows one to discard the constraint $\sum_{i=1}^{L} \lambda_i^2 = 1$. We define U by

$$U(\mathbf{x}, \theta_1, \ldots, \theta_{L-1}) = \sum_{i=1}^{L-1} P_i \cos^2(\theta_i) U_i + P_L U_L,$$

where the λ parameters are replaced by θ. The P_i are defined as $P_i = \prod_{j<i} \sin^2(\theta_j)$. In order to calculate $\mathrm{d}p_{\theta_i}/\mathrm{d}t$ conveniently, let us further introduce

$$\mathscr{U}_{k+1}(\mathbf{x}, \theta_1, \ldots, \theta_{L-1}) = \sum_{i>k}^{L-1} Q_i^k \cos^2(\theta_i) U_i + Q_L^k U_L, \quad Q_i^k = \prod_{k<j<i} \sin^2(\theta_j).$$

In terms of $\mathscr{U}_k$, a new expression for U can be obtained

$$U = \sum_{i<k} P_i \cos^2(\theta_i) U_i + P_k \Big[\cos^2(\theta_k) U_k + \sin^2(\theta_k) \mathscr{U}_{k+1} \Big].$$

In the previous equation, the dependence on a given θ_k has been isolated. This allows an equation of motion for p_{θ_i} to be derived:

$$\frac{\mathrm{d}p_{\theta_i}}{\mathrm{d}t} = -\sin(2\theta_i) P_i \, \Delta\mathscr{U}_{i+1}, \quad \Delta\mathscr{U}_{i+1} = \mathscr{U}_{i+1} - U_i.$$

This approach simplifies λ dynamics simulations by removing the need for a constraint on λ_i.

4.11 Conclusion

In this chapter, we have discussed various methods to calculate the PMF and alchemical transformation.

We described various approaches based on TI, which attempts to calculate the derivative of A by computing a mean force acting on ξ [see (4.2)]. Historically, this was first done by running a constrained simulation and averaging the constraint force $\lambda \nabla \xi$. This is the force that needs to be applied to the system in order to keep ξ constant [see (4.28)]. A more recent technique, the ABF, does not require a constraint to be applied and yields improved sampling by removing the mean force acting on ξ [see (4.33)]. The resulting system randomly samples conformations along ξ with uniform probability. This leads to very good convergence. ABF is an adaptive method, in which the external force is constantly modified as additional samples are gathered and more-accurate averages are obtained. In the limit of a large number of sample points we converge to a Hamiltonian system and a uniform sampling along ξ. ABF is a very efficient method when sampling the conformational space is rendered difficult by the presence of energy barriers. It works well with a single or many reaction coordinates.

Many other techniques have been developed for the PMF such as SMD, which applies an external force to drive the system along a certain path or metadynamics which biases the system away from free energy minima in an attempt to sample more efficiently transition states and rare events.

At the end of the chapter, techniques for alchemical transformations were presented. We showed that, in order to avoid rapid changes in free energy and improve the efficiency of the calculation, the parametrization of the Hamiltonian is critical and soft-core potentials should be used [see (4.50)]. A popular approach is the technique of λ dynamics which leads to an improved sampling. In this approach λ is a variable in the Hamiltonian system [see (4.51)]. Umbrella sampling, metadynamics or ABF can be used to reduce the cost of λ dynamics simulations.

Appendix

A Proof of the Constraint Force Equation

Consider the constrained Hamiltonian with the Fixman potential

$$\mathscr{H}^{\mathrm{F}} = \mathscr{H} + \frac{1}{2\beta} \ln Z_\xi + \lambda^{\mathrm{F}}(\xi - \xi(\mathbf{x})).$$

This Hamiltonian is distinct from $\mathscr{H}_\xi^{\mathrm{F}}$ since it is a function of $(\mathbf{x}, \mathbf{p}_x)$ (set of Cartesian coordinates) whereas $\mathscr{H}_\xi^{\mathrm{F}}$ is a function of $(\mathbf{q}, \mathbf{p}_q)$ (generalized coordinates).

We use the Lagrangian formulation since the constraint $\dot{\xi} = 0$ can be naturally expressed in this setting. The Lagrangian $\mathscr{L}^{\mathrm{F}}$ is associated with $\mathscr{H}^{\mathrm{F}}$

$$\mathscr{L}^{\mathrm{F}} = \mathscr{L} - \frac{1}{2\beta} \ln Z_\xi(\mathbf{q}) - \lambda^{\mathrm{F}}(\xi - \xi(\mathbf{x})).$$

Lagrange's equations are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathscr{L}^{\mathrm{F}}}{\partial \dot{\xi}}\right) \stackrel{\mathrm{def}}{=} \left(\frac{\partial \mathscr{L}^{\mathrm{F}}}{\partial \xi}\right) = \frac{\partial \mathscr{L}}{\partial \xi} - \frac{1}{2\beta}\frac{\partial}{\partial \xi}\ln Z_{\xi} + \lambda^{\mathrm{F}}.$$

Since $\dot{\xi} = 0$, we furthermore have

$$\frac{\partial \mathscr{L}}{\partial \xi} = \frac{\partial}{\partial \xi}\left(\frac{1}{2}\left(\frac{\dot{\xi}}{\dot{\mathbf{q}}}\right)^{t} M^{\mathrm{G}}\left(\frac{\dot{\xi}}{\dot{\mathbf{q}}}\right) - U\right)$$

$$= \frac{\partial}{\partial \xi}\left(\frac{1}{2}\dot{\mathbf{q}}^{t} M_{q}^{\mathrm{G}} \dot{\mathbf{q}} - U\right) = \frac{\partial \mathscr{L}_{\xi}}{\partial \xi},$$

where $\mathscr{L}_{\xi}$ is the Lagrangian of the constrained system $(\mathbf{q}, \mathbf{p}_{q})$ [see (4.22)]. We therefore get the following result for λ^{F}:

$$\lambda^{\mathrm{F}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathscr{L}^{\mathrm{F}}}{\partial \dot{\xi}}\right) - \frac{\partial \mathscr{L}_{\xi}}{\partial \xi} + \frac{1}{2\beta}\frac{\partial}{\partial \xi}\ln Z_{\xi}.$$

Since our expression for $\mathrm{d}A/\mathrm{d}\xi$ involves $\mathscr{H}_{\xi}^{\mathrm{F}}$ (4.26), we need to replace the Lagrangian $\mathscr{L}_{\xi}$ by $\mathscr{H}_{\xi}^{\mathrm{F}}$.

For this, we show that in general for a set $(q_{1}, \ldots, q_{N}, p_{1}, \ldots, p_{N})$

$$\left.\frac{\partial \mathscr{H}}{\partial q_{1}}\right|_{q_{i}, i\neq 1, p} = -\left.\frac{\partial \mathscr{L}}{\partial q_{1}}\right|_{q_{i}, i\neq 1, \dot{q}} \tag{4.52}$$

To see this, we use the definition of $\mathscr{H}$ and $\mathscr{L}$

$$\mathscr{H}(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^{t} Z \mathbf{p} + U,$$

$$\mathscr{L}(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\dot{\mathbf{q}}^{t} Z^{-1} \dot{\mathbf{q}} - U$$

with $Z\mathbf{p} = \dot{\mathbf{q}}$. Since

$$\frac{\partial Z^{-1}}{\partial q_{1}} = -Z^{-1}\frac{\partial Z}{\partial q_{1}}Z^{-1}$$

we see that $\partial \mathscr{H}/\partial q_{1} = -\partial \mathscr{L}/\partial q_{1}$.

With (4.52), we have

$$\lambda^{\mathrm{F}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathscr{L}^{\mathrm{F}}}{\partial \dot{\xi}}\right) + \frac{\partial \mathscr{H}_{\xi}}{\partial \xi} + \frac{1}{2\beta}\frac{\partial}{\partial \xi}\ln Z_{\xi} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathscr{L}^{\mathrm{F}}}{\partial \dot{\xi}}\right) + \frac{\partial \mathscr{H}_{\xi}^{\mathrm{F}}}{\partial \xi} \tag{4.53}$$

using the definition (4.25) of $\mathscr{H}_{\xi}^{\mathrm{F}}$.

Using (4.26) for $\mathrm{d}A/\mathrm{d}\xi$, we now have an expression for the derivative which involves the Lagrange multiplier λ^{F} and the Lagrangian $\mathscr{L}^{\mathrm{F}}$

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \left\langle \lambda^{\mathrm{F}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathscr{L}^{\mathrm{F}}}{\partial \dot{\xi}}\right\rangle^{\mathrm{F}}_{\xi, \dot{\xi}} \tag{4.54}$$

The last step consists of proving that $\langle \frac{\mathrm{d}}{\mathrm{d}t}(\partial \mathscr{L}^{\mathrm{F}}/\partial \dot{\xi}) \rangle^{\mathrm{F}}_{\xi,\dot{\xi}} = 0$ For Hamiltonian systems, it is generally true that the phase-space average of the time derivative of any function of $\mathbf{q}$ and $\mathbf{p}$ is equal to zero. For an arbitrary function $f(\mathbf{q},\mathbf{p})$ and a Hamiltonian function $\mathscr{H}$:

$$
\left\langle \frac{df}{dt} \right\rangle = \frac{\displaystyle\int \frac{\mathrm{d}f}{\mathrm{d}t}\, \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}} = \frac{\displaystyle\int \left(\sum_i \frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right) \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}
$$

$$
= \frac{\displaystyle\int \left(\sum_i \frac{\partial f}{\partial q_i}\frac{\partial \mathscr{H}}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial \mathscr{H}}{\partial q_i} \right) \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}
$$

$$
= k_{\mathrm{B}}T\, \frac{\displaystyle\int \left(\sum_i \frac{\partial^2 f}{\partial q_i \partial p_i} - \frac{\partial^2 f}{\partial q_i \partial p_i} \right) \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}}{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}}\, \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}} = 0
$$

using an integration by parts in the last line. Therefore, (4.53) means that, even though $\partial \mathscr{H}^{\mathrm{F}}_{\xi}/\partial \xi$ is different from λ^{F} pointwise, their averages are equal.

Thus, using (4.54), we have proved the following result:

$$
\frac{\mathrm{d}A}{\mathrm{d}\xi} = \langle \lambda^{\mathrm{F}} \rangle^{\mathrm{F}}_{\xi,\dot{\xi}}. \tag{4.55}
$$

B Connection between the Lagrange Multiplier and the Configurational Space Averaging

When using (4.28) with a Lagrange multiplier, the velocity appears explicitly [see (4.20)]

$$
\left\langle Z_{\xi}^{-1/2}\left(\lambda + \frac{1}{2\beta Z_{\xi}} \sum_i \frac{1}{m_i}\frac{\partial \xi}{\partial x_i}\frac{\partial \ln Z_{\xi}}{\partial x_i} \right) \right\rangle_{\xi,\dot{\xi}}
$$

$$
= \left\langle Z_{\xi}^{-3/2}\left(\nabla U^t \cdot M^{-1} \cdot \nabla \xi - \dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}} + \frac{1}{2\beta}\sum_i \frac{1}{m_i}\frac{\partial \xi}{\partial x_i}\frac{\partial \ln Z_{\xi}}{\partial x_i} \right) \right\rangle_{\xi,\dot{\xi}}
$$
$$\tag{4.56}$$

We will now show that this expression reduces to (4.15), in which the average is computed in configurational space only (no velocity). To show that, we need to calculate analytically the contribution of the velocity term which is

$$
-\langle Z_{\xi}^{-1}\dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}} \rangle_{\xi,\dot{\xi}}
$$

Let us integrate over $\mathbf{p}_q$ in this average. With the condition $\dot{\xi} = 0$, we get

$$\int e^{-\beta/2\,(\mathbf{p}_q)^t (M_q^G)^{-1} \mathbf{p}_q}\; \dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}}\; d\mathbf{p}_q$$

$$= \int e^{-\beta/2\,\dot{\mathbf{q}}^t M_{\mathbf{q}}^G \dot{\mathbf{q}}}\; \left(\begin{pmatrix} \dot{\xi} \\ \dot{\mathbf{q}} \end{pmatrix}^t \cdot J^t \mathbf{H} J \cdot \begin{pmatrix} \dot{\xi} \\ \dot{\mathbf{q}} \end{pmatrix} \right)\; |M_q^G|\; d\dot{\mathbf{q}}$$

$$\text{(change of variable from } \mathbf{p}_q \text{ to } \dot{q})$$

$$= \int e^{-\beta/2\,\dot{\mathbf{q}}^t M_q^G \dot{\mathbf{q}}}\; \left(\dot{\mathbf{q}}^t \cdot (J_q)^t \mathbf{H} J_q \cdot \dot{\mathbf{q}} \right)\; |M_q^G|\; d\dot{\mathbf{q}}$$

where M_q^G is defined by (4.23) and J_q is a $N \times (N-1)$ matrix defined by

$$[J_q]_{ij} = \frac{\partial x_i}{\partial q_j}, \quad \text{and in particular } M_q^G = J_q^t M J_q.$$

The integral can be further simplified using the following property of Gaussian integrals for any matrix A and B:

$$\int u^t \cdot B \cdot u\; \exp(-u^t \cdot A \cdot u)\, du = \frac{\text{Tr}(A^{-1}B)}{2} \int \exp(-u^t \cdot A \cdot u)\, du.$$

With this result, we obtain

$$\int e^{-\beta/2\,\dot{\mathbf{q}}^t M_q^G \dot{\mathbf{q}}}\; \left(\dot{\mathbf{q}}^t \cdot (J_q)^t \mathbf{H} J_q \cdot \dot{\mathbf{q}} \right)\; |M_{\mathbf{q}}^G|\; d\dot{q} =$$

$$= k_B T \text{Tr}\left((M_q^G)^{-1}(J_q)^t \mathbf{H} J_q \right) \int e^{-\beta/2\,(\mathbf{p}_q)^t (M_q^G)^{-1} \mathbf{p}_q}\; d\mathbf{p}_q \tag{4.57}$$

$$= k_B T \text{Tr}\left(J_q (M_q^G)^{-1}(J_q)^t \mathbf{H} \right) \int e^{-\beta/2\,(\mathbf{p}_q)^t (M_q^G)^{-1} \mathbf{p}_q}\; d\mathbf{p}_q.$$

A more explicit expression for the matrix $J_q (M_q^G)^{-1}(J_q)^t$ can be obtained

$$\left[J_q (M_q^G)^{-1}(J_q)^t \right]_{ij} = \frac{\delta_{ij}}{m_i} - Z_\xi^{-1} \left(\frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \right) \left(\frac{1}{m_j} \frac{\partial \xi}{\partial x_j} \right). \tag{4.58}$$

This is proved in Appendix C. Going back to (4.57), we can compute the trace

$$\left\langle \dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}} \right\rangle_{\mathbf{p}_q}$$

$$= \left\langle k_B T \left(\sum_i \frac{1}{m_i} \mathbf{H}_{ii} - Z_\xi^{-1} \sum_{ij} \left(\frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \right) \left(\frac{1}{m_j} \frac{\partial \xi}{\partial x_j} \right) \mathbf{H}_{ij} \right) \right\rangle_{\mathbf{p}_q}$$

$$
= \left\langle k_\mathrm{B} T \left(\sum_i \frac{1}{m_i} \frac{\partial^2 \xi}{\partial x_i^2} - Z_\xi^{-1} \sum_{ij} \frac{1}{m_i m_j} \frac{\partial \xi}{\partial x_i} \frac{\partial^2 \xi}{\partial x_i \partial x_j} \frac{\partial \xi}{\partial x_j} \right) \right\rangle_{\mathbf{P}_q}
$$

$$
= \left\langle k_\mathrm{B} T \left(\sum_i \frac{1}{m_i} \frac{\partial^2 \xi}{\partial x_i^2} - \frac{1}{2} \sum_i \frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \frac{\partial \ln Z_\xi}{\partial x_i} \right) \right\rangle_{\mathbf{P}_q}.
$$

Using the previous equation, we now simplify (4.56)

$$
\left\langle Z_\xi^{-1/2} \left(\lambda + \frac{1}{2\beta Z_\xi} \sum_i \frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \frac{\partial \ln Z_\xi}{\partial x_i} \right) \right\rangle
$$

$$
= \left\langle Z_\xi^{-3/2} \left(\nabla U^t \cdot M^{-1} \cdot \nabla \xi - \dot{\mathbf{x}}^t \cdot \mathbf{H} \cdot \dot{\mathbf{x}} + \frac{1}{2\beta} \sum_i \frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \frac{\partial \ln Z_\xi}{\partial x_i} \right) \right\rangle
$$

$$
= \left\langle Z_\xi^{-3/2} \left(\nabla U^t \cdot M^{-1} \cdot \nabla \xi - \sum_i \frac{1}{\beta m_i} \frac{\partial^2 \xi}{\partial x_i^2} - \frac{1}{\beta m_i} \frac{\partial \xi}{\partial x_i} \frac{\partial \ln Z_\xi}{\partial x_i} \right) \right\rangle
$$

$$
= \left\langle Z_\xi^{-3/2} \nabla U^t \cdot M^{-1} \cdot \nabla \xi - Z_\xi^{-1/2} \sum_i \frac{1}{\beta m_i} \frac{\partial}{\partial x_i} \left(Z_\xi^{-1} \frac{\partial \xi}{\partial x_i} \right) \right\rangle.
$$

Let us introduce the vector w

$$
w_i = \frac{1}{m_i} \frac{\partial \xi}{\partial x_i}.
$$

With this notation

$$
\left\langle Z_\xi^{-1/2} \left(\lambda + \frac{1}{2\beta Z_\xi} \sum_i \frac{1}{m_i} \frac{\partial \xi}{\partial x_i} \frac{\partial \ln Z_\xi}{\partial x_i} \right) \right\rangle
$$

$$
= \left\langle Z_\xi^{-1/2} \left(\nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} - k_\mathrm{B} T \nabla \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} \right) \right\rangle.
$$

Therefore, we have the following expression for $\mathrm{d}A/\mathrm{d}\xi$:

$$
\frac{\mathrm{d}A}{\mathrm{d}\xi} = \frac{\left\langle Z_\xi^{-1/2} \left(\nabla U \cdot \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} - k_\mathrm{B} T \nabla \frac{\mathbf{w}}{\mathbf{w} \cdot \nabla \xi} \right) \right\rangle_{\xi, \dot{\xi}}}{\left\langle Z_\xi^{-1/2} \right\rangle_{\xi, \dot{\xi}}}. \tag{4.59}
$$

This is purely an average over configurational space. This equation is similar to (4.15) with a special choice for $\mathbf{w}$ and the additional weighting factor $Z_\xi^{-1/2}$ since this is a constrained simulation.

C Calculation of $J_q(M_q^{\mathrm{G}})^{-1}(J_q)^t$

We prove that (4.58)

$$\left[J_q(M_q^{\mathrm{G}})^{-1}(J_q)^t\right]_{ij} = \frac{\delta_{ij}}{m_i} - Z_\xi^{-1}\left(\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\right)\left(\frac{1}{m_j}\frac{\partial\xi}{\partial x_j}\right). \qquad (4.60)$$

In this equation, J_q are the $N-1$ columns of J corresponding to the variables $q_1, \cdots, q_{N-1}$, of the Jacobian matrix $J = \partial\mathbf{x}/\partial\mathbf{q}$

$$[J_q]_{ij} = \frac{\partial x_i}{\partial q_j}.$$

The first column of J is $\partial\mathbf{x}/\partial\xi$. The matrix M_q^{G} is equal to $J_q^t M J_q$.

To prove (4.60), first, we multiply both sides of the equation by MJ_q and $\nabla\xi$ and show that both sides are equal. Second we show that the set of vectors

$$\left(m_i\frac{\partial x_i}{\partial q_1}, \ldots, m_i\frac{\partial x_i}{\partial q_{N-1}}, \frac{\partial\xi}{\partial x_i}\right)$$

forms a basis. This proves that the matrices on both sides of (4.60) are equal.

Left-hand side. For any $1 \le j \le N-1$

$$\left[J_q(M_q^{\mathrm{G}})^{-1}(J_q)^t\right]MJ(:,j) = J_q(M_q^{\mathrm{G}})^{-1}\left[(J_q)^t MJ(:,j)\right] = J(:,j) \qquad (4.61)$$

$$\sum_k \frac{\partial x_k}{\partial q_i}\frac{\partial\xi}{dx_k} = 0 \;\Rightarrow\; \left[J_q(M_q^{\mathrm{G}})^{-1}(J_q)^t\right]\nabla\xi = J_q(M_q^{\mathrm{G}})^{-1}\left[(J_q)^t\nabla\xi\right] = 0,$$

$$(4.62)$$

where $J(:,j)$ is column j of matrix J.

Right-hand side. Take the vector $m_k\,\partial x_k/\partial q_j$

$$\sum_k\left[\frac{\delta_{ik}}{m_i} - Z_\xi^{-1}\left(\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\right)\left(\frac{1}{m_k}\frac{\partial\xi}{\partial x_k}\right)\right]\left(m_k\frac{\partial x_k}{\partial q_j}\right) = \frac{\partial x_i}{\partial q_j},$$

$$\text{since}\quad \sum_k\frac{\partial x_k}{\partial q_j}\frac{\partial\xi}{dx_k} = 0.$$

This result is identical to (4.61).

Take the vector $\nabla\xi$

$$\sum_k\left[\frac{\delta_{ik}}{m_i} - Z_\xi^{-1}\left(\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\right)\left(\frac{1}{m_k}\frac{\partial\xi}{\partial x_k}\right)\right]\frac{\partial\xi}{\partial x_k}$$

$$= \frac{1}{m_i}\frac{\partial\xi}{\partial x_i} - Z_\xi^{-1}\left(\frac{1}{m_i}\frac{\partial\xi}{\partial x_i}\right)Z_\xi = 0.$$

This is identical to (4.62).

We now show that if $Z_\xi \neq 0$ then the following set of vectors forms a basis:

$$\left(m_i \frac{\partial x_i}{\partial q_1}, \cdots, m_i \frac{\partial x_i}{\partial q_{N-1}}, \frac{\partial \xi}{\partial x_i} \right)$$

The matrix J is invertible, so is MJ and therefore the vectors MJ_q are all independent. We need to prove that $\nabla \xi$ is not a linear combination of the columns of MJ_q. Let us suppose this is the case. Then there exists a vector μ such that $\mu_1 = 0$ and

$$\nabla \xi = MJ\mu.$$

(Recall that J has been ordered such that $\partial \mathbf{x}/\partial \xi$ is the first column.) But

$$\mu_1 = [J^{-1}M^{-1}\nabla \xi]_1 = Z_{11} = Z_\xi \neq 0, \quad (\text{since } Z = J^{-1}M^{-1}(J^{-1})^t).$$

This is a contradiction. Therefore this set of vectors forms a basis.

References

1. Kirkwood, J.G., Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **1935**, *3*, 300–313
2. Torrie, G.M.; Valleau, J.P., Non-physical sampling distributions in Monte-Carlo free-energy estimation – umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199
3. Kumar, S.; Bouzida, D.; Swendsen, R.H.; Kollman, P.A.; Rosenberg, J.M., The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J. Comput. Chem.* **1992**, *13*, 1011–1021
4. Hagan, M.F.; Dinner, A.R.; Chandler, D.; Chakraborty, A.K., Atomistic understanding of kinetic pathways for single base-pair binding and unbinding in DNA, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 13922–13927
5. Tolpekina, T.V.; den Otter, W.K.; Briels, W.J., Influence of a captured solvent molecule on the isomerization rates of calixarenes, *J. Phys. Chem. B* **2003**, *107*, 14476–14485
6. Virnau, P.; Muller, M., Calculation of free energy through successive umbrella sampling, *J. Chem. Phys.* **2004**, *120*, 10925–10930
7. Chipot, C., Rational determination of charge distributions for free energy calculations, *J. Comput. Chem.* **2003**, *24*, 409–415
8. Huang, D.M.; Geissler, P.L.; Chandler, D., Scaling of hydrophobic solvation free energies, *J. Phys. Chem. B* **2001**, *105*, 6704–6709
9. Hendrix, D.A.; Jarzynski, C., A "fast growth" method of computing free energy differences, *J. Chem. Phys.* **2001**, *114*, 5974–5981
10. Hummer, G.; Szabo, A., Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3658–3661
11. Hummer, G., Fast-growth thermodynamic integration: error and efficiency analysis, *J. Chem. Phys.* **2001**, *114*, 7330–7337
12. Liphardt, J.; Dumont, S.; Smith, S.B.; Tinoco, I.; Bustamante, C., Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, *Science* **2002**, *296*, 1832–1835
13. Crooks, G.E., Thesis, U.C. Berkeley, Department of Chemistry, 1999

14. Crooks, G.E., Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, *Phys. Rev. E* **1999**, *60*, 2721–2726
15. Crooks, G.E., Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, *J. Stat. Phys.* **1998**, *90*, 1481–1487
16. Crooks, G.E., Path-ensemble averages in systems driven far from equilibrium, *Phys. Rev. E* **2000**, *61*, 2361–2366
17. den Otter, W.K., The calculation of free-energy differences by constrained molecular-dynamics simulations, *J. Chem. Phys.* **2000**, *112*, 7283–7292
18. Rodriguez-Gomez, D.; Darve, E.; Pohorille, A., Assessing the efficiency of free energy calculation methods, *J. Chem. Phys.* **2004**, *120*, 3563–3578
19. Ciccotti, G.; Ferrario, M., Blue moon approach to rare events, *Mol. Simul.* **2004**, *30*, 787–793
20. Maragliano, L.; Ferrario, M.; Ciccotti, G., Effective binding force calculation in dimeric proteins, *Mol. Simul.* **2004**, *30*, 807–816
21. Sergi, A.; Ciccotti, G.; Falconi, M.; Desideri, A.; Ferrario, M., Effective binding force calculation in a dimeric protein by molecular dynamics simulation, *J. Chem. Phys.* **2002**, *116*, 6329–6338
22. Ciccotti, G.; Ferrario, M., Rare events by constrained molecular dynamics, *J. Mol. Liq.* **2000**, *89*, 1–18
23. Sprik, M.; Ciccotti, G., Free energy from constrained molecular dynamics, *J. Chem. Phys.* **1998**, *109*, 7737–7744
24. Depaepe, J.M.; Ryckaert, J.P.; Paci, E.; Ciccotti, G., Sampling of molecular conformations by molecular dynamics techniques, *Mol. Phys.* **1993**, *79*, 515–522
25. den Otter, W.K., Thermodynamic integration of the free energy along a reaction coordinate in Cartesian coordinates, *J. Chem. Phys.* **2000**, *112*, 7283–7292
26. den Otter, W.K.; Briels, W.J., Free energy from molecular dynamics with multiple constraints, *Mol. Phys.* **2000**, *98*, 773–781
27. den Otter, W.K.; Briels, W.J., The calculation of free-energy differences by constrained molecular-dynamics simulations, *J. Chem. Phys.* **1998**, *109*, 4139–4146
28. Darve, E.; Pohorille, A., Calculating free energies using average force, *J. Chem. Phys.* **2001**, *115*, 9169–9183
29. Darve, E.; Wilson, M.A.; Pohorille, A., Calculating free energies using a scaled-force molecular dynamics algorithm, *Mol. Simul.* **2002**, *28*, 113–144
30. Micheletti, C.; Laio, A.; Parrinello, M., Reconstructing the density of states by history-dependent metadynamics, *Phys. Rev. Lett.* **2004**, *92*, 170601–170604
31. Stirling, A.; Iannuzzi, M.; Laio, A.; Parrinello, M., Azulene-to-naphthalene rearrangement: the Car–Parrinello metadynamics method explores various reaction mechanisms, *Eur. J. Chem. Phys. Phys. Chem.* **2004**, *5*, 1558–1568
32. Iannuzzi, M.; Laio, A.; Parrinello, M., Efficient exploration of reactive potential energy surfaces using Car–Parrinello molecular dynamics, *Phys. Rev. Lett.* **2003**, *90*, 238302–238304
33. Martoňák, R.; Laio, A.; Parrinello, M., Predicting crystal structures: the Parrinello–Rahman method revisited, *Phys. Rev. Lett.* **2003**, *90*, 075503-01
34. Laio, A.; Parrinello, M., Escaping free-energy minima, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 12562–12566
35. Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K., Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation, *Biophys. J.* **1998**, *75*, 662–671
36. Isralewitz, B.; Gao, M.; Schulten, K., Steered molecular dynamics and mechanical functions of proteins, *Curr. Opin. Struct. Biol.* **2001**, *11*, 224–230

37. Lu, H.; Schulten, K., Steered molecular dynamics simulations of force-induced protein domain unfolding, *Proteins: Struct. Funct. Gene.* **1999**, *35*, 453–463

38. Weinan, E.; Vanden-Eijnden, E., in *Multiscale Modelling and Simulation*, LNCSE 39, 2004 (chapter: Metastability, conformation dynamics, and transition pathways in complex systems)

39. Ciccotti, G.; Kapral, R.; Vanden-Eijnden, E., Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics, *Chem. Phys. Chem.* **2005**, *6*, 1809–1814

40. Andersen, H.C., RATTLE: a 'velocity' version of the SHAKE algorithm for molecular dynamics calculations, *J. Comput. Phys.* **1983**, *52*, 24–34

41. Schlitter, J.; Klahn, M., The free energy of a reaction coordinate at multiple constraints: a concise formulation, *Mol. Phys.* **2003**, *101*, 3439–3443

42. Swegat, W.; Schlitter, J.; Kruger, P.; Wollmer, A., MD simulation of protein–ligand interaction: formation and dissociation of an insulin–phenol complex, *Biophys. J.* **2003**, *84*, 1493–1506

43. Schlitter, J.; Klahn, M., A new concise expression for the free energy of a reaction coordinate, *J. Chem. Phys.* **2003**, *118*, 2057–2060

44. Schlitter, J.; Swegat, W.; Mulders, T., Distance-type reaction coordinates for modelling activated processes, *J. Mol. Mod.* **2001**, *7*, 171–177

45. Carter, E.A.; Ciccotti, G.; Hynes, J.T.; Kapral, R., Constrained reaction coordinate dynamics for the simulation of rare events, *Chem. Phys. Lett.* **1989**, *156*, 472–477

46. Benjamin, I.; Pohorille, A., Isomerization reaction dynamics and equilibrium at the liquid–vapor interface of water. A molecular dynamics study, *J. Chem. Phys.* **1993**, *98*, 236–242

47. Pohorille, A.; Wilson, M.A. Isomerization reactions at aqueous interfaces. in *Reaction Dynamics in Clusters and Condensed Phases – The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, Jortner, J.; Levine, R.; Pullman, B., Eds., vol. 26. Kluwer: Dordrecht, 1993, p. 207

48. Pohorille, A.; Chipot, C.; New, M.; Wilson, M.A., Molecular modeling of protocellular functions. in *Pacific Symposium on Biocomputing '96*, Hunter, L.; Klein, T., Eds. World Scientific: Singapore, 1996, pp. 550–569

49. Pohorille, A.; Wilson, M.A.; Chipot, C., Interaction of alcohols and anesthetics with the water–hexane interface: a molecular dynamics study, *Prog. Coll. Polym. Sci.* **1997**, *103*, 29–40

50. Chipot, C.; Wilson, M.A.; Pohorille, A., Interactions of anesthetics with the water–hexane interface. A molecular dynamics study, *J. Phys. Chem. B* **1997**, *101*, 782–791

51. Pohorille, A.; New, M.H.; Schweighofer, K.; Wilson, M.A., Computer simulations of small molecules in membranes. in *Membrane permeability: 100 years since Ernst Overton*, Deamer, D., Ed., Current Topics in Membranes. Academic: San Diego, 1999, pp. 49–76

52. Hénin, J.; Chipot, C., Overcoming free energy barriers using unconstrained molecular dynamics simulations, *J. Chem. Phys.* **2004**, *121*, 2904–2914

53. Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K., Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality, *J. Chem. Phys.* **2003**, *119*, 3559–3566

54. Hénin, J.; Pohorille, A.; Chipot, C., Insights into the recognition and association of transmembrane α-helices. The free energy of α-helix dimerization in glycophorin A, *J. Am. Chem. Soc.* **2005**, *127*, 8478–8484

55. Beutler, T.C.; Mark, A.E.; van Schaik, R.C.; Gerber, P.R.; van Gunsteren, W.F., Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations, *Chem. Phys. Lett.* **1994**, *222*, 529–539

56. Kong, X.; Brooks III, C.L., λ-dynamics: a new approach to free energy calculations, *J. Chem. Phys.* **1996**, *105*, 2414–2423

57. Guo, Z.Y.; Brooks III, C.L., Rapid screening of binding affinities: application of the lambda-dynamics method to a trypsin-inhibitor system, *J. Am. Chem. Soc.* **1998**, *120*, 1920–1921

58. Guo, Z.; Brooks III, C.L.; Kong, X., Efficient and flexible algorithm for free energy calculations using the λ-dynamics approach, *J. Phys. Chem. B* **1998**, *102*, 2032–2036

59. Bitetti-Putzer, R.; Yang, W.; Karplus, M., Generalized ensembles serve to improve the convergence of free energy simulations, *Chem. Phys. Lett.* **2003**, *377*, 633–641

60. Ota, N.; Stroupe, C.; Ferreira da Silva, J.M.S.; Shah, S.A.; Mares-Guia, M.; Brunger, A.T., Non-Boltzmann thermodynamic integration (NBTI) for macromolecular systems: relative free energy of binding of trypsin to benzamidine and benzylamine, *Proteins* **1999**, *37*, 641–653

61. Ota, N.; Brunger, A.T., Overcoming barriers in macromolecular simulations: non-Boltzmann thermodynamic integration, *Theor. Chem. Acc.* **1997**, *98*, 171–181

62. Boczko, E.M.; Brooks, C.L., Constant-temperature free-energy surfaces for physical and chemical processes, *J. Phys. Chem.* **1993**, *97*, 4509–4513

5

Nonequilibrium Methods for Equilibrium Free Energy Calculations

Gerhard Hummer

5.1 Introduction and Background

In this chapter, we will show how *nonequilibrium* methods can be used to calculate *equilibrium* free energies. This may appear contradictory at first glance. However, as was shown by Jarzynski [1, 2], nonequilibrium perturbations can be used to obtain equilibrium free energies in a formally exact way. Moreover, Jarzynski's identity also provides the basis for a quantitative analysis of experiments involving the mechanical manipulation of single molecules using, e.g., force microscopes or laser tweezers [3–6].

But before proceeding to nonequilibrium averages, we briefly review two closely related and previously introduced methods for free energy calculations, thermodynamic integration and free energy perturbation theory (see Chaps. 2 and 4 for detailed discussions of these methods). In thermodynamic integration, a coupling parameter is used to describe the transition between two states, and equilibrium averages are used to evaluate derivatives of the free energy with respect to that order parameter. One then integrates the free energy derivatives along a continuous path connecting the initial and final states to obtain the free energy difference between them. As an example, Born [7] chose the charge of the ion as a coupling parameter to estimate the ion-solvation free energy, with the average solvent-induced electrostatic potential at the ion site as the corresponding free energy derivative. In free energy perturbation theory, one avoids (at least in principle) intermediate equilibrium averages and instead uses only an equilibrium ensemble of configurations in the initial state. Each of these configurations is converted instantaneously (i.e., without relaxation of the atom positions, etc.) to the final state and the resulting difference in energy is evaluated. A Boltzmann average of that energy difference then yields the free energy difference between the initial and final states.

These two seemingly distinct approaches of thermodynamic integration and perturbation can be seen as the limiting cases of a more general formalism in which the transformation between the two states proceeds at a finite rate. Seen in this light, one might also hope to obtain free energies from a transformation that converts the initial to the final state neither infinitely slowly (as in thermodynamic

integration) nor infinitely fast (as in free energy perturbation theory). Rigorous free energy calculations using transformations at a finite rate are indeed possible through Jarzynski's identity [1]. As it turns out, the simulation algorithm implementing the necessary calculations had already been used extensively before, but the simulation data were analyzed in only an approximate manner and under the assumption of near-equilibrium conditions [8–11].

In the following paragraph, free energy perturbation theory and thermodynamic integration will be briefly described to introduce the necessary notation and provide a framework for the subsequent outline of nonequilibrium free energy calculations. We consider a system with phase-space coordinates $\mathbf{z}$ (i.e., positions and momenta of all atoms) and a Hamiltonian $\mathscr{H}(\mathbf{z}; \lambda)$ that depends parametrically on a coupling parameter λ. When charging a Born ion [7], for instance, λ would be proportional to the charge of the ion, $q = \lambda e$. The Helmholtz free energy A is then a function of λ. In the canonical ensemble, it is given by

$$A(\lambda) = -\beta^{-1} \ln \int \exp[-\beta \mathscr{H}(\mathbf{z}; \lambda)] \mathrm{d}\mathbf{z}. \qquad (5.1)$$

In the canonical partition function of (5.1), we have for simplicity ignored combinatorial prefactors. *Free energy perturbation theory* [12] relies on evaluating effectively the ratio of the partition functions to obtain the free energy difference between the initial and final states corresponding to coupling parameters $\lambda = 1$ and 0 (see also Chap. 2),

$$\exp\{-\beta[A(1) - A(0)]\} = \frac{\int \exp[-\beta \mathscr{H}(\mathbf{z}; 1)] \mathrm{d}\mathbf{z}}{\int \exp[-\beta \mathscr{H}(\mathbf{z}; 0)] \mathrm{d}\mathbf{z}}$$

$$= \langle \exp\{-\beta[\mathscr{H}(\mathbf{z}; 1) - \mathscr{H}(\mathbf{z}; 0)]\}\rangle_0$$

$$= \langle \exp\{+\beta[\mathscr{H}(\mathbf{z}; 1) - \mathscr{H}(\mathbf{z}; 0)]\}\rangle_1^{-1}, \qquad (5.2)$$

where the forward and backward averages $\langle \cdots \rangle_0$ and $\langle \cdots \rangle_1$ are evaluated in the reference states $\lambda = 0$ and 1, respectively. As discussed in Chap. 4, free energy differences can also be obtained from *thermodynamic integration* by calculating derivatives of the free energy with respect to the coupling parameter. The derivative can be expressed as an ensemble average

$$\frac{\partial A(\lambda)}{\partial \lambda} = \frac{\int \frac{\partial \mathscr{H}}{\partial \lambda} \exp[-\beta \mathscr{H}(\mathbf{z}; \lambda)] \mathrm{d}\mathbf{z}}{\int \exp[-\beta \mathscr{H}(\mathbf{z}; \lambda)] \mathrm{d}\mathbf{z}}$$

$$= \left\langle \frac{\partial \mathscr{H}}{\partial \lambda} \right\rangle_\lambda, \qquad (5.3)$$

where the average $\langle \cdots \rangle_\lambda$ is in state λ. To obtain the difference in free energy between states defined by coupling parameter values of 0 and 1, one can then integrate (5.3)

$$A(1) - A(0) \equiv \Delta A = \int_0^1 \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_\lambda d\lambda . \tag{5.4}$$

Without explicit analytical expressions for the free energy derivative (as could be obtained for the case of a Born ion in a dielectric medium), the integral has to be evaluated numerically by simulation.

As an alternative approach to thermodynamic integration, we now consider creating an equilibrium configuration in the state $\lambda = 0$ and then slowly changing λ from 0 to 1 [8, 9]. As the coupling parameter is advanced, the system continues to sample phase space (e.g., by molecular dynamics or Monte Carlo simulations), but under an explicitly time-dependent Hamiltonian (because of the evolving coupling parameter). In the limit of a very slow transformation, with some caveats for Hamiltonian dynamics [13], the system will remain close to equilibrium. The integral in (5.4) can then be evaluated by changing λ continuously

$$\Delta A = \lim_{\tau \to \infty} \int_0^\tau \left. \frac{\partial \mathcal{H}}{\partial \lambda} \right|_{\lambda = \lambda(t)} \dot{\lambda}(t) dt , \tag{5.5}$$

where the time derivative of the coupling parameter λ is denoted by a dot. The coupling parameter λ moves along a continuous prescribed path $\lambda = \lambda(t)$ connecting the initial $[\lambda(0) = 0]$ to the final state $[\lambda(\tau) = 1]$. The time interval of the transformation measures the actual time (in molecular dynamics) or the number of steps (e.g., in Monte Carlo simulations). In (5.5), the limit of $\tau \to \infty$ ensures that the transformation is performed infinitely slowly, and thus reversibly. The right-hand side of (5.4) correspondingly defines the 'reversible work' done on the system during the transformation.

If the system is instead transformed between the initial and final states over a finite time interval τ, the free energy obtained from (5.5) without the $\tau \to \infty$ limit will only be approximate. In the corresponding *slow-growth* method of free energy calculations [8, 9], the system will not be able to sample the phase space exhaustively at each value of λ, rendering the transformation *irreversible*. As the transformation proceeds, the system will be gradually driven out of equilibrium, causing hysteresis effects. From the second law of thermodynamics, we expect that the work $W(\tau)$ performed during the nonequilibrium transformation is on average larger than or equal to the free energy difference between the two states

$$\langle W(\tau) \rangle \geq \Delta A, \tag{5.6}$$

with the difference between the work and free energy accounting for heat-dissipation effects. The work $W(\tau)$ performed on the system is the accumulated energetic cost (or gain) required to change the system

$$\boxed{W(\tau) = \int_0^\tau \left. \frac{\partial \mathcal{H}[\mathbf{z}(t); \lambda]}{\partial \lambda} \right|_{\lambda = \lambda(t)} \dot{\lambda}(t) dt .} \tag{5.7}$$

The average $\langle \cdots \rangle$ in (5.6) is over many repetitions of the transformation initiated from an equilibrium ensemble of configurations at $\lambda = 0$, with the same prescribed

time evolution of the coupling parameter $\lambda(t)$. The equality in (5.6) will normally be achieved only if the transformation is infinitely slow, $\tau \to \infty$. Otherwise, the work performed during the transformation will, on average, exceed the free energy difference between the two states. For paths of finite length, the amount of dissipated work, $\langle W(\tau) \rangle - \Delta A \geq 0$, will depend on the chosen transformation path $\lambda(t)$.

In the following, we first derive Jarzynski's identity by exploiting its close relation to the Feynman–Kac identity for path integrals (Sects. 5.2 and 5.3). In addition, we present explicit derivations for Hamiltonian dynamics, and illustrate the theory for the analytically solvable case of a moving harmonic oscillator. We then introduce and derive the Crooks identity (Sects. 5.4 and 5.5) in which the perturbation is performed in both forward and backward directions. Section 5.6 covers the implementation of nonequilibrium free energy estimates in computer simulations, including time integration, choice of order parameters, creation of initial conditions, and allocation of computer time. The analysis of the simulation results and the choice of free energy estimators are discussed in Sect. 5.7. A simple nonequilibrium perturbation is used for illustration (Sect. 5.8). Section 5.9 describes how nonequilibrium methods such as steered molecular dynamics simulations can be used to determine potentials of mean force. The chapter concludes with a discussion of selected applications and a summary.

5.2 Jarzynski's Identity

Jarzynski has shown that, even for nonequilibrium paths, the inequality (5.6) can be turned into an equality [1]. Jarzynski's identity states that

$$\boxed{\langle \exp[-\beta W(\tau)] \rangle = \exp(-\beta \Delta A)} \tag{5.8}$$

where a prescribed path $\lambda(t)$ connects the initial and final states, with $\lambda(0) = 0$ and $\lambda(\tau) = 1$, and $\Delta A = A(1) - A(0)$ the free energy difference between the two states. The average $\langle \cdots \rangle$ in (5.8) requires some explanation. It is a combination of an ensemble average over initial conditions, and a path average over trajectory realizations. Initial conditions are chosen according to the equilibrium Boltzmann probability in the $\lambda(0)$ state. The path average samples all realizations of dynamic paths, weighted by their respective path action, under the time evolution of the system with an explicitly time-varying Hamiltonian. For deterministic dynamics, only a single trajectory exists for any given initial condition, making the path average unnecessary. For stochastic dynamics, the path average is over realizations of noise. If the stochastic trajectories are generated using standard integrators, then the trajectory average is over different sequences of uncorrelated random numbers.

Jarzynski's identity, (5.8), immediately leads to the second law in the form of (5.6) because of Jensen's inequality, $\langle e^{-x} \rangle \geq e^{-\langle x \rangle}$. Moreover, in the limit of an *infinitely fast* transformation, $\tau \to 0$, we recover free energy perturbation theory. In that limit, the configurations will not relax during the transformation. The average in

(5.8) is then only over the initial conditions, and the work is simply the change in the Hamiltonian when going from the initial to the final state,

$$\lim_{\tau \to 0} W(\tau) = \mathscr{H}(\mathbf{z}(0); \lambda = 1) - \mathscr{H}(\mathbf{z}(0); \lambda = 0) , \qquad (5.9)$$

making (5.2) a limiting case of (5.8). In the opposite limit of $\tau \to \infty$, i.e., an *infinitely slow* transformation, the system is changed reversibly, and we recover thermodynamic integration in the form of (5.5). In the following, a brief 'derivation' of Jarzynski's identity will be given.

5.3 Derivation of Jarzynski's Identity

As shown in the following, Jarzynski's identity follows almost immediately [3] from the Feynman–Kac theorem for path integrals [14]. A brief and more 'pedagogical' discussion of the Feynman–Kac identity in the context of nonequilibrium free energy calculations as well as chemical kinetics, line-shape theory, and quantum mechanics can be found in [15]. The derivation of [3] also leads to an extension of Jarzynski's identity, which will allow us to calculate *potentials of mean force* for fluctuating observables, in addition to free energy differences between states described by control parameters, i.e., coupling parameters under external control. A free energy relation for fluctuating variables (i.e., variables that are not under external control, such as the instantaneous distance between two particles) proved central for the analysis of certain nonequilibrium experiments and computer simulations. Following [3], we assume that the phase-space density $\rho(\mathbf{z}, t)$ of the system evolves according to a Liouville-type equation:

$$\frac{\partial \rho(\mathbf{z}, t)}{\partial t} = L_t \rho(\mathbf{z}, t), \qquad (5.10)$$

where L_t is an explicitly time-dependent evolution operator. The time evolution governed by the Liouville operator L_t is consistent with the Hamiltonian $\mathscr{H}(\mathbf{z}, t)$. As an example, L_t for classical dynamics will be given below explicitly in terms of the Hamiltonian $\mathscr{H}$ (5.20), $\mathscr{H}$ depends parametrically on time, thus in effect absorbing the parametric dependences of $\mathscr{H}$ on λ, and of λ on t. As an example, for Smoluchowski diffusion (i.e., Brownian dynamics) on a time-dependent potential energy surface $V(\mathbf{z}, t)$, the Liouville operator would be $L_t = D \nabla e^{-\beta V(\mathbf{z},t)} \nabla e^{\beta V(\mathbf{z},t)}$, where D is the diffusion coefficient and $\nabla = \partial/\partial \mathbf{z}$.

We assume further that the Boltzmann distribution is stationary, $L_t e^{-\beta \mathscr{H}(\mathbf{z},t)} = 0$. This requirement is satisfied, for instance, by systems that undergo Newtonian, Langevin, or Metropolis Monte Carlo dynamics. For diffusion, $\mathscr{H}(\mathbf{z}, t)$ would just be the potential energy surface $V(\mathbf{z}, t)$, and the corresponding Boltzmann distribution is indeed stationary under the L_t given in the preceding paragraph. To avoid confusion, note that the stationarity of the Boltzmann distribution under the Liouville operator, $L_t e^{-\beta \mathscr{H}(\mathbf{z},t')} = 0$, is in general only valid at equal times $t = t'$, i.e., $L_t e^{-\beta \mathscr{H}(\mathbf{z},t')} \neq 0$ for $t \neq t'$ in most cases. In particular, if one starts at time $t = 0$

from a Boltzmann distribution corresponding to $\mathscr{H}(\mathbf{z}, 0)$ and evolves that distribution under L_t, then at some time $t > 0$ that distribution will no longer be stationary, $L_t \mathrm{e}^{-\beta \mathscr{H}(\mathbf{z},0)} \neq 0$, and the system will be driven out of equilibrium.

Now consider the un-normalized Boltzmann distribution at time t

$$\rho(\mathbf{z}, t) = \frac{\mathrm{e}^{-\beta \mathscr{H}(\mathbf{z},t)}}{\displaystyle\int \mathrm{e}^{-\beta \mathscr{H}(\mathbf{z}',0)} \mathrm{d}\mathbf{z}'} \, . \tag{5.11}$$

Since this distribution is stationary with respect to L_t [i.e., $L_t \rho(\mathbf{z}, t) = 0$], and since $\partial \rho / \partial t = -\beta(\partial \mathscr{H} / \partial t)\rho$, it follows that the above $\rho(\mathbf{z}, t)$ is a solution of the sink (or birth–death) equation

$$\frac{\partial \rho}{\partial t} = L_t \rho - \beta \frac{\partial \mathscr{H}}{\partial t} \rho \, , \tag{5.12}$$

as can be verified by direct substitution. If $\partial \mathscr{H} / \partial t \equiv 0$, (5.12) is the Liouville equation. Otherwise, the 'sink' term locally adds phase-space density where $\partial \mathscr{H}(\mathbf{z}, t)/\partial t$ is negative, and removes density where it is positive. Qualitatively, one may think of the addition and subtraction of phase-space density as compensating for the fact that trajectories evolving under the time-dependent Liouville operator (Hamiltonian) equilibrate too slowly to maintain the time-dependent Boltzmann distribution (5.11).

The solution of the sink equation (5.12), starting from an equilibrium distribution at time $t = 0$, can also be expressed as a path integral by using the Feynman–Kac theorem,

$$\rho(\mathbf{z}, t) = \left\langle \delta(\mathbf{z} - \mathbf{z}(t)) \mathrm{e}^{-\beta \int_0^t \frac{\partial \mathscr{H}}{\partial t'}(\mathbf{z}(t'),t') \mathrm{d}t'} \right\rangle , \tag{5.13}$$

where $\delta(x)$ is Dirac's delta distribution. A motivation of the Feynman–Kac theorem [14] can be found in [15]. In essence, the Feynman–Kac theorem provides a way to reweight trajectories generated without consideration of the sink term in (5.12) to take the sink term into account. In the path average of (5.13), trajectories are generated according to the Liouville operator L_t, without consideration of the sink (or birth–death) term $-\beta(\partial \mathscr{H} / \partial t)\rho$, starting from an initial distribution $\rho(\mathbf{z}, 0)$. The sink term is taken into account by weighting trajectory end points $\mathbf{z}(t)$. Consider $S(t)$, the relative weight of a trajectory end point at time t. Along a given trajectory, that weight will change during the time interval $(t, t + \mathrm{d}t)$ by $\mathrm{d}S = -\beta(\partial \mathscr{H} / \partial t)S(t)\mathrm{d}t$ according to the sink term in (5.13). With $S(0) = 1$, the aggregate weight accumulated along the trajectory is then $S(t) = \exp[-\beta \int_0^t (\partial \mathscr{H} / \partial t')\mathrm{d}t']$, the weight factor in (5.13).

Equating the two different solutions of (5.12), (5.11), and (5.13) immediately gives

$$\boxed{\frac{\mathrm{e}^{-\beta \mathscr{H}(\mathbf{z},t)}}{\displaystyle\int \mathrm{e}^{-\beta \mathscr{H}(\mathbf{z}',0)} \mathrm{d}\mathbf{z}'} = \left\langle \delta(\mathbf{z} - \mathbf{z}(t)) \mathrm{e}^{-\beta \int_0^t \frac{\partial \mathscr{H}}{\partial t'}(\mathbf{z}(t'),t') \mathrm{d}t'} \right\rangle} \tag{5.14}$$

This *identity* [3, 15] between a weighted average of nonequilibrium trajectories (r.h.s.) and the equilibrium Boltzmann distribution (l.h.s.) is implicit in the work of Jarzynski [2], and is given explicitly by Crooks [16]. The average $\langle\ldots\rangle$ is over an ensemble of trajectories starting from the equilibrium distribution at $t = 0$ and evolving according to (5.10). Each trajectory is weighted with the Boltzmann factor of the external work $W(t)$ done on the system,

$$W(t) = \int_0^t \frac{\partial\mathscr{H}(\mathbf{z}(t'), t')}{\partial t'}\mathrm{d}t' \ .$$ (5.15)

By integrating both sides of (5.14) with respect to $\mathbf{z}$, we obtain Jarzynski's identity [1, 2]

$$\mathrm{e}^{-\beta\varDelta A(t)} \equiv \frac{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}(\mathbf{z}, t)}\mathrm{d}\mathbf{z}}{\displaystyle\int \mathrm{e}^{-\beta\mathscr{H}(\mathbf{z}, 0)}\mathrm{d}\mathbf{z}} = \left\langle \mathrm{e}^{-\beta W(t)}\right\rangle \ ,$$ (5.16)

between the Boltzmann-averaged work $W(t)$ and the equilibrium free energy difference $\varDelta A(t)$ between times t and 0.

The above derivation shows that Jarzynski's identity is an immediate consequence of the Feynman–Kac theorem. This connection has not only theoretical value, but is also useful in practice. First, it forms the basis for an equilibrium thermodynamic analysis of nonequilibrium pulling experiments [3, 15]. Second, it helps in deriving a Jarzynski identity for dynamics using thermostats. Moreover, this derivation clarifies an important aspect: trajectories can be thought of as mapping initial conditions ($t = 0$) to trajectory endpoints, and the Boltzmann factor of the accumulated work reweights that map to give the desired Boltzmann distribution. Finally, it can be easily extended to transformations between steady states [17] in which non-Boltzmann distributions are stationary.

5.3.1 Hamiltonian Dynamics

In the following, we will show explicitly that the identity (5.8) holds for deterministic Newtonian dynamics. This derivation builds on the arguments given by Jarzynski in his original paper [1]. The essence of the derivation is: (1) integration of a trajectory for a fixed time t uniquely maps an initial phase point $\mathbf{z}(0)$ to a final phase point $\mathbf{z}(t)$; (2) an initial phase-space density remains unchanged along trajectories because of Liouville's theorem: $\rho(\mathbf{z}(t), t) = \rho(\mathbf{z}(0), 0)$; (3) the work along the trajectory is exactly the change in energy, $W(t) = \mathscr{H}(\mathbf{z}(t), t) - \mathscr{H}(\mathbf{z}(0), 0)$; such that (4) weighting the trajectory end-points by $\exp(-\beta W(t))$ transforms an initial phase-space distribution $\rho(\mathbf{z}(0), 0) \propto \exp(-\beta\mathscr{H}(\mathbf{z}(0), 0))$ to a distribution proportional to $\exp(-\beta\mathscr{H}(\mathbf{z}(t), t))$, i.e., a Boltzmann distribution for the final Hamiltonian.

Let us consider a system described by an explicitly time-dependent Hamiltonian $\mathscr{H}(\mathbf{p}, \mathbf{q}, t)$ where $(\mathbf{p}, \mathbf{q}) = \mathbf{z}$ is a point in phase space. Hamilton's equation of motion are

$$\dot{\mathbf{p}} = -\frac{\partial\mathscr{H}}{\partial\mathbf{q}} \ ; \quad \dot{\mathbf{q}} = \frac{\partial\mathscr{H}}{\partial\mathbf{p}} \ ; \quad \frac{\mathrm{d}\mathscr{H}}{\mathrm{d}t} = \frac{\partial\mathscr{H}}{\partial t}$$ (5.17)

The probability density $\rho(\mathbf{p}, \mathbf{q}, t)$ in phase space satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \mathbf{p}} (\rho \dot{\mathbf{p}}) + \frac{\partial}{\partial \mathbf{q}} (\rho \dot{\mathbf{q}}) = 0. \qquad (5.18)$$

From Hamilton's equations, we find that $\partial \dot{\mathbf{p}}/\partial \mathbf{p} = -\partial \dot{\mathbf{q}}/\partial \mathbf{q}$, such that

$$\frac{\partial \rho}{\partial t} + \dot{\mathbf{p}} \frac{\partial \rho}{\partial \mathbf{p}} + \dot{\mathbf{q}} \frac{\partial \rho}{\partial \mathbf{q}} = 0 . \qquad (5.19)$$

This is Liouville's equation, with the Liouville operator

$$L_t = \dot{\mathbf{p}}(\partial/\partial \mathbf{p}) + \dot{\mathbf{q}}(\partial/\partial \mathbf{q}) = (\partial \mathscr{H}/\partial \mathbf{p})(\partial/\partial \mathbf{q}) - (\partial \mathscr{H}/\partial \mathbf{q})(\partial/\partial \mathbf{p}). \quad (5.20)$$

Equation (5.19) states that the flow in phase space is incompressible. In particular, following along a trajectory starting in $(\mathbf{p}(0), \mathbf{q}(0))$, we find that

$$\frac{d\rho(\mathbf{p}(t), \mathbf{q}(t), t)}{dt} = \frac{\partial \rho}{\partial t} + \dot{\mathbf{p}} \frac{\partial \rho}{\partial \mathbf{p}} + \dot{\mathbf{q}} \frac{\partial \rho}{\partial \mathbf{q}} = 0 , \qquad (5.21)$$

where we have used the chain rule with respect to $\mathbf{p}$ and $\mathbf{q}$, and (5.19). This means that the phase-space density along a trajectory remains unchanged

$$\rho(\mathbf{p}(t), \mathbf{q}(t), t) = \rho(\mathbf{p}(0), \mathbf{q}(0), 0) . \qquad (5.22)$$

From the definition of the work, (5.15), and Hamilton's equations, (5.17), it follows that the work performed along the trajectory is

$$\begin{aligned} W(t) &= \int_0^t \frac{\partial \mathscr{H}(\mathbf{p}(t'), \mathbf{q}(t'), t')}{\partial t'} dt' \\ &= \mathscr{H}(\mathbf{p}(t), \mathbf{q}(t), t) - \mathscr{H}(\mathbf{p}(0), \mathbf{q}(0), 0) . \end{aligned} \qquad (5.23)$$

If we choose our initial density according to a Boltzmann distribution, $\rho(\mathbf{p}(0), \mathbf{q}(0), 0) = \exp(-\beta \mathscr{H}(\mathbf{p}(0), \mathbf{q}(0), 0))/\int d\mathbf{p} \, d\mathbf{q} \exp(-\beta \mathscr{H}(\mathbf{p}, \mathbf{q}, 0))$, the density at the trajectory end point at time t is identical to that at the initial point, $\rho(\mathbf{p}(t), \mathbf{q}(t), t) = \exp(-\beta \mathscr{H}(\mathbf{p}(0), \mathbf{q}(0), 0))/\int d\mathbf{p} \, d\mathbf{q} \exp(-\beta \mathscr{H}(\mathbf{p}, \mathbf{q}, 0))$. If we weight each trajectory end point by the Boltzmann factor of the accumulated work, $\exp(-\beta W(t))$, and use (5.23) for $W(t)$, we obtain a 'Jarzynski-modified' density ρ_J given by

$$\begin{aligned} \rho_J(\mathbf{p}(t), \mathbf{q}(t), t) &= \rho(\mathbf{p}(t), \mathbf{q}(t), t) \exp[-\beta W(t)] \\ &= \frac{\exp[-\beta \mathscr{H}(\mathbf{p}(t), \mathbf{q}(t), t)]}{\displaystyle\int d\mathbf{p} \, d\mathbf{q} \exp[-\beta \mathscr{H}(\mathbf{p}, \mathbf{q}, 0)]} . \end{aligned} \qquad (5.24)$$

That is, we have recovered a Boltzmann distribution according to the Hamiltonian at time t, equivalent to (5.14). Jarzynski's identity (5.8) then follows simply by integration over phase space $(\mathbf{p}, \mathbf{q})$.

This derivation addresses a potentially confusing point, namely that the final state at time t is *not* at equilibrium and may *not* have a well-defined temperature. As is clear from the derivation, temperature here is only a parameter that is once used to specify the initial condition, and then again in the Boltzmann weight of the integrated work. From this viewpoint, the reweighting is a convenient mathematical 'trick.'

5.3.2 Moving Harmonic Oscillator

To further illustrate the theory, we apply Jarzynski's identity to the analytically solvable example of a 1D moving harmonic oscillator with Hamiltonian

$$\mathscr{H}(p, q, t) = \frac{p^2}{2m} + \frac{k}{2}(q - vt)^2, \tag{5.25}$$

where m is the mass, k is the spring constant, and v is the pulling velocity. Because the time dependence of the Hamiltonian simply induces a translational shift, the free energy remains unchanged

$$A(t) = -\beta^{-1} \int dp \int dq \, e^{-\beta \mathscr{H}(p,q,t)} \equiv A(0). \tag{5.26}$$

In the following, we will show explicitly that the correct result is obtained if Jarzynski's identity is used to evaluate the free energy difference, $A(t) - A(0) = \langle \exp(-\beta W) \rangle = 0$.

For given initial conditions, trajectories in phase space satisfying Hamilton's equation of motion, (5.17), are given by

$$\begin{pmatrix} p(t) \\ q(t) \end{pmatrix} = \begin{pmatrix} \cos(\alpha t) & -(km)^{1/2} \sin(\alpha t) \\ (km)^{-1/2} \sin(\alpha t) & \cos(\alpha t) \end{pmatrix} \cdot \begin{pmatrix} p(0) \\ q(0) \end{pmatrix}$$
$$+ \begin{pmatrix} 2mv \sin^2\left(\dfrac{\alpha t}{2}\right) \\ vt - \dfrac{vm^{1/2}}{k^{1/2}} \sin(\alpha t) \end{pmatrix}, \tag{5.27}$$

where $\alpha = (k/m)^{1/2}$. As required by the Liouville theorem (5.22), the map of initial conditions $(p(0), q(0))$ to phase points $(p(t), q(t))$ at time t has a Jacobian equal to one, $|\partial(p(t), q(t))/\partial(p(0), q(0))| = 1$. According to (5.15) and (5.23), the integrated work along a trajectory is given by

$$W(t) = \int_0^t \frac{\partial \mathscr{H}}{\partial t'} dt' = v \left[mv - p(0)\right] \left[1 - \cos\left(\frac{k^{1/2}t}{m^{1/2}}\right)\right]$$
$$- v(km)^{1/2} q(0) \sin\left(\frac{k^{1/2}}{m^{1/2}} t\right), \tag{5.28}$$

which is a linear function with respect to the initial phase-space coordinates $(p(0), q(0))$. Note that so far, temperature has not appeared.

Now we specify an initial phase-space distribution as the Boltzmann distribution for the Hamiltonian at time $t = 0$

$$\rho(q = q(0), p = p(0), t = 0) = \beta \frac{\exp\left(-\dfrac{\beta p^2}{2m} - \dfrac{\beta k q^2}{2}\right)}{2\pi(m/k)^{1/2}}. \tag{5.29}$$

With this distribution being Gaussian in both $p(0)$ and $q(0)$, and the work being linear in $p(0)$ and $q(0)$, the distribution of the work obtained for Boltzmann-weighted initial conditions is Gaussian as well, with a mean and variance of

$$\langle W \rangle = 2mv^2 \sin^2\left(\frac{k^{1/2}t}{2m^{1/2}}\right), \tag{5.30}$$

$$\sigma^2(W) = 2\beta^{-1}\langle W \rangle. \tag{5.31}$$

For a Gaussian distribution of the work, second-order perturbation theory is exact, and $\langle \exp(-\beta W) \rangle = \exp[-\beta\langle W \rangle + \beta^2\,\sigma^2(W)/2]$. We thus find that the free energy difference for the moving oscillator according to (5.8) is $A(t) - A(0) = \langle W \rangle - \beta\,\sigma^2(W)/2 \equiv 0$, which is indeed the desired result (expected from translation invariance).

We note that for the harmonic Hamiltonian in (5.25) the variance of the work approaches zero in the limit of an infinitely slow transformation, $v \to 0$, $\tau \to \infty$, $v\tau = \text{const}$. However, as shown by Oberhofer et al. [13], this is not the case in general. As a consequence of adiabatic invariants of Hamiltonian dynamics, even infinitely slow transformations can result in a non-delta-like distribution of the work. Analytically solvable examples for that unexpected behavior are, for instance, harmonic Hamiltonians with time-dependent spring constants $k = k(t)$.

5.4 Forward and Backward Averages: Crooks Relation

So far, we have only considered perturbations in the forward direction. As in conventional free energy calculations, powerful relations can be derived if forward and backward perturbations are combined. With the free energy being a state function, we can reverse the path direction. This leads to several useful relations derived originally by Crooks [16, 18, 19]. In particular, we obtain immediately

$$\langle \exp[-\beta\underline{W}(\tau)] \rangle = \exp\{-\beta[A(0) - A(1)]\} \equiv \exp(\beta\Delta A), \tag{5.32}$$

where $\underline{W}(\tau)$ is the work accumulated on the reversed path, $\underline{\lambda}(t) = \lambda(\tau - t)$

$$\underline{W}(\tau) = \int_0^\tau \frac{\partial\mathscr{H}[\mathbf{z}(t); \lambda]}{\partial\lambda}\bigg|_{\lambda=\underline{\lambda}(t)} \dot{\underline{\lambda}}(t)\mathrm{d}t. \tag{5.33}$$

In combination with (5.6), this leads to an upper and lower bound for the free energy difference

$$-\langle \underline{W}(\tau) \rangle \leq \Delta A \leq \langle W(\tau) \rangle. \tag{5.34}$$

These bounds are the nonequilibrium equivalents of the Gibbs–Bogoliubov bounds discussed in Chap. 2. Having the free energy now bounded from above and below already demonstrates the power of using both forward and backward transformations. Moreover, as was shown by Crooks [18, 19], the distribution of work values from forward and backward paths satisfies a relation that is central to histogram methods in free energy calculations

$$\boxed{\frac{p_\mathrm{f}[w = W(\tau)]}{p_\mathrm{b}[w = -\underline{W}(\tau)]} = \exp[\beta(w - \Delta A)]} \tag{5.35}$$

where $p_{\mathrm{f}}[w = W(\tau)]$ and $p_{\mathrm{b}}[w = -\underline{W}(\tau)]$ are the probability densities of the work values for forward and reversed transformation paths (with a sign change in the work of the reverse path). Both are normalized, i.e., $\int p_{\mathrm{f}}(w)\mathrm{d}w = \int p_{\mathrm{b}}(w)\mathrm{d}w = 1$.

For computer simulations, (5.35) leads to accurate estimates of free energies. It is also the basis for higher-order cumulant expansions [20] and applications of Bennett's optimal estimator [21–23]. We note that Jarzynski's identity (5.8) follows from (5.35) simply by integration over w because the probability densities are normalized to 1:

$$\int p_f(W)\mathrm{e}^{-\beta W}\mathrm{d}W = \int \mathrm{e}^{-\beta \Delta A} p_b(W)\mathrm{d}W. \tag{5.36}$$

Because of the normalization condition, the right-hand side is equal to $\exp(-\beta \Delta A)$, and Jarzynski's identity follows.

5.5 Derivation of the Crooks Relation (and Jarzynski's Identity)

The Crooks relation follows from an elegant derivation of Jarzynski's identity using path-sampling ideas [18]. For instructive purposes, that derivation is briefly summarized here. Consider generating a discrete trajectory $\mathbf{z}_0 \overset{\mathscr{H}_1}{\to} \mathbf{z}_1 \overset{\mathscr{H}_2}{\to} \ldots \overset{\mathscr{H}_N}{\to} \mathbf{z}_N$, where each step $\mathbf{z}_{i-1} \overset{\mathscr{H}_i}{\to} \mathbf{z}_i$ proceeds according to Hamiltonian $\mathscr{H}_i$, for instance by using Metropolis Monte Carlo, Newtonian, or Langevin dynamics 'integrators.' After each 'time' step, the Hamiltonian is changed in a prescribed way. If each of the steps is Markovian in the full phase space (i.e., it does not depend explicitly on the preceding path), then the probability P_f of generating a particular trajectory can be factorized

$$P_f\left(\mathbf{z}_0 \overset{\mathscr{H}_1}{\to} \mathbf{z}_1 \overset{\mathscr{H}_2}{\to} \ldots \overset{\mathscr{H}_N}{\to} \mathbf{z}_N\right) = p_0(\mathbf{z}_0)\prod_{i=1}^{N} p_i(\mathbf{z}_i|\mathbf{z}_{i-1}), \tag{5.37}$$

where $p_0(\mathbf{z}_0) = \exp[-\beta(\mathscr{H}_0(\mathbf{z}_0) - A_0]$ is the normalized equilibrium Boltzmann probability density at the initial state [with $A_0 = A(0)$ its free energy], and $p_i(\mathbf{z}_i|\mathbf{z}_{i-1})$ is the transition probability from $\mathbf{z}_{i-1}$ to $\mathbf{z}_i$ under the influence of Hamiltonian $\mathscr{H}_i$. Now if these transition probabilities satisfy detailed balance (as they do under Newtonian, Langevin, or Metropolis Monte Carlo dynamics)

$$\frac{p_i(\mathbf{z}_i|\mathbf{z}_{i-1})}{p_i(\mathbf{z}_{i+1}|\mathbf{z}_i)} = \mathrm{e}^{-\beta[\mathscr{H}_i(\mathbf{z}_i) - \mathscr{H}_i(\mathbf{z}_{i-1})]} \tag{5.38}$$

then the probability P_b of picking the time-reversed path under the influence of the corresponding time-dependent Hamiltonian can be related to the probability P_f of the forward path:

$$\frac{P_f}{P_b} = \frac{P\left(\mathbf{z}_0 \overset{\mathscr{H}_1}{\to} \mathbf{z}_1 \overset{\mathscr{H}_2}{\to} \ldots \overset{\mathscr{H}_N}{\to} \mathbf{z}_N\right)}{P\left(\mathbf{z}_N \overset{\mathscr{H}_N}{\to} \mathbf{z}_{N-1} \overset{\mathscr{H}_{N-1}}{\to} \ldots \overset{\mathscr{H}_1}{\to} \mathbf{z}_0\right)} = \frac{p_0(\mathbf{z}_0)\prod\limits_{i=1}^{N} p_i(\mathbf{z}_i|\mathbf{z}_{i-1})}{p_N(\mathbf{z}_N)\prod\limits_{i=1}^{N} p_i(\mathbf{z}_{i-1}|\mathbf{z}_i)}. \tag{5.39}$$

If we now use (5.38) and substitute the equilibrium Boltzmann distributions for $p_0(\mathbf{z}_0)$ and $p_N(\mathbf{z}_N)$, we obtain

$$\frac{P_f}{P_b} = \exp\left[\beta \sum_{i=0}^{N-1} [\mathscr{H}_{i+1}(\mathbf{z}_i) - \mathscr{H}_i(\mathbf{z}_i)] - \beta(A_N - A_0)\right] = \mathrm{e}^{\beta(W - \Delta A)}, \quad (5.40)$$

where we have used that the work W is the accumulated change in the energy

$$W = \sum_{i=0}^{N-1} [\mathscr{H}_{i+1}(\mathbf{z}_i) - \mathscr{H}_i(\mathbf{z}_i)]. \quad (5.41)$$

A_N and A_0 are the free energies corresponding to Hamiltonians $\mathscr{H}_N$ and $\mathscr{H}_0$, with $\Delta A = A_N - A_0$. Equation (5.40) generalizes detailed balance to nonequilibrium trajectories. For every forward path, there is an equivalent backward path whose path probability differs exactly by a factor $\exp[\beta(W - \Delta A)]$. By construction, the work accumulated on that backward path is $\underline{W} = -W$. The distribution of the work is obtained by sampling all paths with their appropriate weights. In particular, by virtue of (5.40), we can obtain the distribution of work on forward paths also by sampling reweighted backward paths. Because the reweighting factor depends only on the work (and on ΔA), we can apply it to the work distribution directly to obtain the Crooks relation, (5.35). We note that for 'reversible' paths, the work distributions will be delta functions, $p_f(W) = p_b(W) = \delta(W - \Delta A)$, and the forward and backward paths have the same path probability.

5.6 Implementation

5.6.1 Dynamics

Implementing Jarzynski's identity in a free energy calculation is relatively straightforward. Many simulation packages already contain code to perform slow-growth thermodynamic integration. Two considerations are important. First, one is now integrating Newton's equations of motion for a time-dependent Hamiltonian, such that the errors in the trajectories $\mathbf{z}(t)$ can be larger. Integration errors are particularly relevant for fast transformations. Second, one needs to evaluate the work along a trajectory. This will normally require an approximation to the time integral, (5.15), such as a trapezoidal rule

$$W(\tau) \approx \sum_{i=0}^{N-1} \frac{t_{i+1} - t_i}{2} \left(\left.\frac{\partial \mathscr{H}}{\partial t}\right|_{t=t_i} + \left.\frac{\partial \mathscr{H}}{\partial t}\right|_{t=t_{i+1}} \right)$$

with $t_0 = 0$ and $t_N = \tau$.

In tests using the moving 1D Hamiltonian harmonic oscillator, (5.25), a velocity Verlet integrator [24] combined with trapezoidal integration of $W(t)$ performed well when compared to the analytic solution. An interesting analysis of how

long time steps can be used to accelerate the calculation without sacrificing accuracy can be found in [25].

Alternatively, one can break up the trajectory dynamics and the change in the coupling parameter into two substeps. In the first half-step, the phase point is advanced by using the short-time expansions of the propagators of the corresponding Liouville or Fokker–Planck equation [24]. In the second half step, the coupling parameter is changed at a fixed configuration, and the corresponding work accumulated,

$$W(t + \Delta t) = W(t) + \mathscr{H}[\mathbf{z}(t), \lambda(t + \Delta t)] - \mathscr{H}[\mathbf{z}(t), \lambda(t)]. \qquad (5.42)$$

This procedure follows, in effect, the derivation of Jarzynski's identity in discrete time [2, 18], as outlined in Sect. 5.5. Finally, for Hamiltonian dynamics, one can use (5.23) and calculate the work directly from the difference in total energy between trajectory start and end points.

5.6.2 Choice of Coupling Parameter

A critical element for the success of the calculations is the choice of a suitable coupling parameter λ and a path $\lambda(t)$ that connects the initial and final states. If the path is poorly chosen, the system will be driven out of equilibrium quickly and the amount of dissipated work, $\langle W(\tau) \rangle - \Delta A \geq 0$, will be large. As a result, the free energy estimate will likely be inaccurate. Finding good transformation paths may require some intuition about how one can connect states smoothly, i.e., without encountering high energy barriers on typical paths. The same general rules apply as in other free energy calculations. Specifically in the context of nonequilibrium free energy calculations, Reinhardt and coworkers [26, 27] have used variational optimization to minimize the amount of dissipated work. A seemingly different approach would be to determine a good (kinetically relevant) 'reaction coordinate' [28], and use that to define $\lambda(t)$.

5.6.3 Creation of Initial Conditions

Before trajectories can be started, an ensemble of initial conditions has to be created. Having a good representation of relevant conformations in the initial state will be particularly important if the transformation is performed relatively fast. In contrast, if the transformation is performed slowly, the system will sample relevant phase space regions for $\lambda \approx 0$ as it slowly moves away from the initial state. Ensembles of equilibrium initial conditions can be taken from a sufficiently long equilibrium simulation, possibly with accelerated sampling using methods discussed in other chapters. An important consideration is that the initial configurations are sufficiently uncorrelated. Otherwise, one may expect correlations among the resulting work values that interfere with applications of 'optimal' free energy estimators [21, 22]. A similar issue of creating representative initial conformations arises, for instance, in the calculation of rate coefficients where equilibrium configurations along the transition state surface have to be generated [29, 30]. One simple way of extending an ensemble of configurations is by resampling velocities from a Maxwell–Boltzmann distribution.

5.6.4 Allocation of Computer Time

The allocation of computer time requires some care. In particular, one must decide whether to use many short simulation runs ('fast growth') or few long runs ('slow growth'). In general, one expects the latter to give more-accurate results for a given amount of computer time [20]. As a rule of thumb, the runs should be slow enough that the standard deviation in the resulting work values is about k_BT. If the standard deviation is smaller, the statistical error in the free energy estimate is only insignificantly smaller than that from a single long run, but from multiple runs one also obtains an estimate of that error. If the standard deviation of the work is much larger, in either the forward or backward direction [31], sampling errors will arise because the estimate of $\langle \exp(-\beta W) \rangle$ will be dominated by only one or a few trajectories with low W values, as discussed in the following. Of relevance in this context is a recent result by Jarzynski [31] who showed that, among forward and backward transformations, the process with the *wider* work distribution provides a better estimate of the free energy (just as test particle insertion usually gives better estimates than particle removal in calculations of chemical potentials [32]).

5.7 Analysis of Nonequilibrium Free Energy Calculations

5.7.1 Exponential Estimator – Issues with Sampling Error and Bias

One of the challenges in using Jarzynski's identity for free energy calculations is the exponential weight $\exp(-\beta W(t))$ given to trajectories. If the work has a broad distribution with respect to k_BT, only a few trajectories at the lower tail of the work distribution will contribute significantly to the weighted average, with all others having 'exponentially' small relative weights. In more physical terms, when the transformation is conducted rapidly, most trajectories do not sample relevant regions of phase space. Rapid crossing of barriers under tension, and crossing over atypically high barriers result in hysteresis effects, and a broadened work distribution that reflects an increasing relevance of dissipation,

$$\langle W(\tau) \rangle - \Delta A \approx \beta \sigma^2(W)/2 \ . \tag{5.43}$$

This relation for the dissipated work is exact for a Gaussian work distribution of variance $\sigma^2(W)$.

Broad work distributions have two important consequences: first, the statistics will be poor; and, second, a bias in the estimator of the free energy change, $A(t) - A(0) = -\beta^{-1}\ln\langle\exp(-\beta W(t))\rangle$, will result in free energy estimates that deviate systematically from the correct free energy difference [10]. This will be discussed in depth in Chap. 6. Specifically, if the free energy is estimated from N work values W_i drawn at random from the work distribution $p_f(W)$,

$$\Delta A_{\mathrm{est}} = -\beta^{-1}\ln\left(N^{-1}\sum_{i=1}^{N}\mathrm{e}^{-\beta W_i}\right) \tag{5.44}$$

then the average $\overline{\Delta A_{\mathrm{est}}}$ over repeated sampling of work values does not converge to the correct free energy: $|\overline{\Delta A_{\mathrm{est}}} - \Delta A| > 0$. In essence, broad work distributions result in large statistical and systematic errors. Various estimates of these errors have been discussed [33–35]. As a rough correction, the bias can be estimated by repeatedly drawing the same number N of work values from a Gaussian (or another appropriate distribution) of the same width as the simulation work distribution, and subtracting the numerical estimate of the free energy obtained from (5.44) from the analytical value. That estimated bias can then be added to the estimator (5.44) for the simulation data [15]. Alternatively, the statistics with which rare, low W values appear can be used to obtain improved estimators [36, 37].

To reduce the width of the work distributions, the transformation can be staged by breaking it up into segments, as is commonly done in regular free energy perturbation theory (see Chap. 2). Such an approach has been used, for instance, to calculate the potential of mean force between two solutes in water [20] and similarly in [38]. However, dividing up the transformation into multiple segments requires re-equilibration at intermediate values of the coupling parameter.

5.7.2 Cumulant Estimators

Instead of estimating $-\beta^{-1} \ln\langle \exp(-\beta W(t)) \rangle$ directly using (5.44), one can use cumulant expansion approaches, as in regular free energy perturbation theory (see e.g., [20, 39] for combining cumulant expansions about the initial and final states). Unbiased estimators for cumulants can be used. Probably the most useful relations involve averages and variances of the work:

$$\Delta A \approx \langle W \rangle - \beta \, \sigma^2(W)/2 \tag{5.45}$$

$$\approx -\langle \underline{W} \rangle + \beta \, \sigma^2(\underline{W})/2. \tag{5.46}$$

Results from forward and backward averages can also be combined [20] into symmetric perturbation formulas:

$$\Delta A \approx (\langle W \rangle - \langle \underline{W} \rangle)/2 \tag{5.47}$$

$$\approx (\langle W \rangle - \langle \underline{W} \rangle)/2 - \frac{\beta}{12} \left(\sigma^2(W) - \sigma^2(\underline{W}) \right). \tag{5.48}$$

Note that in this 'optimal' estimator, the variances are not just averaged, but their difference is divided by 12 [20, 39]. In general, these formulas will be useful if the work distributions are nearly Gaussian and, correspondingly, the variances of forward and backward work are nearly identical, $\sigma^2(W) \approx \sigma^2(\underline{W})$. As with other perturbation expressions, the lower-order formula using only the two average work values will be more robust, but may result in comparatively larger systematic errors; higher-order formulas, in contrast, yield larger statistical errors. Nevertheless, higher cumulants have been used successfully in some free energy calculations (see, e.g., [39]).

5.7.3 Histogram Analysis

Potentially more-accurate free energies can be obtained by using histogram techniques. In particular, the Crooks relation, (5.35), allows us to use Bennett's method of overlapping histograms [21]. If the work distributions from forward and backward paths overlap, or can be extrapolated into the overlap region, we can use (5.35) to extract free energies. A detailed description of the analogous formalism for equilibrium perturbations can be found in Chap. 2. In brief, one collects two histograms $p_f(w = W)$ and $p_b(w = -\underline{W})$ of the work values, one each for forward and backward perturbations, with identical bin locations W_i and widths ΔW. The histograms should be properly normalized such that $\sum_i p_f(W_i)\Delta W = \sum_i p_b(W_i)\Delta W = 1$. The logarithm of the ratio of the histogram values, $\ln[p_f(w)/p_b(w)] = \beta(w - \Delta A)$ plotted against the work w should then give a straight line of slope β that intercepts the $w = 0$ line at a value of $-\beta\Delta A$. This result also provides us with an important test: if the slope deviates significantly from β, the ensembles of forward and backward paths are not mutually consistent, pointing either to sampling issues or other problems.

5.7.4 Bennett's Optimal 'Acceptance-Ratio' Estimator

Beyond histogram analysis, one can also use integrated forms of the Crooks relation, (5.35), to estimate free energies. Formally, we can rewrite (5.35) as an average

$$\int f(W; \Delta A)e^{-\beta W} p_f(W)dW = \int f(W; \Delta A)e^{-\beta \Delta A} p_b(W)dW \quad (5.49)$$

by multiplying both sides of (5.35) by an arbitrary function $f(W; \Delta A)$, bringing the denominator of the left-hand side to the right, and integrating over W. Equation (5.49) is an implicit equation for ΔA, and Bennett [21] determined the function $f = [e^{-\beta(W - \Delta A)}/N_f + 1/N_b]^{-1}$ that minimized the average squared error of the estimated free energy (with N_f and N_b the number of trajectories forward and backward, respectively). An identical result was recently obtained by using a maximum-likelihood approach [22]. The free energy then satisfies the implicit relation

$$\sum_{i=1}^{N_f} \frac{1}{1 + \dfrac{N_f}{N_b}e^{\beta(W_i - \Delta A)}} = \sum_{i=1}^{N_b} \frac{1}{1 + \dfrac{N_b}{N_f}e^{\beta(\underline{W}_i + \Delta A)}} \quad (5.50)$$

Note that on the right-hand side, the work $\underline{W}$ is that of the reversed path [i.e., has the opposite sign; see (5.33)]. Equation (5.50) has to be solved numerically (e.g., by using a Newton–Raphson solver) for the free-energy difference ΔA.

A word of caution should be made here: the assumption in both Bennett's original calculation [21] and in the maximum-likelihood formalism [22] is that all work values are statistically independent. This may not be the case, for instance, if consecutive initial configurations are close in time along an equilibrium trajectory. If results

from either the forward or reverse path are more strongly correlated than those from the opposite path direction, distortions of the free energy estimate seem likely. A typical case where this problem arises is particle insertion versus particle removal [32]. It is generally easier to obtain good statistics for particle insertion than for removal, because large numbers of particles can be inserted into existing configurations but only a given, finite number of particles can be removed. Therefore, the insertion results tend to be strongly correlated, but their large number could overwhelm the removal data. In such cases, care is necessary, and (5.50) may not be applicable.

5.7.5 Protocol for Free Energy Estimates from Nonequilibrium Work Averages

Following is a schematic protocol summarizing the main steps in estimating free energies from nonequilibrium work averages.

1. *Create an equilibrium ensemble of starting configurations.* To create N initial conformations representative of the equilibrium ensemble for Hamiltonian $\mathscr{H}(\mathbf{z}, \lambda = 0)$, one can, for instance, save conformations at regular intervals during a long equilibrium simulation. In some cases, accelerated sampling procedures may be necessary.
2. *Perform nonequilibrium simulations.* During the N simulations of duration τ, one for each of the N initial conformations, the Hamiltonian is changed from $\mathscr{H}(\mathbf{z}, \lambda(0) = 0)$ to $\mathscr{H}(\mathbf{z}, \lambda(\tau) = 1)$.
3. *Determine the work distribution.* After every time (or Monte Carlo) step during the simulations, change the Hamiltonian according to the prescribed path $\lambda(t)$ and accumulate the work performed as part of that change. Collect the work values from the N runs for analysis.
4. *Perform backward perturbation to improve accuracy.* If possible, repeat the procedure starting from $\mathscr{H}(\mathbf{z}, \lambda = 1)$.
5. *Analyze the work distributions.* Use the histogram analysis, optimal and cumulant estimators of Sect. 5.7 to determine the free energy difference.
6. *Check.* Use the Crooks relation (5.35) to check whether the forward and backward work distributions are consistent. Check for consistency of free energies obtained from different estimators. If the amount of dissipated work is large, caution may be necessary. If cumulant expressions are used, the work distributions should be nearly Gaussian, and the variances of the forward and backward perturbations should be of comparable size [as required by (5.35) for Gaussian work distributions]. Systematic errors from biased estimators should be taken into consideration. Statistical errors can be estimated, for instance, by performing a block analysis.

5.8 Illustrating Example

In the following, we will briefly illustrate the application of nonequilibrium free energy calculations for a simple 1D model system. Shown in Fig. 5.1 are the potential energy surfaces
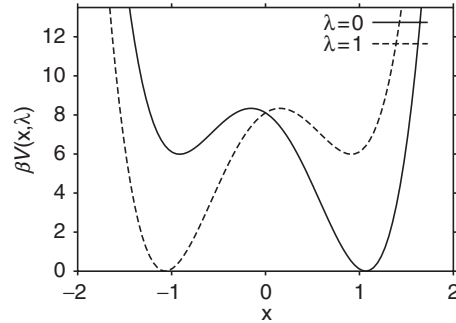
Fig. 5.1. Potential energy surface for $\lambda = 0$ (initial state; *solid line*) and $\lambda = 1$ (final state; *dashed line*)

$$\beta V(x, \lambda) = 5(x^2 - 1)^2 + 6(\lambda - 1/2)x \qquad (5.51)$$

for $\lambda = 0$ and 1. The free energy is defined as $\beta A(\lambda) = -\ln \int \exp[-\beta V(x, \lambda)]\mathrm{d}x$. Because of the symmetry with respect to λ [i.e., $V(x, \lambda) = V(-x, 1 - \lambda)$], the free-energy difference between the two states $\lambda = 0$ and $\lambda = 1$ vanishes, $\Delta A = A(1) - A(0) = 0$.

We have performed Brownian dynamics simulations (i.e., overdamped Langevin dynamics) on the potential surface $V(x, \lambda)$ with a time-dependent λ. In the simulations, the position x was updated as

$$x(t + \Delta t) = x(t) - D\beta \left.\frac{\partial V(x, t)}{\partial x}\right|_{x=x(t)} \Delta t + (2D\Delta t)^{1/2}g_t, \qquad (5.52)$$

where g_t are uncorrelated Gaussian random numbers of zero mean and unit variance. The work was accumulated as

$$W(t + \Delta t) = W(t) + V[x(t), t + \Delta t] - V[x(t), t] \qquad (5.53)$$

with $W(0) = 0$. The diffusion coefficient was chosen as $D = 1$, with a time step of $\Delta t = 0.001$.

The simulations were started from an equilibrium Boltzmann distribution on the free energy surface for $\lambda = 0$. During a time $t = 1$, λ was changed linearly in time from 0 to 1. We also performed simulations in the backward direction. However, because of the symmetry of V with respect to λ, backward transformations are equivalent to forward transformations. Along the resulting trajectories, the work βW was accumulated. Figure 5.2 shows the probability distributions of the work on the forward direction, and on the backward direction multiplied by $\exp(-\beta W)$. As expected from (5.35) for $\Delta A = 0$, the two distributions agree nicely.

The work distribution is bimodal, with a dominant ($>99\%$) peak at high work values of $\beta W \approx 6$, and a small second peak at low work values of $\beta W \approx -6$. The high-work peak is mainly due to trajectories that start in the low-energy well near
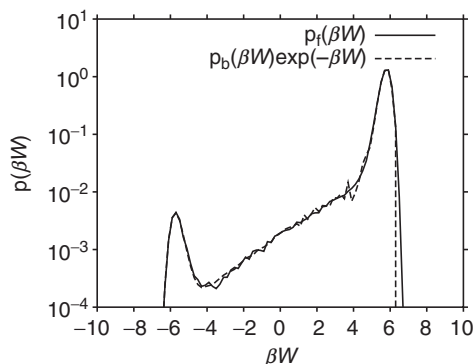
Fig. 5.2. Distribution of the work accumulated when transforming the potential surface from $\lambda = 0$ to 1 during a time $t = 1$. The *solid* and *dashed lines* show the distribution of the work along the forward and backward paths, respectively. The work distribution for backward paths was multiplied by $\exp(-\beta W)$. The distributions are plotted on a semi-logarithmic scale

$x = 1$ (Fig. 5.1). The low-work peak is due to trajectories that start already in the high-energy well at $x \approx -1$.

We used the simulation results to test various estimators for the free energy difference, as shown in Fig. 5.3. Sample sizes N ranged from 100 to 10,000. At least 1,000 independent samples of that size were used to determine the distribution of estimated free energies ΔA_{est}. The resulting histograms of ΔA_{est} are plotted in Fig. 5.3 for different sample sizes and different estimators. Ideally, the histograms should all be delta functions centered at $\Delta A_{\text{est}} = 0$. However, for $N = 100$ we find, for instance, a bimodal distribution of free energies estimated by applying the straightforward exponential average (5.44) and averaging results from forward and backward perturbations (Fig. 5.3a). The asymmetry of the distribution of estimated free energies reflects the strong bias of the exponential estimator, and its width reflects the large statistical error because few events can dominate the estimate. Only for sample sizes of 1,000 to 10,000 does the distribution of ΔA_{est} become centered around zero.

We also used the cumulant estimators of reference [20] in which the means (5.47) and variances (5.48) of forward and backward paths are symmetrically combined. As shown in Figs. 5.3b and c, the resulting distributions of estimated free energy differences are sharply centered around $\Delta A_{\text{est}} = 0$, thus giving excellent free energy estimates. However, this good agreement is somewhat surprising because the underlying work distributions are highly non-Gaussian, as shown in Fig. 5.2. Indeed, the good performance is due to cancellation of errors because of the symmetry of the problem (such that on average all cumulants exactly cancel each other); for 'asymmetric' perturbations from $\lambda = 0$ to 2, the cumulant estimators produce substantial systematic errors (Fig. 5.4).

Shown in Fig. 5.3d are free energies estimated from the same forward and backward simulation runs using Bennett's optimal estimator, obtained by solving (5.50) using a Newton–Raphson method. Unlike the direct exponential estimator (which
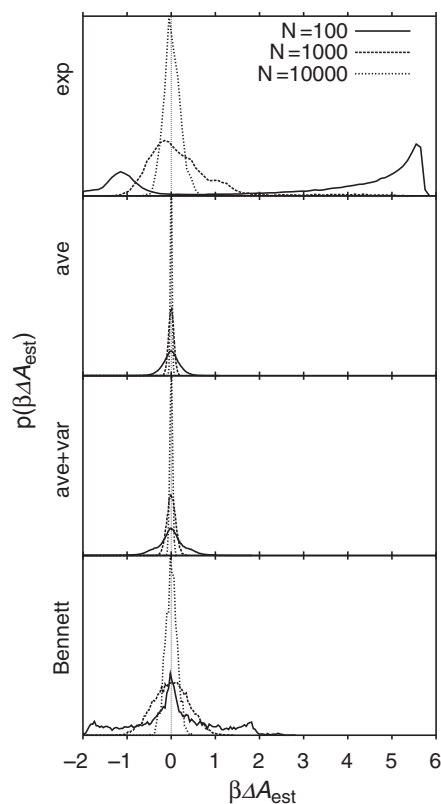
Fig. 5.3. Comparison of different free energy estimators. Plotted are distributions of estimated free energies using sample sizes (i.e., number of independent simulation runs) of $N = 100$ simulations (*solid lines*), as well as $N = 1,000$ (*long dashed*) and $N = 10,000$ simulations (*short dashed lines*). (**a**) Exponential estimator, (5.44). (**b**) Cumulant estimator using averages from forward and backward paths, (5.47). (**c**) Cumulant estimator using averages and variances from forward and backward paths, (5.48). (**d**) Bennett's optimal estimator, (5.50)

is strongly biased, as discussed in Chap. 6), we find the optimal estimator to be centered around $\Delta A_{\text{est}} = 0$. However, the distribution of the estimated free energies is considerably wider than that of the cumulant estimators, with symmetric wings at the smallest sample size of $N = 100$ (i.e., 50 forward and backward paths each).

In summary, using work collected from forward and backward paths greatly improves the accuracy of the estimates, and for the symmetric system studied here eliminates the bias. In our particular example, the cumulant estimators using forward and backward work data produce the most precise free energy estimates, followed by Bennett's optimal estimator. However, this somewhat poorer performance of the 'optimal' estimator is caused in part by the high degree of symmetry of the system studied.
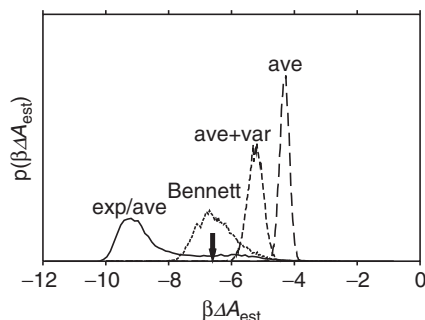
Fig. 5.4. Comparison of different free energy estimators for 'asymmetric' perturbation from $\lambda = 0$ to 2 within $t = 1$. Shown are distributions of free energies estimated using the direct exponential average, (5.44), averaged over forward and backward perturbations (*solid line*); averages (5.47) from forward and backward paths (*long dashed line*); averages and variances (5.48) from forward and backward paths (*short dashed line*); and Bennett's optimal estimator, (5.50), (*dotted line*). In all cases, free energies were estimated from $N = 1,000$ simulations. The vertical arrow indicates the actual free energy difference of $\beta\Delta A \approx -6.6$

5.9 Calculating Potentials of Mean Force

One of the objectives in free energy calculations is to obtain potentials of mean force $G(r)$ along a chosen coordinate $r = r(\mathbf{z})$ defined as

$$e^{-\beta G(r)+C} = \int d\mathbf{z}\, e^{-\beta \mathcal{H}_0(\mathbf{z})} \delta(r - r(\mathbf{z})), \qquad (5.54)$$

where C is an arbitrary constant. Usually, r is a dynamic, fluctuating variable that depends on the coordinates of the particles in the system. In computer simulations, we can normally use Jarzynski's identity directly by making r a non-fluctuating, externally controlled coupling parameter, i.e., $\lambda \equiv r$. However, in single-molecule pulling experiments and in their simulation equivalent, steered molecular dynamics [40–43], we may not be able to (or want to) control r directly. In those experiments (or simulations), the system is usually coupled to a moving harmonic spring that drives the coordinate r. An example would be the end-to-end extension of a molecule which is attached to a pulling spring (atomic force microscope, laser optical tweezer, etc.).

When calculating the potential of mean force along a fluctuating coordinate r, we can at best *observe* r (e.g., the instantaneous molecular extension), but we do not set it explicitly. Therefore, r is no longer an externally controlled coupling parameter, and Jarzynski's identity does not immediately apply. However, as was shown in [3], an extension produces the desired result.

For simplicity, let us consider a molecular system with a Hamiltonian $\mathcal{H}_0(\mathbf{z})$ that is coupled to a harmonic spring with spring constant k and a time-dependent minimum $r(t)$. The explicitly time-dependent Hamiltonian of the complete system is then

$$\mathscr{H}(\mathbf{z}, t) = \mathscr{H}_0(\mathbf{z}) + \frac{k}{2}\left(r(\mathbf{z}) - r(t)\right)^2, \tag{5.55}$$

where $r(\mathbf{z})$ is the molecular extension and $V(r, t) = k(r(\mathbf{z}) - r(t))^2/2$ is the potential energy of the spring.

We can now use the Feynman–Kac relation to obtain the potential of mean force. By multiplying both sides of (5.14) by $\exp[\beta k(r(\mathbf{z}) - r(t))^2/2]\delta(r - r(\mathbf{z}))$ and integrating over $\mathbf{z}$, we obtain an expression for the potential of mean force

$$\boxed{\mathrm{e}^{-\beta G(r)} = \left\langle \delta\left[r - r\left(\mathbf{z}(t)\right)\right] \mathrm{e}^{-\beta\left(\int_0^t \frac{\partial V\{r[\mathbf{z}(\tau)], \tau\}}{\partial \tau} \mathrm{d}\tau - V\{r[\mathbf{z}(t)], t\}\right)} \right\rangle}$$

$$\tag{5.56}$$

Formally, by weighting the trajectory end points $r(\mathbf{z}(t))$ with the Boltzmann factor of the accumulated work (minus the energy stored in the pulling spring), we can produce an unbiased histogram of r proportional to $\exp(-\beta G(r))$. Practically, r values will cluster near $r(t)$ because of the coupling to the spring, and the histogram will have good statistics only there. However, histograms obtained at different times t can be combined to obtain results over a broad range of r. This approach can be thought of as 'dynamic umbrella sampling.' By adapting the histogram-reweighting procedure of Ferrenberg and Swendsen [44] (see Chap. 3) for the nonequilibrium averages [3, 15], we obtain the following expression:

$$\mathrm{e}^{-\beta G(r)} = \frac{\sum_t \dfrac{\langle \delta[r - r(t)]\exp(-\beta W(t))\rangle}{\langle \exp(-\beta W(t))\rangle}}{\sum_t \dfrac{\exp[-\beta V(r, t)]}{\langle \exp(-\beta W(t))\rangle}}. \tag{5.57}$$

Here, the sums are over different time slices t at which coordinates r and corresponding work values have been saved.

5.9.1 Approximate Relations for Potentials of Mean Force

In some cases, one may not be able to collect enough trajectories for a histogram analysis. Instead, one can use moments [15]. In particular, the derivative of the free energy is approximately given by the average force (weighted by the Boltzmann factor of the accumulated work)

$$G_0'(\overline{r(t)}) \approx -\left\langle \frac{\partial V(r, t)}{\partial r} \right\rangle_{r = \overline{r(t)}}. \tag{5.58}$$

The average coordinate r is defined as:

$$\overline{r(t)} = \frac{\left\langle r(t)\mathrm{e}^{-\beta W(t)} \right\rangle}{\left\langle \mathrm{e}^{-\beta W(t)} \right\rangle}. \tag{5.59}$$

This approximation is valid if $r(t)$ is approximately Gaussian distributed.

Above, we obtained the potential of mean force by collecting work and r values. Formally, we can also use the free energy $\Delta A(t)$ measured at all times t to extract the underlying potential of mean force. If we multiply (5.56) by $\exp[-\beta k(r(\mathbf{z}) - r(t))^2/2]$ and integrate over r, we obtain

$$\mathrm{e}^{-\beta \Delta A(t)} = \int_{-\infty}^{\infty} \mathrm{d}r \, \mathrm{e}^{-\beta G(r)} \mathrm{e}^{-\beta k(r - r(t))^2/2}. \qquad (5.60)$$

Effectively, this constitutes a Fredholm integral equation of the first kind for $\exp[-\beta G(r)]$ where we 'know' the left-hand side, $\exp(-\beta \Delta A(t)) = \langle \exp(-\beta W(t)) \rangle$, from (5.8) and the kernel is given by the Boltzmann factor of the pulling potential. In principle, we could solve the integral equation, for instance, by projection onto orthogonal functions to turn it into a linear equation. However, such approaches are notoriously unstable, and thus not too promising in practice.

For very stiff pulling springs (where r is *almost* a coupling parameter under external control), we can instead pursue the so-called 'stiff-spring approximation' of Park et al. [45]. A Fourier representation of the spring Boltzmann factor on the right-hand side of (5.60) results in

$$\mathrm{e}^{-\beta \Delta A(t)} = \int \mathrm{d}s \int \mathrm{d}r \frac{\mathrm{e}^{-\beta G(r) - s^2/(2\beta k) - is(r - r(t))}}{(2\pi \beta k)^{1/2}}. \qquad (5.61)$$

To lowest order in $1/k$ (with $k \to \infty$), we ignore the s^2 in the exponent and the free-energy offset from the k-dependent numerator, and obtain simply

$$\Delta A(t) \approx G(r = r(t)) . \qquad (5.62)$$

This is the result expected if r is effectively a control parameter. Taylor expansion with respect to $1/k$ introduces the correction terms as derivatives. To first order in $1/k$, we obtain

$$\Delta A(t) \approx G(r = r(t)) + \frac{1}{2k} \left(\frac{G''(r)}{\beta} - G'(r)^2 \right)_{r = r(t)}. \qquad (5.63)$$

To order $1/k$, this relation can be inverted, giving the 'stiff-spring approximation' derived originally in [45]

$$\Delta G(r = r(t)) \approx A(t) - \frac{1}{2kv^2} \left(\frac{\ddot{A}(t)}{\beta} - \dot{A}(t)^2 \right), \qquad (5.64)$$

where we assumed for simplicity that $r(t) = vt$ with v a constant pulling velocity. (Otherwise, there will be correction terms including $\ddot{z}(t)$.)

One of the major advances in the application of Jarzynski's identity to the calculation of free energies came from coupling it to path sampling [46, 47]. In a typical application with fast switching, the system is rapidly driven out of equilibrium as the coupling parameter is changed, and nearly all trajectories are essentially

irrelevant for the free energy estimate (i.e., the Boltzmann factor of the trajectory endpoint is small compared to typical values in an equilibrium ensemble). Effectively, those trajectories are wasted. In the biased path sampling of Sun [46], new trajectories are created using a Monte Carlo sampling approach by 'perturbing' existing trajectories. To bias the sampling toward important but poorly sampled small work values, an appropriate bias is added to the acceptance criterion. However, despite this appealing feature, a recent efficiency analysis suggests that this biased search in the optimal case is less efficient than an optimally biased Zwanzig-type free energy perturbation [13]. A thorough discussion of path sampling can be found in Chap. 7.

5.10 Applications

After some initial studies for model systems [20, 38, 48–51], nonequilibrium work averages have so far been applied most extensively in the context of steered molecular dynamics simulations. In steered molecular dynamics [40–43], the molecular system is coupled to a moving harmonic spring, usually through a group of atoms. This approach is inspired by single-molecule pulling experiments using atomic force microscopes or laser optical tweezers. As the spring is moved, tension builds up in the molecular system, eventually forcing a molecular transition. Examples include the unfolding of proteins [42, 52], dissociation of ligands [40, 41], or dragging molecules through a channel [53, 54]. Usually, one is interested in the potential of mean force along the pulling direction, which can be obtained using the dynamic umbrella sampling approach of Sect. 5.9 and [3] or approximations to it.

5.11 Summary

Jarzynski's identity [1] provides a new and potentially powerful tool for the determination of free energy differences in computer simulations. As an immediate application, it puts slow-growth free energy calculations [8, 9] on a fully quantitative footing. More importantly, with the help of (5.8), free energy differences can be calculated at arbitrary rates of transformation. Probably the most important applications to date are in the analysis of single-molecule pulling (and related) experiments where nonequilibrium conditions cannot be easily avoided [3–6]. Extensions of Jarzynski-type relations between work distributions and free energy difference to quantum dynamics have been derived [55–57].

Computational efficiency remains a central question. For a given amount of computer time, how 'good' are nonequilibrium estimates of free energy differences compared to estimates from equilibrium methods? Overall, evidence is mounting that nonequilibrium methods are less efficient than equilibrium methods [13, 20, 38]. However, new approaches have been suggested that use long time steps [25]. For relatively slow transformations, it has been shown [20] that for a given amount of computer simulation time, one obtains more-accurate results for few slow transformations than for many fast transformations. At the other extreme, i.e., in the limit of

near-optimally work-biased fast transformations [13], instantaneous (Zwanzig-type [12]) perturbations appear to be optimal. However, repeated calculations appear to offer some advantages. For instance, if multiple distinct pathways over different 'saddles' connect the two states, and equilibration between those paths is slow, then one could imagine that an umbrella sampling approach might get stuck in one 'transition path' valley (in particular, if runs in a new umbrella window are initiated from a configuration of the preceding window). By using multiple independent runs, nonequilibrium methods seem more likely to explore all possible pathways, and that could well lead to more accurate free energies when compared to naive applications of standard approaches. However, such vague arguments deserve a more quantitative analysis. Moreover, nonequilibrium methods based on Jarzynski's identity are still a very recent addition to the arsenal for free energy calculations, and one may expect considerable improvements in their implementation as well as new developments [23, 25, 58–65].

The identities of Jarzynski and Crooks, and related expressions, allow us to determine free energy differences rigorously from nonequilibrium simulations. In those simulations, the Hamiltonian is changed continuously (or step wise) from the initial ($\lambda = 0$) to the final energy surface ($\lambda = 1$). Free energies are obtained from the distribution of work accumulated during repeated transformations. The resulting formalism connects free energy perturbation theory with thermodynamic integration (as infinitely fast and slow perturbations, respectively). Moreover, it puts slow-growth methods on a sound theoretical basis. In practical applications, choosing reliable estimators and sampling procedures are critical for success. The Crooks identity, (5.35), provides a powerful test of sampling and implementation by probing the consistency of work distributions obtained on forward ($0 \rightarrow 1$) and backward ($1 \rightarrow 0$) perturbations. The usefulness of nonequilibrium work methods is most evident in cases where nonequilibrium cannot be avoided easily, for instance in pulling experiments or in simulations where the transformation has to be completed within a given amount of time. In other cases, there is evidence that near-equilibrium methods produce more-accurate results for a given amount of computer time. However, nonequilibrium methods are an area of much ongoing investigation, and we can expect a steady stream of new insights relevant both fundamentally and practically.

References

1. Jarzynski, C., Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **1997**, *78*, 2690–2693
2. Jarzynski, C., Equilibrium free energy differences from nonequilibrium measurements. A master-equation approach, *Phys. Rev. E* **1997**, *56*, 5018–5035
3. Hummer, G.; Szabo, A., Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3658–3661
4. Liphardt, J.; Dumont, S.; Smith, S. B.; Tinoco, I.; Bustamante, C., Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, *Science* **2002**, *296*, 1832–1835

5. Noy, A., Direct determination of the equilibrium unbinding potential profile for a short DNA duplex from force spectroscopy data, *Appl. Phys. Lett.* **2004**, *85*, 4792–4794

6. Trepagnier, E. H.; Jarzynski, C.; Ritort, F.; Crooks, G. E.; Bustamante, C. J.; Liphardt, J., Experimental test of Hatano and Sasa's nonequilibrium steady-state equality, *Proc. Natl Acad. Sci. USA* **2004**, *101*, 15038–15041

7. Born, M., Volumen und Hydratationswärme der Ionen, *Z. Phys.* **1920**, *1*, 45–48

8. Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R., Thermodynamics of cavity formation in water. A molecular dynamics study, *Faraday Symp. Chem. Soc.* **1982**, *17*, 55

9. Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M., Free energy of hydrophobic hydration. A molecular dynamics study of noble gases in water, *J. Chem. Phys.* **1986**, *85*, 6720–6727

10. Wood, R. H.; Mühlbauer, W. C. F.; Thompson, P. T., Systematic errors in free energy perturbation calculations due to a finite sample of configuration space. Sample-size hysteresis, *J. Phys. Chem.* **1991**, *95*, 6670–6675

11. Hermans, J., Simple analysis of noise and hysteresis in (slow-growth) free energy simulations, *J. Phys. Chem.* **1991**, *95*, 9029–9032

12. Zwanzig, R. W., High-temperature equation of state by a perturbation method. I. Non-polar gases, *J. Chem. Phys.* **1954**, *22*, 1420–1426

13. Oberhofer, H.; Dellago, C.; Geissler, P. L., Biased sampling of nonequilibrium trajectories. Can fast switching simulations outperform conventional free energy calculation methods, *J. Phys. Chem. B* **2005**, *109*, 6902–6915

14. Roepstorff, G., *Path Integral Approach to Quantum Physics*, Springer: Berlin, Heidelberg, New York, 1994

15. Hummer, G.; Szabo, A., Free energy surfaces from single-molecule force spectroscopy, *Acc. Chem. Res.* **2005**, *38*, 504–513

16. Crooks, G. E., Path-ensemble averages in systems driven far from equilibrium, *Phys. Rev. E* **2000**, *61*, 2361–2366

17. Hatano, T.; Sasa, S., Steady-state thermodynamics of Langevin systems, *Phys. Rev. Lett.* **2001**, *86*, 3463–3466

18. Crooks, G. E., Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, *J. Stat. Phys.* **1998**, *90*, 1481–1487

19. Crooks, G. E., Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, *Phys. Rev. E* **1999**, *60*, 2721–2726

20. Hummer, G., Fast-growth thermodynamic integration error and efficiency analysis, *J. Chem. Phys.* **2001**, *114*, 7330–7337

21. Bennett, C. H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **1976**, *22*, 245–268

22. Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S., Equilibrium free energies from non-equilibrium measurements using maximum-likelihood methods, *Phys. Rev. Lett.* **2003**, *91*, 140601

23. Shirts, M. R.; Pande, V. S., Comparison of efficiency and bias of free energies computed by exponential averaging. The Bennett acceptance ratio and thermodynamic integration, *J. Chem. Phys.* **2005**, *122*, 144107

24. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon: Oxford, UK, 1987

25. Lechner, W.; Oberhofer, H.; Dellago, C.; Geissler, P. L., Equilibrium free energies from fast-switching trajectories with large time steps, *J. Chem. Phys.* **2006**, *124*, 044113

26. Reinhardt, W. P.; Hunter III, J. E., Variational path optimization and upper and lower bounds for the free energy via finite time minimization of the external work, *J. Chem. Phys.* **1992**, *97*, 1599–1601

27. Miller, M. A.; Reinhardt, W. P., Efficient free energy calculations by variationally optimized metric scaling concepts and applications to the volume dependence of cluster free energies and to solid–solid phase transitions, *J. Chem. Phys.* **2000**, *113*, 7035–7046

28. Best, R. B.; Hummer, G., Reaction coordinates and rates from transition paths, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6732–6737

29. Chandler, D., Statistical mechanics of isomerization dynamics in liquids and the transition state approximation, *J. Chem. Phys.* **1978**, *68*, 2959–2970

30. Hummer, G., From transition paths to transition states and rate coefficients, *J. Chem. Phys.* **2004**, *120*, 516–523

31. Jarzynski, C., Rare events and the convergence of exponentially averaged work values, *Phys. Rev. E* **2006**, *73*, 046105

32. Frenkel, D., Free-energy computation and first-order phase transitions. In *Molecular Dynamics Simulations of Statistical Mechanical Systems. Proceedings of the Enrico Fermi Summer School, Varenna, 1985* (Amsterdam, 1986), Ciccotti, G.; Hoover, W. G., Eds., North-Holland, pp. 151–188

33. Gore, J.; Ritort, F.; Bustamante, C., Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 12564–12569

34. Zuckerman, D. M.; Woolf, T. B., Theory of a systematic computational error in free energy differences, *Phys. Rev. Lett.* **2002**, *89*, 180602

35. Wu, D.; Kofke, D. A., Asymmetric bias in free-energy perturbation measurements using two Hamiltonian-based models, *Phys. Rev. E* **2004**, *70*, 066702

36. Zuckerman, D. M.; Woolf, T. B., Overcoming finite-sampling errors in fast-switching free-energy estimates. Extrapolative analysis of a molecular system, *Chem. Phys. Lett.* **2002**, *351*, 445–453

37. Ytreberg, F. M.; Zuckerman, D. M., Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems, *J. Comp. Chem.* **2004**, *25*, 1749–1759

38. Rodriguez-Gomez, D.; Darve, E.; Pohorille, A., Assessing the efficiency of free energy calculation methods, *J. Chem. Phys.* **2004**, *120*, 3563–3578

39. Hummer, G.; Szabo, A., Calculation of free energy differences from computer simulations of initial and final states, *J. Chem. Phys.* **1996**, *105*, 2004–2010

40. Grubmüller, H.; Heymann, B.; Tavan, P., Ligand binding molecular mechanics calculation of the streptavidin biotin rupture force, *Science* **1996**, *271*, 997–999

41. Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K., Molecular dynamics study of unbinding of the avidin-biotin complex, *Biophys. J.* **1997**, *72*, 1568–1581

42. Paci, E.; Karplus, M., Forced unfolding of fibronectin type 3 modules. An analysis by biased molecular dynamics simulations, *J. Mol. Biol.* **1999**, *288*, 441–459

43. Park, S.; Schulten, K., Calculating potentials of mean force from steered molecular dynamics simulations, *J. Chem. Phys.* **2004**, *120*, 5946–5961

44. Ferrenberg, A. M.; Swendsen, R. H., Optimized Monte Carlo data analysis, *Phys. Rev. Lett.* **1989**, *63*, 1195–1198

45. Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K., Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality, *J. Chem. Phys.* **2003**, *119*, 3559–3566

46. Sun, S. X., Equilibrium free energies from path sampling of nonequilibrium trajectories, *J. Chem. Phys.* **2003**, *118*, 5769–5775

47. Ytreberg, F. M.; Zuckerman, D. M., Single-ensemble nonequilibrium path-sampling estimates of free energy differences, *J. Chem. Phys.* **2004**, *120*, 10876–10879

48. Hendrix, D. A.; Jarzynski, C., A fast growth method of computing free energy differences, *J. Chem. Phys.* **2001**, *114*, 5974–5981

49. Hummer, G., Fast-growth thermodynamic integration results for sodium ion hydration, *Mol. Simul.* **2002**, *28*, 81–90

50. Hu, H.; Yun, R. H.; Hermans, J., Reversibility of free energy simulations slow growth may have a unique advantage with a note on use of Ewald summation, *Mol. Simul.* **2002**, *28*, 67–80

51. Darve, E.; Pohorille, A., Calculating free energies using average force, *J. Chem. Phys.* **2001**, *115*, 9169–9183

52. Marszalek, P. E.; Lu, H.; Li, H. B.; Carrion-Vazquez, M.; Oberhauser, A. F.; Schulten, K.; Fernandez, J. M., Mechanical unfolding intermediates in titin modules, *Nature* **1999**, *402*, 100–103

53. Jensen, M. O.; Park, S.; Tajkhorshid, E.; Schulten, K., Energetics of glycerol conduction through aquaglyceroporin glpf, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 6731–6736

54. Amaro, R.; Luthey-Schulten, Z., Molecular dynamics simulations of substrate channeling through an alpha-beta barrel protein, *Chem. Phys.* **2004**, *307*, 147–155

55. Mukamel, S., Quantum extension of the Jarzynski relation analogy with stochastic dephasing, *Phys. Rev. Lett.* **2003**, *90*, 170604

56. Jarzynski, C.; Wojcik, D. K., Classical and quantum fluctuation theorems for heat exchange, *Phys. Rev. Lett.* **2004**, *92*, 230602

57. De Roeck, W.; Maes, C., Quantum version of free-energy-irreversible-work relations, *Phys. Rev. E* **2004**, *69*, 026115

58. Atilgan, E.; Sun, S. X., Equilibrium free energy estimates based on nonequilibrium work relations and extended dynamics, *J. Chem. Phys.* **2004**, *121*, 10392–10400

59. Ytreberg, F. M.; Zuckerman, D. M., Peptide conformational equilibria computed via a single-stage shifting protocol, *J. Phys. Chem. B* **2005**, *109*, 9096–9103

60. Chernyak, V.; Chertkov, M.; Jarzynski, C., Dynamical generalization of nonequilibrium work relation, *Phys. Rev. E* **2005**, *71*, 025102

61. Rodinger, T.; Pomès, R., Enhancing the accuracy the efficiency and the scope of free energy simulations, *Curr. Opin. Struct. Biol.* **2005**, *15*, 164–170

62. De Koning, M., Optimizing the driving function for nonequilibrium free-energy calculations in the linear regime. A variational approach, *J. Chem. Phys.* **2005**, *122*, 104106

63. Lua, R. C.; Grosberg, A. Y., Practical applicability of the Jarzynski relation in statistical mechanics. A pedagogical example, *J. Phys. Chem. B* **2005**, *109*, 6805–6811

64. Adib, A. B., Entropy and density of states from isoenergetic nonequilibrium processes, *Phys. Rev. E* **2005**, *71*, 056128

65. Collin, D.; Ritort, F.; Jarzynski, C.; Smith, S.B.; Tinoco, I.; Bustamante, C. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **2005**, *437*, 231–234

6

Understanding and Improving Free Energy Calculations in Molecular Simulations: Error Analysis and Reduction Methods

Nandou Lu and Thomas B. Woolf

6.1 Introduction

Free energy simulations have been widely applied to the computation of physical, chemical, and biological systems. Yet the reliability and efficiency of the approach has not been considered in depth. This is despite the many free energy calculation methods that have been developed over the decades. Thus, many fundamental questions remain unanswered, e.g., what method is best for evaluating the free energy? Is the free energy estimate reliable and what is the error in it? How can one assess the quality of the free energy result when the true answer is unknown? What is the minimal computational effort needed to get a decent result? How can one improve the efficiency and reliability of free energy calculations?

There are many practical issues related to obtaining converged free energy results without numerical errors. Since the calculation is computationally intensive and the goal is a useful result in a reasonable time, it remains a significant challenge for molecular simulations of large and complex systems [1]. This is because, despite the wide application of free energy simulations, the fundamental sampling and convergence behavior of free energy methods is less well understood. Fundamental understanding of a free energy method's behavior, such as its limitations and its advantages, its calculation efficiency and its reliability, is important not only for simulation practice but also for the development of novel methodologies. The key towards increasing understanding is to characterize the free energy error of a method.

In this chapter, we will examine in depth the characteristic errors of two free energy techniques and present improved methods based on a better understanding of their behavior. The two techniques examined are free energy perturbation (FEP) [2] and nonequilibrium work (NEW) based on Jarzynski's equality [3–6]. These techniques are discussed in Chaps. 2 and 5. The FEP method is one of the most popular approaches for computing free energy differences in molecular simulation; see, e.g., [1, 7–10]. The recently developed NEW method, which is closely related to FEP, is gaining popularity in both simulation [11–18] and experimental applications [19–21].

6.1.1 Sources of Free Energy Error

The error in a free energy calculation is due to the many numerical approximations in a computer simulation. These approximations can be understood as arising from two major sources [22–24]: the representation of the system and finite sampling. Examples of errors that may arise from the system representation include the choice of the empirical force field, the choice of interaction cutoff distances, the treatment of electrostatic effects are treated, the choice of system size, and the choice of model representation (for example coarse graining or the use of an implicit solvent representation). These system choices can lead to a systematic error, creating a computer model with an unrealistic representation of the system of interest that will not converge to the correct free energy. This type of systematic error will not be reduced by adopting improved free energy calculation algorithms or by increasing the simulation time.

Imperfect sampling contributes the other major source of error found in free energy calculations. Although statistical mechanics provides a formally exact equation for computing free energy – with knowledge of the entire phase space of a system – in practice, a computer simulation is finite in time, and thus free energy has to be computed using a finite set of configurations. Therefore, finite sampling errors are an intrinsic part of free energy simulation and can be an even more critical part of the analysis due to rough energy landscapes for many biophysically interesting systems. But, unlike the systematic error that may occur due to the choice of representation, finite sampling error generally decreases as more sampling is conducted and better free energy and/or sampling algorithms are used.

In a free energy simulation the measured quantity is the relative free energy between two systems. As has already been discussed in Chaps. 2 and 5, both FEP and NEW start from an equilibrated reference system, and perturb to a target system. The difference between the potential energy changes, ΔU, for FEP or work values, W, for NEW of these two systems are measured and collected on the computer to estimate the free energy difference. Finite sampling errors of two types are involved in the calculation: (1) sampling within the phase space of the reference system, and (2) sampling the phase space along the transitions from the reference system towards the target system. The second type of error can be described in terms of the sampling of perturbations (potential energy change ΔU or NEW W) during the simulation [14, 17, 24–28]. These two types of sampling error are generally considered separately. Even when the conformations of the reference system are perfectly sampled (zero sampling error of the first type), free energy calculations may still have statistical error due to incomplete sampling of ΔU or W. Unfortunately, for finite-length FEP and NEW calculations, the important parts of $\{\Delta U\}$ and $\{W\}$ are rarely sampled, and simulation error due to poor sampling can be significant – see later in this chapter for more details. Note that this type of error can be reduced by increasing the sample size, or (often more effectively) by adopting better free energy methods. Understanding and characterizing the finite sampling error between the two endpoints of the free energy change are the focus of this chapter. That is, we do not consider systematic errors due to a possibly inappropriate model of a system, and

we further assume that incomplete sampling of the reference system is not a major source of error in the calculation.

6.1.2 Accuracy and Precision: Bias and Variance Decomposition

How should we best describe the error of a measurement result (an estimate of the true value)?

Suppose N independently repeated samples $\{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N\}$ are obtained to measure the value of a quantity X, where $\hat{x}$ represents an estimate of its true value x. Usually the best estimate of x is provided by the mean or average of $\{\hat{x}\}$, $\hat{\bar{x}}(N)$. The reliability of the estimator $\hat{\bar{x}}$ can be characterized by the *mean square error*, (e.g., [29, 30])

$$m^2(N) = \frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x)^2. \tag{6.1}$$

Following a straightforward derivation, we have (assuming the absence of sampling noise)

$$m^2(N) = \sigma^2(N) + \delta^2(N), \tag{6.2}$$

where the mean square error is decomposed into two terms: the sample *variance*, σ^2, and the square of the finite sampling *bias*, δ^2. The variance represents the statistical part of the mean square error and is defined solely by the sampled values

$$\sigma^2(N) = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{x}_i - \hat{\bar{x}}(N)\right)^2. \tag{6.3}$$

Note that the variance does not depend on the true value x, and the mean estimator $\hat{\bar{x}}$ has the least variance. The finite sampling bias is the difference between the estimate $\hat{\bar{x}}$ and the true value x, and represents the finite sampling systematic part of the generalized error

$$\delta(N) = \hat{\bar{x}}(N) - x. \tag{6.4}$$

In this chapter we use the terms *precision* and *accuracy* in relation to the finite sampling variance and bias, respectively. Also, we describe the overall quality of an estimator – the mean square error – by the term *reliability*. Note the difference between our terminology and that in some statistics literature where accuracy is used to describe the overall quality (i.e., the reliability in this chapter). The decomposition of the error into the variance and bias allows us to use different approaches for studying the behavior of each term.

6.1.3 Dominant Errors

In the equations above, the mean square error, the sample variance, and the finite sampling bias are all explicitly written as functions of the sample size N. Both the variance and bias diminish as $N \rightarrow \infty$ (infinite sampling). However, the variance

and bias generally have different dependencies on the sample size N, and the dominant part of the mean square error may change as N changes. This is especially true for free energy calculations [17, 24, 25, 31, 32]. By definition, the variance decays as $N^{-1/2}$. For both FEP and NEW calculations, the decay of the bias is proportional to $N^{-\gamma}$ with γ ranging from 0 (small-N limit) to 1 (large-N limit) [14, 25]. As will be discussed in more detail later in this chapter, the samples (of ΔU or W) important to the accuracy of the FEP or NEW calculation correspond to the rare events in a computer simulation. Missing such important samples will cause significant bias in the free energy estimate obtained in practical simulations, in which the sample size has to be finite. In fact, the FEP and NEW calculation formulae are biased in any finite-length simulations [3, 11, 12, 17, 25, 27, 32]. In short, in the finite-N region, the bias caused by the missing of important but rare samples is usually more significant than the variance and thus is the dominant error in the free energy estimate. As the sample size N approaches infinity ($N \rightarrow \infty$), the sample variance may surpass the finite sampling bias and become the dominant source of error [17, 25].

The sample size in a real simulation is always finite, and usually relatively small. Thus, understanding the error behavior in the finite-size sampling region is critical for free energy calculations based on molecular simulation. Despite the importance of finite sampling bias, it has received little attention from the community of molecular simulators. Consequently, we would like to emphasize the importance of finite sampling bias (accuracy) in this chapter.

The usual nature of statistics is to consider only the mean and first moments of the distributions. This cannot fully capture the complexity of error analysis in free energy calculations, where poor sampling of important but rare events is the major source of error. We further note that in the finite sampling region (and/or the region where the nonequilibrium paths are far from reversibility), techniques such as series expansion [12–15, 33, 34] and bootstrap analysis may become less effective for the analysis of free energy errors. This is because of the nature of the finite-sampled region – too much data may be missing to enable a resampling (bootstrap) or an expansion in powers to fill in the missing information [35]. At the same time, for situations with sufficient data, bootstrap or other extrapolation methods can help to gain insight into the errors. In this chapter, different approaches will be presented for examining the systematic (bias) and statistical (variance) errors for both the small- and large-N regions. In addition, principles and novel methods for improving free energy calculations are presented.

6.1.4 Outline

This chapter is organized as follows: a brief overview of FEP and NEW methods, together with important concepts used in later analysis, is presented in Sect. 6.2. In Sect. 6.3 we discuss how the considerations of important phase space and probability distribution functions can be used to understand and model free energy error. In Sect. 6.4 we proceed with the analysis of both types of sampling error. The principles of optimal multistage FEP (MFEP) calculations are laid out in Sect. 6.5. Methods for better free energy calculation, such as overlap sampling (OS) and extrapolation are

given in Sects. 6.6 and 6.7, followed by concluding remarks and perspectives for future development in Sect. 6.8.

6.2 Overview of the FEP and NEW Methods

6.2.1 Free Energy Perturbation

FEP is one of the most popular methods for calculating free energy differences in molecular simulation and is discussed in detail in Chap. 2. Using the FEP formula, the free energy difference between two systems 0 and 1 is computed according to

$$\exp(-\beta \Delta A) = \langle \exp(-\beta \Delta \mathcal{H}) \rangle_0 , \tag{6.5}$$

where $\Delta A = A_1 - A_0$ is the Helmholtz free energy difference, $\Delta \mathcal{H} = \mathcal{H}_1 - \mathcal{H}_0$ is the Hamiltonian difference (the perturbation); the angle brackets represent a canonical ensemble average performed on an equilibrated system designated by the subscript. Usually the kinetic component of the Hamiltonian is not included in the free energy calculation, and $\Delta \mathcal{H}_{01}$ can be replaced by the potential energy difference $\Delta U = U_1 - U_0$.

Forward and Reverse FEP

In an FEP calculation, one of the two systems, e.g., the 0 system (6.5), governs the simulation (usually Monte Carlo or molecular dynamics); this system is called the reference while the other (system 1) is the target. At different configurations of the equilibrated reference system, repeated perturbation trials are performed by switching the Hamiltonian of the system to that of the target from the reference, and perturbations $\{\Delta U\}$ are collected to compute the ensemble average. Either system 0 or 1 can be chosen as the reference, thus for any given system pair, the FEP can be conducted in two opposite perturbation directions (from reference to target): $0 \rightarrow 1$ or $1 \rightarrow 0$ [26, 32, 36, 37]. These two directions are usually (arbitrarily) referred to as the *forward* and *reverse* (or *backward*) directions, respectively. They have the same working formula (6.5), except for the exchange of the subscripts "0" and "1." However, in this chapter, we always define the difference of a quantity between the two systems as $\Delta X = X_1 - X_0$, regardless of which system is taken as the reference; for example, the free energy and potential energy changes are defined as $\Delta A = A_1 - A_0$ and $\Delta U = U_1 - U_0$, respectively. With this definition, the working equations of the forward and reverse FEP calculations can be written as

$$\boxed{\exp(-\beta \Delta A) = \langle \exp(-\beta \Delta U) \rangle_0 ,} \tag{6.6}$$

$$\boxed{\exp(+\beta \Delta A) = \langle \exp(+\beta \Delta U) \rangle_1 ,} \tag{6.7}$$

respectively. Note that with these definitions the "+" sign in (6.7) is necessary due to the definition of 'difference.' This use allows us to consider ΔU and W throughout

without changing sign, at the expense of having formulae that differ in sign depending on the direction of the work events or perturbations.

In principle, the forward and reverse calculations should produce identical free energy estimates. However in real simulations these estimates usually differ. Furthermore, as we explore finite sampling errors, we will see that the reliability (the error in each direction) of these estimates also differs [24, 26, 38].

As the difference between the 0 and 1 systems becomes larger, the free energies obtained from (6.6) or (6.7) show slower convergence and poorer reliability. To handle large perturbations (as is the case for many real problems), stratified (multistage) FEP approaches [22, 31, 39–42] are usually adopted instead of methods directly based on (6.6) or (6.7), which are, correspondingly, referred to as single-stage FEP methods.

Multistage FEP Calculations

An MFEP calculation, sometimes also referred to as a multi-window calculation, has already been introduced in Chap. 2. The idea behind this method is to split a large perturbation into a series of smaller ones which are suitable for handling by the single-stage FEP calculations of (6.6) or (6.7). This is reasonable, since free energy is a state variable, and the overall ΔA is given by the sum of the changes for each substage. To split the perturbation, one or more intermediates (designated by M) are constructed besides the initial and final systems 0 and 1. The introduction of intermediates provides additional degrees of freedom for conducting FEP calculations, such as the definition of the intermediates, the number of intermediates (stages), the simulation lengths for each stage. Also, for each successive state pair, the FEP can be conducted in two directions. As a result, an MFEP calculation can usually be conducted in many different ways [24, 26, 31, 32]. Even for the simple case of a two-stage FEP calculation, there are four possible different implementations, as listed in Table 6.1.

Fine-tuning the degrees of freedom in an MFEP calculation can improve the efficiency and reliability of the free energy calculation. However, the optimization of an MFEP calculation is usually not a straightforward task [41]. Because single-stage FEP calculations are the basic components of a multistage calculation, the principles of optimal multistaging design can be obtained by studying single-stage FEP behavior. Thus, first understanding the optimal manner of performing a single-stage

Table 6.1. Example of four staging strategies (perturbation directions) for a two-stage FEP calculation involving the systems 0 and 1, and the intermediate M

Calculation strategy	Free energy formula
$0 \rightarrow M \rightarrow 1$	$\Delta A = \Delta A_{0M} + \Delta A_{M1}$
$0 \leftarrow M \leftarrow 1$	$\Delta A = -\Delta A_{M0} - \Delta A_{1M}$
$0 \rightarrow M \leftarrow 1$	$\Delta A = \Delta A_{0M} - \Delta A_{1M}$
$0 \leftarrow M \rightarrow 1$	$\Delta A = -\Delta A_{M0} + \Delta A_{M1}$

The best choice of strategy depends on the problem under study

calculation and using it to provide insight into multistage calculation is a reasonable route to follow.

The intermediates of an MFEP calculation are usually defined using Hamiltonians related to those of system 0 ($\mathscr{H}_0$) and 1 ($\mathscr{H}_1$). This is generally achieved with a parameter-scaling approach (see Sect. 2.6). Linear scaling is the simplest form

$$\mathscr{H}(\lambda) = (1 - \lambda)\mathscr{H}_0 + \lambda\mathscr{H}_1, \tag{6.8}$$

where λ has a value between 0 and 1. For $\lambda = 0$ and 1, we recover $\mathscr{H}_0$ and $\mathscr{H}_1$, respectively, and other values of λ define intermediate states. Although, as discussed in Chap. 2, linear scaling may not be optimal, it is often favored for its simplicity. In an MFEP calculation, the perturbation between two successive states defined by λ_i and λ_{i+1} can be computed as

$$\Delta = U(\lambda_{i+1}) - U(\lambda_i) = -\Delta\lambda_i U_0 + \Delta\lambda_i U_1, \tag{6.9}$$

where $\Delta\lambda_i = \lambda_{i+1} - \lambda_i$. In many applications, the increment $\Delta\lambda$ is chosen to be equal for all i, which is a convenient but not necessarily an optimal choice.

6.2.2 Nonequilibrium Work Free Energy Methods

An important recent advance in free energy calculations comes from nonequilibrium statistical mechanics. The free energy difference between two states 0 and 1 of a classical system in contact with a heat reservoir can be estimated by perturbing the system so that NEW is performed during the transition between these two states. It is well known that, for repeated trials of such transitions, the average work done for the transitions satisfies $\langle W \rangle \geq \Delta A$, where equality is achieved only when the transition is infinitely slow (i.e., following a reversible path). As discussed in detail in Chap. 5, Jarzynski [3, 4] proved an equality between the free energy difference and the work done along the transition trajectories

$$\boxed{\exp(-\beta\Delta A) = \langle\exp(-\beta W)\rangle_0 .} \tag{6.10}$$

This equality holds even if the transitions from one state to the other (also referred to as switching) take a finite time. The NEW method for calculating the free energy using Jarzynski's equality (6.10) has rapidly gained popularity; it can be applied to a broad range of processes, including both computer simulations and experimental studies.

6.3 Understanding Free Energy Calculations

Two concepts are helpful in understanding and characterizing relative free energy computational errors: *phase space relationships* and *probability distribution functions of perturbations*.

6.3.1 Important Phase Space

The classical phase space Γ of an N-body system consists of a $3N$-dimensional configuration phase and a $3N$-dimensional momentum phase. Generally, the kinetic component of free energy is not considered in simulations. In common practice the two terms *phase space* and *configuration space* are used equivalently, as we do here. One goal for molecular simulation is to explore all representative phase space regions of a system. Remember that not all the configurations contribute to the partition function to the same extent. The contribution of a phase space point Γ_i to the partition function is described by its Boltzmann factor $\exp[-\beta U(\Gamma_i)]$, where $U(\Gamma_i)$ is the potential energy at configuration Γ_i. Significant contributions come from those phase space points with low (negative) U, or, rarely, the set of phase space points with relatively high energies but having extremely high probability to be visited by the system. Here we will call those configurations Γ_i making significant contributions to the partition function the *important phase space points* or the *important configurations* of a system, and the regions containing these points are denoted Γ^*. We can define $\Gamma_i \in \Gamma^*$ if $U(\Gamma_i)$ is less than a characteristic energy, e.g., the most-likely energy or the average energy of the system. Note that this definition may not apply to all distributions, e.g., certain very narrow or very broad distributions may not easily be described by this type of division. Still, for many cases of biophysical interest, regions of phase space that contribute most strongly can be identified either by their energies and Boltzmann factors or by cluster analysis and inspection. Thus, in general, Γ^* is the set of representative configurations and needs to be sampled well in a simulation to measure correctly the ensemble-averaged properties of the system [24, 26, 31, 32, 43, 44]. Note that with this definition the set of points may be very large or small, depending on the relative roughness of the potential energy surface. Regardless of the total size of this important phase space region, it is critical that it be well sampled for the relative free energies to be computed accurately.

Usually the 0 and 1 systems in a free energy calculation share the same phase space, but their important phase space regions (Γ_0^* and Γ_1^*, respectively) differ. In an FEP calculation, it is not sufficient to obtain a good free energy estimate by sampling from only the important configurations of the reference (e.g., Γ_0^*). Instead, both of the important regions for the reference and the target (through trial perturbations) should be well sampled because the ensemble average of the exponential in the perturbation ΔU is taken to compute ΔA. Suppose that a configuration Γ_i is sampled in system 0 (thus $\Gamma_i \in \Gamma_0^*$ assuming sampling in Γ_0^* is accessible) but $\Gamma_i \notin \Gamma_1^*$. Such a perturbation results in a large positive ΔU, and thus has a near-zero contribution to ΔA. If the Γ_1^* region is not sampled in the simulation, the free energy estimate will be inaccurate. In short, in terms of phase space, finite sampling error comes from missing regions of the configuration space that are important for both the reference and the target systems. A good FEP simulation strategy will require that there is no restriction to visiting important phase space regions of the target system by sampling the phase space using the Hamiltonian of the reference system.

So how is it possible to ensure this? Examining the phase space overlap relationship of the reference and the target systems will give insight towards answering this question.

Phase Space Relationship and Sampling Barriers

A two-dimensional cartoon is helpful in understanding overlap relationships between the important phase spaces Γ_0^* and Γ_1^*. As illustrated in Fig. 6.1, there are four possible ways that Γ_0^* and Γ_1^* can be related: (a) Γ_1^* can form a subset of Γ_0^*; (b) Γ_0^* and Γ_1^* may almost coincide, i.e., have complete overlap; (c) Γ_0^* and Γ_1^* may have partial overlap; and (d) Γ_0^* and Γ_1^* may have no overlap. For simplicity, in Fig. 6.1 we have only considered the case in which Γ^* is continuous in space; but the same principle applies to more-complicated situations, in which regions of Γ^* are separated in space.

Consider a single-stage FEP calculation with system 0 as the reference. For the situation illustrated in Fig. 6.1a, where Γ_1^* lies inside the Γ_0^*, configurations in Γ_1^* are
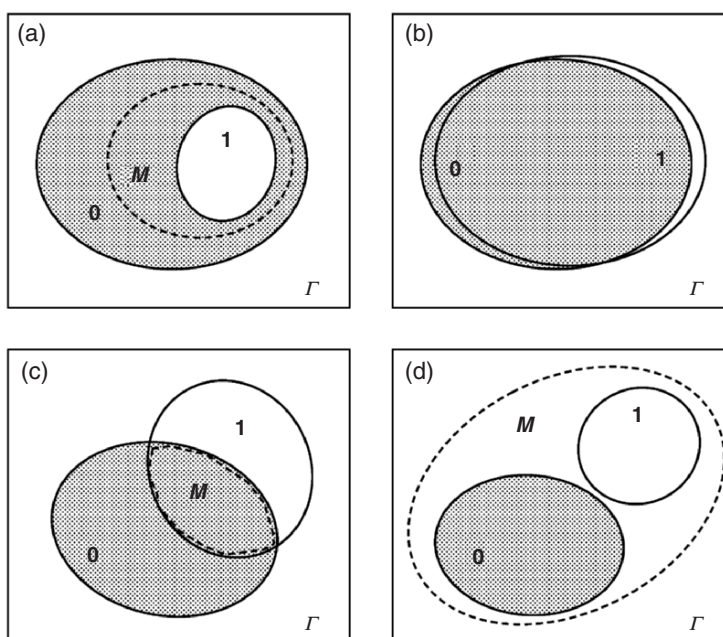
Fig. 6.1. A cartoon depiction of the phase space Γ and important phase space regions and their relationship. The important phase space regions of systems 0 (Γ_0^*) and 1 (Γ_1^*) are abstractly represented by *shaded and open oval shapes*, respectively. These regions can be related in four ways: (**a**) subset, (**b**) coincidence, (**c**) partial overlap, and (**d**) no overlap. Also sketched is the important phase space region of the intermediate M in a two-stage calculation (see section on "Multiple-Stage Design"). The appropriate staging strategy differs according to the different overlap relationships between Γ_0^* and Γ_1^*

sampled when a Monte Carlo or molecular dynamics simulation explores the phase space of Γ_0^*. The probability that Γ_1^* is encountered is proportional to the relative size (the number of configurations) of Γ_1^* and Γ_0^* – one may say that the sampling barrier in this case is *entropic*. If the simulation is long enough, both phase spaces Γ_1^* and Γ_0^* can be well sampled, and an accurate free energy estimate can be obtained. Clearly, in a finite-length simulation, the error in the calculated free energy difference and its convergence rate will depend on the relative sizes of Γ_1^* and Γ_0^*. This, in turn, depends on the perturbation magnitude (entropy difference [24, 26, 32]). There are two possible extreme cases: (1) Γ_1^* is tiny so that it is a pinhole in Γ_0^*, and (2) Γ_1^* and Γ_0^* are almost identical. This latter situation is the case (b) of Fig. 6.1 where the finite sampling error of the FEP is usually not a concern. In this case, a single-stage forward FEP calculation usually yields an accurate free energy estimate, given sufficient computational effort. In general, for effective calculation of ΔA one wants to ensure that Γ_1^* and Γ_0^* do not differ too much, which implies that an MFEP calculation may be required.

An example of how important the degree of phase space overlap is for the calculation can be illustrated by further consideration of Figs. 6.1a and 6.2. In particular, if system 1 instead of system 0 is chosen as the reference, as in Fig. 6.1a. Now, most of the time is spent sampling configurations inside the Γ_1^* region. To visit the region belonging to Γ_0^* but outside of Γ_1^* – which is important for an accurate free energy estimate – a region of low probability has to be sampled, this is proportional to $\exp(-\beta U)$ where U is once again the potential energy. Thus for the reverse FEP
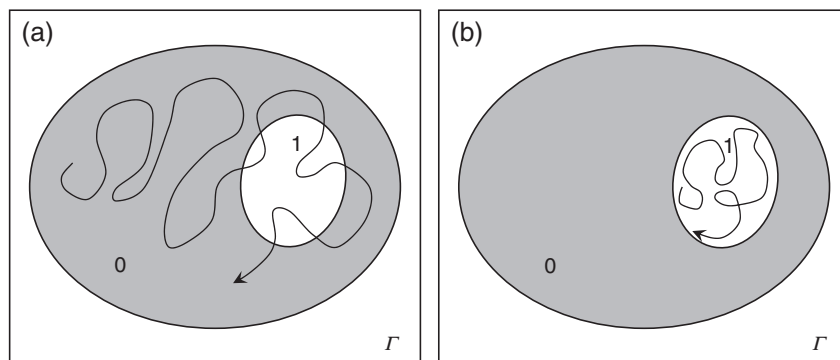
Fig. 6.2. Schematic illustration of the phase space sampling for the case presented in Fig. 6.1a, i.e., the important phase space regions of system 1 (Γ_1^*) is a subset of that of system 0 (Γ_0^*). The sampling behavior depends on which system is chosen as the reference in the simulation. The sampling process is shown schematically as the *curve with an arrow*. (**a**) System 0 is the reference. During the simulation, the phase space region important to both systems 0 and 1 is encountered by chance, which is proportionally to the *relative size* of Γ_1^* and Γ_0^*. The sampling barrier to Γ_0^* is entropic. (**b**) System 1 is the reference. The sampling of the Γ_0^* region outside Γ_1^* is limited by the high energy relative to the Γ_1^* region, thus many configurations important to the free energy calculation cannot be sampled due to the high energy restriction

calculation the sampling is limited by the most-likely regions of phase space and thus *energetics*, in contrast to the entropic restriction to good sampling in the forward FEP calculation. Missing these important configurations with the reference point in Γ_1^* is almost certain in a finite-length reverse FEP calculation, and thus will produce an unreliable result. This also illustrates the importance of examining both the variance and bias of a free energy calculation: the free energy result can be very precise, as independent simulations will produce consistent answers (related to their own average); however, the bias may be large because of the inability to sample important regions of phase space.

Applying the same analysis one can see that for the case described by Fig. 6.1c, neither the forward nor the reverse FEP calculation is reliable; and for the case in Fig. 6.1d, their results will be completely wrong. Thus, a single-stage FEP calculation can produce accurate free energy estimates only for the cases in Fig. 6.1a (in the $0 \rightarrow 1$ direction) and Fig. 6.1b. In conclusion, 'to guarantee a reliable free energy estimate the important phase space of the target state should be a subset of that of the reference state.' One way to achieve this is to choose the higher-entropy system as the reference (in Fig. 6.2a rather than Fig. 6.2b [24–27, 32], see Sect. 6.4 for more details).

Multiple-Stage Design

An MFEP calculation may be helpful in improving the efficiency and accuracy of free energy estimates for the situations illustrated in Fig. 6.1. However, for the cases of Fig. 6.1c, d, it is required in order to produce a reliable free energy result. The subset phase space relationship provides a framework for multiple-stage design [24, 26, 43]. The concept of placing intermediates in different situations is laid out in the following and details for constructing the intermediates will be discussed later in this chapter. For simplicity, a two-stage FEP calculation is illustrated as an example. Thus we have initial and final states 0 and 1, and one intermediate M.

For the case in Fig. 6.1a, the important configurations of an appropriately constructed intermediate, Γ_M^*, should be a subset of Γ_0^* and at the same time a superset of Γ_1^*, i.e., $\Gamma_0^* \in \Gamma_M^* \in \Gamma_1^*$. Thus, two separate FEP calculations should be performed for the transformations $0 \rightarrow M$ and $M \rightarrow 1$, yielding an overall free energy difference of $\Delta A \equiv \Delta A_{0 \rightarrow 1} = \Delta A_{0 \rightarrow M} + \Delta A_{M \rightarrow 1}$.

For the case Fig. 6.1c, in which Γ_0^* and Γ_1^* partially overlap, a good strategy is to construct M that Γ_M^* is located in the overlap region of Γ_0^* and Γ_1^*, i.e., $(\Gamma_M^* \in \Gamma_0^*) \cap (\Gamma_M^* \in \Gamma_1^*)$. Two separate FEP calculations are required for the transformations $0 \rightarrow M$ and $1 \rightarrow M$, and the corresponding free energy difference is $\Delta A = \Delta A_{0 \rightarrow M} - \Delta A_{1 \rightarrow M}$. This procedure is referred to as *overlap sampling*.

Finally, for the case Fig. 6.1d, in which Γ_0^* and Γ_1^* have no overlap, one may construct an intermediate whose important phase space regions contain both Γ_0^* and Γ_1^*, i.e., $(\Gamma_0^* \in \Gamma_M^*) \cap (\Gamma_1^* \in \Gamma_M^*)$, and conduct two FEP simulations for $M \rightarrow 0$ and $M \rightarrow 1$. Correspondingly, the overall free energy difference is given

by $\Delta A = \Delta A_{M \to 1} - \Delta A_{M \to 0}$. This approach is equivalent to the popular umbrella sampling (US) technique [45, 46], described in Chap. 3.

Note that for the cases Fig. 6.1c, d, depending on the difference between the 0 and 1 systems, it may be better to go with higher-order stage calculations. For example, for case (d) one can construct an intermediate M_B bridging Γ_0^* and Γ_1^*, so that the relationship of $\Gamma_{M_B}^* \sim \Gamma_0^*$ and $\Gamma_{M_B}^* \sim \Gamma_1^*$ are both described by case (c), and then perform overlap sampling calculations between 0 and M_B as well as 1 and M_B.

NEW Calculation: Down the Funnel

An argument analogous to that presented above for FEP can be made in connection with the NEW calculation, because one finite-time switching step in NEW is the equivalent to a perturbation trial in FEP. The difference is that in practice a NEW calculation involves multiple switching steps along the transition path from 0 to 1. At switching step i the important phase space region is Γ_i^*. As the transition moves one step forward, the system evolves to a state with important phase space Γ_{i+1}^*. The work increment for this switching step is calculated as $w_i = U_{i+1}(\Gamma_i) - U_i(\Gamma_i)$. Certainly small w_i makes a strong contribution to the total W along the entire path if Γ_i belongs to both Γ_i^* and Γ_{i+1}^*. If the system evolves along a path well connected in all Γ_i^*, we obtain a work value W that provides an important contribution to ΔA [cf. (6.10)]. This is clearly the goal of the calculation. However, if during the transition, the path encounters a configuration that belongs to Γ_i^* but not Γ_{i+1}^*, a large (positive) w_i for that step will result, which leads to a large W. Therefore the particular trajectory with its entire switching path will make a near-zero contribution to ΔA.

To ensure the accuracy of the free energy estimate by sampling the most important set of trajectories, we choose the sequence of systems so that *each* successive state obeys a phase space subset relationship with the one that preceded it. This situation is illustrated schematically in Fig. 6.3. We say that a path following such a trajectory moves *down the funnel* [43].

As the initial and final states are set by the problem under study, their important phase space relationship could be any one of the cases illustrated in Fig. 6.1. For cases Fig. 6.1c, d, it is impossible to construct a funnel path from 0 to 1 directly. To satisfy the funnel requirement, similar to the MFEP calculation, a staged NEW calculation can be performed. For example, in the case Fig. 6.1c, one can first construct an intermediate in the common region of Γ_0^* and Γ_1^*, then perform two separate NEW calculations following the paths $0 \to M$ and $1 \to M$, respectively. This NEW-overlap sampling (NEW-OS) technique will be discussed in detail in Sect. 6.6.

Note that the funnel requirement is attenuated to the degree that the path approaches reversibility. Actually for a reversible path, the phase spaces of successive states along the path will be almost identical (cf. Fig. 6.1b).

Entropy

Entropy can be thought of as describing the fraction of configuration space accessible to a particular energy. Thus, the entropy difference ΔS between two phase
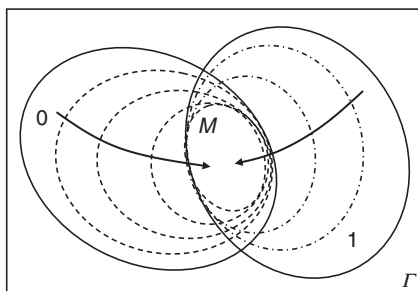
Fig. 6.3. To ensure the accuracy of a nonequilibrium work free energy calculation, the switching paths should go *down the funnel*. The important phase space regions for the intermediate states along the ideal funnel paths are illustrated in this plot, for the case where Γ_0^* and Γ_1^* are partially overlapped. Two funnel paths need to be constructed to transfer the systems from both 0 and 1 to a common intermediate M where Γ_M^* is inside the Γ_0^* and Γ_1^* overlap region. The construction of such paths is discussed in Sect. 6.6

space regions is an appropriate quantity to characterize differences in size between the important phase space regions of systems 0 and 1 [24, 26, 27, 32]. As discussed above, to ensure that free energy calculations are accurate by avoiding energetic sampling restrictions a subset relationship for Γ_0^* and Γ_1^* is required. To do so one should start the perturbation from the system having the larger important phase space region. This means that the reference system should be that with the higher entropy, and the transformation should proceed in the direction in which the entropy change ΔS is negative. In fact, as will be shown later in this chapter, entropy plays a key role in determining both the statistical and the systematic errors of a free energy calculation. Again, the requirement is attenuated as a NEW path approaches reversibility.

Summary

The concept of important phase space regions provides a way to understand the nature of FEP and NEW calculations. Studying the phase space overlap relationship for the reference and target systems helps to establish the principles for reliable free energy calculations. In order to guarantee a reliable FEP free energy estimate, the important phase space of the target state should be a subset of that of the reference state. Phrased in another way, for a single-stage calculation, both the starting and ending points need to be well sampled as well as the regions where the two optimally overlap. When this principle is not satisfied automatically by the problem at hand, appropriate stratification (multistage) design can ensure that this condition is fulfilled for the FEP calculation at each substage. Similarly, for NEW calculations, a finite-time switching path can be constructed to follow a funnel trajectory to ensure a correct free energy result.

Currently, our understanding and the guidelines extracted from phase space considerations are largely limited to a qualitative level. Further development is needed to

quantify these phase space overlap relationships between two systems. However even with solely the concept itself, these considerations of phase space are important and are helpful tools for understanding and studying free energy calculations.

6.3.2 Probability Distribution Functions of Perturbations

The perturbation process can be quantified using distribution functions of the potential energy change ΔU (for FEP) or work W (for NEW). The procedure is the same for both the FEP and NEW calculations, thus we use the term *perturbation* and the notation x to unify both.

As discussed in Chap. 2, the ensemble averages in (6.6) and (6.10) can be rewritten in terms of a one-dimensional integral weighted by the probability of encountering a specific x in a simulation. The probability of encountering a perturbation with value x during sampling for an FEP or NEW calculation is given by

$$f(x) = \langle \delta\left[x(\Gamma) - x\right] \rangle_0, \quad g(x) = \langle \delta\left[x(\Gamma) - x\right] \rangle_1, \tag{6.11}$$

for forward and reverse calculations, respectively, where δ is the Dirac delta function applied to the perturbation x. Integrating over these distribution functions, we have

$$\exp(-\beta\Delta A) = \int_{-\infty}^{\infty} \exp(-\beta x) f(x) \mathrm{d}x \tag{6.12}$$

for the forward calculation, and

$$\exp(+\beta\Delta A) = \int_{-\infty}^{\infty} \exp(+\beta x) g(x) \mathrm{d}x \tag{6.13}$$

for the reverse calculation. Note that in (6.12) and (6.13), both $f(x)$ and $g(x)$ are normalized, and, as before, the "+" signs in (6.13) is due to our definition of the *difference*.

An important feature for these distribution functions is that, instead of being independent of each other, they are coupled for a given pair of systems 0 and 1. For example, in an FEP calculation, we have [36, 37]

$$
\begin{aligned}
g(\Delta U) &= \langle \delta\left[\Delta U(\Gamma) - \delta U\right] \rangle_1 \\
&= \frac{1}{Q_1} \int \mathrm{d}\Gamma \delta\left[\Delta U(\Gamma) - \Delta U\right] \exp\left\{-\beta\left[U_0(\Gamma) + \Delta U\right]\right\} \\
&= \frac{Q_0}{Q_1} \frac{\exp(-\beta\Delta U)}{Q_0} \int \mathrm{d}\Gamma \delta\left[\Delta U(\Gamma) - \Delta U\right] \exp\left[-\beta(U_0(\Gamma))\right] \\
&= \exp(+\beta\Delta F) \exp(-\beta\Delta U) f(\Delta U).
\end{aligned}
\tag{6.14}
$$

In the right-hand side of the second equality, we have replaced $U_1(\Gamma)$ with $U_0(\Gamma) + \Delta U$. It can be shown that for a NEW calculation, $f(W)$ and $g(W)$ obey the same relationship [7, 43]. Thus, in unified notation, we have

$$\boxed{f(x)\exp(\beta\Delta A) = g(x)\exp(\beta x).}$$ (6.15)

This equation relates the free energy difference between two systems to the individual perturbations x and the f and g distribution functions. The relationship (6.15) is important in both the characterization of free energy error and the development of improved free energy methods.

An example plot of f and g distributions is shown in Fig. 6.4. Note that in general, the f and g distribution functions are *not* identical [cf. (6.15); one exception is the case of a NEW calculation with a reversible transition path, see Sect. 6.4.2]. They have different peak positions and widths (σ_f^2 and σ_g^2) characterizing the distributions.

Important Perturbation Tails

The perturbation distribution is a powerful tool for studying free energy errors [24, 26, 27, 38]. As one can see in (6.12) and (6.13), the contribution of an individual x to ΔA is weighted by the Boltzmann factor; thus it is strongly nonlinear. For a forward calculation, the significant contributions, weighted by $\exp(-\beta x)$, to the free energy come from those x located in the low-x (negative) tail of the distribution $f(x)$; similarly, for a reverse calculation, the important contributions, weighted by $\exp(+\beta x)$, come from the high-x (positive) tail of $g(x)$. To emphasize this concept, f and g and the corresponding Boltzmann weighting factors are shown in Fig. 6.4 together with the distributions.

Thus, the cause of the finite sampling error in free energy estimates from FEP or NEW simulations is the poor sampling of the important low-x tail of the f distribution and the high-x tail of the g distribution, for a forward and a reverse calculation, respectively. Sampling of these important tails corresponds to the sampling of the
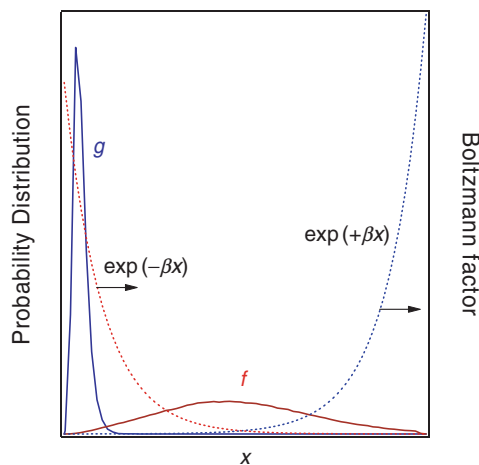
Fig. 6.4. A plot of typical f and g probability distributions as functions of the perturbation x. Both distributions are normalized. The associated Boltzmann factor integrands for the f and g distributions are also shown schematically in the plot, as the *dashed curves*

important phase space regions Γ_0^* and Γ_1^*. Due to the strong weights given to the important but rare perturbations, the FEP and NEW calculations can only provide a biased free energy estimate in practice, and some systematic error is an intrinsic feature of these calculations. However, using the distribution functions it is possible to conduct a quantitative error analysis, which will be given in the following sections.

Traditional Ways to Overcome Errors

Here we discuss a set of traditional methods for combining both the forward and reverse perturbation samples to enhance free energy calculations.

It has been observed that the errors in the forward and reverse calculations have different signs, i.e., one of the calculations usually overestimates the free energy difference, while the other underestimates [39, 42, 47–50]. Such observations have led to a common practice (in the context of FEP) for trying to overcome free energy errors: taking the average of the forward and reverse results as an estimate of ΔA. This approach is referred to as the *direct averaging* method. The double wide sampling method [39], introduced in Chap. 2, can be considered as a time-saving implementation of direct averaging in an MFEP calculation. However, recent advances, discussed later in this chapter, have led to the conclusion that the magnitude of error in the forward and reverse calculations is not, in general, the same and thus the statistical and systematic errors do not cancel out by taking an average from these calculations [23, 24, 26]. The direct averaging method can thus be seen as not reliable, and in some cases it is actually better to use the free energy result from a calculation in a single direction, rather than the average, since the error in one direction can be significantly worse than in the other.

Another old, but unfortunately less popular, technique – often referred to as the *overlap distribution* or *overlap histogram* method – deals with the error of FEP calculation in a more reliable way than the direct averaging counterpart. It utilizes the f and g distributions (in histogram form) and their relationship, given in (6.15) [31, 51–53]. It follows from (6.15) that when $f(x) = g(x)$, $x = \Delta A$, i.e., at the crossing point of $f(x)$ and $g(x)$, the value of x itself gives the free energy difference. In practice, such an estimate is not very reliable because of the large uncertainty that may be involved in identifying this crossover point. Instead, one takes the average of ΔA_i computed over the region in which the f and g histograms overlap, as the free energy estimate

$$\Delta A = \frac{1}{K} \sum_{i=1}^{K} \left[\beta^{-1} \ln \frac{g(x_i)}{f(x_i)} + x_i \right], \tag{6.16}$$

where K is the number of bins in the overlapping region of the f and g histograms. The quality of the free energy estimate can be further improved by using different weights for each i [31, 43, 52]. Graphical analysis is usually adopted to identify the f and g overlapping regions [7, 37].

The overlap distribution method is more reliable than the direct averaging or single-direction FEP/NEW calculation. For example, it is capable of producing a good estimate of the chemical potential of dense fluids where the latter methods fail

[32]. However, it does have some disadvantages. The f and g histograms obtained from a finite-length simulation contain errors themselves. The histogramming and graphical analysis require additional steps in the calculation, and can introduce new sources of errors [54]. Better alternatives, such as Bennett's method and the overlap sampling method, will be discussed further in this chapter.

For an FEP or NEW calculation, it is also helpful to plot the f and g distributions collected from simulation data. An insufficient overlap between f and g is usually a warning sign: it suggests that the perturbation is too large and additional stratification is needed. If f and g are not smooth in the overlapping region it usually means that the sample size is too small and longer simulations are required.

Summary

The perturbation distribution functions, f and g for forward and reverse FEP or NEW calculations, respectively, are very useful for characterizing free energy errors. In terms of perturbation distribution functions, the major finite sampling errors in an FEP or NEW calculation come from poor sampling of the important distribution tails – the low-perturbation tail of f and high-perturbation regions of g. The straight FEP and NEW free energies are biased, but modified methods based on the relationships of the f and g distributions can be used to reduce the finite sampling error.

6.4 Modeling Free Energy Errors

Perhaps the most challenging part of analyzing free energy errors in FEP or NEW calculations is the characterization of finite sampling systematic error (bias). The perturbation distributions f and g enable us to carry out the analysis of both the finite sampling systematic error (bias) and the statistical error (variance).

6.4.1 Accuracy of Free Energy: A Model

As discussed in Sect. 6.1, the bias due to finite sampling is usually the dominant error in free energy calculations using FEP or NEW. In extreme cases, the simulation result can be precise (small variance) but inaccurate (large bias) [24, 32]. In contrast to precision, assessing the systematic part (accuracy) of finite sampling error in FEP or NEW calculations is less straightforward, since these errors may be due to choices of boundary conditions or potential functions that limit the results systematically.

We again use a unified notation x to represent the potential energy change Δu for an FEP calculation and for NEW values W for a NEW calculation. In terms of perturbation distribution functions, the accuracy of free energy estimates depends on how well the $f(x)$ or $g(x)$ distributions are sampled. In a long but finite-length simulation, the region near the peak of a distribution may be well sampled, whereas the sampling of both the high-x and low-x tails may be poor, and may never happen.

Consider the forward calculation as an example. Earlier in this chapter, we concluded that the most important contribution to a forward free energy calculation comes from the low-x tail of $f(x)$, and its poor sampling results in the major systematic error of the calculation. To simplify further the analysis we assume that there is a *limit-perturbation* x_f such that $x > x_f$ is well sampled and $x < x_f$ is *never* sampled. This is illustrated in Fig. 6.5.

In this model, the finite sampling systematic error is due to the missed sampling of the important region $x < x_f$. The free energy estimate given by the model is

$$\exp(-\beta \Delta \hat{A}_{\mathrm{fwd}}) = \frac{\displaystyle\int_{x_f}^{\infty} f(x) \exp(-\beta x) \, \mathrm{d}x}{\displaystyle\int_{x_f}^{\infty} f(x) \, \mathrm{d}x}. \tag{6.17}$$

The denominator on the right-hand side is the renormalization factor, which is usually close to unity for large sample sizes and thus can be safely ignored (otherwise, the analysis presented below requires minor modification). Then we have

$$\exp(-\beta \Delta \widehat{A}_{\mathrm{fwd}}) = \int_{x_f}^{\infty} f(x) \exp(-\beta x) \mathrm{d}x. \tag{6.18}$$

Meanwhile, the exact free energy difference ΔA is given by

$$\exp(-\beta \Delta A) = \int_{-\infty}^{x_f} f(x) \exp(-\beta x) \mathrm{d}x + \int_{x_f}^{\infty} f(x) \exp(-\beta x) \mathrm{d}x. \tag{6.19}$$
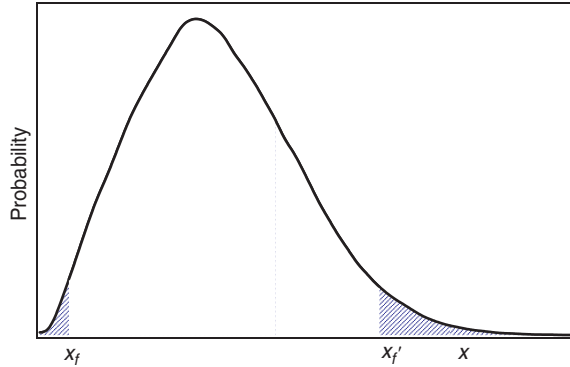
Fig. 6.5. Graphical illustration of the inaccuracy model and the relative free energy error in forward and reverse free energy calculations. A limit-perturbation x_f is adopted to (effectively) describe the sampling of the distribution: the regions above x_f are assumed to be perfectly sampled while regions below it (*shaded area*) are never sampled. We may also put a similar upper limit x'_f for the high-x tail, where there is no sampling for regions above it. However, this region (in a forward calculation) makes almost zero contribution to the free energy calculation and its error. Thus for simplicity we do not apply such an upper limit here

By subtracting (6.18) from (6.19), we obtain the finite sampling systematic error in $\exp(-\beta\Delta\widehat{A}_{\mathrm{fwd}})$

$$\delta E_{\mathrm{fwd}} \equiv \exp(-\beta\Delta\widehat{A}_{\mathrm{fwd}}) - \exp(-\beta\Delta A)$$

$$= -\int_{-\infty}^{x_f} f(x)\exp(-\beta x)\mathrm{d}x. \tag{6.20}$$

Similarly, for the reverse direction we can adopt a (highest) limit-perturbation x_g on the high-x tail which separates the regions of perfect sampling $(x \leq x_g)$ and zero sampling $(x > x_g)$. The same analysis leads to

$$\delta E_{\mathrm{rvs}} \equiv \exp(+\beta\Delta\widehat{A}_{\mathrm{rvs}}) - \exp(+\beta\Delta A)$$

$$= -\int_{x_g}^{\infty} g(x)\exp(+\beta x)\mathrm{d}x. \tag{6.21}$$

Let us further look at the relative systematic errors (δe). For a forward calculation

$$\delta e_{\mathrm{fwd}} = \frac{\delta E_{\mathrm{fwd}}}{\exp(-\beta\Delta A)}$$

$$= -\int_{-\infty}^{x_f} f(x)\exp(-\beta x)\exp(+\beta\Delta A)\mathrm{d}x$$

$$= -\int_{-\infty}^{x_f} g(x)\mathrm{d}x. \tag{6.22}$$

In the last equality of (6.22), the relationship (6.15) is used. Similarly, for a reverse calculation, we have:

$$\delta e_{\mathrm{rvs}} = \frac{\delta E_{\mathrm{rvs}}}{\exp(+\beta\Delta A)} = -\int_{x_g}^{\infty} f(x)\mathrm{d}x. \tag{6.23}$$

The errors in (6.20)–(6.23) are given for the exponential form of the free energy difference, and the inaccuracy in $\Delta\widehat{A}_{\mathrm{fwd}}$ and $\Delta\widehat{A}_{\mathrm{rvs}}$ can be obtained from them easily. Note that when δe is small, (6.22) and (6.23) give the *absolute* systematic error in $\beta\Delta\widehat{A}$ itself (through the Taylor expansion of δe to the second order).

Note that the effective limit-perturbations x_f and x_g are functions of the sample size. More sampling will 'push' x_f and x_g further toward the low-x and high-x tails, respectively, thereby reducing the error. We can think of these limits as reflecting the major bias restrictions to convergence in the calculation. For an x_f or x_g that represents the average of the limit-perturbations observed in independently repeated simulations (with the same sample size), this inaccuracy model gives an estimate of the average bias defined in (6.4). However, we prefer to use the most-likely behavior to describe the free energy error, and the most-likely bias is estimated by using the most-likely x_f and x_g in (6.22) and (6.23), respectively.

Equations (6.22) and (6.23) show that the relative systematic error in free energy due to finite sampling is simply given by the area under the tail region (marked by the effective limit-perturbation x_f or x_g) of the complementary distribution – see Fig. 6.6. Note that δe is always negative in both (6.22) and (6.23), which indicates that $\Delta \widehat{A}_{\mathrm{fwd}} \geq \Delta A$ (the exact value) and $\Delta \widehat{A}_{\mathrm{rvs}} \leq \Delta A$, i.e., the free energy difference is overestimated in a forward calculation, and underestimated in a reverse calculation. This conclusion agrees with the common observation regarding errors in the forward and reverse FEP calculations. With the equations presented above, it is possible to estimate inaccuracy in an FEP or NEW calculation from the estimated $f(x)$ and $g(x)$ obtained from simulation data and from the knowledge of the limit-perturbations x_f and x_g. We stress that, in general, the magnitudes of the forward and reverse FEP or NEW errors are *different*.

Most-Likely Inaccuracy

What level of inaccuracy can be expected for a simulation with a certain sample size N? This question can be transformed to another one: what is the effective limit-perturbation x_f or x_g in the inaccuracy model [(6.22) or (6.23)]? To assess the error in a free energy calculation using the model, one may histogram f and g using the perturbations collected in the simulations, and plot x in the tail of the distribution. However, if x_f is taken too small the accuracy is overestimated, and the assessed reliability of the free energy is therefore not ideal. In the following, we discuss the most-likely analysis, which provides a more systematic way to estimate the accuracy of free energy calculations.

The goal is to find the most-likely limit-perturbations x_f^* and x_g^*. We use a probability density function $P(x)$ to describe the spread of effective limit-perturbations
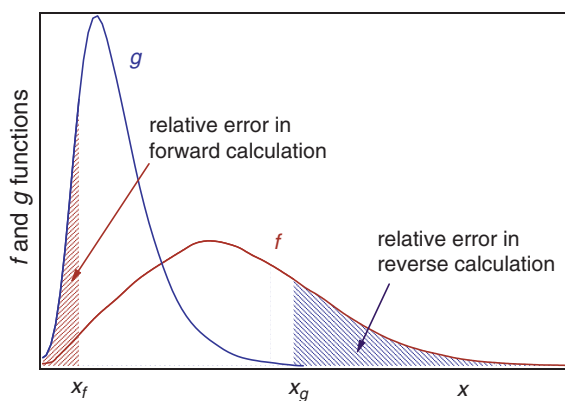
Fig. 6.6. The relative inaccuracy in $\exp(-\beta\Delta A)$ of the forward (f) calculation is given by the (*shaded*) area with x less than the limit-perturbation x_f under the g distribution, while the relative inaccuracy of the reverse (g) calculation is given by the area with x larger than x_g under the f distribution

$\{x_f\}$ or $\{x_g\}$ in independently repeated simulations with the same sample size N. The peak (mode) of $P(x)$ corresponds to the *most-likely* limit-perturbation x^*. As mentioned above, the inaccuracy model with x^* gives an estimate of the most-likely systematic error observed in any simulation with sample size N.

Consider a forward calculation. The probability that a particular x_f is the effective lowest perturbation encountered in a simulation with N independent perturbation trials is given by [25]

$$P(x_f) = f(x_f) \left[\int_{x_f}^\infty f(x) \mathrm{d}x \right]^{N-1}. \tag{6.24}$$

The peak of $P(x_f)$ is located by maximizing $P(x_f)$ as a function of x_f. Following a straightforward series of steps, we have (for large but finite N)

$$\left. \frac{\partial \ln f(x)}{\partial x} \right|_{x=x_f^*} = N f(x_f^*). \tag{6.25}$$

Similarly, the most-likely limit-perturbation x_g^* for a reverse calculation with sample size N is given by

$$\left. \frac{\partial \ln g(x)}{\partial x} \right|_{x=x_g^*} = -N g(x_g^*). \tag{6.26}$$

The minus sign on the right-hand side of (6.26) corresponds to the negative slope on the high-x tail of $g(x)$.

Note the relationship between x^* and N. A larger sample size N pushes the limit-perturbation further down the tail (smaller x_f^* and larger x_g^*), thus reducing the free energy error. With (6.25) and (6.26), the most-likely values of x_f^* and x_g^* can be estimated using the f and g histograms obtained from simulations. This gives us a means to assess the finite sampling systematic error in free energy without knowing the true ΔA.

Free Energy Bounds From Inaccuracy

From the discussion above, we know that the free energy estimates $\Delta \widehat{A}$ in a finite-length forward and reverse calculations form bounds for the true free energy difference, ΔA, i.e., [27]

$$\boxed{\Delta \widehat{A}_{\mathrm{rvs}} \leq \Delta A \leq \Delta \widehat{A}_{\mathrm{fwd}}.} \tag{6.27}$$

These bounds originate from the systematic errors (biases) due to the finite sampling in free energy simulations and they differ from other inequalities such as those based on mathematical statements or the second law of thermodynamics. The bounds become tighter with more sampling. It can be shown that, statistically, in a forward calculation $\Delta A(M) \leq \Delta A(N)$ for sample sizes M and N and $M > N$. In a reverse calculation, $\Delta A(M) \geq \Delta A(N)$. In addition, one can show that the inequality (6.27) presents a tighter bound than that of the second law of thermodynamics

$(\Delta A \leq \langle W \rangle)$ for a NEW calculation as well as the Gibbs–Bogoliubov inequality for an FEP calculation, discussed in [27] and Chap. 2. However, in contrast to Gibbs–Bogoliubov bounds, the inequality (6.27) is only a statement of the likely outcome of a simulation and, therefore, it may be violated in individual cases. The violation becomes increasingly probable as the free energy estimate approaches the correct value, i.e., as the equality limit is reached and the noise (imprecision) in the calculation becomes comparable to the free energy difference itself. In this case, precision becomes of more concern than accuracy.

Entropy and its Contribution to Free Energy Inaccuracy

In practice it is helpful to know the order of magnitude of the sample size N needed to reach a reasonably accurate free energy. The inaccuracy model described above presents an effective way to relate the sample size N and the finite sampling error through perturbation distribution functions. Alternatively, one can develop a heuristic that does not involve distribution functions and is determined by exploring the common behavior of free energy calculations for different systems [25]. Although only FEP calculations are considered in this section, the analysis extends to NEW calculations.

Common sense suggests that the magnitude of the perturbation between two systems of interest is a key factor in determining the free energy error. As discussed above, this can also be characterized in terms of entropy differences. As before, we assume that the forward direction corresponds to the negative entropy difference $\Delta S = S_1 - S_0$. Consider the value of N that yields a relative error of 50% in the free energy calculation. According to this accuracy model, a 50% fractional error in the forward FEP calculation occurs when the most-likely limit energy, u_f^*, lies at the median of $g(\Delta U)$. This is a reasonable choice, since it reflects sampling in perturbation energy space up to a certain level on the distribution. The analysis could be made more exact to any level by assuming general shapes for the distributions and different samples along those distributions. To simplify the analysis we assume $g(\Delta U)$ is a symmetric distribution; this is appropriate for many FEP cases [note this is usually not a good assumption for $f(\Delta U)$]. Thus, the median is identical to the peak value. Applying the f and g relationship, (6.25) becomes

$$\left. \frac{\partial \ln g(\Delta U)}{\partial \Delta U} \right|_{\Delta U = u_f^*} + \beta = N g(u_f^*) \exp\left[-\beta \left(\Delta A - u_f^* \right) \right]. \qquad (6.28)$$

Note that $\beta \Delta A = \beta \Delta U_{01} - \Delta S / k_B$, where ΔU_{01} is the potential energy difference between systems 0 and 1, i.e., $\Delta U = \langle U_1 \rangle_1 - \langle U_0 \rangle_0$. Since a 50% fractional error, u_f^* lies at the peak of $g(\Delta U)$ and (6.28) becomes

$$\beta \Delta A = N_{1/2} \exp(\Delta S / k_B) g(u_f^*) \exp\left[+\beta(u_f^* - \Delta U_{01}) \right], \qquad (6.29)$$

where the subscript '1/2' indicates 50% accuracy. The height of the peak of $g(\Delta U)$, i.e., $g(u_f^*)$, is proportional to σ_g^{-1} where σ_g is the standard deviation of g (e.g., for

a Gaussian distribution, $g(u_f^*) = \left(\sqrt{2\pi}\sigma_g\right)^{-1}$). If ΔU_{01} is approximated by u_f^* (the mean ΔU of $g(\Delta U)$), we have:

$$N_{1/2} \sim \exp(-\Delta S/k_\mathrm{B})\beta\sigma_g. \tag{6.30}$$

Note that more sampling is needed as the magnitude of the entropy difference ($\Delta S < 0$) increases, and, much less so, as the g distribution widens.

The primary observation from this analysis is the importance of the quantity $N\exp(\Delta S/k_\mathrm{B})$ in characterizing the expected systematic error in FEP calculations. The expression $N\exp(\Delta S/k_\mathrm{B})/(\beta\sigma_g)$ provides a better gauge of inaccuracy, but we take $N\exp(\Delta S/k_\mathrm{B})$ to emphasize the key factors. This is consistent with the common understanding that the accuracy of an FEP calculation is influenced by both the sampling size and the magnitude of the perturbation. Now we know that the latter is in the form $\exp(\Delta S/k_\mathrm{B})$. Actually the relationship between $N\exp(\Delta S/k_\mathrm{B})$ and the relative free energy error δe is found to be common and consistent in a number of different simulations for different types of systems (see Figs. 6.7 and 6.8). Further, $\delta e \sim [N\exp(\Delta S/k_\mathrm{B})]^{-\gamma}$, where γ is not a constant, but instead takes a value between 0 and 1 and increases with $N\exp(\Delta S/k_\mathrm{B})$. This indicates that inaccuracy decays relatively rapidly for sufficiently large N. The decay rate should be compared with and exceed $N^{-1/2}$, which is the decay rate for imprecision. This leads to the conclusion that, for sufficiently large N, inaccuracy eventually overcomes imprecision and thus the statistical error becomes dominant.

Simulation tests for different systems indicate that, for an FEP calculation to approach the accuracy level $\delta e < 5\%$, $N\exp(\Delta S/k_\mathrm{B})$ should be between 100 and
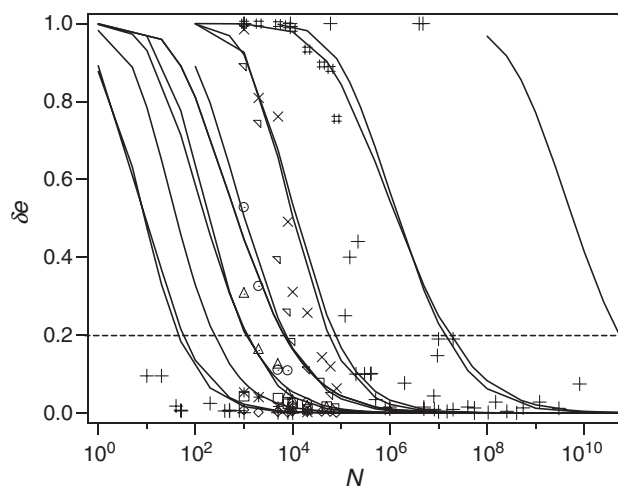
Fig. 6.7. The relative error versus the number of samples in several model systems. Plotting the error versus sample size directly creates a distribution of convergence rates. See [25] for more details on the systems
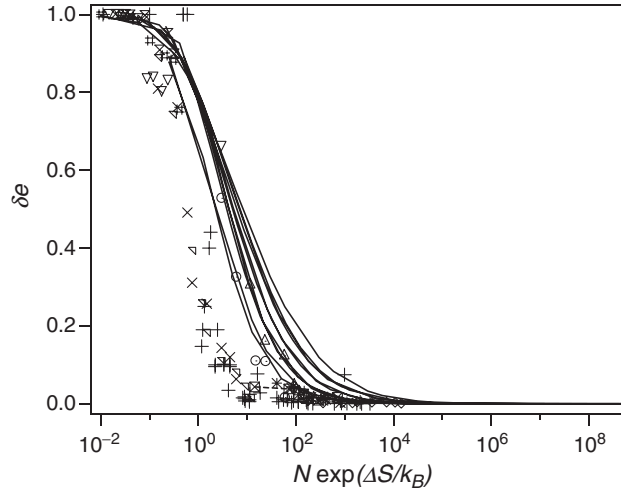
Fig. 6.8. The relative error versus the number of samples in several model systems. Plotting the error versus sample size directly creates a distribution of convergence rates. When the same data is plotted versus $N \exp(\Delta S/k_{\mathrm{B}})$ the curves superimpose. See [25] for more details on the systems

1,000 [25]. This means that the necessary sample size will be very large if the systems 0 and 1 differ too much. For example, for $\Delta S/k_{\mathrm{B}} = -15$, $N > 3 \times 10^8$ is needed to reach a 95% accuracy level. In this case, it is better to introduce intermediates and perform multistage calculations instead.

6.4.2 Variance in Free Energy Difference

The variance characterizes the spread of ΔA if an infinite number of independent simulations are carried out, each with a finite sample of size N. In practice, usually only one estimate (or a small number of repeats) of free energy differences are taken, and the variance in free energy must be estimated. One way to compute the variance is to use the error propagation formula (for a forward calculation)

$$\sigma^2_{\Delta A} = \frac{\exp(+2\beta\Delta A)}{\beta^2} \frac{\sigma^2_{\exp(-\beta x)}}{N}. \tag{6.31}$$

For simplicity, we have used ΔA in the subscript instead of $\Delta\widehat{A}$. For correlated samples $\{x\}$, the block averaging technique [37] can be used to improve the accuracy of the variance estimate. In this technique, the sequence of $\{x\}$ is broken up into blocks, each containing a certain number of samples, and the averages from each block are used to compute the variance. Unlike finite sampling bias, estimating the variance is not difficult in practice, although the accuracy of the estimated variance may be a concern. Nevertheless, modeling the variance is helpful for understanding free energy calculations.

The *scaled variance* of $\beta\Delta A$, defined as $N\sigma^2_{\beta\Delta A}$, can be related to the probability distribution function of the perturbation x [26]:

$$
\begin{aligned}
N\sigma^2_{\beta\Delta A} &= \mathrm{e}^{+2\beta\Delta A}\sigma^2_{\exp(-\beta x)} \\
&= \mathrm{e}^{+2\beta\Delta A}\left\{\int f(x)\exp(-2\beta x)\mathrm{d}x - \left[\int f(x)\exp(-\beta x)\mathrm{d}x\right]^2\right\} \quad (6.32) \\
&= \mathrm{e}^{+2\beta\Delta A}\int f(x)\exp(-\beta x)\mathrm{d}x - 1.
\end{aligned}
$$

Applying the relationship between f and g, we have

$$
N\sigma^2_{\beta\Delta A} = \exp(+\beta\Delta A)\int g(x)\exp(-\beta x)\mathrm{d}x - 1. \qquad (6.33)
$$

Similarly, for a reverse calculation:

$$
N\sigma^2_{\beta\Delta A} = \exp(-\beta\Delta A)\int f(x)\exp(+\beta x)\mathrm{d}x - 1. \qquad (6.34)
$$

Then we can rewrite (6.33) as

$$
N\sigma^2_{\beta\Delta A} = \exp\left[+\beta(\Delta A - \bar{x}_g)\right]\int g(x)\exp\left[-\beta\left(x - \bar{x}_g\right)\right]\mathrm{d}x - 1, \qquad (6.35)
$$

where $\bar{x}_g$ is the average of the perturbations x according to the complementary distribution $g(x)$. Applying the Taylor series expansion to $\exp\left[-\beta(x - \bar{x}_g)\right]$ near $\bar{x}_g$, we have

$$
\begin{aligned}
N\sigma^2_{\beta\Delta A} &= \exp\left[+\beta\left(\Delta A - \bar{x}_g\right)\right]\left[1 + \frac{1}{2}\int g(x)\beta^2\left(x - \bar{x}_g\right)^2\mathrm{d}x + \cdots\right] - 1 \\
&\approx \exp\left[+\beta\left(\Delta A - \bar{x}_g\right)\right]\left(1 + \frac{\beta^2}{2}\sigma^2_{x,g}\right) - 1,
\end{aligned}
$$
$$(6.36)$$

where $\sigma^2_{x,g}$ is the variance of the perturbation distribution $g(x)$. The expansion in (6.36) can be truncated at the second order.

NEW Calculations Near the Equilibrium Region

The approximation in (6.36) becomes exact in a NEW calculation near the equilibrium region, i.e., when the switch is sufficiently slow that the path is close to reversible. In the near-equilibrium region, according to the central limit theorem, the work distribution $f(W)$ is a Gaussian [5]. Consequently, the distribution $g(W)$ is also a Gaussian and the higher-order (>2) terms in the expansion of (6.36) are zero. In fact, if one of the work distributions is Gaussian, the other should also be Gaussian, and they have the same variance

$$
\sigma^2_{W,f} = \sigma^2_{W,g} \equiv \sigma^2_W. \qquad (6.37)
$$

This can be concluded from the Jarzynski equality (in the distribution function form) and the relationship between the f and g distributions. To repeat

$$\exp(-\beta\Delta A) = \int f(W)\exp(-\beta W)\mathrm{d}W,$$

$$\exp(+\beta\Delta A) = \int g(W)\exp(+\beta W)\mathrm{d}W,$$

$$(6.38)$$

and

$$\boxed{f(W)\exp(\beta\Delta A) = g(W)\exp(\beta W).}$$

$$(6.39)$$

One can further conclude that that these two Gaussian distributions are symmetrically located on the upper and lower sides of ΔA, and the free energy difference ΔA, the mean work $\bar{W}$ ($\overline{W_{0\to 1}}$ for the forward and $-\overline{W_{1\to 0}}$ for the reverse transformation) and the variance of work σ_W^2 obey the following relationships:

$$\Delta A = \overline{W_{0\to 1}} - \frac{1}{2}\beta\sigma_W^2 \tag{6.40}$$

and

$$\Delta A = -\overline{W_{1\to 0}} + \frac{1}{2}\beta\sigma_W^2. \tag{6.41}$$

Note that $\beta\sigma_W^2/2$ is the average dissipated work for the nonequilibrium process. As discussed in Chap. 5, (6.40) and (6.41) can also be obtained using the cumulant expansion of $\exp(-\beta\Delta A)$ [12–14, 17]. The expansion technique applies to non-Gaussian cases too, and equations such as (6.40) and (6.41) may be used to estimate ΔA from the mean and variance of the work distributions instead of Jarzynski's equality directly. These methods, however, work well only when the system is near equilibrium during switching, because higher terms in the expansion are negligible. For most applications, the inaccuracy model described in the last section provides a better description of the systematic errors. More-reliable methods (overlap sampling) will be discussed in Sect. 6.6.

With (6.37) and (6.40) or (6.41), (6.36) becomes (for the Gaussian distributions)

$$N\sigma_{\beta\Delta A}^2 = \exp\left(-\frac{1}{2}\beta^2\sigma_W^2\right)\left(1 + \frac{1}{2}\beta^2\sigma_W^2\right) - 1. \tag{6.42}$$

Now the variance in free energy difference is described in terms of the variance of work. The analysis above also indicates that the Gaussian distributions $f(W)$ and $g(W)$ are related

$$f(W + \beta\sigma_W^2) = g(W). \tag{6.43}$$

Combining (6.35), (6.41), and (6.43), we have:

$$N\sigma_{\beta\Delta A}^2 = \exp(\beta^2\sigma_W^2) - 1, \tag{6.44}$$

which is equivalent to (6.42) in this particular case.

An example of a Gaussian distribution pair is shown in Fig. 6.9. As the switching path approaches reversibility, $f(W)$ and $g(W)$ becomes closer to each other and their variance decreases. Both the bias and variance of the free energy estimate also decrease. Finally, at reversibility, the two distributions coincide at $x = W = \Delta A$, and converge at a single point ($x = \Delta A$, $f(x) = g(x) = 1$), as predicted from the second law of thermodynamics.

The Role of Entropy

Generally, a NEW calculation has finite sampling and is usually far from the 'near-equilibrium' region. Thus, the perturbation distributions are not Gaussian, and the conclusion of (6.42) or (6.44) is no longer valid. However, (6.36) is still useful.

Note that $\Delta A - \bar{x}_g \geq 0$ in (6.36) is the average dissipated work of the NEW calculation in the complementary (g) direction. Equation (6.36) shows that, given the same number of switching paths (N), a simulation with a slower switch leads to a smaller variance in the free energy, since the value of $\Delta A - \bar{x}_g$ becomes smaller. Also note that $\Delta A - \bar{x}_g$ has its maximum value in the FEP case (instantaneous switch), and is zero when the switching path is reversible.

The behavior of the variance can be further understood by separating the entropy and energy components of the free energy [26]. Equation (6.36) becomes

$$N\sigma^2_{\beta\Delta A} = \exp(-\Delta S/k_\mathrm{B})\exp(\beta z)\left(1 + \frac{1}{2}\beta^2\sigma^2_{x,g}\right) - 1, \qquad (6.45)$$

where $z = \Delta U - \bar{x}_g$, which decreases as the switch approaches equilibrium. Also, switching determines the variance of the perturbation $\sigma^2_{x,g}$. To ensure the appropriate
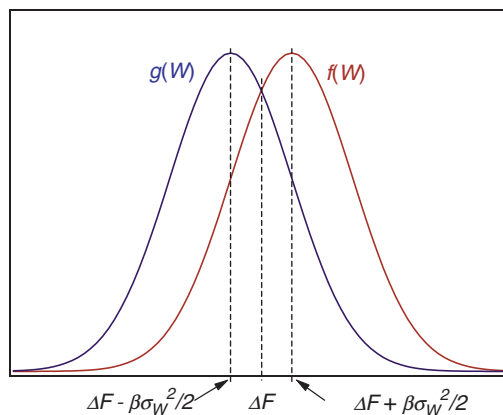
Fig. 6.9. When one of the probability distribution functions $f(W)$ and $g(W)$ is a Gaussian, the other must also is a Gaussian with the same variance (σ^2_W). These two density functions peak at $\Delta A + \beta\sigma^2_W$ and $\Delta A - \beta\sigma^2_W$, respectively. Their crossing point gives the free energy difference ΔA

sampling in the forward direction, we again choose it as the direction in which entropy decreases, i.e., $\Delta S < 0$.

Equation (6.45) reveals that there are three factors working together to determine the variance of the free energy estimate. One is the magnitude of the perturbation between the states 0 and 1 as described by $\exp(-\Delta S/k_{\mathrm{B}})$. A smaller perturbation (smaller $|\Delta S/k_{\mathrm{B}}|$) is preferable for reducing the variance, as one can predict from common sense. Thus, for systems with large perturbations, stratification is helpful in reducing the variance. In fact, the term $\exp(-\Delta S/k_{\mathrm{B}})$ describes the precision with which the entropy difference itself is measured. The second factor is the finite-time switching process, as described by $\exp(\beta z)\left(1 + \frac{1}{2}\beta^2\sigma_{x,g}^2\right)$. A slower switch leads to smaller values of z and $\sigma_{x,g}^2$, thus yielding a more precise free energy estimate. Note that the switching process affects the variance through z (exponential form) more than $\sigma_{x,g}^2$, thus $\exp(\beta z)$ is the key factor. The third factor is the sample size N as described by $1/N$. Compared to the entropy difference and z (both in exponential form), increasing the sample size is a less effective way to improve the precision.

In practice, the computation time is another factor affecting the level of precision a simulation can reach. More computational effort is required to perform a NEW calculation with more stages, slower switching paths, and/or longer sampling. An optimal calculation requires a balance among all these factors, as well as a balance between the precision and accuracy of the result. With a slower switching process, $\exp(\beta z)$ increasingly cancels out the effect of $\exp(-\beta\Delta S/k_{\mathrm{B}})$. Thus a near-equilibrium NEW calculation works regardless of the perturbation magnitude between the reference and the target state and their phase space relationship (also see discussion in Sect. 6.3.1). However, such a calculation may not be possible in practice.

6.5 Optimal Staging Design

With the analysis above, we can answer some important practical questions related to MFEP calculations. Should an MFEP calculation be used at all? How many stages are needed? How should the intermediates be formulated? What is the necessary sample size for each stage?

Through the phase space analysis we have arrived at the principle for choosing perturbation directions in a multistage calculation: the reference and target systems in each stage of calculation should form a subset relation in their important phase space regions, and the perturbation should be carried out from the system with the superset important phase space to the one with the subset. In terms of entropy, along the appropriate FEP calculation direction, ΔS should be negative (from the larger entropy system to the smaller one).

Thus, we consider a multistage calculation consisting of a total of n intermediates

$$(0) \to (m_1) \to (m_2) \to \cdots \to (m_n) \to (1).$$

The overall free energy difference is given by

$$\Delta A = \sum_{i=0}^{n} \Delta A_i. \tag{6.46}$$

The overall bias of the calculation is the sum of those for each stage (absolute value) and the overall variance is given as

$$\left(\sigma^2\right)_{\text{tot}} = \sum_{i=0}^{n} \left(\sigma_i^2\right). \tag{6.47}$$

Now consider the optimal placement of intermediates leading to the least overall variance. The criterion of minimal variance is obtained by inserting (6.45) into the right-hand side of (6.47) for each stage, and minimizing the overall error subject to

$$\Delta S_{\text{tot}} = \sum_{i=0}^{n} \Delta S_i, \tag{6.48}$$

where $\Delta S_i = S(m_{i+1}) - S(m_i)$ is the entropy difference for stage i. We have

$$\Delta(\Delta S/k_{\text{B}})_{ij} \equiv \Delta S_j/k - \Delta S_i/k = \ln\left(\zeta_j/\zeta_i\right), \quad (i = 0 \ldots n-1, j = i+1), \tag{6.49}$$

where $\zeta_i = \exp(\beta z_i)\left(1 + \frac{1}{2}\beta^2 \sigma_{x,g}^2\right)$. If we apply the same finite-time switching schedules for all the stages, we would expect that ζ_j/ζ_i is close to 1 and thus

$$\Delta(\Delta S)_{ij} \approx 0. \tag{6.50}$$

This means that, to obtain the least variance in a multistage calculation, the intermediates should be constructed to have *equal entropy difference* for all stages. This criterion differs from the often used but unjustified rule of thumb that free energy differences should be equal in all stages [22, 42]. Simulation tests show that the entropy criterion leads to a great improvement in calculation precision compared to its free energy counterpart [26]. The same optimization criterion holds for calculation of entropy and enthalpy differences [44].

Now consider the finite sampling systematic error. As discussed in Sect. 6.4.1, the fractional bias error in free energy is related to both the sample size and entropy difference: $\delta e \sim N \exp(-\Delta S/k_{\text{B}})$. With intermediates defined so that the entropy difference for each substage is the same (i.e., $\Delta S/n$), the sampling length N_i required to reach a prescribed level of accuracy is the same for all stages, and satisfies

$$N_i \exp(\Delta S/nk) = c, \tag{6.51}$$

where c is a constant. The total number of FEP samples N is (not considering the pre-equilibrium simulations for each reference system)

$$N = N_i n = nc \exp(-\Delta S/nk). \tag{6.52}$$

Minimizing N with respect to n, we have:

$$\boxed{-\Delta S_i/k = \frac{-\Delta S/k_{\mathrm{B}}}{n_{\mathrm{opt}}} = 1.} \tag{6.53}$$

That is, the optimal number of stages corresponds to unit entropy difference per stage.

The same analysis can be repeated for precision. From the analysis in Sect. 6.4.2, we have (for a single-stage calculation)

$$\sigma_{\Delta A}^2 \sim N^{-1} \exp(-\Delta S/k_{\mathrm{B}}). \tag{6.54}$$

The overall variance of an n-stage calculation is

$$\sigma_{\Delta A}^2 \sim n \left(\frac{N}{n}\right)^{-1} \exp(-\Delta S/nk) = \frac{n^2}{N} \exp(-\Delta S/nk). \tag{6.55}$$

Minimizing the variance with respect to n yields

$$\boxed{-\Delta S_i/k = \frac{-\Delta S/k_{\mathrm{B}}}{n_{\mathrm{opt}}} = 2.} \tag{6.56}$$

The same criterion for the optimal number of stages is obtained if instead we approach the problem by minimizing the overall sample size N for fixed precision and accuracy. The criteria for optimizing precision and accuracy differ slightly. In practice, the simulation should be long enough that the inaccuracy is not a major source of error. This means that we take (6.56) for designing optimal stages. Regarding the minimum requirement for the sample size at each stage of calculation, the heuristic presented in Sect. 6.4.1 yields $N_i > 100 \exp(\Delta S_i/k) = 100 \exp$ (6.2) $\approx 1,000$ independent samples.

6.6 Overlap Sampling Techniques

It would be valuable if one could proceed with a reliable free energy calculation without having to be too concerned about the important phase space and entropy of the systems of interest, and to analyze the perturbation distribution functions. The OS technique [35, 43, 44, 54] has been developed for this purpose. Since this is developed from Bennett's acceptance ratio method, this will also be reviewed in this section. That is, we focus on the situation in which the two systems of interest (or intermediates in between) have partial overlap in their important phase space regions. The partial overlap relationship should represent the situation found in a wide range of real problems.

6.6.1 Overlap Sampling in FEP

As concluded in Sect. 6.3.1, when the important phase spaces Γ_0^* and Γ_1^* have partial overlap, the appropriate staging strategy is to construct an intermediate M whose

important phase space is inside the overlapping region, which means that Γ_M^* is a subset of both Γ_0^* and Γ_1^*. Then, two separate FEP calculations for $0 \to M$ and $1 \to M$ are performed, yielding the free energy difference $\Delta A = \Delta A_{0M} - \Delta A_{1M}$.

A reasonable starting point for the potential energy of an OS intermediate M is

$$U_M = -\beta^{-1} \ln w(\Delta u) + (U_0 + U_1)/2, \tag{6.57}$$

where w is a weighting function that allows for fine-tuning of M. The free energy difference between systems 0 and 1 is given by

$$
\begin{aligned}
\exp(-\beta \Delta A) &= \frac{Q_1}{Q_0} = \frac{Q_M}{Q_0} \Big/ \frac{Q_M}{Q_1} \\
&= \frac{\langle w(\Delta u) \exp(-\beta \Delta u/2)\rangle_0}{\langle w(\Delta u) \exp(+\beta \Delta u/2)\rangle_1} \,,
\end{aligned}
\tag{6.58}
$$

where Q is the partition function of a canonical ensemble. Equation (6.58) is the general form of an overlap sampling working formula.

No sampling on the intermediate M in (6.58) is required. This feature offers the flexibility of optimization, as $w(\Delta u)$ can be adjusted freely without imposing any changes or a redesign of the simulation processes. In practice, the sampling requirement for the OS is the same as performing two separate FEP simulations: one in the forward direction ($0 \to 1$) and the other in the reverse direction ($1 \to 0$); and the same sets of forward and reverse perturbation data $\{\Delta u\}$ are used.

The accuracy of the OS calculation can be characterized using the f and g distributions, in a similar manner to that described in Sect. 6.4.1 for modeling the systematic error in the free energy. Assume that sampling of the f distribution is perfect for regions with ΔU greater than Δu_f, and no sampling below Δu_f is available. Similarly, assume that the sampling g is perfect up to Δu_g, but no sampling is available beyond. Then the relative systematic error of $\exp(-\beta \Delta A)$ due to finite sampling is

$$\delta_e = \frac{\displaystyle\int_{-\infty}^{\Delta u_f} w(\Delta U)\,[f(\Delta U)g(\Delta U)]^{1/2}\,\mathrm{d}\Delta U - \int_{\Delta u_g}^{\infty} w(\Delta U)\,[f(\Delta U)g(\Delta U)]^{1/2}\,\mathrm{d}\Delta U}{\displaystyle\int_{-\infty}^{\infty} w(\Delta U)\,[f(\Delta U)g(\Delta U)]^{1/2}\,\mathrm{d}\Delta U}.$$

$$\tag{6.59}$$

Interestingly, unlike the error in a single-direction calculation [cf. (6.22) and (6.23)], the systematic error in the OS calculation depends on the *product* of the f and g distributions. Thus the systematic error is related to the degree of overlap between f and g. The integral over the product $f(\Delta u)g(\Delta u)$ will be smaller than the integral over the distributions themselves, which suggests that the error in the OS method is naturally smaller. In addition, it follows from (6.59) that errors from inadequate sampling in the forward and reverse directions will tend to cancel each other. As one can see in (6.59), the size of the perturbation and the sampling size play important roles in determining the error in ΔA. Smaller perturbations yield better overlap between f and g. Longer sampling pushes the limit energies Δu_f and Δu_g further down the tails.

The weighting function $w(\Delta U)$ provides the most flexible way for reducing the finite sampling error. We have two apparent choices for $w(\Delta U)$. The first one is to use a truncated $w(\Delta U)$ function: $w(\Delta U) = 0$ for $\Delta U < \Delta u_f$ and $\Delta U > \Delta u_g$. In this case, (6.59) gives a zero error. The functional form of $w(\Delta U)$ for ΔU ranges between Δu_f and Δu_g will affect the precision of the calculation, but not its accuracy. The disadvantage of doing so is that we need to identify Δu_f and Δu_g by studying the f and g distributions (see Sect. 6.4 for the method to obtain these quantities).

The second choice does not require such additional analysis. In addition, it is helpful for considering the precision and the accuracy of free energy differences simultaneously. The idea is that one can reduce the systematic error by balancing $\int_{-\infty}^{\Delta u_f} w(\Delta U) \left[f(\Delta U) g(\Delta U) \right]^{1/2} d\Delta U$ and $\int_{\Delta u_g}^{\infty} w(\Delta U) \left[f(\Delta U) g(\Delta U) \right]^{1/2} d\Delta U$, to make both of them as small as possible, and then increase weight on the regions, in which $f(\Delta U)$ and $g(\Delta U)$ overlap well.

The minimization of the statistical error can be done by examining the variance of $\exp(-\beta \Delta A)$ directly. Bennett [55] has studied this problem by combining the forward and reverse FEP simulations, and the same analysis can be followed for the OS case. In Bennett's analysis, the weighting function w is placed to balance the forward and reverse FEP contributions [7, 55]

$$\beta \Delta A = \ln \langle w \exp(-\beta U_0) \rangle_1 - \ln \langle w \exp(-\beta U_1) \rangle_0 . \tag{6.60}$$

The variance of the free energy calculation is (with error propagation)

$$\sigma_{\beta \Delta A}^2 = \frac{\left\langle [w \exp(-\beta U_1)]^2 \right\rangle_0 - \langle w \exp(-\beta U_1) \rangle_0^2}{n_0 \langle w \exp(-\beta U_1) \rangle_0^2} \\ + \frac{\left\langle [w \exp(-\beta U_0)]^2 \right\rangle_1 - \langle w \exp(-\beta U_0) \rangle_1^2}{n_1 \langle w \exp(-\beta U_0) \rangle_1^2}, \tag{6.61}$$

where n_0 and n_1 are the sample sizes for the forward and reverse FEP calculations. The variance can be minimized with respect to w using Lagrange multipliers, from which we have

$$w = K \left[\frac{Q_0}{n_0} \exp(-\beta U_1) + \frac{Q_1}{n_1} \exp(-\beta U_0) \right]^{-1}, \tag{6.62}$$

where K is an arbitrary constant and Q is the partition function. Substituting (6.62) into (6.60), we obtain

$$\exp(-\beta \Delta A) = \frac{\left\langle \{1 + \exp[\beta(\Delta u - C)]\}^{-1} \right\rangle_0}{\langle \{1 + \exp[-\beta(\Delta u - C)]^{-1} \rangle_1} \exp(-\beta C), \tag{6.63}$$

where $\exp(-\beta C) \equiv (Q_1 n_0)/(Q_0 n_1)$. Note that (6.63) is valid for any choice of C, however, this particular choice of C is optimal.

For the OS calculation, (6.58), the same optimal solution is reached by choosing a Gaussian-like hyperbolic secant function for $w(\Delta u)$:

$$w(\Delta u) = 1/\cosh[\beta(\Delta u - C)/2]. \tag{6.64}$$

Equation (6.58) becomes identical to that of the Bennett method (also referred to as the acceptance ratio method), i.e.

$$\exp(-\beta\Delta A) = \frac{\left\langle \{1 + \exp[\beta(\Delta u - C)]\}^{-1}\right\rangle_0}{\left\langle \{1 + \exp[-\beta(\Delta u - C)]^{-1}\right\rangle_1} \exp(-\beta C). \tag{6.65}$$

Clearly, the value of C is related to the free energy difference ΔA:

$$\beta\Delta A = \beta C - \ln\frac{n_1}{n_0}. \tag{6.66}$$

When the same sampling size is taken for both forward and reverse calculations, it becomes

$$C = \Delta A. \tag{6.67}$$

Since we do not know the value of C in advance, the optimal C and thus the free energy difference ΔA can be solved in practice by iterating self-consistently (6.65) and (6.66) or (6.67). A convenient way to do so is to record all the perturbation data during the simulation, then compute C and ΔA in a postsimulation analysis. This method is also referred to as Bennett's method or the acceptance ratio method.

However, this analysis has been performed from a purely statistical perspective, leading to the minimal statistical error for the calculation. The phase space relationship, the staging scheme (conceptual intermediate M), and thus the accuracy of the calculation are not included in Bennett's picture. However, it turns out that the calculation is also optimal from the accuracy point of view. With this optimal choice of $C = \Delta A$, the weight function $w(\Delta u)$ given by (6.64) has its peak exactly at the crossover between f and g, where $\Delta U = \Delta A$ [cf. (6.15)]. In contrast, the weights for the low-Δu tail of f and high-ΔU tail of g are diminished, thus resulting in small systematic error.

Simple Overlap Sampling

Arbitrary weighting functions can be used for the OS method. The simplest choice is $w(\Delta U) = 1$ for all ΔU, thus the OS formula reduces to:

$$\exp(-\beta\Delta A) = \frac{\langle\exp(-\beta\Delta U/2)\rangle_0}{\langle\exp(+\beta\Delta U/2)\rangle_1}. \tag{6.68}$$

We refer to (6.68) as the simple overlap sampling (SOS) method. The optimization feature is lost in the SOS method, but it usually produces a very good estimate of ΔA and can be used as a simple alternative to Bennett's form for those who do not like the small additional work of solving for C. However, we note that the application of the OS method with $C = 0$ in (6.64) is not recommended because its performance is uneven [35].

The calculation effort involved in the SOS is identical to the direct averaging method $[\Delta A = (\Delta A_{\mathrm{fwd}} + \Delta A_{\mathrm{rvs}})/2]$. However, the reliability of the corresponding free energy estimates can differ significantly [35, 44, 54]. For comparison, the working equation for a direct average method is

$$\exp(-\beta\Delta A) = \frac{\langle\exp(-\beta\Delta U)\rangle_0^{1/2}}{\langle\exp(+\beta\Delta U)\rangle_1^{1/2}}. \qquad (6.69)$$

Overlap sampling, especially its optimized form with Bennett's weighting function, can greatly improve the reliability (both precision and accuracy) of the free energy estimate over conventional FEP techniques. By solving for the optimal C (and thus ΔA) OS with Bennett's weighting function actually locates the crosspoint of f and g distributions automatically. The SOS provides a simple but reliable alternative to the direct averaging method. With the capability of optimization and automation and without the need for histogram analysis, OS is a better method than the overlap histogram approach discussed in Sect. 6.3.2 [54].

A certain degree of overlap between f and g is required for a reliable OS calculation [a lack of overlap leads to a small denominator on the right-hand side of (6.59), producing a large error]. Bennett's form of OS method will fail to locate the crossover point if it does not exist at all. But the requirement for overlap, which can be measured by $\int [f(u)g(u)]^{1/2}\,du$, is usually small [35]. In fact, the overlap sampling technique can handle much larger perturbations than single-direction FEP counterparts, which benefits efficiency. For very different systems, a multistage OS calculation can be applied, in which a single-stage OS formula is applied to obtain the free energy difference between two successive (intermediate) systems. It is helpful to employ time-saving tricks, such as perturbing both forward and backwards from each intermediate state.

Performance comparisons between overlap sampling (including Bennett's method and SOS) and conventional FEP methods are shown in Fig. 6.10. As one can see, overlap sampling methods converge to the correct free energy much faster than direct FEP methods, and have smaller statistical errors. The performance of Bennett's method is amazing. Also, since more than one form of $w(\Delta U)$ can be chosen for a given set of forward and reverse perturbation data, multiple free energy estimates can be made from the same set of simulations. The consistency of these free energy results can provide an additional assurance for the quality of the estimate, since there should be only one correct free energy difference. In short, without the requirement of modifying the FEP simulation, the overlap sampling technique improves both the efficiency and reliability of free energy calculations.

6.6.2 Overlap and Funnel Sampling in NEW Calculations

Down the Funnel: Phase Space Consideration

To repeat briefly, the NEW method is related to free energy differences between systems 0 and 1 through Jarzynski's identity
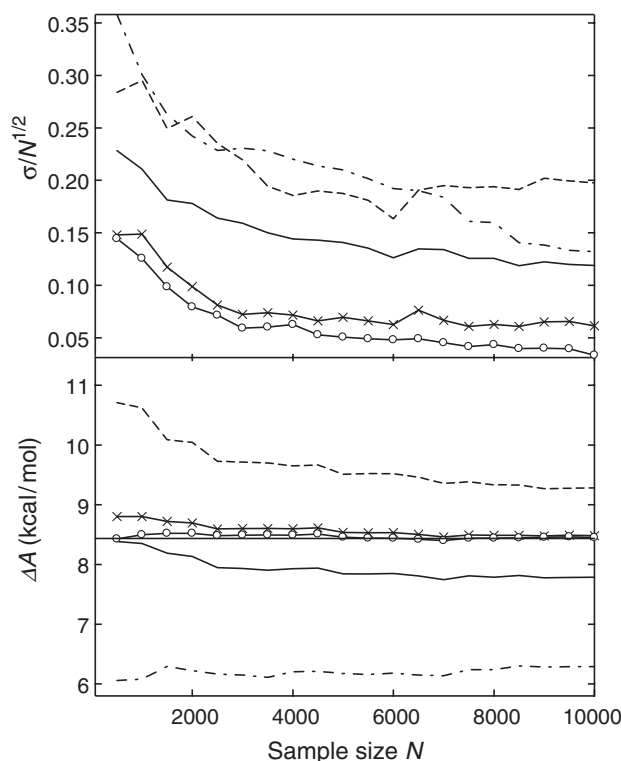
Fig. 6.10. Comparison of overlap sampling and FEP calculation results for the free energy change along the mutation of an adenosine in aqueous solution (between $\lambda = 0.05$ and 0.45) in a molecular dynamics simulation. The results represent the average behavior of 14 independent runs. (MD time step.) The sampling interval is 0.75 ps. The upper half of the plot presents the standard deviation of the mean (with gives statistical error) for ΔA as a function of sample size N; the lower half of the plot gives the estimate of ΔA – for comparison of the accuracy, the correct value of ΔA is indicated by the *bold horizontal line*
Key: *dashed curve* – forward FEP, *dash-dotted curve* – reverse FEP, *solid curve* – direct FEP averaging, *solid curve with crosses* – simple overlap sampling, *solid curve with open circles* – overlap sampling with the optimal Bennett's weights. Data have units of kcal mol^{-1}

$$\exp(-\beta \Delta A) = \langle \exp(-\beta W_{0 \rightarrow 1}) \rangle_0 \, , \tag{6.70}$$

where $W_{0 \rightarrow 1}$ is the nonequilibrium (finite-time) work for switching system from 0 to 1 along a path at a finite rate, and the ensemble average is taken over the equilibrium system 0. The NEW calculation can be conducted in the reverse direction with a path from 1 to 0. Note that, according to the definition of the work, $W_{0 \rightarrow 1} = -W_{1 \rightarrow 0}$. The finite sampling error of a NEW calculation is due to missed sampling of important but rare W. As concluded in Sect. 6.3.1, to ensure appropriate sampling, the finite-time switching path should go down the funnel, i.e., the sequence of systems traversed in a NEW calculation must proceed such that *each* successive system obeys a phase space subset relation with the ones that precedes it.

The linear parameter scaling approach of (6.9) is usually used to evolve the system from one state to the other. One disadvantage of such a scaling approach is that it offers no control of the phase space a path traverses. Instead, a similar parameter γ ($\gamma \in [0, 1]$, $\gamma = 0$ and 1 for the '0' and '1' systems, respectively) can be used to describe progress along the path, e.g., $\gamma = 0 \rightarrow \gamma_1 \rightarrow \cdots \rightarrow \gamma_i \rightarrow \gamma_{i+1} \rightarrow \cdots \rightarrow \gamma_n = 1$ where n is the total number of 'steps' of the switch. The set of important configurations for the (intermediate) system defined by γ_i is denoted $\Gamma^*_{\gamma_i}$.

Depending on the relationship between the important phase space regions Γ^*_0 and Γ^*_1, different strategies can be used to construct funnel paths. It may be better to set up the free energy calculation in stages, and construct funnel paths for each stage of the NEW calculation. In the following we consider the case in which the 0 and 1 systems have a partial overlap in their important phase space regions. As discussed before, for this situation overlap sampling is the technique of choice for free energy calculations.

Overlap Sampling with Funnel Paths

One key element of overlap sampling is an intermediate M such that Γ^*_M is inside the overlapping region of Γ^*_0 and Γ^*_1. Then, the overall free energy calculation is separated into two calculations, corresponding to free energy changes for $0 \rightarrow M$ and $1 \rightarrow M$ (cf. Fig. 6.3). In this case, we perform two separate NEW calculations, and the overall free energy difference is

$$\exp(-\beta \Delta A) = \frac{\langle \exp(-\beta W_{0 \rightarrow M}) \rangle_0}{\langle \exp(-\beta W_{1 \rightarrow M}) \rangle_1}. \tag{6.71}$$

Now we proceed with the construction of the intermediate M and the switching paths $0 \rightarrow M$ and $1 \rightarrow M$. The key considerations are: (1) both paths should follow the funnel sampling path to eliminate the systematic error due to the inaccessibility of important phase space, and (2) M is the optimal choice for minimizing the variance of the calculation. The optimal intermediate (for minimal variance) can be defined as follows [43]:

$$\exp(-\beta U_M) = \{\exp(-\beta U_0) + \exp[-\beta(U_1 - \Delta A)]\}^{-1}. \tag{6.72}$$

We can define the potential energy of a state along the path as follows:

$$\exp(-\beta U_\gamma) = [(1 - \gamma)\exp(+\beta U_0) + \gamma \exp(+\beta U_1)]^{-1}. \tag{6.73}$$

Then, systems 0 and 1 are recovered at $\gamma = 0$ and 1, respectively.

Now let us examine the phase space of an intermediate γ defined by (6.73). Consider a phase space point Γ_i. If Γ_i is outside of Γ^*_0 but inside Γ^*_1 (i.e., $\exp[-\beta U_0(\Gamma_i)]$ is large but $\exp[-\beta U_1(\Gamma_i)]$ is small), we can expect that the value of U_0 will be small (negative) or moderate and that of U_1 will be very large (positive). Consequently, the right-hand side of (6.73) will be small, which indicates that U_γ is a large positive number. Therefore, this specific Γ_i is unlikely to be part of Γ^*_γ, the important phase

space of γ system. The same conclusion can be reached if Γ_i is part of Γ_1^*, but not of Γ_0^*. However, if Γ_i is a subset to both Γ_0^* and Γ_1^* (both $\exp[-\beta U_0(\Gamma_i)]$ and $\exp[-\beta U_1(\Gamma_i)]$ are small), we can expect that $U_\gamma(\Gamma_i)$ makes an important contribution to its partition and thus Γ_i is part of Γ_γ^*. Therefore, at intermediate γ, configurations important to *both* 0 and 1 systems are increasingly important. Thus for an intermediate M defined by an appropriate γ, we should expect that the path taken from 0 and 1 to M will proceed down a funnel.

In addition, by comparing (6.72) and (6.73) we obtain γ^*, which defines the optimal common destination M:

$$\frac{\gamma^*}{1 - \gamma^*} = \exp(-\beta \Delta A). \tag{6.74}$$

Equations (6.71), (6.73), and (6.74) complete the NEW-OS method. To summarize, in the method, an intermediate M is introduced and the free energy difference is computed as $\Delta A = \Delta A_{0 \to M} + \Delta A_{1 \to M}$ (6.71). In this approach the free energy components are computed using two NEW calculations taking paths $0 \to M$ and $1 \to M$. The funnel sampling paths and the optimal intermediate M [(6.73) and (6.74)] are used to ensure appropriate sampling and to minimize both statistical and systematic error.

The optimal parameter γ^* and thus the free energy difference ΔA can be obtained by solving (6.71), (6.73), and (6.74) self-consistently. With a predefined set of $\{\gamma_i\}$, which is usually nonuniformly spaced and fixed during the simulations, the NEW-OS calculation proceeds as follows. Starting from the equilibrated system 0, perform the NEW calculation from 0 to 1 (with the predefined set $\{\gamma_i\}$) that follows the path defined by (6.73). Separately perform the same NEW calculation for $1 \to 0$ using the same set of γ_i, but in the reversed order, starting from the equilibrated system 1. During the NEW calculations, for each switch step along the path, record the partial work values at each point of γ_i (e.g., the work W_{0i} involved in switching the system from 0 to γ_i). Thus after the simulations we obtain not only the ensemble of the total work W_{01} and W_{10}, but also ensembles of partial work W_{0i} and W_{i0} for all i. Ensemble averages are then computed for all γ_i using (6.71), and optimal γ (and, therefore, ΔA) is selected to satisfy (6.74).

Softening Funnel Paths

One practical problem for the funnel path (6.73) is related to the exponential operations for both U_0 and U_1. For two states 0 and 1 that differ significantly, it is likely that the work collected from many switching paths will make near-zero contribution to the free energy calculated from (6.71), except for those paths that happen to start in the Γ_M^* region. That is, the transition from 0 and 1 to M may follow *pinhole* funnel paths. Also, the optimal γ^* would be very close to 0 or 1, making it difficult to locate, and therefore to optimize the calculation. The problem may be relieved by two approaches: (1) bias the sampling of the paths toward those making large contributions to the free energy while satisfying the *down the funnel* principle, and (2) modify the

'shape' of the funnel to ensure smoother and broader $0 \rightarrow M$ and $1 \rightarrow M$ paths. In the following we briefly discuss the second approach by introducing additional parameters.

The finite-time switching path of (6.73) can be modified with two constants α and D

$$\exp(-\beta U_\gamma) = \exp(-\beta U_0) \left[(1 - \gamma) + \gamma \exp\left[+\alpha(U_1 - U_0) - D \right] \right]^{-\beta/\alpha}. \quad (6.75)$$

The new parameter α affects the softness of the transition from 0 and 1 to M: by choosing an appropriate value for α, paths starting from many configurations in Γ_0 and Γ_1 can make important contribution to the free energy calculation. A good choice for α is $\alpha = (U_1 - U_0)^{-1}$ for typical values of U_0 and U_1. Indeed, an appropriate choice of α makes U_γ a more linear function of γ. Note that α is a fixed parameter and does not change during the simulation. One disadvantage is that once α is introduced, the path will no longer lead to the rigorously optimized intermediate as defined by (6.74). Then, a modified version of (6.74) is used as the criterion for selecting γ^*

$$\frac{\gamma^*}{1 - \gamma^*} = \exp\left[-\beta(\Delta A - D) \right]. \quad (6.76)$$

The parameter D is employed simply to make the identification of γ^* more convenient in practice. As one can see, with the 'ideal' choice of $D = \beta \Delta A$, the value of γ^* in (6.76) is simply 0.5.

Generalized Acceptance Ratio Method

Following Bennett, Crooks proposed the *generalized acceptance ratio* (GAR) method to combine the forward and reverse NEW calculations to minimize the statistical error of the relative free energy [56]

$$\exp(-\beta \Delta A) = \frac{\left\langle \left\{ 1 + \exp[\beta(W_{0 \rightarrow 1} - C)] \right\}^{-1} \right\rangle_0}{\left\langle \left\{ 1 + \exp[\beta(W_{1 \rightarrow 0} + C)]^{-1} \right\rangle_1} \exp(-\beta C), \quad (6.77)$$

where C is a constant and the statistical error in ΔA is minimized when $C = \Delta A$.

Equation (6.77) involves only work values for the forward ($W_{0 \rightarrow 1}$) and reverse ($W_{1 \rightarrow 0}$) NEW calculations but no details regarding the path that the finite-time switches traverse. Note that for the FEP case, the optimal OS calculation [cf. (6.59)] becomes identical to the acceptance ratio method. This is because the switch between 0 and 1 in the FEP calculation is instantaneous and the intermediate states along the path are 'eliminated.' However this is no longer true for a general NEW calculation, and the acceptance ratio method is not equivalent to overlap sampling. The GAR method provides no control of the systematic error due to finite sampling, since it lacks considerations of the phase space relations between the states along the path (down the funnel). However, the GAR method does offer an improvement over the conventional NEW methods (forward or reverse NEW or their direct average), when

the important values of W have been sampled. To analyze existing simulation data, we suggest that it is better to apply the GAR method if the only available work samples are $W_{0\rightarrow1}$ and $W_{1\rightarrow0}$, and therefore optimal OS calculations cannot be carried out. In this case, one should use at least the direct extension of the SOS. The appropriate SOS formula is obtained by replacing the potential energy change with work in (6.68)

$$\exp(-\beta\Delta A) = \frac{\langle\exp(-W_{0\rightarrow1}/2)\rangle_0}{\langle\exp(-W_{1\rightarrow0}/2)\rangle_1}. \tag{6.78}$$

An example of a NEW-OS calculation is shown in Fig. 6.11.

6.6.3 Umbrella Sampling and Weighted Histogram Analysis

Umbrella Sampling (US)

The US technique [45, 46] is often used to compute the potential of mean force (PMF), which is the free energy change along a chosen order parameter. US has been already discussed in Chap. 3, so here we remind the reader only of the main idea.

Umbrella sampling is aimed at overcoming the sampling problem by modifying the potential function so that different favorable states separated by energetic barriers are sufficiently sampled. An artificial umbrella potential is introduced:

$$U'(\mathbf{x}) = U(\mathbf{x}) + U_B(\mathbf{x}), \tag{6.79}$$
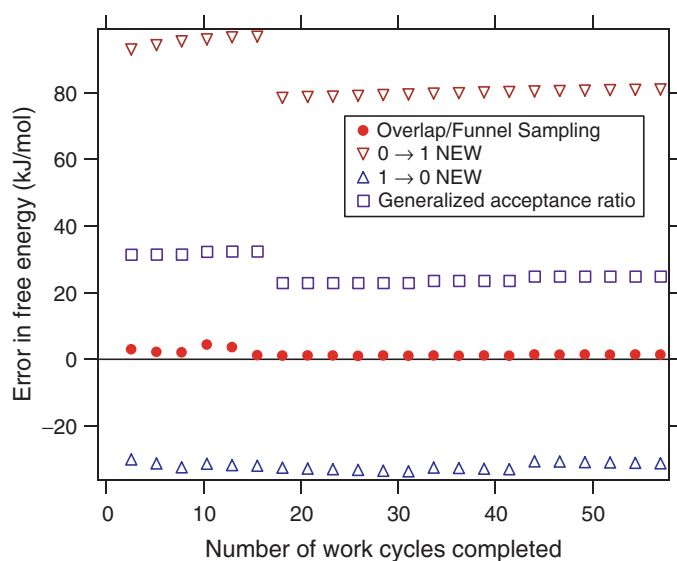
Fig. 6.11. The error in the free energy measured by several NEW implementations. Results are from Monte Carlo simulations of ion charging in water at 298 K. System 0 consists of a single Lennard-Jones atom with charge of +1e and 216 SPC water molecules, and system 1 is the same but with the charge turned off. One 'work cycle' contains 100 nonuniform steps in γ from 0 to 1 and back. For a detailed description of the simulation, see [43]

where $\mathbf{x}$ is the configurational space, U' is the umbrella potential, and U_B is the weighting function (also referred to as the biasing or perturbing potential). Then, the average of a quantity F becomes

$$\langle F \rangle = \frac{\langle F(\mathbf{x}) \exp(+\beta U_B(\mathbf{x})) \rangle_B}{\langle \exp(+\beta U_B(\mathbf{x})) \rangle_B}, \tag{6.80}$$

where the subscript B indicate that the system is sampled according to the umbrella potential.

For the calculation of the free energy difference ΔA between two states 0 and 1, the umbrella sampling utilizes the biasing potential W to ensuring sampling of phase space important to both 0 and 1 (cf. Fig. 6.1d). Then, F in (6.80) is taken as $\exp[-\beta(U_1 - U_0)]$.

Weighted Histogram Analysis

The weighted histogram analysis method (WHAM) is often used to combine the information of different simulations in an optimal way. The central idea and details of WHAM are given in Chap. 3 of this book. Here we only present a brief overview. Usually a number of simulations are performed at different values of the coupling parameter (6.8). An umbrella biasing potential U_{Bi} can be used for each of the simulations. At simulation i ($i = 1 \ldots n$) the unbiased probability distribution $p_i(\xi)$ can be reconstructed from the biased one, $p_i'(\xi)$ and represented as a normalized histogram obtained from the data collected in the umbrella sampling simulation

$$p_i(\xi) = \exp\{-\beta[f_i - U_{Bi}(\xi)]\}\, p_i'(\xi), \tag{6.81}$$

where f_i is the free energy difference between the biased state $U_{Bi}(\xi)$ and the reference state [without $U_{Bi}(\xi)$]. The goal is to combine all the $p_i(\xi)$ to construct an (unbiased) probability distribution $p_0(\xi)$, from which the free energy can be computed as a function of ξ. After minimizing the statistical error in the total probability distribution, i.e, $\partial \sigma^2(p_0)/\partial p_i = 0$ for all i, we have (in histogram form)

$$\begin{aligned} p_0(\xi) &= \sum_{i=1}^{n} \frac{N_i \exp[-\beta(U_{Bi} - f_i)]}{\displaystyle\sum_{j=1}^{n} N_j \exp[-\beta(U_{Bj} - f_j)]} p_i(\xi) \\ &= \sum_{i=1}^{n} \frac{N_i}{\displaystyle\sum_{j=1}^{n} N_j \exp[-\beta(U_{Bj} - f_j)]} p_i'(\xi), \end{aligned} \tag{6.82}$$

where n is the number of simulations and N_i is the number of samples (number of configurations collected) on simulation i.

In the simulations, the free energy parameters $\{f_i\}$ are unknown. However they can be obtained by solving (6.82) and the following equation self-consistently:

$$\exp(-\beta f_i) = \sum_{j=1}^{N_i} \frac{\exp\left(-\beta U_{Bi}\right)}{\sum_{k=1}^{n} N_k \exp\left[-\beta\left(U_{Bk} - f_k\right)\right]},$$ (6.83)

for all $i = 1 \ldots n$.

In multiple simulations, both overlap sampling and umbrella sampling + WHAM analysis can handle a wide range of problems that have different phase space relationships (also discussed in Sect. 6.3.1). In the umbrella + WHAM technique, the systematic error of the calculation is controlled by employing appropriate umbrella biasing potentials, and the statistical error is reduced using the WHAM analysis. However, the histogram analysis itself may produce both systematic and statistical errors that depend on the number of bins and the bin size of the histograms [57]. One drawback of umbrella sampling is that choosing the optimal weighting function may be difficult [7]. Since the simulation is performed according to the biased umbrella potential, an entirely new set of simulations needs to be conducted once the weighting function is found to be badly chosen and needs to be changed. In contrast, the OS technique offers full flexibility in optimizing the calculation without requiring the simulations to be rerun. In fact, multiple estimates of the free energy difference can be obtained from the same simulation data set, which provides an additional check of convergence of the calculated free energies [35].

6.7 Extrapolation Methods

6.7.1 Block Averaging Analysis

Straightforward application of the Jarzynski formula leads to a nonlinear estimate of relative free energy changes as the sample size increases. This is simply related to the relatively rare occurrence of tail events that dominate the distribution of work values. The jagged appearance of a graph of the running estimate of relative free energy versus number of work events makes it difficult to imagine extrapolating to a relative free energy estimate from a finite-sampled set of work values.

Thus, while the probability of a particular work value and the distribution of work values can give some estimate of the relative error in the free energy (for example by performing bootstrap or subsampling analysis over the full data set), there is no inherent way to extrapolate from the full finite data set to a larger (better converged) estimate.

The approach of block averaging provides some suggestions for how estimates of the free energy from NEW can be extrapolated with some estimates for the error in the extrapolation. The general idea is to transform the set of work values into a smooth curve by considering subsets of the data (blocks) and the subset estimates of relative free energy. By extrapolating along the smooth curves generated by this block averaging method, it becomes possible to imagine estimates of NEW that may be both efficient and labeled with an estimate of convergence quality [14, 15, 28, 58].

Block averaging comes from a set of N work values $\{W_1, W_1, \ldots, W_N\}$ generated in a NEW simulation. The free energy difference ΔA can be estimated by computing the block average with different block sizes n

$$\Delta A_n = \frac{n}{N} \sum_{j=1}^{N/n} -\beta^{-1} \ln \langle \exp(-\beta W) \rangle_{n,j}, \tag{6.84}$$

where the ratio N/n denotes the largest integer less than or equal to the literal fraction, and the individual block averages are defined by

$$\langle \exp(-\beta W) \rangle_{n,j} = \frac{1}{n} \sum_{i=(j-1)n+1}^{jn} \exp(-\beta W_i). \tag{6.85}$$

The result of this type of transformation is shown in Fig. 6.12. The jagged uneven curve is from the running estimate for the relative free energy from the Jarzynski formula. This running estimate would be very hard to use for extrapolation, due to the large changes in the estimate as the rare events are sampled (remember that the curve is based on the exponentially weighted distribution).

Note that the true free energy difference is $\Delta A = \Delta A_\infty$, and the other limit gives the average work, $\langle W \rangle = \Delta A_1$. The second law of thermodynamics indicates that $\Delta A_\infty = \Delta A \leq \langle W \rangle = \Delta A_1$. In fact, due to the monotonic behavior of the Boltzmann factor in (6.84), one has the general inequality

$$\Delta A_{n+1} \leq \Delta A_n. \tag{6.86}$$

The uncertainty in the finite-data free energy values, $\delta \Delta A$, can be estimated as twice the standard error of the mean, $\delta \Delta A = 2\sigma_n / \sqrt{N/n}$
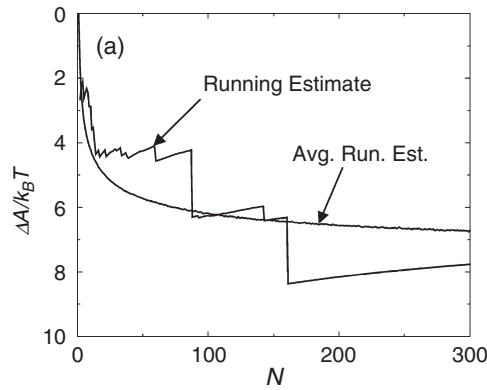
Fig. 6.12. Block averaging of an FEB or NEW calculation can be used to create a smoothed running profile. This contrasts with the jagged steps of a running average based on the same data. The running steps occur with the relatively rare sampling of events in the tails of the distribution that dominate the relative free energy estimate. The smoothed running curve uses the block procedure described in the text

$$\sigma_n^2 = \frac{1}{N/n} \sum_{j=1}^{N/n} (\Delta A_{n,j} - \langle \Delta A_n \rangle)^2, \tag{6.87}$$

which gives roughly a 90% confidence interval.

The inequality (6.86) suggests that the free energy difference ΔA may be obtained by extrapolation of ΔA_n to $n \to \infty$. While an extrapolation towards infinity is well defined in a general sense, an improved estimate may be obtained by consideration of an extrapolation in $1/n$ where the infinite limit is reached by moving to the left on the x-axis, rather than to the right. In the following, two extrapolation schemes based on these ideas are discussed.

6.7.2 Linear Extrapolation

Observations suggest that $\Delta A \sim n$ curves can be fit to a power law [15]:

$$\Delta A_n = \Delta A_\infty + \phi_1 (1/n)^{\tau_1}. \tag{6.88}$$

In a related form, ΔA_n is represented as a power series

$$\Delta A_n = \Delta A_\infty + \sum_{k=1}^{k_{\max}} \phi_k (1/n)^{k\tau_0}, \tag{6.89}$$

where the parameter τ_0 can be obtained by fitting to the data or in some other way such as examining the leading $1/n$ behavior. Empirical tests suggest that $\tau_0 = 0.266$ is a good starting point. The estimate of ΔA is given as $1/n^\tau \to 0$ (i.e., n approaches ∞); this corresponds to the intercept at $1/n^\tau = 0$ in a plot of $\Delta A \sim 1/n^\tau$.

This linear extrapolation scheme has one drawback: if the data do not include the leading $1/n$ behavior, the fits will not exhibit optimal extrapolative power. And the uncertainty of the extrapolated ΔA increases as $1/n^\tau$ approaches zero. Note that the
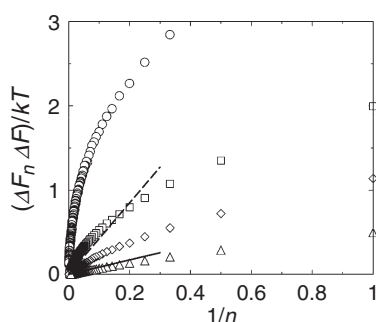
Fig. 6.13. Extrapolation to a free energy estimate based on block averages can best be analyzed in a $1/N$ plot. In this type of plot the large-N limit is towards the origin, rather than increasing to the right. By fitting the data to an extrapolating estimate of the expansion (see text) an estimate of the free energy difference can be made

Taylor expansion of (6.89) is not based on a physical model, and more-complicated nonlinear forms could be used. However, nonlinear extrapolation methods seem to offer little improvement over their linear counterparts [58]. Larger improvements appear to be achieved using the cumulative integral extrapolation method discussed in Sect 6.7.3.

6.7.3 Cumulative Integral Extrapolation

Since the linear and related expansion formulas depend on fits to regions of the curve that are statistically less and less reliable, it makes sense to find a measure for extrapolation that depends on the relative accuracy of the relative free energy estimate for all points along the curve. The cumulative integral extrapolation method is one approach to this idea.

The general idea of this approach is to consider the relative free energy to be a smooth function of ΔA_n from the inverse $\chi \equiv 1/n^\tau$ running from 0 to 1. At the zero limit this reflects infinite sampling and is thus the direction of extrapolation. Note that the value of τ that optimizes this transformation must be determined as part of the procedure.

In more detail, the most precise ΔA_n values are obtained for smaller n, or larger $\chi \equiv 1/n^\tau \approx 1$. The linear extrapolation scheme relies exclusively on small χ values. More precise free energy estimate can be obtained by including all values of χ during the extrapolation [58].

Now, consider treating ΔA_n as a smooth function $\Delta A_n(\chi)$, from $\chi = 0$ ($n = \infty$) to 1 ($n = 1$). The area under this function is given by the integral

$$\int_0^1 \mathrm{d}\chi \Delta A_n(\chi) = \int_0^1 \mathrm{d}\chi \frac{\mathrm{d}\Delta A_n(\chi)}{\mathrm{d}\chi} + \Delta A_n(\chi = 0), \qquad (6.90)$$

in which the second step involves integration by parts. Here, $\Delta A_n(\chi = 0)$ is the extrapolated free energy estimate, ΔA_{rci}. From (6.90) it follows that:

$$\Delta A_{rci} = \int_0^1 \mathrm{d}\chi \left[\Delta A_n(\chi) - (1 - \chi)\frac{\mathrm{d}\Delta A_n(\chi)}{\mathrm{d}\chi} \right]. \qquad (6.91)$$

The *cumulative integral* function can be defined by

$$CI(\chi) = \int_0^1 \mathrm{d}\chi' \left[\Delta A_n(\chi') - (1 - \chi')\frac{\mathrm{d}\Delta A_n(\chi')}{\mathrm{d}\chi'} \right], \qquad (6.92)$$

where the integral is performed in the reverse direction $\chi' = 1$ (where the data is most precise) to $\chi' = \chi$.

To obtain an extrapolated value for ΔA, consider the case where one has more than enough data to obtain ΔA exactly. In this situation, *for a carefully chosen τ*, $CI(\chi)$ will have nearly zero slope for small χ, since accumulating more χ values will not change the estimate. Thus, one can hope to extrapolate ΔA by simply finding a value of τ for which the slope $\mathrm{d}[CI(\chi)]/\mathrm{d}\chi$ is the smallest for all χ, then the extrapolated free energy ΔA_{rci} will be the value of $CI(\chi)$ for the smallest value of χ available, $\chi_{\min}$.

A fully automated procedure can be used to choose the value of τ for which the slope of the small-χ tail of $CI(\chi) \sim \chi$ is minimized. Then the free energy difference is estimated by $\Delta A_{rci} = CI(\chi)$. Simulation tests show that one can obtain ΔA estimates from the CI extrapolation using much less data than the direct Jarzynski estimates [58].

6.8 Concluding Remarks

In principle, methods for calculating free energies that are based on rigorous statistical mechanics equations, such as the FEP or the NEW method, which relies on Jarzynski's identity, are able to produce the *exact* relative free energies. However, in these methods significant weight is put on important but rare samples, and as a result, free energy estimates from both FEP and NEW are biased. In a practical simulation, in which only a finite sample of configurations is collected, the free energy estimate unavoidably suffers finite sampling error. These errors are usually reproducible, so that they may not be easily recognized by repeating simulations (gauging precision). Statistical methods, such as bootstrap analyses, are unable to recover the sampling bias error if the important samples are missing in the first place. Understanding the behavior of a free energy method in the finite sampling region becomes a key for the development of reliable and efficient free energy methods and their application to chemistry and biology.

In this chapter a detailed analysis of both the systematic and statistical errors due to finite sampling is presented. The concept of important phase space is introduced to understand characteristic features of FEP and NEW sampling processes. Sampling may be problematic because of energetic or entropic barriers (or both), depending on the overlap relationships of the important phase spaces for the system pairs under study, as well as due to the choice of perturbation directions. The energetic barrier is much more difficult to overcome in simulations. Then, it is more reliable to start the calculation from the system with the larger important phase space regions, so that important configurations to the free energy calculation can be well sampled (given sufficient but finite-length sampling). To ensure an accurate estimate of the free energy difference, one should choose appropriate perturbation directions, and set up appropriate intermediate stages or switching paths for the calculation.

Quantitative models characterizing the finite sampling systematic error and statistical error are developed using the perturbation probability distribution functions in both the forward and reverse directions. Without assuming any specific functional form for the distributions, these models effectively describe the FEP or NEW calculation behavior in the region of finite sample size. The sample size, the entropy difference between the two systems of interest, and the finite-time switching process (in NEW) are related to both the systematic and statistical errors of the free energy. It is revealed that it is the entropy, rather than the free energy itself, that plays a more important role in determining the reliability of free energy calculations. The analysis also leads to some helpful heuristics for estimating the reliability level of a simulation result, as well as the principles of optimal staging design in a multiple-stage FEP calculation, where the entropy difference between two systems is a key quantity.

Methods based on understanding of the behavior of the FEP and NEW methods are proposed to improve calculation efficiency and reliability. In particular, the overlap sampling technique offers a simple (no requirement for modifying the simulation) and effective (more reliable and efficient) way to improve the free energy estimate. In the overlap sampling, a conceptual intermediate is constructed inside the overlapped region of the important phase space of the two end systems (reference and target). Then, two separate FEP or NEW calculations are performed in opposite directions and sampling information from both of these is combined. Unlike the direct averaging method, which takes the direct average of the forward and reverse results as the free energy estimate, the overlap sampling is able to overcome the finite sampling error and produce reliable free energy results. Further, its built-in feature allows for optimization of the free energy calculation to minimize both statistical and systematic errors. For FEP, the optimal overlap sampling is identical to Bennett's acceptance ratio method. In the NEW case, however, it leads to the overlap funnel sampling procedure, in which the finite-time transformation of the system follows funnel-like paths from both end states to the common overlap sampling intermediate.

The bias in the large sampling region is characterized using the expansion of the free energy difference with respect to the sample size. The block averaging technique, in which the logarithm of the average of the exponential in Jarzynski's identity is computed with different block sizes, provides rich information about the free energy difference. In particular, the block-averaged free energy difference is a smooth, monotonically decreasing function of the block size, and can be fit to a power series. This further allows the extrapolation of the free energy difference to the limit of infinite sampling, i.e., where the free energy difference has its true value. Both the linear extrapolation technique and the improved reverse cumulative integral extrapolation are presented. Compared to the original Jarzynski formula, these techniques allow a better estimate of the free energy difference from a smaller set of sample data.

Finally, the material presented in this chapter paves the way for further improvements in free energy calculations. Two promising directions for future studies are: improving the methods for sampling phase space so that is satisfies the most effective overlap and/or subset relationships and developing better techniques for averaging samples and extrapolating from finite-sampled sets.

References

1. Leach, A.R., *Molecular Modelling, Principles and Applications*, Prentice Hall: London, 2001
2. Zwanzig, R.W., High-temperature equation of state by a perturbation method, *J. Chem. Phys.* **1954**, *22*, 1420–1426
3. Jarzynski, C., Equilibrium free-energy differences from nonequilibrium measurements: a master equation approach, *Phys. Rev. E* **1997**, *56*, 5018–5035
4. Jarzynski, C., Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **1997**, *78*, 2690–2693
5. Crooks, G.E., Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, *Phys. Rev. E* **1999**, *60*, 2721–2726

6. Crooks, G.E., Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, *J. Stat. Phys.* **1998**, *90*, 1481–1487

7. Frenkel, D., Smit, B., *Understanding Molecular Simulation: From Algorithms to Applications*, Academic: San Diego, 2002

8. Beveridge, D.L., DiCapua, F.M., Free energy via molecular simulation: applications to chemical and biomolecular systems, *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431–492

9. Kollman, P., Free energy calculations: applications to chemical and biochemical phenomena, *Chem. Rev.* **1993**, *32*, 2395–2417

10. Wang, W. et al., Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein–ligand, protein–protein, and protein–nucleic acid noncovalent interactions, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211–243

11. Hendrix, D.A., Jarzynski, C., A 'fast growth' method of computing free energy differences, *J. Chem. Phys.* **2001**, *114*, 5974–5981

12. Hummer, G., Fast-growth thermodynamic integration: error and efficiency analysis, *J. Chem. Phys.* **2001**, *114*, 7330–7337.

13. Hummer, G., Fast-growth thermodynamics integration: results for sodium ion hydration, *Mol. Simul.* **2002**, *28*, 81–90

14. Zuckerman, D.M., Woolf, T.B., Theory of a systematic computational error in free energy differences, *Phys. Rev. Lett.* **2002**, *89*

15. Zuckerman, D.M., Woolf, T.B., Overcoming finite sampling errors in fast-switching free-energy estimates: extrapolative analysis of a molecular system, *Chem. Phys. Lett.* **2002**, *351*, 445–453

16. Sun, S.X., Equilibrium free energies from path sampling of nonequilibrium trajectories, *J. Chem. Phys.* **2003**, *118*, 5769–5775

17. Gore, J., Ritort, F., Bustamante, C., Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 12564–12569

18. Park, S. et al., Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality, *J. Chem. Phys.* **2003**, *119*, 3559–3566

19. Liphardt, J. et al., Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, *Science* **2002**, *296*(5574), 1832–1835

20. Hummer, G., Szabo, A., Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3658–3661

21. Lavery, R. et al., Structure and mechanics of single biomolecules: experiment and simulation, *J. Phys. Condens. Matter* **2002**, *14*, R383–R414

22. Radmer, R.J., Kollman, P.A., Free energy calculation methods: a theoretical and empirical comparison of numerical errors and a new method for qualitative estimates of free energy changes, *J. Comput. Chem.* **1997**, *18*, 902–919

23. Lu, N.D., *Accuracy and precision of free-energy calculations via molecular simulation*. Department Chemical Engineering, University of Buffalo, State University of New York: Buffalo, NY, 2002

24. Lu, N.D., Kofke, D.A., Accuracy of free-energy perturbation calculations in molecular simulation. I. Modeling, *J. Chem. Phys.* **2001**, *114*, 7303–7311

25. Lu, N.D., Kofke, D.A., Accuracy of free-energy perturbation calculations in molecular simulation. II. Heuristics, *J. Chem. Phys.* **2001**, *115*, 6866–6875

26. Lu, N.D., Kofke, D.A., Optimal intermediates in staged free-energy calculations, *J. Chem. Phys.* **1999**, *111*, 4414–4423

27. Lu, N.D., Kofke, D.A., Adhikari, J., Variational formula for the free energy based on incomplete sampling in a molecular simulation, *Phys. Rev. E* **2003**, *68*, 026122

28. Zuckerman, D.M., Woolf, T.B., Systematic finite sampling inaccuracy in free energy differences and other nonlinear quantities, *J. Stat. Phys.* **2004**, *114*, 1303–1323

29. Bendat, J.S., Piersol, A.G., *Random Data: Analysis and Measurement Procedures*, Wiley: New York, 1971

30. Efron, B., Tibshirani, R.J., *An Introduction to the Bootstrap*, Chapman & Hall/CRC: New York, 1993

31. Kofke, D.A., Cummings, P.T., Quantitative comparison and optimization of methods for evaluating the chemical potential by molecular simulation, *Mol. Phys.* **1997**, *92*, 973–996

32. Kofke, D.A., Cummings, P.T., Precision and accuracy of staged free-energy perturbation methods for computing the chemical potential by molecular simulation, *Fluid Phase Equilibria* **1998**, *150*, 41–49

33. Wood, R.H., Muhlbauer, W.C.F., Thompson, P.T., Systematic errors in free energy perturbation calculations due to a finite sample of configuration space: sample-size hysteresis, *J. Phys. Chem.* **1991**, *95*, 6670–6675

34. Hummer, G., Calculation of free-energy differences from comptuer simulations of initial and final states, *J. Chem. Phys.* **1996**, *105*, 2004

35. Lu, N.D., Woolf, T.B., Kofke, D.A., Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods, *J. Comput. Chem.* **2004**, *25*, 28–39

36. Shing, K.S., Gubbins, K.E., The chemical potential in dense fluids and fluid mixtures via computer simulation, *Mol. Phys.* **1982**, *46*, 1109–1128

37. Allen, M.P., Simulation and phase diagrams. In: *Proceedings of the Euroconference on Computer simulation in Condensed Matter Physics and Chemistry*, Binder, K., Ciccotti, G., Eds. European Union, **1996**, pp. 255–284

38. Lu, N.D., Kofke, D.A., Simple model for insertion/deletion asymmetry of free-energy calculations. In: *Foundations of Molecular Modeling and Simulation*, Cummings, P., Westmoreland, P., Eds. AIChE Symposium Series, **2001**, pp. 146–149

39. Jorgensen, W.L., Ravimohan, C., Monte Carlo simulation of differences in free energies of hydration, *J. Chem. Phys.* **1985**, *83*, 3050–3054

40. Pearlman, D.A., A comparison of alternative approaches to free energy calculations, *J. Phys. Chem.* **1994**, *98*, 1487–1493

41. Pearlman, D.A., Govinda, R., Free energy calculations: methods and applications. In: *Encyclopedia of Computational Chemistry*, Schleyer, P., Ed. Wiley: Chichester, 1998

42. Pearlman, D.A., Kollman, P.A., A new method for carrying out free energy perturbation calculations: dynamically modified windows, *J. Chem. Phys.* **1989**, *90*, 2460–2470

43. Lu, N. et al., Using overlap and funnel sampling to obtain accurate free energies from nonequilibrium work measurements, *Phys. Rev. E* **2004**

44. Lu, N.D., Singh, J.K., Kofke, D.A., Appropriate methods to combine forward and reverse free energy perturbation averages, *J. Chem. Phys.* **2003**, *118*, 2977–2984

45. Torrie, G.M., Valleau, J.P., Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199

46. Valleau, J.P., Torrie, J.A., A guide for Monte Carlo for statistical mechanics. In: *Statistical Mechanics, Part A*, Berne, B., Ed. Plenum: New York, **1977**, pp. 169–194

47. Henchman, R.H., Essex, J.W., Free energies of hydration using restrained electrostatic potential derived charges via free energy perturbations and linear response, *J. Comput. Chem.* **1999**, *20*, 499–510

48. Wood, R.H., Estimation of errors in free energy calculations due to the lag between the Hamiltonian and the system configuration, *J. Phys. Chem.* **1991**, *95*, 4838–4842

49. Jorgensen, W.L. et al., Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water, *J. Chem. Phys.* **1988**, *89*, 3742–3746
50. Chipot, C. et al., Molecular dynamics free energy perturbation calculations: influence of nonbonded parameters on the free energy of hydration of charged and neutral species, *J. Phys. Chem.* **1994**, *98*, 11362–11372
51. Jacucci, G., Quirke, N., Monte-Carlo Calculation of the free-energy difference between hard and soft core diatomic liquids, *Mol. Phys.* **1980**, *40*, 1005–1009
52. Mezei, M., Test of overlap ratio metho on the calculation of the aqueous hydration free energy difference between acetone and dimethyl amine, *Mol. Phys.* **1988**, *65*, 219–223
53. Deitrick, G.L., Scriven, L.E., Davis, H.T., Efficient molecular simulation of chemical potentials, *J. Chem. Phys.* **1989**, *90*, 2370–2385
54. Lu, N.D., Woolf, T.B., Overlap perturbation methods for computing alchemical free energy changes: variants, generalizations and evaluations, *Mol. Phys.* **2004**, *102*, 173–181
55. Bennett, C.H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **1976**, *22*, 245–268
56. Crooks, G.E., Path-ensemble averages in systems driven far from equilibrium, *Phys. Rev. E* **2000**, *61*, 2361–2366
57. Kobrak, M.N., Systematic and statistical error in histogram-based free energy calculations, *J. Comput. Chem.* **2003**, *24*, 1437–1446
58. Ytreberg, F.M., Zuckerman, D.M., Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems, *J. Comput. Chem.* **2004**, *25*, 1749–1759

7

Transition Path Sampling and the Calculation of Free Energies

Christoph Dellago

7.1 Rare Events and Free Energy Landscapes

Many interesting physical, chemical, and biological processes occur on time scales that exceed those accessible by molecular dynamics simulation by orders of magnitude. Often, these long time scales are due to high free energy barriers (large compared to $k_\mathrm{B}T$) that the system must cross when moving between long-lived stable states. Examples for rare barrier-crossing events include chemical reactions, the nucleation of first-order phase transitions and the folding of proteins. One approach to treat such rare events in computer simulations is based on selecting a putative reaction coordinates and then determining the free energy as a function of that coordinate using the methods described in Chap. 4. By combining the free energy calculation with information obtained from dynamical trajectories initiated on a dividing surface separating the long-lived stable states, reaction rates can also be calculated [1–3].

To define a good reaction coordinate it is necessary to identify those degrees of freedom that capture the essential physics of the process. In complex systems, however, such information is often unavailable and this approach fails. In order to illustrate this problem in greater detail let us consider, as an example, a volume of pure liquid water carefully cooled below the freezing temperature. Although under such conditions the solid (ice) is the more stable phase, the system can remain liquid for hours or days even for strong undercooling. The reason for this behavior is that the phase transition from the liquid to the solid proceeds through the formation of a small crystalline nucleus somewhere in the liquid. The nucleus can then grow and the crystalline region eventually encompasses the whole sample. While overall the free energy of the system decreases during the transition, the initial stages of the freezing process are free-energetically uphill. This free-energetic cost is associated with the formation of an interface between the solid and the liquid.

Within classical nucleation theory the free energy of the system as a function of the radius r of a spherical crystallite can be expressed as

$$A(r) = 4\pi r^2 \gamma + (4/3)\pi r^3 \rho_s \Delta\mu. \tag{7.1}$$

Here, γ is the surface tension of the solid–liquid interface, ρ_s is the particle number density of the solid and $\Delta\mu = \mu_s - \mu_l$ is the difference in chemical potential between the solid and the liquid phase. The first term on the right-hand side of the above equation is the surface free energy of the nucleus and it is proportional to the surface area of the crystallization nucleus. The second term is proportional to the volume of the nucleus and, since the chemical potential of the solid is lower than that of the liquid, gives a negative contribution to the total free energy $A(r)$. While for growing crystallite size r the volume term will eventually prevail, for small sizes the free energy $A(r)$ is dominated by the surface free energy. As a consequence, the free energy as a function of r increases initially. Then, it reaches a maximum at a certain critical size r^* and rapidly decreases after that. It is this free energy barrier that prevents the metastable liquid from rapidly crystallizing and allows supercooled water to stay liquid for long times. As recently shown by Moroni, ten Wolde, and Bolhuis [4], however, the size of the critical nucleus is not sufficient to characterize critical nuclei for the crystallization of a liquid. In addition to the size, cluster shape and structure play an important role in the freezing process. Such a failure to include important degrees of freedom in the definition of the reaction coordinate can lead to problems illustrated in Fig. 7.1.

The *transition path sampling* method, developed by Chandler and coworkers [5] building on earlier work of Pratt [6], is a computer simulation technique designed to overcome this kind of problem. Based on a statistical, reaction coordinate-free description of pathways connecting long-lived stable states, it focuses on those segments of the time evolution of the system on which the rare but important barrier-crossing event actually occurs. To do this a priori knowledge of the reaction mechanism in the form of a reaction coordinate is not required. Rather, it is sufficient to describe the initial and final state of the system. In a transition path sampling simulation the ensemble of transition pathways is sampled with a Monte Carlo procedure. As a result one obtains a set of dynamical pathways which can then be further analyzed to obtain information about the reaction mechanism (or mechanisms).

In the next two sections we will briefly review the theoretical foundation of the transition path sampling method and explain its practical application. For a detailed description of the transition path sampling approach we refer the reader to recent review articles [7–9]. After these introductory sections we will discuss how transition path sampling can be used for the calculation of the free energy as a function of a predefined reaction coordinate. In a transition path sampling simulation one considers only trajectories connecting certain regions in configuration space. Due to the bias introduced by this requirement, configurations on transition pathways are not distributed according to the equilibrium distribution of the system. Rather, configurations with low weight in the equilibrium ensemble may have a much larger weight in the transition path ensemble if they belong to regions in configuration space that must be traversed by the system as it moves from the initial to the final region. Due to this bias conventional free energy profiles as a function of some variable can not be calculated in a straightforward way from pathways sampled according the transition path ensemble. Under certain circumstances, however, one can correct for this bias and use transition path sampling methodologies to calculate free energy profiles
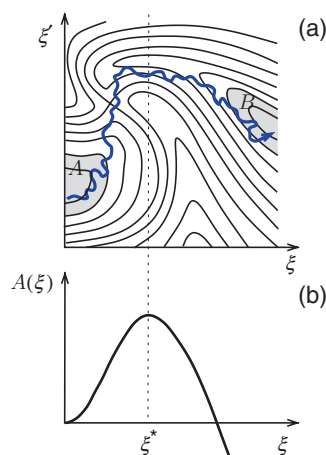
Fig. 7.1. (a) Hypothetical free energy landscape $A(\xi, \xi')$ as a function of two selected degrees of freedom ξ and ξ'. Such a free energy surface might result, for instance, for a system with a first-order phase transition. The variable ξ could then be the size of a cluster of the stable phase forming in the metastable phase and ξ' could be some other, important degree of freedom. Due to a rare fluctuation a system initially in the metastable state $\mathscr{A}$ can overcome the nucleation barrier following a pathway (thick solid line) that crosses the transition state corresponding to the critical nucleus. After passing the transition state the system then relaxes in the stable state $\mathscr{B}$. Although the cluster size ξ can be used to tell wether the system is in state $\mathscr{A}$ or $\mathscr{B}$ it fails to capture all important features of the transition because during the transition system-atic motion along ξ' must occur. **(b)** The failure of ξ to include the essential physics of the transition becomes apparent when we determine the free energy of the system as a function of the variable ξ only, $A(\xi) = \int d\xi' \exp\{-\beta A(\xi, \xi')\}$. This function, shown in the lower panel of the figure, displays a barrier separating the ξ-values corresponding to the region $\mathscr{A}$ from those from region $\mathscr{B}$. The top of the barrier, located at ξ^*, does, however, not coincide with the transition region. Rather, configurations with $\xi = \xi^*$ (distributed along the *dotted line* in the upper panel) will most likely belong to the basins of attraction of regions $\mathscr{A}$ or $\mathscr{B}$

along a given reaction coordinate [10]. Of course, it is also trivially possible to gen-erate an equilibrium distribution in configuration space by applying path sampling techniques to a path probability in which the requirement that paths start and end in certain regions has been removed. In this case free energies can be simply calculated from configurations lying on the sampled pathways [11]. For such an unconstrained ensemble of pathways additional precautions must be taken to guarantee that the rare event of interest occurs at least on some of the trajectories collected in the simulation. These issues are discussed in Sect. 7.4.

Transition path sampling can also be helpful in the calculation of free ener-gies in the context of fast-switching methods described in Chap. 5. As shown by Jarzynski [12], equilibrium free energies can be computed from the work performed on a system in repeated transformations carried out arbitrarily far from equilib-rium. From a computational point of view, this remarkable theorem is attractive because it promises efficient free energy calculations due to the reduced cost of

short nonequilibrium trajectories. A straightforward application of the Jarzynski theorem in the fast-switching regime, however, is plagued by severe statistical problems. These difficulties stem from the fact that trajectories with important work values are generated only very rarely, causing averages to converge very slowly. Transition path sampling techniques can be used to alleviate this difficulty [13–16]. Whether such a combination of the fast-switching approach with the transition path sampling methodology can be superior to conventional free energy calculation methods such as umbrella sampling or thermodynamic integration is still being debated [16]. The application of transition path sampling to the biased sampling of fast-switching trajectories is discussed in Sect. 7.5.

In Sect. 7.6 we then describe how reaction rate constants can be calculated within the transition path sampling approach. Reaction rate constants can be extracted from time correlation functions of the population functions associated with the initial and final regions of the reaction [17]. The mathematical expressions of such time correlation functions are isomorphic to the equations for equilibrium free energy differences. Thus, free energy calculation methods such as thermodynamic integration [18] (see Chap. 4), umbrella sampling [17] (see Chap. 3) and even novel fast-switching approaches [19] (see Chap. 5) can be adapted to calculate 'free energies' in trajectory space from which kinetic coefficients can then be determined. Such a generalized free energy can be viewed as the reversible work necessary to transform an ensemble of nonreactive trajectories into an ensemble of reactive ones.

7.2 Transition Path Ensemble

The basis of the transition path sampling method is the statistical description of dynamical pathways in terms of a probability distribution. To define such a distribution consider a molecular system evolving in time and imagine that we take snapshots of this system at regularly spaced times t_i separated by the time step Δt. Each of these snapshots, or states, consists of a complete description z of the system in terms of the positions $\mathbf{q} = \{q_1, q_2, \cdots, q_N\}$ and momenta $\mathbf{p} = \{p_1, p_2, \cdots, p_N\}$ of all N atoms in the system, $z = \{\mathbf{q}, \mathbf{p}\}$. If we follow the system for a total time $\mathscr{T}$ we obtain an ordered sequence of $L = \mathscr{T}/\Delta t + 1$ states

$$z(\mathscr{T}) \equiv \{z_0, z_{\Delta t}, z_{2\Delta t}, \ldots, z_{\mathscr{T}}\}. \tag{7.2}$$

This sequence of states is a discrete representation of the continuous dynamical trajectory starting from z_0 at time $t = 0$ and ending at $z_{\mathscr{T}}$ at time $t = \mathscr{T}$. Such a discrete trajectory may, for instance, result from a molecular dynamics simulation, in which the equations of motion of the system are integrated in small time steps. A trajectory can also be viewed as a high-dimensional object whose description includes time as an additional variable. Accordingly, the discrete states on a trajectory are also called *time slices*.

The probability to observe a particular sequence of states depends on the distribution of the initial conditions and the dynamical rule describing the time evolution

of the system. If the dynamics is Markovian,[1] i.e., if the probability to move from z_t to $z_{t+\Delta t}$ after one time step Δt depends only on z_t and not on the history of the system prior to t, the total path probability can be written as the product of the single-time-step transition probabilities $p(z_t \rightarrow z_{t+\Delta t})$,

$$\mathscr{P}[z(\mathscr{T})] = \rho(z_0) \prod_{i=0}^{\mathscr{T}/\Delta t - 1} p(z_{i\Delta t} \rightarrow z_{(i+1)\Delta t}). \tag{7.3}$$

The first factor on the right-hand side of the above equation, $\rho(z_0)$, is the distribution of initial conditions z_0, which, in many cases, will just be the equilibrium distribution of the system. For a system at constant volume in contact with a heat bath at temperature T, for instance, the equilibrium distribution is the canonical one

$$\rho(z) = \exp\{-\beta\mathscr{H}(z)\}/Z \tag{7.4}$$

and

$$Z = \int \mathrm{d}z \exp\{-\beta\mathscr{H}(z)\} \tag{7.5}$$

is the partition function normalizing the canonical distribution. The transition probability $p(z_t \rightarrow z_{t+\Delta t})$ is simply the conditional probability to find the system in an infinitesimal region around $z_{t+\Delta t}$ at time $t + \Delta t$ provided the system was in state z_t a short time Δt earlier.

The specific form of the short-time transition probability depends on the type of dynamics one uses to describe the time evolution of the system. For instance, consider a single, one-dimensional particle with mass m evolving in an external potential energy $V(q)$ according to a Langevin equation in the high-friction limit

$$m\gamma\dot{q} = -\frac{\partial V(q)}{\partial q} + \mathscr{R}. \tag{7.6}$$

Here, γ is the friction coefficient and $\mathscr{R}$ is a Gaussian random force uncorrelated in time satisfying the fluctuation dissipation theorem, $\langle\mathscr{R}(0)\mathscr{R}(t)\rangle = 2m\gamma k_{\mathrm{B}}T\delta(t)$ [21], where $\delta(t)$ is the Dirac delta function. The random force is thought to stem from fast and uncorrelated collisions of the particle with solvent atoms. The above equation of motion, often used to describe the dynamics of particles immersed in a solvent, can be solved numerically in small time steps, a procedure called *Brownian dynamics* [22]. Each Brownian dynamics step consists of a deterministic part depending on the force derived from the potential energy and a random displacement δq_R caused by the integrated effect of the random force

$$q(t + \Delta t) = q(t) - \frac{\Delta t}{\gamma m}\frac{\partial V}{\partial q} + \delta q_R. \tag{7.7}$$

The random displacement δq_R is a Gaussian random variable with zero mean and variance

---

[1] See, for instance [20] for a detailed discussion of the Markov property and of Markov processes.

$$\sigma^2 = \frac{2k_{\mathrm{B}}T}{m\gamma}\Delta t. \tag{7.8}$$

From the statistics of this random displacement, a Gaussian short-time transition probability follows:

$$p(q_t \to q_{t+\Delta t}) = \frac{1}{\sqrt{2\pi\sigma_R^2}}\exp\left\{-\frac{\left[q(t+\Delta t) - q(t) + \frac{\Delta t}{\gamma m}\frac{\partial V}{\partial q}\right]^2}{2\sigma_R^2}\right\}. \tag{7.9}$$

The probability of a complete Brownian path is then obtained as the product of such single-time-step transition probabilities. For other types of dynamics, such as Newtonian dynamics, Monte Carlo dynamics or general Langevin dynamics, other appropriate short-time-step transition probabilities need to be used [5, 8].

In the transition path sampling method we are interested in trajectories that start in a certain region of configuration space, which we will call region $\mathscr{A}$, and end in another region, $\mathscr{B}$. We call such trajectories *reactive*. Accordingly, we restrict the probability distribution from (7.3) to reactive trajectories only (see Fig. 7.2)

$$\mathscr{P}_{\mathscr{A}\mathscr{B}}[z(\mathscr{T})] \equiv h_{\mathscr{A}}(z_0)\mathscr{P}[z(\mathscr{T})]h_{\mathscr{B}}(z_{\mathscr{T}})/Z_{\mathscr{A}\mathscr{B}}(\mathscr{T}). \tag{7.10}$$

Here, $h_{\mathscr{A}}(z)$ is the characteristic function for the reactant region $\mathscr{A}$, equal to unity if state z is in this region and that vanishes otherwise. The characteristic function $h_{\mathscr{B}}(z)$ for the product region $\mathscr{B}$ is defined similarly. The factor

$$Z_{\mathscr{A}\mathscr{B}}(\mathscr{T}) \equiv \int \mathscr{D}z(\mathscr{T})\, h_{\mathscr{A}}(z_0)\mathscr{P}[z(\mathscr{T})]h_{\mathscr{B}}(z_{\mathscr{T}}) \tag{7.11}$$

normalizes the path distribution. The notation $\int \mathscr{D}z(\mathscr{T})$, familiar from path integral theory, implies a summation over all pathways $z(\mathscr{T})$ of length $\mathscr{T}$. For the discrete pathways considered in the transition path sampling method this integration extends over all variables z_i at all times t_i with $i = 0, \ldots, L$. Through the multiplication of the unrestricted path probability $\mathscr{P}[z(\mathscr{T})]$ by the characteristic functions of regions $\mathscr{A}$ and $\mathscr{B}$ in (7.10) all paths either not starting in $\mathscr{A}$ or not ending in $\mathscr{B}$ (or both) are assigned a vanishing weight. The ensemble of all reactive pathways statistically described by the distribution (7.10) is called the *transition path ensemble*.
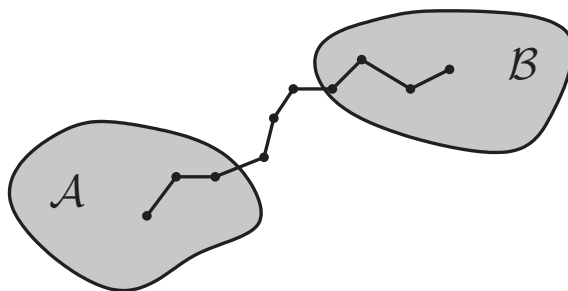
Fig. 7.2. Transition pathway connecting the stable regions $\mathscr{A}$ and $\mathscr{B}$

7.3 Sampling the Transition Path Ensemble

The path ensemble defined in (7.10) is a statistical description of *all* dynamical pathways connecting regions $\mathscr{A}$ and $\mathscr{B}$ whithin time $\mathscr{T}$. Our goal now is to generate trajectories according to this weight function, or, in other words, to *sample* the transition path ensemble. We can achieve this by carrying out a Monte Carlo procedure acting on entire trajectories.

7.3.1 Monte Carlo Sampling in Path Space

Monte Carlo sampling of trajectories works in a way analogous to the Monte Carlo simulation of, say, an atomic liquid [22, 23]. In that case the simulation proceeds by repetition of a basic step. In each step, a new configuration is first generated from an old one, for instance by displacing a randomly selected particle by a small random amount. Then, the new configuration is either accepted or rejected according to how the probability of the new configuration compares to that of the old one. Most often in Monte Carlo simulations a convenient acceptance criterion is derived from the detailed balance condition satisfied by the Metropolis rule [24]. By repeated application of this acceptance/rejection procedure the system performs a random walk through the configuration space, visiting configurations with a frequency proportional to their statistical weight under the particular conditions (density, temperature, etc.) considered in the simulation.

The Metropolis Monte Carlo algorithm can be applied very generally to any kind of probability distribution function. In particular, we can use Metropolis Monte Carlo to sample the transition path ensemble, in which case the random walk is carried out in the space of trajectories instead of configuration space. The general procedure, however, remains the same. In each path sampling Monte Carlo step, a new trajectory is generated from an old one. This new trajectory is then accepted or rejected according to how the statistical weight of the new trajectory compares to that of the old one. If the new trajectory is not connecting $\mathscr{A}$ with $\mathscr{B}$, it has a weight of zero in the transition path ensemble and it must be rejected. But if the new trajectory is reactive, it is accepted with a certain nonvanishing probability. By repeating this basic Monte Carlo step over and over again, one carries out a random walk in path space visiting trajectories with a likelihood proportional to their weight in the transition path ensemble. Note that, while pathways are sampled with a Monte Carlo procedure, the trajectories are fully dynamical pathways generated according to the rules of the underlying dynamics. The pathways collected in this way can then be further analyzed to find the reaction mechanism.

Just as in a conventional Monte Carlo simulation, correct sampling of the transition path ensemble is enforced by requiring that the algorithm obeys the *detailed balance* condition. More specifically, the probability $\pi[z^{(o)}(\mathscr{T}) \rightarrow z^{(n)}(\mathscr{T})]$[2] to move from an 'old' path $z^{(o)}(\mathscr{T})$ to a new one $z^{(n)}(\mathscr{T})$ in a Monte Carlo step must be exactly balanced by the probability of the reverse move from $z^{(n)}(\mathscr{T})$ to $z^{(o)}(\mathscr{T})$

---

[2] In the theory of Markov chains the probability $\pi[z^{(o)}(\mathscr{T}) \rightarrow z^{(n)}(\mathscr{T})]$ is called the *transition matrix* [20].

$$\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{o})}(\mathscr{T})]\pi[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] =$$
$$\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{n})}(\mathscr{T})]\pi[z^{(\mathrm{n})}(\mathscr{T}) \to z^{(\mathrm{o})}(\mathscr{T})]. \qquad (7.12)$$

This detailed balance condition makes sure that the path ensemble $\mathscr{P}_{\mathscr{A}\mathscr{B}}[z(\mathscr{T})]$ is stationary under the action of the Monte Carlo procedure and that therefore the correct path distribution is sampled [23, 25]. The specific form of the transition matrix $\pi[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})]$ depends on how the Monte Carlo procedure is carried out. In general, each Monte Carlo step consists of two stages: in the first stage a new path $z^{(\mathrm{n})}(\mathscr{T})$ is generated from an old one with a certain generation probability $P_{\mathrm{gen}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})]$. For simplicity, this so-called *trial move* is often carried out such that the generation probability is symmetric, i.e., such that the probability to generate $z^{(\mathrm{n})}(\mathscr{T})$ from $z^{(\mathrm{o})}(\mathscr{T})$ equals the probability to generate $z^{(\mathrm{o})}(\mathscr{T})$ from $z^{(\mathrm{n})}(\mathscr{T})$

$$P_{\mathrm{gen}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] = P_{\mathrm{gen}}[z^{(\mathrm{n})}(\mathscr{T}) \to z^{(\mathrm{o})}(\mathscr{T})]. \qquad (7.13)$$

In the second stage of each Monte Carlo step the new (or *trial*) pathway is accepted with a certain *acceptance probability* $P_{\mathrm{acc}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})]$. The total probability $\pi[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})]$ to move from $z^{(\mathrm{o})}(\mathscr{T})$ to $z^{(\mathrm{n})}(\mathscr{T})$ in a Monte Carlo step is the product of the generation and the acceptance probability

$$\pi[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] =$$
$$P_{\mathrm{gen}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] \times P_{\mathrm{acc}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})]. \quad (7.14)$$

The detailed balance condition (7.12) can now be satisfied by selecting an appropriate acceptance probability. By inserting the product in (7.14) into the detailed balance condition we find that for a symmetric generation probability the acceptance probabilities for the forward and the reverse move must be related by

$$\frac{P_{\mathrm{acc}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})]}{P_{\mathrm{acc}}[z^{(\mathrm{n})}(\mathscr{T}) \to z^{(\mathrm{o})}(\mathscr{T})]} = \frac{\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{n})}(\mathscr{T})]}{\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{o})}(\mathscr{T})]}. \qquad (7.15)$$

This condition can be satisfied using the Metropolis rule [24]

$$P_{\mathrm{acc}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] = \min\left\{1, \frac{\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{n})}(\mathscr{T})]}{\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{o})}(\mathscr{T})]}\right\}. \qquad (7.16)$$

Here, the minimum function returns the smaller of the arguments. According to this rule a new path is accepted, if a random number chosen from a uniform distribution in the interval $[0, 1]$ is smaller than $\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{n})}(\mathscr{T})]/\mathscr{P}_{\mathscr{A}\mathscr{B}}[z^{(\mathrm{o})}(\mathscr{T})]$. To implement the Monte Carlo procedure described above one must specify how to create a new pathway from an old one. How to do that is the subject of Sect. 7.3.2.

7.3.2 Shooting and Shifting

The efficiency of a transition path sampling simulation crucially depends on how new pathways are generated from old ones. Various schemes to do that are possible.

Here, we will explain the most effective of them, the so-called *shooting algorithm*. For simplicity we will focus on how to do it for deterministic trajectories such as those obtained from a molecular dynamics simulation. We note, however, that very similar algorithms can be applied to stochastic trajectories [5, 7, 8].

Consider a molecular system consisting of N atoms with Hamiltonian $\mathscr{H}(\mathbf{q}, \mathbf{p})$ = $K(\mathbf{p}) + V(\mathbf{q})$. $K(\mathbf{p})$ and $V(\mathbf{q})$ are the kinetic and potential energy, respectively. Since the system evolves according to Hamilton's equations of motion

$$\dot{q}_i = \frac{\partial \mathscr{H}}{\partial p_i}, \tag{7.17}$$

$$\dot{p}_i = -\frac{\partial \mathscr{H}}{\partial q_i}, \tag{7.18}$$

the time evolution of the system is *deterministic*, i.e., an entire trajectory $z(\mathscr{T})$ is completely determined by its initial condition z_0 (or any other point on the trajectory). We can therefore write the state of the system z_t at time t as a unique function of the state of the system at time 0

$$z_t = \phi_t(z_0). \tag{7.19}$$

Here, the function $\phi_t(z)$ is also called the propagator of the system. Applying the propagator ϕ_t to a given state z yields the state of the system at a time t later. Although the specific form of the propagator is in general not known analytically, it is possible to obtain good approximations for it by numerical integration of the equations of motion [22].

For deterministic dynamics the state $z_{t+\Delta t}$ at time $t + \Delta t$ is of course completely determined by the state of the system z_t a time step Δt earlier. Therefore, the single-time-step transition probability $p(z_t \rightarrow z_{t+\Delta t})$ can be written in terms of a delta function

$$p(z_t \rightarrow z_{t+\Delta t}) = \delta[z_{t+\Delta t} - \phi_{\Delta t}(z_t)]. \tag{7.20}$$

Then the path probability from (7.3) consists of a product of such delta functions. Due to the singular nature of such a path probability it is more convenient to view the entire deterministic trajectory as represented by its initial state z_0. In this case the transition path ensemble from (7.10) reduces to a distribution of initial conditions z_0 yielding pathways connecting $\mathscr{A}$ with $\mathscr{B}$

$$\mathscr{P}_{\mathscr{A}\mathscr{B}}(z_0) \equiv h_{\mathscr{A}}(z_0)\rho(z_0)h_{\mathscr{B}}(z_{\mathscr{T}})/Z_{\mathscr{A}\mathscr{B}}(\mathscr{T}), \tag{7.21}$$

where $z_{\mathscr{T}} = \phi_{\mathscr{T}}(z_0)$. The normalizing factor $Z_{\mathscr{A}\mathscr{B}}(\mathscr{T})$ is now simply obtained by integration over phase space,

$$Z_{\mathscr{A}\mathscr{B}}(\mathscr{T}) \equiv \int \mathrm{d}z_0 \, h_{\mathscr{A}}(z_0)\rho(z_0)h_{\mathscr{B}}(z_{\mathscr{T}}). \tag{7.22}$$

Thus, the transition path ensemble is now represented by a distribution of initial conditions in phase space (we have, in effect, 'integrated out' all path variables except

those belonging to the initial state z_0). For this ensemble of initial conditions the Monte Carlo procedure described above is (for a symmetric generation probability)

$$P_{\mathrm{acc}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] = h_{\mathscr{A}}(z_0^{(\mathrm{n})}) h_{\mathscr{B}}(z_{\mathscr{T}}^{(\mathrm{n})}) \min \left\{ 1, \frac{\rho[z_0^{(\mathrm{n})}]}{\rho[z_0^{(\mathrm{o})}]} \right\}, \qquad (7.23)$$

where $z_0^{(\mathrm{o})}$ is the initial point of the old path and $z_0^{(\mathrm{n})}$ is that of the new one. Clearly, pathways either not beginning in $\mathscr{A}$ or not ending in $\mathscr{B}$ have a vanishing acceptance probability and are always rejected. A new pathway connecting $\mathscr{A}$ and $\mathscr{B}$, on the other hand, is accepted with a probability depending solely on the relative weight of the initial conditions of the old and the new path.

So far we have developed a general Monte Carlo procedure that can be used to sample a given path distribution such as the transition path ensemble (7.10). To implement this procedure we must specify how a new pathway is created from an old one. The efficiency of a transition path sampling simulation will crucially depend on the particular way such a trial move is carried out. To obtain high efficiency one needs to make sure that a newly generated pathway is as different as possible from the old path. At the same time, it is important that the new pathway has a sufficient probability to be accepted. These two requirements often oppose each other: an algorithm in which the trial path is drastically different from the old path may be very inefficient due to an extremely low acceptance probability. On the other hand, algorithms with a high acceptance probability can be inefficient because the trial path differs only slightly from the old path and so diffusion through path space is slow. In the *shooting* and *shifting* procedures these two aspects are balanced by using the propagation rules of the underlying dynamics to generate a new trajectory from an old one. In the following we will explain how shooting and shifting moves can be carried out for Hamiltonian systems, but we stress that very similar procedures can be developed for other types of dynamics. For simplicity we will explain only the general idea. For further details the reader is referred to [8].

In a shooting move a trial path is generated from an old path as follows. First, one of the L states of the old path $z^{(\mathrm{o})}(\mathscr{T})$ is selected at random with a uniform probability, i.e., all states on the path have the same probability to be selected.[3] The selected state $z_t^{(\mathrm{o})}$ consists of the positions $\mathbf{q}_t^{(\mathrm{o})}$ and momenta $\mathbf{p}_t^{(\mathrm{o})}$ of all particles. Next, new momenta are obtained by adding a small perturbation $\delta\mathbf{p}$ to the old momenta

$$\mathbf{p}_t^{(\mathrm{n})} = \mathbf{p}_t^{(\mathrm{o})} + \delta\mathbf{p}. \qquad (7.24)$$

The momentum displacements $\delta\mathbf{p}$ are, for instance, selected from a Gaussian distribution whose width can be used to tune the acceptance probability. Note that it is not strictly necessary to draw the momentum displacements from a Gaussian distribution; any other isotropic distribution can be used as well for this purpose. Starting from the new state $z_t^{(\mathrm{n})}$, which consists of the old positions and the new momenta,

---

[3] It is, of course, possible to bias the selection toward certain segments of the path. In this case, the bias must be properly taken into account in the acceptance probability.

the equations of motion are integrated forward to time $\mathscr{T}$ and backward to time 0. For the integration of the equations of motion any suitable molecular dynamics algorithm can be used [22, 23]. Due to the added perturbation the new trajectory $z^{(n)}(\mathscr{T})$ will differ from the old trajectory $z^{(o)}(\mathscr{T})$. By how much depends on the amplitude of the momentum displacement and the chaoticity of the underlying dynamics.

The new path must now be accepted or rejected. According to (7.23) the new path can be accepted only if its initial state $z_0^{(n)}$ lies in region $\mathscr{A}$ and its final state $z_{\mathscr{T}}^{(n)}$ lies in region $\mathscr{B}$. If this is not the case, i.e., if the newly generated trajectory does not connect $\mathscr{A}$ and $\mathscr{B}$, the acceptance probability vanishes and the new path is rejected. If, on the other hand, the trajectory is reactive, it is accepted with probability $\min\{1, \rho[z_0^{(n)}]/\rho[z_0^{(o)}]\}$. Thus, for a canonical distribution of initial conditions the acceptance probability depends only on the energy difference of the initial states of the old and the new trajectory. Some computing time can be saved by first integrating the equations of motion backward to time $t = 0$ and then testing whether the new initial point is in $\mathscr{A}$. As pathways not starting in A are rejected, the forward segment of the trajectory needs to be computed only if the new initial point is in region $\mathscr{A}$.

For a microcanonical distribution of initial conditions all initial conditions have the same energy and the acceptance probability becomes

$$P_{\mathrm{acc}}[z^{(o)}(\mathscr{T}) \to z^{(n)}(\mathscr{T})] = h_{\mathscr{A}}[z_0^{(n)}]h_{\mathscr{B}}[z_{\mathscr{T}}^{(n)}]. \tag{7.25}$$

Thus, every trial trajectory that connects $\mathscr{A}$ and $\mathscr{B}$ is accepted. In the microcanonical case momentum displacements $\delta\mathbf{p}$ must be selected such that the system does not change its total energy. How to do that and how to take other constraints such as conserved total linear and angular momentum into account is explained in detail in [8].

Shooting moves can be complemented with shifting moves. In this computationally inexpensive move a new path is generated from an old one by translating the trajectory forward or backward in time. More specifically, in a forward shifting move a trajectory segment of length δt is first removed from the beginning of the path. The time length δt for this operation is selected from a random distribution (note that δt needs to be a multiple of the step size Δt). Then, a segment of the same length is grown at the end of the path by integrating the equations of motion for $\delta t/\Delta t$ time steps starting from the final point of the old path. As a result of this procedure the new path overlaps with the old one over a length $\mathscr{T} - \delta t$ and the initial point of the new path is given by $z_0^{(n)} = z_{\delta t}^{(o)}$. Of course, the segment at the end of the path needs to be grown only if $z_0^{(n)}$ lies inside $\mathscr{A}$. Otherwise the new pathway is rejected without any need to integrate the equation of motion. For a backward-shifting move one proceeds in an analogous way. First one removes a trajectory segment of length δt from the end of the old path and then regrows a new path segment of the same length by integrating the equations of motion backward in time starting from the initial point of the old path. Similar to the forward shift, one needs to grow the backward segment only if the endpoint of the new path is inside $\mathscr{B}$. The shifting procedure effectively

translates the path in time in a way reminiscent of the *reptation* motion of a polymer in a dense melt [26].

For Hamiltonian dynamics with a canonical or microcanonical distribution of initial conditions the acceptance probability for pathways generated with the shifting algorithm is particularly simple. Provided forward and backward shifting moves are carried out with the same probability the acceptance probability from (7.23) reduces to

$$P_{\text{acc}}[z^{(\text{o})}(\mathscr{T}) \to z^{(\text{n})}(\mathscr{T})] = h_{\mathscr{A}}[z_0^{(\text{n})}]h_{\mathscr{B}}[z_t^{(\text{n})}], \qquad (7.26)$$

implying that any new path that still connects regions $\mathscr{A}$ and $\mathscr{B}$ is accepted. Similarly simple acceptance probabilities can also be derived for pathways generated with stochastic instead of deterministic dynamics [8].

Although new pathways generated by the shifting algorithm differ little from the corresponding old pathways, especially in the transition region, shifting moves can improve the convergence of averages taken over the transition path ensemble. Since shifting moves just translate a trajectory forward and backward in time, they must be combined with other types of moves such as shooting moves to achieve an ergodic sampling of the transition path ensemble.

In addition to the shooting and shifting algorithms, other path moves have been devised for stochastic dynamics [8]. For instance, in the *local algorithm* a new pathway is generated by modifying one single time slice of the old path [5]. In the Crooks–Chandler algorithm [27] global path moves are carried out by local changes in the sequence of random numbers used to generate the pathways. The Crooks–Chandler algorithms is particularly useful for the simulation of rare fluctuations in systems driven far away from equilibrium [27]. Exploiting the formal similarity of dynamical pathways with polymers, the transition path ensemble can also be sampled with a variant of the configurational bias Monte Carlo scheme (CBMC) [5, 23, 28]. In this approach a guiding field biases the growth of pathways such that they are likely to be reactive. Pathways are modified globally also in the *dynamical algorithm*, in which the path distribution is sampled by integration of fictitious equations of motion [5]. Since all these algorithms do not make use of the natural propensity of the dynamics to converge toward the stable regions, they lack the efficiency of the shooting and shifting algorithms.

7.3.3 Efficiency

The efficiency of a transition path sampling simulation depends on how rapidly the relevant regions of path space are sampled. For Hamiltonian trajectories, the sampling speed can be optimized by tuning the magnitude of the momentum displacement $\delta \mathbf{p}$. For very small momentum displacements a newly generated pathway closely resembles the old pathway. Although the average acceptance probability is high in this case (similar pathways most likely have similar weights in the transition path ensemble), sampling progresses slowly because of the similarity of consecutive pathways. This is seen most clearly in the case of a vanishing momentum displacement for which each new pathway is, trivially, exactly identical to the old pathway

and so the simulation does not move at all although the acceptance probability is equal to one. If, on the other hand, the momentum displacement is very large each new pathway may be very different from the old one. Nevertheless, the sampling speed can be slow due to an excessively small average acceptance probability. Most likely, a new path that is drastically different from the old one does not connect regions $\mathscr{A}$ and $\mathscr{B}$ and, therefore, is rejected.

To optimize the efficiency of a transition path sampling simulation one has to select momentum displacements that reconcile these two opposing effects. A quantitative efficiency measure can be obtained by calculating autocorrelation functions of path properties as a function of the number of sampling cycles [17]. Fast decay of this autocorrelation function is indicative of rapid sampling and hence high efficiency. In practical transition path sampling applications, it is rarely possible to carry out a systematic efficiency analysis and then to select momentum displacements minimizing correlations. Instead, one can tune the magnitude of the momentum displacements to obtain a certain intermediate average acceptance probability. Transition paths sampling simulations carried out for a simple model system indicate that, as a rule of thumb, optimum efficiency is obtained for average acceptance rates in the range 40–60% [17].[4]

7.3.4 Initial Pathway and Definition of the Stable States

To start the transition path sampling procedure an initial path connecting region $\mathscr{A}$ with region $\mathscr{B}$ must be available. Such a first reactive trajectory can be generated in various ways that depend on the particular problem of interest. No general procedure to generate initial trajectories for transition path sampling simulations is available. Of course, a straightforward molecular dynamics simulation started from one of the stable states will eventually produce a suitable reactive trajectory. In most interesting cases, however, the CPU time required to generate even one single reactive trajectory exceeds by far the available resources. In some cases, a temperature increase may help. For instance, to study the folding of a protein one may carry out a molecular dynamics simulation starting from the folded state. The folded state is destabilized by increasing the temperature until the protein unfolds. The resulting high-temperature reactive trajectory can then be used as an initial trajectory to start the path sampling procedure. Another way to generate an initial pathway consists of using a bias function to guide a molecular dynamics simulation from one stable state to the other. The reactive trajectory obtained in this way can be used as an initial pathway in a transition path sampling *without* bias. Alternatively, artificially constructed sequences of states connecting $\mathscr{A}$ and $\mathscr{B}$ can be used to start the transition path sampling simulation. Most of these procedures to generate initial trajectories produce pathways that are reactive but otherwise carry a very low weight in the transition path ensemble. During the path sampling procedure any initial path will quickly relax toward more likely regions of path space.

---

[4] As in the case of conventional Monte Carlo simulations depending on the cost of rejected moves the optimum acceptance probability may be lower than 40% [23].

Another important issue in transition path sampling applications is the definition of the stable regions $\mathscr{A}$ and $\mathscr{B}$. It is convenient to define this regions with the help of a (possibly multidimensional) order parameter $\xi(\mathbf{q})$ depending on the atomic positions $\mathbf{q}$.[5] For instance, for the nucleation of the first-order phase transition discussed in the introduction, $\xi(\mathbf{q})$ could be the size of the largest cluster of the stable phase. In the case of chemical reactions the order parameter $\xi(\mathbf{q})$ might be a bond length, a bond angle or a dihedral angle or a function thereof. Regions $\mathscr{A}$ and $\mathscr{B}$ can then be defined by requiring that the order parameter $\xi(\mathbf{q})$ lies within certain limits. These limits need to be chosen such that the stable regions are large enough to accommodate most equilibrium fluctuations of the system such that the system is located in $\mathscr{A}$ or $\mathscr{B}$ most of the time. Excursions of the system outside these regions should occur only rarely. At the same time, it is important that regions $\mathscr{A}$ and $\mathscr{B}$ do not overlap. Even more stringently, each of the two stable regions should not include configurations that belong to the *basin of attraction* of the other regions. Here, the basin of attraction of, say, $\mathscr{A}$ is defined to consist of all configurations that will quickly evolve into region $\mathscr{A}$. If one stable state contains points belonging to the basin of attraction of the other stable state, the transition path sampling procedure will produce pathways that are not true reactive trajectories. Although specifying the stable regions requires some care, appropriate definitions can usually be found with some trial and error [8].

7.4 Free Energies from Transition Path Sampling Simulations

As discussed in previous chapters, one is often interested in calculating the free energy as a function of a given *reaction coordinate ξ*. Such a free energy profile $A(\xi)$ is defined as

$$A(\xi) = -k_{\mathrm{B}}T \ln P(\xi), \tag{7.27}$$

where $P(\xi)$ is the probability distribution of the reaction coordinate ξ in the equilibrium ensemble with distribution function $\rho(\mathbf{q})$

$$P(\tilde{\xi}) = \int \mathrm{d}V \rho(\mathbf{q})\delta[\tilde{\xi} - \xi(\mathbf{q})]. \tag{7.28}$$

Here, the integration extends over the entire configuration space. In computer simulations probability distribution functions are approximated by histograms that are computed by determining how often the reaction coordinate $\xi(\mathbf{q})$ falls within the various histogram bins. Of course, to obtain the correct histogram, configurations classified in the histogram must be sampled according to the equilibrium distribution $\rho(\mathbf{q})$. This can be done efficiently with various molecular dynamics and Monte Carlo simulation methods, as discussed in Chap. 3.

---

[5] In most applications of transition path sampling it is sufficient to define the stable regions $\mathscr{A}$ and $\mathscr{B}$ in terms of configurational coordinates without reference to the momenta. The transition path sampling formalism, however, can be also applied to situations in which $\mathscr{A}$ and $\mathscr{B}$ also depend on the atomic momenta.

If some values of the reaction coordinate are rarely sampled but are nevertheless important, special techniques are needed as discussed extensively in several chapters of this book. Umbrella sampling and other non-Boltzmann sampling methods may be used to access low-probability regions of configuration space and to accurately calculate strongly varying free energy profiles. Since the transition path sampling method is designed to generate trajectories traversing exactly such low-probability regions associated with rare events, it is tempting to use this approach for the calculation of free energy profiles. In the transition path ensemble, however, pathways are required to start and end in particular regions of configuration space. Due to this requirement (or constraint) the distribution of configurations along such pathways deviates from the equilibrium distributions $\rho(\mathbf{q})$. Therefore, one cannot calculate the probability distribution $P(\xi)$ from the set of configurations on trajectories harvested in a transition path simulation.

Although it is not possible to determine free energy profiles directly from a transition path sampling simulation in a straightforward way, path sampling techniques such as shooting and shifting can be useful for this purpose.[6] To generate configurations according to the equilibrium distribution one can sample the path distribution (7.3) without the requirement that the pathway starts in $\mathscr{A}$ and ends in $\mathscr{B}$ [11]. In the shooting and shifting procedure new pathways can then be accepted with a probability depending on the weight of the initial conditions alone

$$P_{\mathrm{acc}}[z^{(\mathrm{o})}(\mathscr{T}) \to z^{(\mathrm{n})}(\mathscr{T})] = \min\left\{1, \frac{\rho[z_0^{(\mathrm{n})}]}{\rho[z_0^{(\mathrm{o})}]}\right\}. \tag{7.29}$$

In this case the shooting and shifting procedure may be viewed as a particular move in a Monte Carlo simulation similar to hybrid Monte Carlo [30].

For Newtonian dynamics and a canonical distributions of initial conditions one can reject or accept the new path before even generating the trajectory. This can be done because Newtonian dynamics conserves the energy and the canonical phase-space distribution is a function of the energy only. Therefore, the ratio $\rho[z_0^{(\mathrm{n})}]/\rho[z_0^{(\mathrm{o})}]$ at time 0 is equal to the ratio $\rho[z_{t'}^{(\mathrm{n})}]/\rho[z_{t'}^{(\mathrm{o})}]$ at the shooting time and the new trajectory needs to be calculated only if accepted. For a microcanonical distribution of initial conditions all phase-space points on the energy shell have the same weight and therefore all new pathways are accepted. The same is true for Langevin dynamics with a canonical distribution of initial conditions.

Provided the underlying dynamics conserves the equilibrium distribution, as does Newtonian dynamics, the phase-space points lying on the harvested trajectories are distributed according to the equilibrium distribution and can be used to calculate equilibrium averages such as the free energy profile $A(\xi)$. If some important ranges of the reaction coordinate are rarely visited, biasing procedures may help to achieve ergodic sampling of all important regions in configuration space. One may,

---

[6] Note that the transition path sampling method can be also used for the calculation of *activation energies* (as opposed to activation *free* energies) [29]. This approach is useful for systems in which it is not possible to identify transition states.

for instance, employ a variant of the umbrella sampling method and divide the order parameter range of interest into several overlapping windows $\mathscr{W}_i$ [11]. Then, for each window a path sampling simulation is carried out with path weight

$$\mathscr{P}_{\mathscr{W}_i}[z(\mathscr{T})] \equiv \mathscr{H}_{\mathscr{W}_i}[z(\mathscr{T})]\mathscr{P}[z(\mathscr{T})]/Z_{\mathscr{W}_i}(\mathscr{T}). \qquad (7.30)$$

Here, $\mathscr{H}_{\mathscr{W}_i}[z(\mathscr{T})]$ is a function that is unity if the path $z(\mathscr{T})$ has at least one configuration with the order parameter in the window $\mathscr{W}_i$ and vanishes otherwise. The factor:

$$Z_{\mathscr{W}_i}(\mathscr{T}) \equiv \int \mathscr{D}z(\mathscr{T})\,\mathscr{H}_{\mathscr{W}_i}[z(\mathscr{T})]\mathscr{P}[z(\mathscr{T})] \qquad (7.31)$$

normalizes the path ensemble $\mathscr{P}_{\mathscr{W}_i}[z(\mathscr{T})]$. The function $\mathscr{H}_{\mathscr{W}_i}$ appearing in the above path ensemble makes sure that each generated pathway includes at least one configuration with order parameter in the window $\mathscr{W}_i$. All the configuration satisfying this requirement can be used to calculate the order parameter distribution and, from it, the free energy profile in the particular window. Just as in a regular umbrella sampling simulations the free energy profiles in the various windows are then matched to obtain a continuous curve over the whole range of interest. While such a path sampling procedure is possible and practical [11], further studies will be necessary to determine whether its efficiency is competitive with other free energy calculation methods discussed in this book.

Free energy profiles can also be evaluated within the partial path transition interface sampling method (PPTIS), a path sampling technique designed for the calculation of reaction rate constant in systems with diffusive barrier-crossing events [31, 32]. In this approach, the reaction rate is expressed in terms of transitions probabilities between a series of nonintersecting interfaces located between regions $\mathscr{A}$ and $\mathscr{B}$. The interfaces can be defined by requiring that the reaction coordinate $\xi(\mathbf{q})$ takes particular values. The intervals between these values then correspond to the windows defined above. The transition probabilities needed to determine the overall transition rate constant are calculated in path sampling simulations based on the shooting algorithm in which short trajectories are required to cross at least two adjacent interfaces. Since this requirement introduces a bias in the distribution of configuration along such pathways, order parameter distributions, which are equilibrium averages, cannot be computed directly in a PPTIS simulation. By a clever comparison of path ensembles belonging to neighboring interfaces, however, it is possible to correct for this bias [10]. As a result, one can calculate a free energy profile $A(\xi)$ as a function of a given reaction coordinate ξ at no additional cost.

7.5 The Jarzynski Identity: Path Sampling of Nonequilibrium Trajectories

As discussed in detail in Chap. 5, free energy differences can be calculated from the statistics of the work carried out during nonequilibrium transformations. The basis for this method, also known as 'fast switching', is an identity derived in 1997 by

Jarzynski [12]. Although this identity is an exact result, statistical sampling problems arise if the transformation moves the system too far from equilibrium. In this section we will explain the origin of these difficulties and show how transition path sampling can be used to overcome them.

Consider a thermodynamic system with an external parameter (or constraint) λ that can be used to control the state of the system. When changing the control parameter λ a certain amount of work is performed on the system. According to the second law of thermodynamics the average work necessary to do that is smaller than the Helmholtz free energy difference between the two equilibrium states corresponding to the initial and final values of the constraint [33]

$$\langle W \rangle \geq \Delta A. \tag{7.32}$$

The equality is valid if the control parameter is changed reversibly, i.e., if the system is in equilibrium at all times during the transformation. Equivalently, this result can be stated as the *maximum work theorem* [34]: the amount of work *delivered* by a system during a transformation from a specific initial to a specific final state is always smaller than the free energy difference between the initial and final states. The work is maximum and equal to the free energy difference for a reversible process, hence the term *reversible work* for the equilibrium free energy.

The maximum work theorem can be used to calculate free energies. In the thermodynamic integration scheme, for instance, free energy differences are determined by calculating the reversible work required to change a control parameter reversibly, that is very slowly (see Chap. 4). This reversibility requirement, however, can be relaxed within a recently developed fast-switching approach. In a remarkable theorem, Jarzynski has proven that under very general conditions the so-called Clausius inequality [(7.32)] can be turned into an equality by considering an exponential of the work instead of the work itself [12, 35, 36]

$$\langle \exp(-\beta W) \rangle = \exp(-\beta \Delta A). \tag{7.33}$$

The angular brackets $\langle \cdots \rangle$ in (7.33) denote an average over an ensemble of non-equilibrium transformation processes initiated from states z distributed according to a canonical distribution. The Jarzynski identity (7.33) is valid for nonequilibrium transformations carried out at arbitrary speed.

The Jarzynski identity can be used to calculate the free energy difference between two states 0 and 1 with Hamiltonians $\mathscr{H}_0(z)$ and $\mathscr{H}_1(z)$. To do that we consider a Hamiltonian $\mathscr{H}(z, \lambda)$ depending on the phase-space point z and the control parameter λ. This Hamiltonian is defined in such a way that λ_0 corresponds to the Hamiltonian of the initial state, $\mathscr{H}(z, \lambda_0) = \mathscr{H}_0(z)$, and λ_1 to the Hamiltonian of the final state, $\mathscr{H}(z, \lambda_1) = \mathscr{H}_1(z)$. By changing λ continuously from λ_0 to λ_1 the Hamiltonian of the initial state is transformed into that of the final state. The free energy difference:

$$\Delta A = -k_{\mathrm{B}} T \ln \frac{\int \mathrm{d}z \, \exp\{-\beta \mathscr{H}_1(z)\}}{\int \mathrm{d}z \, \exp\{-\beta \mathscr{H}_0(z)\}} = -k_{\mathrm{B}} T \ln \frac{Z_0}{Z_1} \tag{7.34}$$

can now be calculated by first generating initial conditions distributed according to the canonical distribution $\exp\{-\beta\mathscr{H}_0(z)\}/Z_0$. Then, dynamical trajectories of a certain length $\mathscr{T}$ in time are initiated at these initial conditions. While the system evolves in time, the control parameter is changed from λ_0 at time $t = 0$ to λ_1 at time $t = \mathscr{T}$ according to a certain arbitrary protocol. By changing the external parameter λ we perform the work W on the system, and this work may be different for each trajectory. Averaging $\exp(-\beta W)$ over all trajectories we obtain an estimate of $\exp(-\beta\Delta A)$ and hence of ΔA. In the path integral notation introduced in Sect. 7.2 this average can be expressed as

$$\exp(-\beta\Delta A) = \int \mathscr{D}z(\mathscr{T})\mathscr{P}[z(\mathscr{T}),\lambda(\mathscr{T})]\exp\{-\beta W[z(\mathscr{T}),\lambda(\mathscr{T})]\} \quad (7.35)$$

where $\lambda(\mathscr{T})$ denotes the complete history of the control parameter λ from $t = 0$ to $\mathscr{T}$. The path probability $\mathscr{P}[z(\mathscr{T}),\lambda(\mathscr{T})]$, which includes the probability $\exp\{-\beta\mathscr{H}_0(z)\}/Z_0$ of the initial conditions and the work $W[z(\mathscr{T}),\lambda(\mathscr{T})]$ performed along the path, depends on the path $z(\mathscr{T})$ itself as well as on the progression $\lambda(\mathscr{T})$ of the control parameter.

The Jarzynski identity (7.33) is an exact result and applies to transformations of arbitrary length $\mathscr{T}$. From a computational point of view this property seems very attractive because it implies that free energy differences can be calculated from short and therefore computationally inexpensive trajectories. However, the convergence of the exponential average from (7.33) quickly deteriorates if the transformation (or the *switching*) is carried out too rapidly.

This statistical problem, which can easily offset the gain originating from the low computational cost of short trajectories, is best understood by rewriting the Jarzynski identity as an average over the work distributions $P(W)$

$$\exp(-\beta\Delta A) = \int dW \, P(W)\exp(-\beta W). \quad (7.36)$$

Here, $P(W)$ is the probability density for observing a work value W. If the transformation is carried out slowly, the work distribution is a function narrowly peaked near the free energy difference ΔA. In the reversible limit of infinitely slow switching each trajectory yields the same work W equal to ΔA and $P(W)$ is a delta function centered at ΔA.[7] For increasing switching rate the work distribution becomes wider and is shifted toward work values that are large compared to the free energy difference. In this case, the work distribution $P(W)$ and the integrand of (7.36), $P(W)\exp(-\beta W)$, can be peaked at very different work values and have little overlap. Thus, most trajectories have work values that essentially do not contribute to the exponential average. Only work values from the low-W wing of the work distribution generate significant contributions, but these work values are rarely

---

[7] For isolated Hamiltonian systems the width of the work distribution remains finite even in the limit of infinitely slow switching. This is a consequence of the so-called *adiabatic invariants* [16].

generated in a straightforward fast-switching simulation. As a consequence, free energies estimated from a finite set of fast-switching trajectories can have large statistical errors [16, 36, 37]. These difficulties are familiar from applications of thermodynamic perturbation theory and have been discussed in Chaps. 2 and 6.

For straightforward fast-switching simulations this statistical problem limits the switching rates to values for which the average work does not deviate from the free energy difference by more than the thermal energy $k_{\mathrm{B}}T$ [37]. A systematic analysis of the statistical error in the estimated free energy shows that in this regime the fast-switching method does not bring any computational gain. In other words, the error in the free energy calculated from one single long trajectory or many shorter ones is the same [37].[8]

A way to circumvent the statistical problems related to the work statistics of short trajectories was recently suggested by Sun [13]. The basic idea of this approach, which can be thought of as a generalization of thermodynamic integration to trajectory space, is to devise a sampling scheme that favors the rare trajectories with work values mostly contributing to the exponential average of the Jarzynski identity. Sun achieves this by introducing a parameter α into the exponential average

$$\exp\{-\beta\Delta\tilde{A}(\alpha)\} = \int \mathscr{D}z(\mathscr{T})\mathscr{P}[z(\mathscr{T})] \exp\{-\beta\alpha W[z(\mathscr{T})]\}. \qquad (7.37)$$

To simplify the notation we have dropped $\lambda(\mathscr{T})$ from the argument list of the path probability $\mathscr{P}[z(\mathscr{T})]$ and the work $W[z(\mathscr{T})]$. Of course, in this definition the new free energy difference $\Delta\tilde{A}(\alpha)$ also acquires a dependence on the parameter α. The parameter-dependent free energy $\Delta\tilde{A}(\alpha)$ differs from the original free energy difference ΔA. For $\alpha = 0$ the integral on the right-hand side of the above equation is unity and $\Delta\tilde{A}(\alpha) = 0$. For $\alpha = 1$, however, the original free energy is recovered, $\Delta\tilde{A}(1) = \Delta A$. We can hence calculate ΔA by taking the derivative of $\Delta\tilde{A}(\alpha)$ with respect to α and then integrating this quantity from 0 to 1

$$\Delta A = \int_0^1 \mathrm{d}\alpha \frac{\mathrm{d}\Delta\tilde{A}(\alpha)}{\mathrm{d}\alpha}. \qquad (7.38)$$

To carry out this integration we need to determine the derivative of $\Delta\tilde{A}(\alpha)$. Differentiating (7.37) we obtain

$$\frac{\mathrm{d}\tilde{A}(\alpha)}{\mathrm{d}\alpha} = \frac{\int \mathscr{D}z(\mathscr{T})\mathscr{P}[z(\mathscr{T})] \exp\{-\beta\alpha W[z(\mathscr{T})]\}W[z(\mathscr{T})]}{\int \mathscr{D}z(\mathscr{T})\mathscr{P}[z(\mathscr{T})] \exp\{-\beta\alpha W[z(\mathscr{T})]\}} = \langle W \rangle_\alpha. \quad (7.39)$$

The notation $\langle \cdots \rangle_\alpha$ used in the last equality of the above equation indicates that this expression can be viewed as the average work in the so-called *work-weighted path ensemble*

$$\boxed{\mathscr{P}_\alpha[z(\mathscr{T})] = \mathscr{P}[z(\mathscr{T})] \exp\{-\beta\alpha W[z(\mathscr{T})]\}/Z_\alpha} \qquad (7.40)$$

---

[8] Although in this slow-switching regime nonequilibrium simulations do not offer any direct reduction of the computational cost of free energy calculation, they have the added advantage of being easily parallelizable and of permitting an estimation of statistical errors.

where the path ensemble is normalized by the partition function

$$Z_\alpha = \int \mathscr{D}z(\mathscr{T})\mathscr{P}[z(\mathscr{T})] \exp\{-\beta\alpha W[z(\mathscr{T})]\}. \tag{7.41}$$

In the work-weighted path ensemble the statistical weight of a particular trajectory $z(\mathscr{T})$ explicitly depends on the work performed on the system along that trajectory. The work distribution in the work-weighted path ensemble for a particular value of α is

$$P_\alpha(W) = \frac{P(W)\exp(-\beta\alpha W)}{\int \mathrm{d}W\, P(W)\exp(-\beta\alpha W)}. \tag{7.42}$$

For $\alpha = 0$ this work distribution is identical to the work distribution obtained in a straightforward fast-switching simulation. In the other limit, at $\alpha = 1$, the work distributions is proportional to the integrand in (7.36). Thus, Sun's procedure guarantees that all important work values are sampled regardless of the length path length. Indeed, it has been demonstrated that the Sun approach can be used to calculate free energy differences using very short pathways [13]. It also follows from these considerations that the free energy difference ΔA can be viewed as the reversible work necessary to transform the ensemble of unconstrained paths (7.3) into the work-weighted path ensemble (7.40).

To calculate the path average $\langle W \rangle_\alpha$ from (7.39) we need to sample the work-weighted path ensemble (7.40). Since the weight of a trajectory in this ensemble explicitly depends on the work W performed along this trajectory, we cannot simply do that by generating suitable initial conditions and growing fast-switching trajectories from them. Instead, the work-weighted path ensemble can be sampled with the transition path sampling procedures described in Sect. 7.3. With the shooting algorithm we can generate a new fast-switching trajectory from an old one by first changing the momenta at a randomly selected time slice. Then the equations of motion are integrated backward and forward starting from the point with modified momenta while the control parameter λ is changed according to the protocol $\lambda(\mathscr{T})$. The new path is then accepted with a probability depending on the work of the new and the old path. By carrying out such a path simulation for several different values of α between 0 and 1 we can determine work averages $\langle W \rangle_\alpha$ that, according to (7.38), are then integrated numerically to obtain the free energy difference ΔA.

Although this path sampling procedure can be used to calculate free energy differences from very short nonequilibrium trajectories, the question arises if such an approach is competitive with conventional free energy calculation methods. As shown by Sun [13], for infinitely fast switching the work-weighted path sampling method reduces to the conventional thermodynamic integration algorithm that does not make use of nonequilibrium trajectories. An error analysis carried out for different example systems indicated that optimum efficiency is obtained in this limit of infinitely short trajectories, implying that conventional thermodynamic integration outperforms work-weighted path sampling. Whether this is true in general or whether there are cases in which the fast-switching path sampling approach is more efficient is an open question.

An alternative path sampling method to evaluate Jarzynski's exponential average has been put forward by Ytreberg and Zuckerman [14] and by Athènes [15]. The basic idea here is the same as in Sun's work-weighted path sampling approach: trajectories with rare but important work values are sampled with enhanced likelihood. This can be achieved by introducing a *bias function* (or *umbrella function*) $\pi[z(\mathscr{T})]$ in the exponential path average [14]

$$\exp\{-\beta\Delta A\} = \frac{\int \mathscr{D}z(\mathscr{T})\, P\left[z\left(\mathscr{T}\right)\right]\pi\left[z\left(\mathscr{T}\right)\right]\left[\dfrac{\exp\left\{-\beta W\left[z\left(\mathscr{T}\right)\right]\right\}}{\pi\left[z\left(\mathscr{T}\right)\right]}\right]}{\int \mathscr{D}z\left(\mathscr{T}\right)\, P\left[z\left(\mathscr{T}\right)\right]\pi\left[z\left(\mathscr{T}\right)\right]\left[\dfrac{1}{\pi\left[z\left(\mathscr{T}\right)\right]}\right]} \quad (7.43)$$

$$= \frac{\langle\exp\{-\beta W[z(\mathscr{T})]\}/\pi[z(\mathscr{T})]\rangle_\pi}{\langle 1/\pi[z(\mathscr{T})]\rangle_\pi}. \quad (7.44)$$

The notation $\langle\cdots\rangle_\pi$ used in the second line of the above equation implies a path average over the biased path distribution

$$\mathscr{P}_\pi[z(\mathscr{T})] = \mathscr{P}[z(\mathscr{T})]\pi[z(\mathscr{T})]/\mathscr{Z}_\pi, \quad (7.45)$$

where

$$\mathscr{Z}_\pi = \int \mathscr{D}z(\mathscr{T})\mathscr{P}[z(\mathscr{T})\pi[z(\mathscr{T})]. \quad (7.46)$$

Since the bias function should enhance the sampling of pathways with important work values it can be made to depend on the work only, $\pi[z(\mathscr{T})] = \pi[W(z(\mathscr{T}))]$. To minimize the statistical error in the free energy difference the bias function needs to be selected such that both the statistical errors of the numerator and the denominator of (7.44) are small. Ideally, the bias function should have a large overlap with both the unbiased work distribution $P(W)$ and the integrand of (7.36), $P(W)\exp(-\beta W)$. Just as Sun's work-biased ensemble $\mathscr{P}_\alpha[z(\mathscr{T})]$, the biased path ensemble $\mathscr{P}_\pi[z(\mathscr{T})]$ can be sampled with the shooting algorithm [14]. Then, the acceptance probability for this move also depends on the bias function $\pi[z(\mathscr{T})]$ for the new and old trajectory. Note that other non-Boltzmann sampling techniques, such as flat-histogram sampling [38], multicanonical sampling [39], or parallel tempering [40], described in Chaps. 3 and 8, can be combined with the path sampling procedure to enhance the convergence of Jarzynski's exponential average. Another way to increase the efficiency of fast-switching simulations is to determine only approximate trajectories by integrating the equations of motion with large time steps [41]. It can be shown that Jarzynski's identity remains valid also for such computationally less expensive trajectories.

Zuckerman and Ytreberg have shown for several model systems that compared to straightforward fast-switching simulations large efficiency gains can be achieved by using the bias function $\pi(W) = \exp(-\beta W/2)$. But as in the case of Sun's method, the question arises whether this biased path sampling approach is computationally competitive with conventional approaches such as umbrella sampling. Expressions for the statistical error in the free energies obtained in these biased path sampling schemes have been developed and used to estimate the error for different example systems [16]. These results indicate that conventional umbrella sampling with a good

bias function is superior to the biased path sampling of nonequilibrium trajectories. More research is necessary, however, to find out if this is true in general.

7.6 Rare Event Kinetics and Free Energies in Path Space

On a phenomenological level, transitions between long-lived stable states can be described in terms of reaction rate constants. Since such reaction rate constants can be measured in experiments, their calculation in computer simulations is of great interest. Pathways harvested in transition path sampling simulations are true dynamical trajectories. Therefore they can be used to calculate reaction rate constants for the transitions between stable states.[9] In this section, we will explain how to calculate reaction rate constants in the framework of transition path sampling by exploiting an isomorphism between time correlation functions and free energy differences. In a sense, the strategy of such an approach is opposite to the free energy calculations described in the previous sections. While path sampling methods were previously used to calculate free energies, here free energy calculation methods are used to calculate reaction rate constants in transition path sampling simulations. Note that reaction rate constants can be also calculated with the very efficient transition interface sampling (TIS) method which is based on path sampling ideas [31, 32]. Although TIS is more efficient than the approach described below it is not discussed here, because it does not make use of the free energy calculation methods to which this book is devoted.

On a phenomenological level, transitions between long-lived stable states can be described in terms of reaction rate constants. Consider, for instance a solution of two well-defined chemical species $\mathscr{A}$ and $\mathscr{B}$ that can interconvert through the unimolecular reaction

$$\mathscr{A} \rightleftharpoons \mathscr{B}. \tag{7.47}$$

The solution is assumed to be sufficiently dilute that the solute molecules do not interact with each other (at the same time we assume that there still is a macroscopic number of solute molecules in the solution). Due to the reaction the concentrations $c_{\mathscr{A}}$ and $c_{\mathscr{B}}$ of molecules of type $\mathscr{A}$ and $\mathscr{B}$, respectively, can change in time. The concentration of $c_{\mathscr{A}}$ decreases when molecules of type $\mathscr{A}$ transform into molecules of type $\mathscr{B}$ and increases due to the inverse reaction. Since, according to the assumptions, the solute molecules are statistically independent from each other, the time evolution of $c_{\mathscr{A}}(t)$ is well described by the phenomenological [33]

$$\frac{\mathrm{d}c_{\mathscr{A}}(t)}{\mathrm{d}t} = -k_{\mathscr{A}\mathscr{B}}c_{\mathscr{A}}(t) + k_{\mathscr{B}\mathscr{A}}c_{\mathscr{B}}(t). \tag{7.48}$$

---

[9] The calculation of reaction rate constants with the transition path sampling methods does not require understanding of the reaction mechanism, for instance in the form of an appropriate reaction coordinate. If such information is available other methods such as the reactive flux formalism are likely to yield reaction rate constants at a lower computational cost than transition path sampling.

An analogous loss–gain equation holds for the concentration $c_{\mathscr{B}}(t)$

$$\frac{dc_{\mathscr{B}}(t)}{dt} = -k_{\mathscr{B}\mathscr{A}}c_{\mathscr{B}}(t) + k_{\mathscr{A}\mathscr{B}}c_{\mathscr{A}}(t). \tag{7.49}$$

In these equation all microscopic details of the dynamics are condensed into the forward and backward reaction rate constants $k_{\mathscr{A}\mathscr{B}}$ and $k_{\mathscr{B}\mathscr{A}}$.

Linear response theory[10] provides a link between the phenomenological description of the kinetics in term of reaction rate constants and the microscopic dynamics of the system [33]. All information needed to calculate the reaction rate constants is contained in the time correlation function

$$C(t) = \frac{\langle h_{\mathscr{A}}(z_0)h_{\mathscr{B}}(z_t)\rangle}{\langle h_{\mathscr{A}}\rangle} \tag{7.50}$$

for a particular molecule. The functions $h_{\mathscr{A}}$ and $h_{\mathscr{B}}$ are unity if a particular molecule is in state $\mathscr{A}$ or $\mathscr{B}$, respectively, and vanish otherwise. The angular brackets denote an equilibrium average over all trajectories (or, for Hamiltonian dynamics, a simple equilibrium phase-space average). The time correlation function $C(t)$ is the conditional probability to find a particular molecule in state $\mathscr{B}$ at time t provided it was in state $\mathscr{A}$ at time zero. According to the fluctuation–dissipation theorem, for long times $C(t)$ behaves like the concentration $c_{\mathscr{B}}(t)$ relaxing from a nonequilibrium state in which only molecules of type $\mathscr{A}$ exists. For long times, the behavior of $C(t)$ is hence completely determined by the values of the reaction rate constants

$$C(t) \approx \langle h_{\mathscr{B}}\rangle(1 - \exp(-t/\tau_{\mathrm{rxn}})), \tag{7.51}$$

where the *reaction time* τ_{rxn} is related to the forward and backward reaction rate constants $k_{\mathscr{A}\mathscr{B}}$ and $k_{\mathscr{B}\mathscr{A}}$ by

$$\tau_{\mathrm{rxn}} = (k_{\mathscr{A}\mathscr{B}} + k_{\mathscr{B}\mathscr{A}})^{-1}. \tag{7.52}$$

For short times, the correlation function $C(t)$ depends on the microscopic details of the dynamics as the system crosses from $\mathscr{A}$ to $\mathscr{B}$. These motions take place on a molecular time scale τ_{mol} essentially equal to the time required to move through the transition region. For times t larger than τ_{mol} but still very small compared to the reaction time τ_{rxn} (if the crossing event is rare $\tau_{\mathrm{rxn}} \gg \tau_{\mathrm{mol}}$ such that such an intermediate time regime exists), $C(t)$ can be replaced by an approximation linear in time. Using the detailed balance condition $k_{\mathscr{B}\mathscr{A}}/k_{\mathscr{A}\mathscr{B}} = \langle h_{\mathscr{B}}\rangle/\langle h_{\mathscr{A}}\rangle$ [33] one then obtains

$$C(t) \approx k_{\mathscr{A}\mathscr{B}}t. \tag{7.53}$$

The slope of $C(t)$ in the time regime $\tau_{\mathrm{mol}} < t \ll \tau_{\mathrm{rxn}}$ is the forward reaction rate constant. Thus, for the calculation of reaction rate constants it is sufficient to determine the time correlation function $C(t)$. In the following paragraphs we will show how to do that in the transition path sampling formalism.

---

[10] It can be shown that the assumption of a weak perturbation central to linear response theory can be relaxed in this case [9]. The equations presented in this section relating the kinetic coefficients with the microscopic dynamics of the system remain valid for arbitrarily strong perturbations.

In principle, the time correlation function $C(t)$ can be calculated from a single, long molecular dynamics simulation. However, such a simulation is impractical because for the rare transitions of interest the computer resources necessary to collect a statistically sufficient number of reaction events are excessive. The transition path sampling method focuses exactly on these reactive events and is therefore not plagued by this problem. To calculate the time correlation function $C(t)$, however, it is not sufficient to consider reactive trajectories only. Rather, it is necessary to determine the total weight of all the reactive trajectories relative to the total weight of all nonreactive trajectories. In other words, we need to calculate how many of all possible trajectories originating in the initial region $\mathscr{A}$ arrive in the final region $\mathscr{B}$ after some time t. Roughly, this corresponds to determining the 'size' in path space of the subensemble of reactive trajectories relative to the size of the set of all trajectories starting in $\mathscr{A}$. But determining relative statistical weights of certain subensembles is exactly what we do in free energy calculations. This analogy suggests to determine reaction rate constants by applying free energy calculation methods to ensembles of pathways.

To make this idea more precise we rewrite the time correlation function $C(t)$ in terms of integrals over pathways

$$C(t) = \frac{\int \mathscr{D}z(t)\, h_{\mathscr{A}}(z_0)\mathscr{P}[z(t)]h_{\mathscr{B}}(z_t)}{\int \mathscr{D}z(t)\, h_{\mathscr{A}}(z_0)\mathscr{P}[z(t)]} = \frac{Z_{\mathscr{A}\mathscr{B}}(t)}{Z_{\mathscr{A}}}. \qquad (7.54)$$

This expression may be interpreted as a ratio of two partition functions. In the denominator we have the partition function $Z_{\mathscr{A}}$ of all trajectories starting in region $\mathscr{A}$ with endpoint anywhere; the integral in the numerator is the partition function $Z_{\mathscr{A}\mathscr{B}}(t)$ of all trajectories starting in $\mathscr{A}$ and ending in $\mathscr{B}$ [this is the normalizing factor of (7.11)]. We can then view the ratio of partition functions as the exponential of the free energy difference between these two ensembles of trajectories

$$C(t) \equiv \exp\{-\Delta \mathsf{A}_{\mathscr{A}\mathscr{B}}(t)\}. \qquad (7.55)$$

Here, we have denoted this path 'free energy' with A typeset in sans serif to make clear that this quantity differs from the conventional Helmholtz free energy A. The free energy difference $\Delta \mathsf{A}_{\mathscr{A}\mathscr{B}}(t)$ can be interpreted as the reversible work necessary to transform the ensemble of trajectories of length t starting in $\mathscr{A}$ without any restriction on the endpoint into the ensemble of pathways starting in $\mathscr{A}$ and ending in $\mathscr{B}$. By virtue of this analogy between regular free energies and 'free energies' in trajectory space we can choose from a number of free energy calculation methods to determine the time correlation function $C(t)$. In the following paragraphs we show how that can be achieved with umbrella sampling. Of course, the path free energy $\Delta \mathsf{A}_{\mathscr{A}\mathscr{B}}(t)$ can also be calculated with other free energy methods such as thermodynamic integration [18] or even with the fast-switching method of Sect. 7.5 [19].

As before, we imagine that we can define the stable regions $\mathscr{A}$ and $\mathscr{B}$ with the help of an order parameter $\xi(z)$. A phase-space point z is in region $\mathscr{A}$ if the order parameter is within a certain range, $\xi_{\mathscr{A}}^{\min} \leq \xi(z) \leq \xi_{\mathscr{A}}^{\max}$, and the order parameters

of configurations belonging to $\mathscr{B}$ are in a distinct range, $\xi_{\mathscr{B}}^{\min} \leq \xi(z) \leq \xi_{\mathscr{B}}^{\max}$. Let us now consider the probability[11] that a trajectory starting in region $\mathscr{A}$ ends at a configuration with an order parameter value of $\tilde{\xi}$

$$P_{\mathscr{A}}(\tilde{\xi}, t) = \int \mathscr{D}z(t) \, h_{\mathscr{A}}(z_0) \mathscr{P}[z(t)] \delta(\tilde{\xi} - \xi(z_t)). \tag{7.56}$$

This probability distribution is similar to the probability distribution from (7.28), except that here we average only over configurations that have evolved from configurations in $\mathscr{A}$ a time t earlier. By integrating this probability distribution over all order parameter values corresponding to region $\mathscr{B}$ we obtain the total probability that a system initially in $\mathscr{A}$ is in $\mathscr{B}$ at time t later. This is nothing other than the time correlation function $C(t)$ and we can write

$$C(t) = \int_{\xi_{\mathscr{B}}^{\min}}^{\xi_{\mathscr{B}}^{\max}} \mathrm{d}\xi \, P_{\mathscr{A}}(\xi, t). \tag{7.57}$$

What remains to do is to actually determine the probability distribution $P_{\mathscr{A}}(\xi, t)$.

If the transition from $\mathscr{A}$ to $\mathscr{B}$ is a rare event, $P_{\mathscr{A}}(\xi, t)$ cannot be calculated by initiating trajectories in $\mathscr{A}$ and averaging over them. Instead we can use the same strategy as employed in Sect. 7.4 for the calculation of the distribution $P(\xi)$ (7.28). Again, we divide the order parameter range between $\mathscr{A}$ and $\mathscr{B}$ into a sequence of overlapping windows. These windows correspond to the regions $\mathscr{W}_i$ in phase space (we call these regions windows as well). In each of the narrow windows we now carry out a separate path sampling simulations with pathways required to start in $\mathscr{A}$ and end in $\mathscr{W}_i$. The corresponding path ensemble that needs to be sampled for window $\mathscr{W}_i$ is

$$\mathscr{P}_{\mathscr{A}\mathscr{W}_i}[z(t)] \equiv h_{\mathscr{A}}(z_0) \mathscr{P}[z(t)] h_{\mathscr{W}_i}(z_t)/Z_{\mathscr{A}\mathscr{W}_i}(t), \tag{7.58}$$

where the function $h_{\mathscr{W}_i}(z_t)$ is unity if z_t is in $\mathscr{W}_i$ and vanishes otherwise. The factor

$$Z_{\mathscr{A}\mathscr{W}_i}(t) \equiv \int \mathscr{D}z(t) \, h_{\mathscr{A}}(z_0) \mathscr{P}[z(t)] h_{\mathscr{W}_i}(z_t) \tag{7.59}$$

normalizes this particular path ensemble. The order parameters at the path endpoints are binned into histograms. If the windows are sufficiently narrow each of the path sampling simulations yields an accurate distribution $P_{\mathscr{A}}(\tilde{q}, t)$ up to a constant factor and restricted to the particular window for which the simulation is carried out. By matching the distributions where they overlap one finally obtains the complete distribution, from which $C(t)$ can be calculate by integration.

Through a procedure such as umbrella sampling we can calculate the correlation function $C(t)$ for a particular time t. For a determination of the reaction rate constants, however, we need the derivative of $C(t)$. Of course, the time derivative of $C(t)$ could be determined by calculating $C(t)$ at different path lengths t and taking the derivative numerically. Fortunately, such a computationally expensive procedure is not necessary. One can derive expressions for the reaction rate constant that

---

[11] Strictly speaking, we need to consider a probability density.

require one single-path free energy calculation [8] yielding considerable computational savings.

7.7 Summary

In this chapter we have reviewed the principles of transition path sampling and we have learned how this methodology, developed to study rare transitions between long-lived stable states in complex systems, can also be used to perform and enhance free energy calculations. As explained in Sect. 7.4, the free energy as a function of a given reaction coordinate can be determined from pathways generated with the shooting algorithm described in Sect. 7.3. In such simulations it is important to correct for the bias introduced by the requirement that the pathways visit particular regions of configuration space.

Transition path sampling methods also offer a practical way to improve the calculation of free energies based on the Jarzynski identity relating the reversible work between two states to the work statistics of nonequilibrium transformations. Due to the exponential averaging required by this approach, sampling problems occur at high switching rates. The biased path sampling techniques described in Sect. 7.5 can help to overcome these difficulties by favoring rare but important trajectories with work values contributing most to the exponential average of the Jarzynski identity. Whether this biased path sampling approach to the evaluation of Jarzynki's exponential average will yield free energy calculations that are computationally competitive with the other methods described in this book is an open question of current research.

Finally, in Sect. 7.6, we have discussed how various free energy calculation methods can be applied to determine 'free energies' of ensembles of pathways rather than ensembles of trajectories. In the transition path sampling framework such path free energies are related to the time correlation function from which rate constants can be extracted. Thus, free energy methods can be used to study the kinetics of rare transitions between stable states such as chemical reactions, phase transitions of condensed materials or biomolecular isomerizations.

References

1. Anderson, J. B., Statistical theories of chemical reactions: distributions in the transition region, *J. Chem. Phys.* **1973**, *58*, 4684–4692
2. Bennett, C. H. Molecular dynamics and transition state theory: the simulation of infrequent events. In *Algorithms for Chemical Computations* (Washington, DC, 1977), Christoffersen, R. E., Ed., Am. Chem. Soc., pp. 63–97
3. Chandler, D., Statistical mechanics of isomerization dynamics in liquids and the transition state approximation, *J. Chem. Phys.* **1978**, *68*, 2959–2970
4. Moroni, D.; ten Wolde, P. R.; Bolhuis, P. G., Interplay between structure and size in a critical crystal nucleus, *Phys. Rev. Lett.* **2005**, *94*, 235703/1–4
5. Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D., Transition path sampling and the calculation of rate constants, *J. Chem. Phys.* **1998**, *108*, 1964–1977

6. Pratt, L., A statistical method for identifying transition states in high dimensional problems, *J. Chem. Phys.* **1986**, *85*, 5045–5048

7. Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L., Transition path sampling: throwing ropes over mountain passes in the dark, *Ann. Rev. Phys. Chem.* **2002**, *53*, 291–318

8. Dellago, C.; Bolhuis, P. G.; Geissler, P. L., Transition path sampling, *Adv. Chem. Phys.* **2002**, *123*, 1–78

9. Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition path sampling methods. In *Computer Simulations in Condensed Matter: From Materials to Chemical Biology; Lecture Notes in Physics* (2006), Ciccotti, G.; Binder, K., Eds., vol. 703, Springer: Berlin, Heidelberg, New York, pp. 337–378

10. Moroni, D.; van Erp, T. S.; Bolhuis, P. G., Simultaneous computation of free energies and kinetics of rare events, *Phys. Rev. E* **2005**, *71*, 056709/1–5

11. Radhakrishnan, R.; Schlick, T., Biomolecular free energy profiles by a shooting/umbrella sampling protocol, "BOLAS", *J. Chem. Phys.* **2004**, *121*, 2436–2444

12. Jarzynski, C., Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **1997**, *78*, 2690–2693

13. Sun, S. X., Equilibrium free energies from path sampling of nonequilibrium trajectories, *J. Chem. Phys.* **2003**, *118*, 5769–5775

14. Ytreberg, F. M.; Zuckerman, D. M., Single-ensemble nonequilibrium path-sampling estimates of free energy differences, *J. Chem. Phys.* **2004**, *120*, 10876–10879

15. Athènes, M., A path-sampling scheme for computing thermodynamic properties of a many-body system in a generalized ensemble, *Eur. Phys. J. B* **2004**, *38*, 651–663

16. Oberhofer, H.; Dellago, C.; Geissler, P. L., Biased sampling of non-equilibrium trajectories: can fast switching simulations beat conventional free energy calculation methods? *J. Phys. Chem. B* **2005**, *109*, 6902–6915

17. Dellago, C.; Bolhuis, P. G.; Chandler, D., On the calculation of rate constants in the transition path ensemble, *J. Chem. Phys.* **1999**, *110*, 6617–6625

18. Dellago, C.; Geissler, P. L. Monte Carlo sampling in path space: calculating time correlation functions by transforming ensembles of trajectories. In *The Monte Carlo Method in the Physical Sciences: Celebrating the 50th anniversary of the Metropolis algorithm* (Melville, New York, 2003), Gubernatis, J. E., Ed., AIP Conference Proceedings 690, pp. 192–199

19. Geissler, P. L.; Dellago, C., Equilibrium time correlation functions from irreversible transformations in trajectory space, *J. Phys. Chem. B* **2004**, *108*, 6667–6672

20. van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*, Elsevier: Amsterdam, 1992

21. Zwanzig, R., *Nonequilibrium Statistical Mechanics*, Oxford University Press: Oxford, 2001

22. Allen, M. P.; Tildesley, D., *Computer Simulation of Liquids*, Clarendon: Oxford, 1987

23. Frenkel, D.; Smit, B., *Understanding Molecular Simulation: From Algorithms to Applications*, Academic: San Diego, 2002

24. Metropolis, N.; Metropolis, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., Equation of state calculations for fast computing machines, *J. Chem. Phys.* **1953**, *21*, 1087–1092

25. Landau, D. P.; Binder, K., *Monte Carlo Simulations in Statistical Physics*, Cambridge University Press: Cambridge, 2000

26. de Gennes, P. G., Reptation of a polymer chain in the presence of fixed obstacles, *J. Chem. Phys.* **1971**, *55*, 572–579

27. Crooks, G. E.; Chandler, D., Efficient transition path sampling for nonequilibrium stochastic dynamics, *Phys. Rev. E* **2001**, *64*, 026109/1–4

28. Csajka, F. S.; Chandler, D., Transition pathways in a many-body system: application to hydrogen-bond breaking in water, *J. Chem. Phys.* **1998**, *109*, 1125–1133

29. Dellago, C.; Bolhuis, P. G., Activation energies from transition path sampling simulations, *Mol. Sim.* **2004**, *30*, 795–799

30. Duane, S.; Kennedy, A.; Pendleton, B. J.; Roweth, D., Hybrid Monte Carlo, *Phys. Lett. B* **1987**, *195*, 216–222

31. van Erp, T. S.; Moroni, D.; Bolhuis, P. G., A novel path sampling method for the calculation of rate constants, *J. Chem. Phys.* **2003**, *118*, 7762–7774

32. Moroni, D.; van Erp, T. S.; Bolhuis, P. G., Rate constants for diffusive processes by partial path sampling, *J. Chem. Phys.* **2004**, *120*, 4055–4065

33. Chandler, D., *Introduction to Modern Statistical Mechanics*, Oxford University Press: Oxford, 1987

34. Callen, H. B., *Thermodynamics and an Introduction to Thermostatistics*, Wiley: New York, 1985

35. Crooks, G. E., Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, *J. Stat. Phys.* **1998**, *90*, 1481–1487

36. Ritort, F., Work fluctuations and transient violations of the second law: Perspectives in theory and experiment, *Sem. Poincare* **2003**, *2*, 193–227

37. Hummer, G., Fast-growth thermodynamic integration: Error and efficiency analysis, *J. Chem. Phys.* **2001**, *114*, 7330–7337

38. Wang, F.; Landau, D. P., Multiple-range random walk algorithm to calculate the density of states, *Phys. Rev. Lett.* **2001**, *86*, 2050–2053

39. Berg, B.; Neuhaus, T., Multicanonical algorithms for first order phase transitions, *Phys. Lett. B* **1991**, *267*, 249–253

40. Geyer, C. J.; Thompson, E. A., Annealing Markov chain Monte Carlo with applications to ancestral inference, *J. Am. Stat. Soc.* **1995**, *90*, 909–933

41. Lechner, W.; Oberhofer, H.; Dellago, C.; Geissler, P. L., Equilibrium free energies from fast-switching simulations with larg time steps, *J. Chem. Phys* **2006**, *124*, 04113/1–12

8

Specialized Methods for Improving Ergodic Sampling Using Molecular Dynamics and Monte Carlo Simulations

Ioan Andricioaei

8.1 Background

One of the most important problems facing free energy calculations by computer simulations for complex systems such as proteins and nucleic acids is the need to enhance the search of their configurational space. One characteristic of such systems is a broad range of energy barriers at many scales, both lower and higher than the thermal energy. The ergodic hypothesis [1] relies on the assumption that equilibrium time averages are equal to the corresponding thermodynamic ensemble averages. As a consequence, every point in the phase space must be accessible from every other point.

For example, when calculating free energy using umbrella sampling (or potentials of mean force along a reaction coordinate), the requirement that neighboring windows should overlap relates to ergodicity in the sense that neighboring regions must be accessible from each other. For complex systems, the ergodic hypothesis is broken on the time scale of conventional simulations since various regions of their configuration space become disconnected (separated by large free energy barriers) and configuration points trapped in such regions have their own invariant probability distributions. Therefore, for such systems there is an acute need for methods with enhanced configurational sampling to solve the problem of broken ergodicity [2].

Most illustrative in this regard is the work of Hodel et al. [3] on the free energy errors caused by insufficient sampling. Even for a relatively small biomolecular system such as a nine-residue peptide loop with anchored ends, improper sampling of conformational substates caused by broken ergodicity was shown to result in errors of the order of $1\,\mathrm{kcal\,mol^{-1}}$, which was estimated to be about 50% of the total free energy difference. Importantly, this error was much larger than the statistical error, the sample-size hysteretic error, and the systematic 'error' due to changing the force field. It then becomes imperative for a reliable estimation of the free energy that underlying equilibrium distributions in the configuration space be generated by employing existing, or devising new, methods that increase the rate of conformational sampling.

An overreaching theme of the present chapter, besides broken ergodicity, has to do with the fact that most of the enhanced sampling methods that we shall discuss address situations in which one cannot clearly identify a reaction coordinate that can be conveniently used to describe the kinetic evolution of the system of interest. While methods for enhanced sampling are designed to yield accurate results faster than regular molecular dynamics or Monte Carlo (MC) methods, it is our belief that there is no 'perfect' method, but that, rather, there are methods that perform better for particular applications. Moreover, it should be noted that, while in instances when a proper reaction coordinate can be identified methods described in other chapters are probably more efficient, they could still benefit by sampling in conformational directions perpendicular to the reaction coordinate.

This chapter begins by defining a means of measuring how efficiently conformational space is explored, then introduces, in the next section, the need for and the general approach to enhanced conformational sampling. The following sections then describe selected methods for enhanced sampling, both for classical and quantum systems. They form by no means an exhaustive list, but rather one that tries to selectively (and subjectively) mix a few established methods with several new developments to give an overview of this active research area and to stimulate further ideas and possible future developments from the reader. We have tried to impart a sense of cohesion in the book by frequent references to other chapters.

8.2 Measuring Ergodicity

As discussed in the example from the work of Hodel et al. [3], one of the most efficient ways to improve the accuracy of free energy calculations with a given force field is to enhance the conformational sampling. Thus, it is important to assess the extent to which phase space is covered.

The focus of this section is to answer the question of how one can know how thoroughly the space of conformations is sampled in a particular simulation. To quantify the extent of sampling, we shall introduce measures in conformational space that use self-averaging. We shall present the fluctuation metric and the energy metric as ergodic measures, and will give examples of the exploration of the conformational space of atomic clusters. In the context of free energy calculations, these ergodic measures are recommended to gauge the contribution of conformational sampling to the convergence of the free energy values.

The ergodic measure estimates the rate of self-averaging in an equilibrium MC or molecular dynamics simulation [4–6]. Self-averaging is a necessary (but not a sufficient) condition for the ergodic hypothesis to be satisfied. The rate of self-averaging for a given property is expected to be proportional to the rate of conformation space sampling. Let us consider the potential energy metric defined for two independent trajectories α and β. We shall use a *move average* employing the potential energy U. If the system of interest is inhomogeneous, as is the case for biomolecules for instance, U is typically chosen to be the nonbonded energy. We define the potential

U for the jth particle along the α trajectory after n moves exploring space according to a known equilibrium distribution to be

$$u_j^\alpha(n) = \frac{1}{\sum\limits_k w(\mathbf{x}_k)} \sum\limits_k^n w(\mathbf{x}_k) U_j(\mathbf{x}_k).$$

(8.1)

Here, $w(\mathbf{x}_k)$ is the weighting factor for any property at a given position on the kth step $\mathbf{x}_k$. For example, for a constant-temperature molecular dynamics or a Metropolis MC run, the weighting factor is unity. However, we wish to leave some flexibility in case we want to use non-Boltzmann distributions; then, the weighting factor will be given by a more complicated function of the coordinates. The ergodic measure is then defined as a sum over N particles

$$d_U(n) = \frac{1}{N} \sum_j \left[u_j^\alpha(n) - u_j^\beta(n) \right]^2.$$

(8.2)

For an ergodic system, if the simulation length $n \to \infty$, then $d_U(n) \to 0$. By analogy with molecular dynamics, for large n we expect the form of the convergence to be [5]

$$d_U(n) = d_U(0) \frac{1}{D_U n}$$

(8.3)

where D_U is a rate for self-averaging of U over the two independent trajectories. In other words, the inverse of the ergodic measure, $d_U(0)/d_U(n)$, goes to infinity linearly with the simulation time n, that is, it goes to infinity diffusively, and the 'diffusion constant', D_U, characterizes how fast a particular sampling is 'diffusing' in conformation space.

Therefore, we associate rapid and effective sampling of phase space with a large value of D_U. The choice of the potential energy as a quantity to self-average in the metric is arbitrary. However, it has been shown to be a good measure of the extent of sampling in a variety of systems. For MC algorithms to calculate free energies (using, for example, umbrella sampling), parameters that maximize D_U should be chosen. Additionally, D_U can be used to compare the efficiency of distinct free energy methods.

8.3 Introduction to Enhanced Sampling Strategies

When calculating free energies, one generates, either by molecular dynamics or MC, configuration space samples distributed according to a probability distribution function (e.g., the Boltzmann distribution in the case of the Helmholtz free energy).

As explained above, simulating systems whose phase space is partitioned by 'broken ergodicity' is a significant challenge. In such cases, any average that is calculated effectively 'breaks' into sums of subaverages, with each subaverage taken over a subset of the phase space. The problem is acute in complex disordered systems, in which the potential energy surface is rugged and regions of configurational

space may be separated by energy barriers much greater than the thermal energy. For many systems of interest, such as proteins and glasses, the time scale for functionally important motions greatly exceeds that of molecular dynamics simulation [7]. Enhanced sampling algorithms that increase the frequency of barrier crossing and, by doing so, allow for an accelerated search of phase space are essential if reliable equilibrium averages are to be computed. Because accurate free energy calculations of the type described in this book invariably require good conformational sampling as a necessary condition for convergence, it is advisable that such enhanced sampling strategies are employed.

An important property of an enhanced sampling distribution is that it should include a significantly enhanced probability of visiting barrier regions or making moves in which the system crosses these barriers. In systems such as biomolecules and glasses, it is difficult to make such nonlocal moves. A number of advances in MC methodology that address the problem of broken ergodicity [8] have been reported in recent years. Rossky et al. [9] proposed the use of Brownian dynamics as a smart way of doing MC simulations. Cao and Berne [10] have developed an *anti-force bias* MC method, in which the system is encouraged to move toward minima in a convex region of the potential surface, or over barriers in a nonconvex region of the potential surface. The algorithm leads to accelerated barrier crossing, which may be an infrequent event. Frantz et al. [11] proposed the J-walking method, which uses a high-temperature MC run to generate trial moves. At high temperatures barriers may be crossed easily, overcoming problems of broken ergodicity. The trial moves are accepted so as to compute averages at a lower temperature of interest. The J-walking method was employed by Tsai and Jordan [12] to examine phase changes in small rare gas and water clusters. A very powerful method, applied widely to a variety of systems, is the parallel tempering method [13–15], in which walks at various temperatures are used. Multicanonical MC [16, 17] and cluster move methods [18] have been developed to address the problems of critical slowing down associated with phase transitions. Some of these methods and many others have been successful in simulations of biomolecular systems [19] with only a moderate computational overhead.

8.4 Modifying the Configurational Distribution: Non-Boltzmann Sampling

We start by reminding the reader of the original and seminal ideas of Torrie and Valleau, introduced in Chap. 3. In doing so, we thereby prepare for the subsequent sections in this chapter.

To handle broken ergodicity in the calculation of *thermodynamic* averages (i.e., time-independent averages of the type $\langle A \rangle \propto \int A(\mathbf{x}) \exp(-\beta E(\mathbf{x})) \mathrm{d}\mathbf{x}$) a host of methods have been devised [2]. In a subset of these, the Boltzmann distribution is altered and replaced with a more delocalized one, $w(E)$. One is thus able to generate (faster) samples distributed according to w and subsequently, one can use importance

sampling manipulation [20] to obtain the average corresponding to the unaltered system. Examples that improve upon using the venerable umbrella sampling technique are the Blue Moon ensemble method [21] or the scaled force algorithm [22]. It is of note that importance sampling manipulations in the context of umbrella sampling have been used for free energy calculation from the early stages of the development of such ideas [23]. The reason was similar, although not identical, to the broken ergodicity argument. When calculating free energy differences between two states, there needs to be a good overlap between the distributions of conformational points corresponding to the two states. During regular Boltzmann sampling, such an overlap is unlikely (unless the two states are quite similar). Therefore, nonphysical sampling was introduced to generate (i.e., broaden) the overlap (of course correcting for the nonphysicality afterwards), much like nonphysical enhanced sampling is used to broaden the distributions and make visiting barrier regions more likely than in the case of physical sampling. There are several ingenious ways to enhance sampling using nonphysical distributions. We present the details of just two methods in the following sections.

8.4.1 Flattening the Energy Distribution: Multicanonical Sampling and Related Methods

As detailed in Chap. 3, the multicanonical MC method [16, 17] has gained widespread interest and has been applied to a variety of systems. While most of the detailed implementation is covered in that chapter, here we review it briefly to frame it in the more general context of generating nonphysical distributions as means for enhanced conformational sampling. Multicanonical MC sampling, and its twin, entropy sampling MC [24], aim to carry out MC simulation with a uniform energy probability distribution. In addition to improving sampling, the approach has the benefit that the temperature dependence of the energy, the entropy, and other physical quantities are obtained at the end of the simulation. Multicanonical samples were originally generated by MC, but molecular dynamics and hybrid MC algorithms can also be used [19, 25]. In multicanonical sampling, conformations with energy E [where E is the potential energy if MC is employed, or the total energy if constant-temperature MD (or hybrid MC) are used], are assigned a 'multicanonical' weight,

$$w(E) \propto 1/n(E) = \exp\left[-S(E)\right], \tag{8.4}$$

instead of the canonical weight $\exp(-\beta E)$. Here, $n(E)$ is the density of states and $S(E)$ is the microcanonical entropy. With this choice of the weights $w(E)$ (which need to be precalculated), the distribution of energies $P(E)$ is given by

$$P(E) \propto n(E)w(E) = \text{const.} \tag{8.5}$$

As a result, in multicanonical simulation one-dimensional free diffusion in the energy space is generated, thereby permitting unhindered escape over energy barriers. As with other nonphysical distribution algorithms, one calculates the thermodynamic average of any physical observable by applying a reweighting factor, in

this case, $\exp(-\beta E)/w(E)$. The free diffusion in the one-dimensional energy space means that the 'nonphysical' probability density function generated through multicanonical sampling is simply a uniform distribution in energy [viz. (8.5)]. Convenient uniform probability distributions in thermodynamic functions other than the energy E have been proposed. Examples are random walks generating uniform distributions in temperature space [26], (i.e., $p(T) = \text{const.}$) and microcanonical-entropy space [27] ($p(S) = \text{const.}$). A comment should be made here regarding the comparison between multicanonical methods and methods relying on modification of the underlying dynamics through the Hamiltonian (e.g., adaptive umbrella sampling; see Chaps. 3 and 4). The latter yield a flat probability distribution function along a predefined order parameter(s), a predefined, low-dimensional manifold. For example, in the multidimensional adaptive umbrella sampling method [28] or in the adaptive biasing force (ABF) method [29], which have been shown to provide rapid and accurate free energy profiles for peptides and small proteins, the underlying Hamiltonian

$$\mathscr{H} = \mathscr{H}^0 + V(\xi_1, \xi_2, \ldots, \xi_s),\tag{8.6}$$

consists of the unperturbed Hamiltonian $\mathscr{H}^0$ of the system plus an umbrella potential V, which is a function of s important degrees of freedom, i.e., $\xi_1, \xi_2, \ldots, \xi_s$, for which we desire to have a uniform sampling distribution. The umbrella potential is determined adaptively during the simulation. The correct potential of mean force along $\xi_1, \xi_2, \ldots, \xi_s$ is recuperated in the adaptive umbrella simulation by the use of the WHAM method [30], but this is not required in the ABF simulations.

As in multicanonical sampling, a general feature of enhanced sampling methods is to 'spread out' the distributions in conformational space, such that the probability of visiting the barrier region relative to minima is enhanced. While such methods are able to achieve significant speed-up in the exploration of space, it is important to note that an optimum compromise between the exploration of space and the convergence of equilibrium properties must be achieved if thermodynamic averaging is desired. If a multidimensional configurational space becomes flat and the system wanders through this space aimlessly, it only occasionally visits its relevant parts.

Here an additional distinction is to be made between thermodynamic averages of a *conformational* observable such as the internal energy, which converges well if potential minima are correctly sampled, and *statistical* properties such as free energies, which depend on the entire partition function.

Consider, for instance, the free energy $A(\xi)$ projected onto important degrees of freedom. This is a potential of mean force along ξ and is obtained by integrating out all the other degrees of freedom, $\{\mathbf{q}\}$. It is related to the probability of finding the system at a particular value of ξ by

$$\mathrm{e}^{-\beta A(\xi)} = \int \mathrm{e}^{-\beta V(\xi, \{\mathbf{q}\})} \mathrm{d}\{\mathbf{q}\}.\tag{8.7}$$

Since the potential of mean force is a *statistical* property, it is insufficient to calculate it directly by importance sampling which, by design, emphasizes potential minima

and samples the maxima less frequently. This can be seen by expressing the potential of mean force in (8.7) for a value ξ_M in a maximum-energy region as an ensemble average, i.e., an average that can be accumulated in a direct sampling run in the canonical ensemble

$$A(\xi_M) = \frac{1}{\beta} \ln \langle e^{\beta V(\xi_M, \{\mathbf{q}\})} \rangle. \tag{8.8}$$

For a value ξ_M in the barrier region, the largest contribution to $A(\xi_M)$ comes from terms with large $V(\xi_M, \{\mathbf{q}\})$, which are exponentially unlikely to occur. To generate correct free energy profiles, it is most efficient to use *indirect* methods that force the system to sample regions which would not be sampled in regular methods that directly sample the Boltzmann distribution at room temperature. In contrast, the poor sampling in the barrier regions exhibited by the direct methods has little effect on *conformational* equilibrium properties as long as the important low-energy regions are sampled with the correct relative weight, such as in the smart darting method described in Sect. 8.6.

8.4.2 Generalized Statistical Sampling

As one of the possible ways to alter the sampling distribution in a manner that is conducive to enhanced sampling, we present a strategy based on probability distributions that arise in a generalization of statistical mechanics proposed by Tsallis [31]. In this formulation, the *generalized entropy* for an N-body system is defined as [31, 32]

$$S_q = \frac{k}{q-1} \int p_q(\mathbf{x})(1 - [p_q(\mathbf{x})]^{q-1}) d\mathbf{x}, \tag{8.9}$$

where q is a real number and S_q tends to the Gibbs–Shannon entropy $S = -k \int p(\mathbf{x}) \ln p(\mathbf{x}) d\mathbf{x}$ when $q = 1$. To derive the configurational probability distribution function the generalized entropy is optimized subject to the constraints

$$\int p_q(\mathbf{x}) d\mathbf{x} = 1 \qquad \text{and} \qquad \int [p_q(\mathbf{x})]^q V(\mathbf{x}) d\mathbf{x} = V_q \tag{8.10}$$

where $V(\mathbf{x})$ is the potential energy and V_q is the Tsallis ensemble average internal energy. The probability of a point in configuration space is found to be

$$p_q(\mathbf{x}) = \frac{1}{Z_q} \left[1 - (1-q)\beta V(\mathbf{x})\right]^{\frac{1}{1-q}} \tag{8.11}$$

where

$$Z_q = \int \left[1 - (1-q)\beta V(\mathbf{x})\right]^{\frac{1}{1-q}} d\mathbf{x} \tag{8.12}$$

is the generalized configurational partition function. Most importantly, note that the generalized probability depends as a power law on the energy, which is weaker than the well-known exponential dependence in the Boltzmann distribution. As a result, for $q > 1$, the generalized ensemble is more delocalized in conformation space than the canonical ensemble.

One can perform a MC simulation based on the acceptance probability

$$p = \min \left[1, \left(\frac{p_q(\mathbf{x}_{\text{new}})}{p_q(\mathbf{x}_{\text{old}})} \right)^q \right]. \tag{8.13}$$

Because it is broader, the equilibrium distribution $[p_q(\mathbf{x})]^q$ with $q > 1$ can be sampled more effectively than in the standard Metropolis MC, which corresponds to $q = 1$ [33]. Subsequently, equilibrium averages can be calculated more effectively using a regular reweighting method based on standard importance sampling manipulations [20]. The reason for enhanced sampling becomes clear if one analyzes the simple one-dimensional example of a harmonic oscillator. Substituting $V = kx^2$ into (8.11), one obtains a Gaussian distribution for $q = 1$, and a broader, longer-tailed Cauchy–Lorenz distribution in x-space for $q = 2$.

Note that by defining the effective potential

$$\bar{V} = \frac{q}{\beta(q-1)} \ln \left[1 - (1-q)\beta V \right], \tag{8.14}$$

the MC acceptance probability equation (8.13) can be written in the familiar form

$$p = \min \left[1, \exp \left(-\beta \Delta \bar{V} \right) \right]. \tag{8.15}$$

The standard Metropolis MC corresponds to the $q = 1$ limit, in which case the probability of accepting a new configuration of the system is

$$p = \min \left[1, \exp \left(-\beta \Delta \mathrm{V} \right) \right], \tag{8.16}$$

where $\Delta V = V(\mathbf{x}_{\text{new}}) - V(\mathbf{x}_{\text{old}})$.

Also note that the definition of the effective potential $\bar{V}$ in (8.14) enables one to conceive a constant-temperature molecular dynamics method (instead of MC) to generate the Tsallis distributions. Given this effective potential, it is possible to define a constant-temperature molecular dynamics algorithm such that the distribution $P_q(\mathbf{x})$ is sampled in the trajectory. The equation of motion takes on the simple and suggestive form

$$m_k \frac{\mathrm{d}^2 \mathbf{x}_k}{\mathrm{d}t^2} = -\nabla_{\mathbf{x}_k} \bar{V} = -\nabla_{\mathbf{x}_k} V(\mathbf{x}) q [1 - (1-q)\beta V(\mathbf{x})]^{-1} \tag{8.17}$$

for a particle of mass m_k and position $\mathbf{x}_k$ and $\bar{V}$ defined by (8.14). The effective force derived from the effective potential $\bar{V}(\mathbf{x})$ has a number of interesting properties. It is

of the form $F_q(\mathbf{x}; \beta) = -\nabla_{\mathbf{x}_k} \bar{V} = F_1(\mathbf{x})\alpha_q(\mathbf{x}; \beta)$, where $F_1(\mathbf{x})$ is the 'exact' force for standard molecular dynamics ($q = 1$) and $\alpha_q(\mathbf{x}; \beta)$ is a scaling function, which is unity when $q = 1$ but can otherwise have a strong effect on the dynamics.

Assume that the potential is defined to be a positive function. In the regime $q > 1$, the scaling function $\alpha_q(\mathbf{x}, \beta)$ is largest near low-lying minima of the potential. In barrier regions, where the potential energy is large, the scaling function $\alpha_q(\mathbf{x}, \beta)$ is small. This has the effect of reducing the magnitude of the force in the barrier regions. Therefore, a particle attempting to pass over a potential energy barrier will meet with less resistance when $q > 1$ than when $q = 1$. At equilibrium, this leads to more-delocalized probability distributions with an increased probability of sampling barrier regions. This argument demonstrates that, when $q > 1$, generalized molecular dynamics trajectories will cross barriers more frequently and explore phase space more efficiently. Here, an interesting connection can be made to the 'scaled force' version of the ABF method [29]. Instead of subtracting the force acting on the reaction coordinate in the ABF method, one can scale it (e.g., by α_q) to achieve exactly the same effect – 'flattening' the configurational space. Here we see that two different methods, which have different theoretical underpinnings, could lead to the same effective result.

Enhanced sampling in conformational space is not only relevant to sampling classical degrees of freedom. An additional reason to illustrate this particular method is that the delocalization feature of the underlying distribution in Tsallis statistics is useful to accelerate convergence of calculations in *quantum* thermodynamics [34]. We focus on a related method that enhances sampling for quantum free energies in Sect. 8.4.2.

In summary, an MC walk generated in algorithms based on sampling generalized distributions (or a constant-temperature molecular dynamics trajectory) allows one to obtain a well-defined statistical distribution that is broader than its Boltzmann counterpart. As usual, appropriately weighted averages lead one to calculate equilibrium thermodynamic averages for the Gibbs–Boltzmann canonical ensemble. We refer the reader to [35] for an application to a test case for conformational sampling, a 13-atom cluster modeled by Lennard-Jones interactions, which demonstrate the enhancement in sampling and efficiency in the computation of equilibrium thermodynamic averages.

Another way to think about the reason for enhancement brought about by generalized sampling is to cast it as a general approach of scaling down the energy barriers. Not just any kind of scaling will do, however. The transformation $V \to \bar{V}$ in (8.14) has the desirable feature of leaving invariant the position of minima and maxima. For reasons of efficiency when converging to thermodynamic averages, one would still like to visit minima more often than higher-energy states, even on the transformed potential. While in thermodynamic integration (TI) and thermodynamic perturbation methods the integration variable does constrain, by design, the system along various high-energy intermediates along the pathway between which the free energy is computed, thorough sampling in the degrees of freedom perpendicular to it still necessitates excursions to the important low-energy states. The desirable feature of preserving the places where sampling is supposed to matter in the case of $V \to \bar{V}$
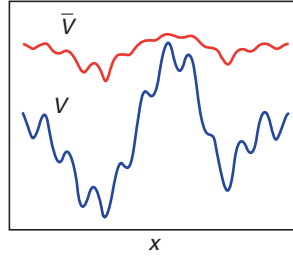
Fig. 8.1. The Tsallis-transformed effective potential $\bar{V}$ is smoother than the physical, untransformed potential V and sampling on it is enhanced; stationary points of any order preserve their x location

translates to preserving the location of all stationary points on the potential energy surface (see Fig. 8.1).

It is pedagogically interesting to compare the Tsallis sampling algorithm with the multicanonical sampling method [16, 17]. In the latter, a random walk in energy is the essential feature used to enhance sampling, with success in a number of applications. It has been shown [36], however, that in the thermodynamic limit the multicanonical algorithm is identical to the regular Metropolis scheme. The reasoning is that, in a system with a large number of particles N, the entropy $S(E)$ is a smooth function of the energy E, and can thus be expanded to first order in ΔE. In this case, the acceptance probability of the multicanonical sampling update becomes

$$\lim_{N \to \infty} \exp(-\Delta S) = \exp\left(-\frac{\partial S}{\partial E}\Delta E\right) = \exp\left(-\frac{\Delta E}{T}\right), \qquad (8.18)$$

where we have used the equality $\partial S/\partial E = 1/T$.

In the thermodynamic limit, the Tsallis updating scheme has the form

$$\lim_{N \to \infty} \exp\left(-\frac{\Delta \bar{E}}{T}\right) = \left(\frac{E}{E + \Delta E}\right)^{\frac{q}{1-q}}, \qquad (8.19)$$

which is basically unity since the energy change ΔE is local (i.e., small) and E is of order N.

It is ironic that in the $N \to \infty$ limit the Tsallis sampling algorithm has the main feature for which the multicanonical algorithm was designed – it performs a random walk in energy – while the multicanonical algorithm loses this feature.

8.5 Methods Based on Exchanging Configurations: Parallel Tempering and Related Strategies

A host of successful methods have been devised that use two or more replicas of the system run in parallel and corresponding to different simulation parameters. The enhanced equilibrium averaging is achieved by Metropolis-type acceptance–rejection

schemes for the swaps between the replicas indexed by these parameters. The swaps can be attempted in a random or systematic (i.e., periodic) fashion. The basic idea of these methods relies on the fact that, at different values of the parameter, the acceptance probabilities (or trajectories) are different. For example, higher acceptance probabilities at larger parameter values can be used as an efficient means of transport over barriers, whereas those at parameter values closer to those of the real physical system provide good local sampling of the minima.

One of the most popular such techniques is parallel tempering in the canonical ensemble, for which the index parameter is temperature [13–15]. While parallel tempering (or replica exchange) strategies had been independently proposed on multiple occasions in various scientific areas, perhaps the earliest seed of the idea can be found in early work by Swendsen and Wang [37].

8.5.1 Theory

For clarity, let us present the details of the swap process in a temperature-based parallel tempering algorithm. Generally, parallel tempering involves the implementation of the following iterative steps:

1. Run M replicas (walkers) in parallel at increasing temperatures

$$T_1, T_2, \ldots, T_i, \ldots, T_j, \ldots, T_M,$$

   where the temperature of interest is usually the lowest one.
2. Pick at random two temperatures T_i and T_j and attempt a swap $\mathbf{x} \rightleftharpoons \mathbf{x}'$ between the configuration $\mathbf{x}$ of a walker at T_i and the configuration $\mathbf{x}'$ of a walker at T_j according to a predefined acceptance probability [cf. (8.22)].
3. If the swap is rejected, count in the current conformation. If the swap is accepted, assign to the configuration of walker i temperature T_j and vice versa.
4. Repeat until convergence.

Assuming that after their previous swap the two walks were sufficiently long to be in the asymptotic regime, this means that transient behavior has elapsed and the system has relaxed to equilibrium for the respective parameters. Then, the joint configurational probability density just before the current swap is simply

$$p^{(2)}(\mathbf{x}, \mathbf{x}') \propto \exp\left(-\frac{V(\mathbf{x})}{k_\mathrm{B} T_i}\right) \exp\left(-\frac{V(\mathbf{x}')}{k_\mathrm{B} T_j}\right). \tag{8.20}$$

The probability to swap, $W(\mathbf{x} \rightarrow \mathbf{x}', \mathbf{x}' \rightarrow \mathbf{x})$, has to obey detailed balance

$$p^{(2)}(\mathbf{x}, \mathbf{x}')W(\mathbf{x} \rightarrow \mathbf{x}', \mathbf{x}' \rightarrow \mathbf{x}) = p^{(2)}(\mathbf{x}', \mathbf{x})W(\mathbf{x}' \rightarrow \mathbf{x}, \mathbf{x} \rightarrow \mathbf{x}') \tag{8.21}$$

and therefore the swap should be accepted with probability

$$\min\left\{1, \exp\left[\left(\frac{1}{k_\mathrm{B} T_i} - \frac{1}{k_\mathrm{B} T_j}\right)(V(\mathbf{x}') - V(\mathbf{x}))\right]\right\}. \tag{8.22}$$

In this way, stationarity of the probability distribution in configuration space with respect to swapping is enforced [38], with the advantage that the mixing resulting from the swaps allows the lowest-temperature walker to equilibrate more rapidly by coupling it to the faster-equilibrating walker at a higher temperature. In practical terms, as can be verified from the acceptance formula above, if a swap is attempted between two widely different temperatures T_i and T_j, the acceptance probability is quite low, so one uses swaps between systems with small temperature differences, usually adjacent to each other, i.e., with $|i - j| = 1$. With adjacent swapping, acceptance of configuration swaps will be significant if the energy probability distributions of the two systems (at T_i and T_{i+1}) overlap significantly. Since for large systems the relative spread of the probability distributions decays as $\sqrt{N}$, the step in temperature $\Delta T_i = T_{i+1} - T_i$ will need to be small, and many replicas are needed [i.e., the number of replicas M is of $\mathcal{O}(\sqrt{N})$]. Simulating M systems of size N for a given number of steps obviously takes M times more effort than simulating one system only. However, despite this M-fold increase in system size, parallel tempering is, in general, still efficient because proper sampling at a single temperature usually takes more than M times longer.

How many temperatures should the replicas be spread over, and what should be the highest? In the close-kin method of simulated annealing, an optimal schedule for changing the temperature from high to low is sought in the hope of finding the global energy minimum without the simulation getting trapped in intermediates. Quite similarly for parallel tempering, it is of substantial importance to find an optimal assignment for the replica temperatures. One approach involves tending to achieve an equal acceptance probability for all swaps. Under the approximation of constant heat capacity, Kofke [39] has shown that the asymptotic value of the acceptance depends solely on the ratio T_{i+1}/T_i, which would imply a geometric progression of temperatures. Predescu et al. [40] have expressed, assuming negligible correlation between successive swaps, the acceptance probability as a function of the effective fraction, defined as the expected probability that a configuration from the lowest-temperature replica successfully reaches the highest-temperature one. This represents an adequate measure of the quality of a parallel tempering technique, as far as swapping is concerned. Refinement of the simulation protocol has also been achieved by adaptively changing the allocation of temperatures during the simulation to obtain a desired trial exchange acceptance probability [41, 42].

As a result of stationarity, the partition function of the replicated ensemble can be written

$$Q(T_1, T_2, \dots T_M) = \prod_{i=1}^{M} (\Lambda_i^3)^{-1} \int d\mathbf{x} \, e^{-V(\mathbf{x})/k_B T_i}, \tag{8.23}$$

where $\Lambda_i = \prod_{k=1}^{N} N(h^2/2\pi m_k k_B T_i)^{1/2}$ involves the products of the thermal de Broglie wavelengths for the N particles, each of mass m_k, that comprise the replica at temperature T_i.

A useful feature of the multiple-temperature dependence of the partition function $Q(T_1, T_2, \dots, T_M)$ in the equation above is that it can be subsequently manipulated

not only to derive thermodynamic properties at the 'physical' temperature of interest, but also to map temperature dependencies across the whole range of utilized temperatures.

Molecular dynamics has also been used to replace the MC moves for conformational advancement [43]. In the molecular dynamics version of parallel tempering, often referred to as replica exchange molecular dynamics, momenta are used in the propagation scheme such that a constant temperature is maintained between the swaps. After the swap in conformational space (with the same acceptance criterion as in the MC implementation), a readjustment in momentum space is also needed. This is done by renewing the momenta for replica i by the transformation $\mathbf{p}_i^{\text{new}} = \sqrt{T^{\text{new}}/T^{\text{old}}}\mathbf{p}_i^{\text{old}}$.

While between the swaps the motion of the system is somewhat realistic, it is important to emphasize that the swaps between two temperatures are nonphysical. This therefore destroys the sequencing of dynamical events (that would be required to calculate, for example, time correlation functions) and renders the dynamics and kinetics artificial.

8.5.2 Extensions

Instead of attempting swaps between adjacent replicas, an alternative exchange strategy for parallel tempering simulations has been proposed. In the all-exchanges parallel tempering method of Calvo [44], the acceptance probabilities of all possible swap moves are calculated a priori. One specific swap move is then selected according to its probability and enforced. The efficiency of the method was illustrated in the case of Lennard-Jones clusters. Judging by the convergence of the caloric curve for Lennard-Jones particles, the scheme appears more than twice as fast as conventional parallel tempering.

The index parameter can be energy as well, which is the case for parallel tempering in the microcanonical ensemble [45]. Other methods have been devised using as the index parameter the chemical potential, as in the hyper-parallel tempering method [46], a delocalization parameter, as in q-jumping [47] (see also Sect. 8.4.2) and generalized parallel tempering [48–50], or suitable modifications of the potential, as in the so-called Hamiltonian replica exchange method [51, 52]. Multicanonical ensembles have also been used in the context of parallel tempering [53, 54], which increased the efficiency of the algorithm relative to regular parallel tempering by decreasing the number of replicas needed to be simulated.

8.5.3 Selected Applications

While applications of parallel tempering to various fields ranging from materials science to chemistry to statistical physics abound, we would like to showcase here its use in biomolecular conformational sampling, and in particular to protein folding. Exciting in its own right, this problem is also prototypical for cases where parallel tempering is expected to be most usefully employed, i.e., when the underlying energy surface of the system is rugged. Biomolecular applications originated with the work

by Hansmann [55] who used MC-based parallel tempering on the Met-enkephalin peptide with encouraging results.

Molecular-dynamics-based parallel tempering methods (also known as replica exchange molecular dynamics) have also been used to address the energetics behind, and the detailed nature of, the protein folding mechanism. An example is the work of Garcia and Onuchic [56], who used replica exchange MD for the study of a three-helix bundle protein at atomic resolution over a wide range of temperatures and sampling both unfolded and folded states, to obtain the free energy, entropy, and enthalpy surfaces along structural folding coordinates. The observed multitude of the transitions between all minima on the free energy surface enabled a quantitative determination of the free energy barriers and the ensemble of configurations associated with the underlying folding intermediates.

Other exciting applications involved using parallel tempering in connection with available experimental data. For example, Falcioni and Deem [57] used X-ray data to refine structures of zeolites, and Haliloglu et al. [58] refined NMR structural data for proteins (in particular using residual dipolar coupling constraints).

8.5.4 Practical Issues

An issue to consider in practical implementations for various systems is size consistency: when the control parameter, in this case the temperature T, is an intensive variable, it has to change less between replicas as the system size increases, so that significant acceptance for swaps is assured. Another observation is that the configurations in the underlying distributions obey the correct (equilibrium) statistics, but dynamics inferred from the walkers may be nonphysical. An important advantage of parallel tempering is that there is no need to define a configurational order parameter. On the other hand, configurational order parameters are useful for approximating kinetics, i.e., they are variables which could be advanced by dynamical equations of motion, thereby imparting a time evolution to the system. This is discussed in depth in Chap. 4.

Another practical limitation in complex applications lies in the fact that, if temperature is used as a control parameter, one needs to worry about the integrity of a system that is heated too much (e.g., water–membrane systems or a protein heated above its denaturation temperature). When issues such as those mentioned above are addressed, parallel tempering can be turned into a powerful and effective means of enhanced conformational sampling for free energies over a range of temperatures for various systems.

For an excellent overview of several other applications of parallel tempering, as well as for details on the pertinent questions to be addressed in practical implementations, the reader is referred to a recent review by Earl and Deem [59].

8.5.5 Related Methods

A method related to parallel tempering is J-walking [60], in which, as in parallel tempering, configurations from a high-temperature walk are used to make a low-temperature walk ergodic. The J-walking strategy involves the high-temperature

walk feeding configurations to the low-temperature walk, rather then the high- and low-temperature walkers exchanging configurations as in parallel tempering. The relationship between J-walking and parallel tempering has been analyzed in [61].

Improvements to parallel-tempering-type algorithms have been proposed. For example, the generalized Tsallis distributions discussed earlier in this chapter have been used to enhance the sampling properties of parallel tempering methods. Other examples are adding additional dimensions for the walker to wander, as in the catalytic tempering of Berne et al. [62], which is reminiscent of Purisima and Scheraga's original ideas on increasing system dimensionality developed for global optimization [63] (see also [64]).

8.6 Smart Darting and Basin Hopping Monte Carlo

In some instances, we have prior knowledge of states of the system that are thermodynamically meaningful. Then we can take advantage of such information and generate the proper samples that allow, for instance, the calculation of the relative free energy of such states. Let us reconsider the partition function for the ensemble of states for N distinguishable particles in three dimensions,

$$Q(\beta) = \Big(\prod_{k=1}^{N} \Lambda_k^3 \Big)^{-1} \int d\mathbf{r}\, e^{-\beta V(\mathbf{r})} = \Big(\prod_{k=1}^{N} \Lambda_k^3 \Big)^{-1} Z(\beta), \qquad (8.24)$$

where $\Lambda_k = (h^2\beta/2\pi m_k)^{1/2}$ is the thermal de Broglie wavelength for a particle of mass m_k, $V(\mathbf{r})$ is the potential energy, and $Z(\beta)$ is the configuration integral. In the inherent structure picture of statistical mechanics, proposed by Stillinger and Weber [65, 66], the $3N$-dimensional configuration space is decomposed into a set of basins of attraction. Any point in configurational space (excluding maxima, saddle points and ridges) will be mapped to a minimum on the potential surface (e.g., either by steepest descent or by some form of annealing to average over thermal motion). Labeling the basins i and calling $\mathscr{B}_i$ the regions of configuration space which form basins of attraction draining to the ith minimum, located at $\mathbf{R}_i$ of energy V_i, the configuration integral may be written

$$Z(\beta) = \sum_i Z_i(\beta) = \sum_i \int_{\mathscr{B}_i} d\mathbf{r}\, e^{-\beta V(\mathbf{r})}. \qquad (8.25)$$

The potential energy in the ith basin can be written $V(\mathbf{r}) = V_i + \Delta_i V(\mathbf{r})$ leading to

$$Z_i(\beta) = e^{-\beta V_i} \int_{\mathscr{B}_i} d\mathbf{r}\, e^{-\beta \Delta_i V(\mathbf{r})}. \qquad (8.26)$$

These are exact expressions for the configuration integrals. Alternatively, we can write the partition function as

$$Q(\beta) = \sum_i \exp[-\beta A_i] \qquad (8.27)$$

where A_i is the Helmholtz free energy of the ith basin. A method that allows for combining local sampling of the basins $\mathscr{B}_i$ with a convenient means of transportation between the basins would constitute a good approach towards the sampling of the configuration space. This is the spirit of the smart darting method described in this section. Using a search method (e.g., high-temperature MC or q-jumping MC [47]), pick configurations from which to perform steepest descents on the potential energy surface, generating a set of M configurations, corresponding to distinct local minima $\{\mathbf{R}_i\}_{i=1,2,\ldots,M}$. From this set of minima construct $M(M-1)$ displacement vectors, or 'darts,' of the type

$$\mathbf{D}_{ij} = \mathbf{R}_j - \mathbf{R}_i, \; i \neq j, \; i,j = 1, 2, \ldots, M \tag{8.28}$$

Also, choose a real number ϵ. Then define an ϵ-sphere around each element of the set $\{\mathbf{R}_i\}$

$$S_\epsilon(\mathbf{R}_i) = \{\mathbf{r} \mid \| \mathbf{r} - \mathbf{R}_i \| < \epsilon\}. \tag{8.29}$$

For efficient sampling, the value of ϵ should be chosen small in the following sense. The difference in the potential energy of any configuration within any $S_\epsilon(\mathbf{R}_i)$ and the configuration $\mathbf{R}_i$ of the local minimum of that ϵ-sphere should be much less than the thermal energy of a degree of freedom. The parameter ϵ should be chosen small enough that no two spheres overlap, otherwise the sampling procedure requires modification. There are two types of steps: jumps and local moves, which occur with probability P and $1 - P$, respectively. Therefore, during MC sampling done locally within a basin, check with probability P whether the current configuration $\mathbf{r}$ is in an ϵ-sphere and do one of the following two things:

(1) If, say, $\mathbf{r} \in S_\epsilon(\mathbf{R}_k)$, then randomly pick another local minimum, say the lth one, and jump to the $S_\epsilon(\mathbf{R}_l)$ sphere by the translation

$$\mathbf{r} \rightarrow \mathbf{r} + \mathbf{D}_{kl}. \tag{8.30}$$

Accept or reject the move according to the Boltzmann criterion.

(2) If $\mathbf{r}$ is outside any ϵ-sphere, then *count again* the current configuration $\mathbf{r}$ (i.e., reject an implicitly attempted jump along D_{kl} because it would land outside the ϵ-sphere for minimum l). The remainder of the simulation steps, on the average a fraction $1 - P$ of them, are local MC steps drawn from a uniform distribution and accepted or rejected according to the Boltzmann criterion.

For this algorithm, one can prove that detailed balance is guaranteed and the exact average of any configuration-dependent property over the accessible space is obtained. Two key issues determine the detailed balance. The first is the fact that the trial probability to pick the displacement vector $\mathbf{D}_{kl}$ to go from the kth to the lth ϵ-sphere equals the trial probability to pick the displacement vector $\mathbf{D}_{lk}$ for the reverse step. The second issue is that the trial probability for a local MC step that moves the walker from a point inside an ϵ-sphere to a point outside that sphere is the same as for the reverse move; i.e., $(1 - P)$ times what it would be in a walk restricted to local moves.

The smart darting method has its roots in the approach previously suggested by the smart walking method [67], and is related to the MC minimization [68] and newer basin-sampling methods such as the puddle jumping strategy [69]. The smart darting method has been recently improved and expanded [70] and has been combined with parallel tempering [71].

In real-world implementations, a practical issue is that, if the walker enters a new basin purely by local moves (as might happen at high temperatures), this must be recognized. For large M, efficient implementation of this monitoring may require neighbor-listing of the minima or some other form of bookkeeping, to prevent frequent searches over all M minima.

In general, methods that couple local to global moves, of the types presented in this section, are expected to depend critically on the ability to exhaustively map all thermodynamically significant basins, which is not always guaranteed. However, these methods are likely to be used profitably in particular applications where only a small, known set of conformations are relevant (e.g., in biomolecular applications where sets of structures are known from X-ray or NMR data).

8.7 Momentum-Enhanced HMC

This section is used to introduce the momentum-enhanced hybrid Monte Carlo (MEHMC) method that in principle converges to the canonical distribution. This ad hoc method uses averaged momenta to bias the initial choice of momenta at each step in a hybrid Monte Carlo (HMC) procedure. Because these average momenta are associated with *essential* degrees of freedom, conformation space is sampled effectively. The relationship of the method to other enhanced sampling algorithms is discussed.

One general approach to enhancing sampling, which is the focus of this section, is based on the fact that both fast and slow dynamical modes contribute to the time evolution of biomolecular systems, but in most cases the motions of primary interest are the slow ones, which typically correspond to the largest structural changes [72, 73].

While the *actual* time spent by an *ensemble* of molecules in the fast manifold is the same as that spent in the slow manifold, the *computer* time needed for convergence of properties in the slow manifold (when simulating a single molecule) is much larger than that spent in the fast manifold. Therefore, identifying the slow manifold and artificially accentuating the motion along it can decrease the amount of computational time required for the events of primary interest to occur and for statistical averages to converge.

The direction of slow motion can be found from normal-mode diagonalization. However, such a procedure would require, for a system of N particles, $\mathscr{O}(N^3)$ calculations. Instead, we would like to find the direction of the slow modes by a method that requires, for each evaluation of the slow direction, at most as many calculations as the evaluation of an energy, roughly an $\mathscr{O}(N^2)$ operation (if no cut-offs are used).

As we shall see below, a useful strategy to identify the slow manifold is to calculate an average of the momentum, $\mathbf{p}$, over τ time units during a molecular dynamics propagation

$$\bar{\mathbf{p}} = \frac{1}{\tau}\int_0^\tau \mathbf{p}(t)\mathrm{d}t. \tag{8.31}$$

In order for the averaged momentum $\bar{\mathbf{p}}$ to point along the slow manifold (i.e., for the components of momentum in the fast manifold to average to zero), one has to choose the averaging time τ so as to be several times longer than the period of the fast modes but shorter than those of the slow modes.

The choice of the momenta, rather than of the forces, is illustrated by the simple example of a two-dimensional harmonic oscillator described by the parametric equations

$$x = X\sin(\omega_x t + \phi_x), \quad y = Y\sin(\omega_y t + \phi_y). \tag{8.32}$$

If $\omega_x \gg \omega_y$, the motion in x is fast and that in y is slow. Differentiation with respect to t yields

$$p_x = P_x\cos(\omega_x t + \phi_x), \quad p_y = P_y\cos(\omega_y t + \phi_y) \tag{8.33}$$

and

$$f_x = F_x\sin(\omega_x t + \phi_x), \quad f_y = F_y\sin(\omega_y t + \phi_y). \tag{8.34}$$

Given equipartition of energy (i.e., $m\omega_x^2 X^2 = m\omega_y^2 Y^2$), the relative amplitudes of the displacements, momenta, and forces are

$$X \ll Y, \quad P_x = P_y, \quad F_x \gg F_y. \tag{8.35}$$

We observe in (8.35) that the magnitudes of the momenta in the fast and slow directions are comparable, in contrast to the magnitudes of the forces. Due to the fact that the instantaneous force in the fast direction is much larger in magnitude than that in the slow direction, the time-averaged $\bar{f}_x$ is comparable to $\bar{f}_y$, where the bar denotes time averaging. In other words, the average force in the fast direction converges to zero poorly because it involves differences of large numbers. In other words the forces in the fast manifold, while taking on both positive and negative values, can have absolute values several orders of magnitude larger than in the slow manifold, and therefore do not cancel out to the same level of precision as if their magnitudes were comparable to the slow ones. As a result, the guiding force does not necessarily emphasize the slow degree of freedom. However, the average momentum does. This is because the magnitude of the fast and slow momenta are comparable, and therefore the fast momenta will average to zero within the precision used for the slow momenta. We shall next present a momentum-based enhanced sampling method that rigorously converges to the canonical distribution under particular limiting assumptions.

To incorporate the guiding momentum in a straightforward manner and to ensure canonical sampling, we adapt an HMC scheme [74, 75] rather than molecular dynamics. At each step in an HMC simulation, random momenta are assigned, several steps of molecular dynamics are performed to move the system, and the move is accepted or rejected according to a Metropolis-like criterion [76] based on the change in total energy (which includes the kinetic energy). Because the time steps are much larger than those employed in MD, the discretization errors lead only to an approximate conservation of the total energy, but not to exact conservation. The implementation of the HMC method is described below together with the variant we introduce here. In brief, in the proposed self-guided scheme, the initial momenta at each step are selected with a bias towards the soft degrees of freedom and the acceptance criterion is modified accordingly to maintain detailed balance.

In the standard HMC method two ingredients are combined to sample states from a canonical distribution efficiently. One is molecular dynamics propagation with a large time step and the other is a Metropolis-like acceptance criterion [76] based on the change of the total energy. Typically, the best sampling of the configuration space of molecular systems is achieved with a time step of about 4 fs, which corresponds to an acceptance rate of about 70% (in comparison with 40–50% for Metropolis MC of pure molecular liquids).

In the standard HMC method, the $3N$ components of the vector $\mathbf{p}$ are usually drawn randomly from a Gaussian (Maxwellian) probability distribution

$$\mathscr{P}_M(\mathbf{p}) = C\mathrm{e}^{-\frac{1}{2}\mathbf{p}^T A\mathbf{p}}, \tag{8.36}$$

where $A = \beta M^{-1}$, $\beta = 1/kT$, M is the mass matrix, and C is a normalization factor. If the slow manifold in the $3N$-dimensional coordinate space were known, one could enhance sampling of that manifold by skewing the matrix A to increase the probability of larger components in the soft directions. Unfortunately, identification of the slow manifold would require frequent diagonalization of the covariance matrix of atomic fluctuations, which is not computationally efficient. However, it is possible to use this idea in an approximate fashion that does not require matrix diagonalization. The Maxwell distribution function (8.36) is a special case of the general multivariate Gaussian distribution

$$\mathscr{P}_G(\mathbf{p}) = C\mathrm{e}^{-\frac{1}{2}\mathbf{p}^T A\mathbf{p}-\mathbf{B}\mathbf{p}}, \tag{8.37}$$

where A is now a positive-definite symmetric matrix which determines the width along the eigendirections and $\mathbf{B}$ is a $3N$-dimensional row vector which determines the position of the maximum. We can choose $A = M^{-1}$ and $\mathbf{B} = \pm\xi\mathbf{p}_0^T$, where ξ is a multiplicative constant and $\mathbf{p}_0(t)$ is the average momentum over the past t_l time units

$$\mathbf{p}_0(t) = \frac{1}{t_l}\int_{t-t_l}^{t}\mathbf{p}\,\mathrm{d}\tau, \tag{8.38}$$

which, in the numerical implementation, is calculated from the velocities during the molecular dynamics propagation. In other words, one draws initial momenta at each step from the *bilobal* distribution

$$\mathscr{P}_B(\mathbf{p}) = C \left(\mathrm{e}^{-\beta \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} - \beta\mathbf{B}\mathbf{p}} + \mathrm{e}^{-\beta \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + \beta\mathbf{B}\mathbf{p}} \right), \qquad (8.39)$$

which is peaked around $\mathbf{B}$ and $-\mathbf{B}$. For $\mathbf{B}$ to point along the slow manifold (i.e., for the components of $\mathbf{B}$ in the fast manifold to be negligible), one has to choose the averaging time t_l to be larger than the period of the fast modes but smaller than those of the slow modes.

After the momenta are selected from the distribution (8.39), the dynamics is propagated by a standard leapfrog algorithm (any symplectic and time-reversible integrator is suitable). The move is then accepted or rejected according to a criterion based on the detailed balance condition

$$\mathscr{P}(\mathbf{x})\mathscr{W}(\mathbf{x} \to \mathbf{x}')\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' = \mathscr{P}(\mathbf{x}')\mathscr{W}(\mathbf{x}' \to \mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}', \qquad (8.40)$$

where $\mathscr{P}(\mathbf{x})$ is the equilibrium (Boltzmann) probability of state $\mathbf{x}$ and $\mathscr{W}(\mathbf{x} \to \mathbf{x}')$ is the transfer matrix element giving the probability that the system will go from $\mathbf{x}$ to $\mathbf{x}'$. The quantity $\mathscr{W}$ can be decomposed into

$$\mathscr{W}(\mathbf{x} \to \mathbf{x}')\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' = \mathscr{S}(\mathbf{x} \to \mathbf{x}')\mathscr{A}(\mathbf{x} \to \mathbf{x}')\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}', \qquad (8.41)$$

where $\mathscr{S}$ is the probability of attempting the move from $\mathbf{x}$ to $\mathbf{x}'$ and $\mathscr{A}$ is the probability of accepting the move. For the deterministic dynamic mapping $(\mathbf{p}, \mathbf{x}) \to (\mathbf{p}', \mathbf{x}')$ obtained from the integration of the equation of motion, the probability of attempting the move $\mathbf{x} \to \mathbf{x}'$ is equal to the probability of picking the specific momentum vector $\mathbf{p}$, which leads (through the dynamic mapping) to $\mathbf{x}'$:

$$\mathscr{S}(\mathbf{x} \to \mathbf{x}')\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' = \mathscr{P}_B(\mathbf{p})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p}. \qquad (8.42)$$

Since the molecular dynamics integrator that we use is time-reversible, the *reverse* move, $\mathbf{x} \to \mathbf{x}'$, is generated if we pick, at $\mathbf{x}'$, exactly the same momentum with which we arrived there, but with opposite sign. In other words, the reverse move is attempted with probability

$$\mathscr{S}(\mathbf{x}' \to \mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' = \mathscr{P}_B(-\mathbf{p}')\mathrm{d}\mathbf{p}'\mathrm{d}\mathbf{x}' = \mathscr{P}_B(\mathbf{p}')\mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{p}'. \qquad (8.43)$$

By substitution, the detailed balance condition becomes

$$\mathscr{P}(\mathbf{x})\mathscr{P}_B(\mathbf{p})\mathscr{A}(\mathbf{x} \to \mathbf{x}')\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} = \mathscr{P}(\mathbf{x}')\mathscr{P}_B(\mathbf{p}')\mathscr{A}(\mathbf{x}' \to \mathbf{x})\mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{p}'. \qquad (8.44)$$

If the dynamic mapping is also area-preserving (i.e., $\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p} = \mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{p}'$), (8.44) is satisfied if we choose

$$\mathscr{A}(\mathbf{x} \to \mathbf{x}') = \min\left(1, \mathrm{e}^{-\beta \Delta V} \frac{\mathscr{P}_B(\mathbf{p}')}{\mathscr{P}_B(\mathbf{p})}\right) = \min\left(1, \mathrm{e}^{-\beta \Delta \mathscr{H}} \frac{\mathrm{e}^{\beta \mathbf{B} \mathbf{p}'} + \mathrm{e}^{-\beta \mathbf{B} \mathbf{p}'}}{\mathrm{e}^{\beta \mathbf{B} \mathbf{p}} + \mathrm{e}^{-\beta \mathbf{B} \mathbf{p}}}\right).$$

$$(8.45)$$

The last equality is obtained by substituting (8.39) for $\mathscr{P}_B$. In the limit of $\mathbf{B} = 0$, we recover the acceptance probability of the standard HMC method [75].

Equations (8.44) and (8.45) guarantee convergence to a canonical distribution only in the case of fixed $\mathbf{B}$. Because $\mathbf{B}$ varies (i.e., the method uses information from momenta sampled in the past in determining the vector $\mathbf{B}$), the evolution is not strictly Markovian. As a consequence, the correlations introduced can lead to the accumulation of systematic errors in the determination of configuration averages [77]. However, these correlations can be broken if the update of $\mathbf{B}$ is not done each step, but with a lower updating frequency. This is analogous to other approximately Markovian procedures employed in MC simulations (e.g., update of the maximum displacements allowed for individual atoms [78]).

The MEHMC method is part of a general class of enhanced sampling methods, in which the dynamics along the soft (slow) degrees of freedom are emphasized relative to those along the stiff (fast) degrees of freedom. One of the conceptually simplest and most widely used such methods is the SHAKE algorithm, in which constraints are applied to selected bond lengths and angles to allow larger time steps to be taken [79]. Another example, in which prior knowledge about the system is used to group atoms into rigid substructures, is the multibody, order-N dynamics [MBO(N)D] algorithm [80]. Faster integration of the equations of motion can also be obtained by decomposing the dynamical propagator into fast and slow components [81]. Other methods in this class seek to distinguish the slow and fast motions automatically by eliminating high-frequency modes. These include large-time-step dynamics based on a stochastic action [82, 83], a generalized moment expansion [84], projection of a generalized Langevin equation onto degrees of freedom identified by mode coupling theory [85], dynamic integration within a subspace of low-frequency eigenvectors [86], and digital filtering of the velocities [87, 88]. It should be stressed that, for all enhanced sampling techniques in this category, one should not overly enhance the sampling. Otherwise, the slow manifold is no longer slow compared to the remaining degrees of freedom. In other words, one needs to give the system enough time to equilibrate along orthogonal degrees of freedom.

Most other dynamic enhanced sampling methods fall into another class, in which the effective energy surface is deformed or non-Boltzmann weighting is introduced to facilitate movement on it. One way in which barriers can be lowered is to average over multiple copies of selected atoms [89–92]. Another way is to introduce delocalization [93–95]. Other methods in this class use non-Boltzmann sampling to increase the probability of high-energy states [16, 28, 96–98]. Parallel tempering [99–101] can be viewed as a member of the latter group.

Consideration of the various methods points to certain issues that could arise in MEHMC simulations of systems other than those studied here. One source for concern is that the efficiency of sampling clearly depends on the choice of averaging period. A similar criticism has been put forward with regard to the *essential*

dynamics method [72]. In this method, the directions of the essential dynamics are identified as the lowest-frequency eigenvectors of the atomic position covariance matrix accumulated during a molecular dynamics simulation; this method is therefore for analysis rather than for the generation of a trajectory. Nevertheless, if the simulation time is not long enough to obtain convergence of the collective motions, the eigenvectors will resemble those of a system with the same number of degrees of freedom executing random diffusion [73, 102]. A second source of concern is that the momentum kick given by the MEHMC method may not be enough to overcome large free enthalpy barriers along the slow manifold.

In such cases, the MEHMC method could be employed in combination with an enhanced sampling method that deforms the effective energy surface (but preserves the location of the potential minima), such as that in [29, 97]. Likewise, it may be worthwhile to explore the use of a reversible multiple-time-scale molecular dynamics propagator [103] with MEHMC to accelerate the dynamical propagation.

In spite of these potential concerns, the MEHMC method is expected to be a useful tool for many applications. One task for which it might be particularly well suited is to generate a canonical ensemble of representative configurations of a biomolecular system quickly. Such an ensemble is needed, for example, to represent the initial conditions for the ensemble of trajectories used in fast-growth free energy perturbation methods such as the one suggested by Jarzynski's identity [104] (see also Chap. 5).

Another interesting application of the MEHMC method could be to introduce faster relaxation of the solvent (or certain portions of the solvent) upon changing the conformation of a solute during a conformational free energy calculation. This is analogous to importance sampling in MC solvation free energy calculations [105]. Such calculations are done using either a thermodynamic perturbation or a TI approach (see Chaps. 2 and 4). In these two approaches, a series of simulations is performed for several values of a coupling parameter λ_i by accumulating equilibrium averages $\langle \exp(-\beta(\mathscr{H}(\lambda_{i+1}) - \mathscr{H}(\lambda_i))) \rangle_i$ or, respectively, $\langle \partial \mathscr{H} / \partial \lambda_i \rangle_i$, in the ensemble corresponding to the perturbed Hamiltonian $\mathscr{H}(\lambda_i)$. It would be of interest to quantify the extent to which the enhanced sampling promoted faster convergence of the equilibrium averages in each λ_i window when the $\mathbf{B}$ vector was applied to either solvent or solute molecules alone.

The potential applications presented in the two paragraphs above can be regarded as special cases of a more general situation. As has already been discussed in free energy perturbation (FEP) (see Chap. 2) and TI (see Chap. 4), a host of methods for calculating the free energy require a thorough sampling of configurations representative of a given ensemble. The MEHMC method may very well help in generating such configurations efficiently.

8.8 Skewing Momenta Distributions to Enhance Free Energy Calculations from Trajectory Space Methods

The identification of the slow manifold introduced in the previous section for the MEHMC method turns out to be effective not only for enhanced thermodynamic

averaging, but also to enhance the generation of trajectories connecting distinct states. A thorough sampling of such trajectories is important for methods that are aimed at computing conformational transition rates or free energy barriers. Efficient trajectory sampling is also important for the estimation of free energy profiles along a reaction coordinate using the Jarzynski identity, i.e., in computing equilibrium free energies from nonequilibrium fast trajectories (see Chap. 5). While the Jarzynski identity is exact in practice, one of its computational drawbacks is that the trajectories that count most are statistically rare. A computational method of preferentially sampling those trajectories is desirable. Some strategies along these lines have already been discussed in Chaps. 5 and 6. Here, we discuss another method that allows for computing the entire potentials of the mean force based on the formalism of Hummer and Szabo [106].

8.8.1 Introduction

There has been considerable recent interest in using approaches that allow the generation of ensembles of dynamical paths to calculate kinetic properties of conformational transitions [107, 108] or to reconstruct entire free energy profiles [106] (see previous chapter). These approaches are expected to be useful in particular for large-dimensional systems (such as proteins or nucleic acids) because there is no need to calculate saddle points, the number of which grows exponentially with the number of degrees of freedom. The fundamental object for these types of calculations is an average of an observable, A, over an infinite ensemble of trajectories that are initiated from the equilibrium distribution of phase space vectors $\mathbf{\Gamma} = (\mathbf{x}, \mathbf{p})$, and which evolve over time t according to a specific dynamical flow, which may or may not conserve an equilibrium distribution. The observable, A, can be a function of the endpoints $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_t$, or a functional of the entire trajectory leading to $\mathbf{\Gamma}_t$. Then, as we shall see next, we are interested in the ensemble average of this quantity,

$$C(t) = \langle A[\mathbf{\Gamma}(t)] \rangle, \tag{8.46}$$

where $\langle \cdots \rangle$ indicates an average over the trajectory ensemble.

For the purposes of the present treatment, we wish to rewrite this trajectory average as an average over the initial, equilibrium distribution. If the system evolves according to deterministic (e.g., Hamiltonian) dynamics, each trajectory is uniquely determined by its initial point, and (8.46) can be written without modification as an average over the canonical phase space distribution.

If the system evolves according to some stochastic scheme, each initial point can lead to a multitude of trajectories. We note, however, that as long as each trajectory is initiated from an equilibrium distribution, (8.46) can still be rewritten as an average over the initial distribution:

$$C(t) = \langle Q(\mathbf{\Gamma}_0; t) \rangle_0, \tag{8.47}$$

where $Q(\mathbf{\Gamma}_0; t)$ is the average of the observable A over all realizations of the dynamics of duration t, which begin from the initial phase space point $\mathbf{\Gamma}_0$. We

have replaced $A[\mathbf{\Gamma}(t)]$, a functional of the trajectory, with a function of the initial conditions. The notation $\langle \cdots \rangle_0$ indicates an average over the equilibrium distribution at $t = 0$, so that $\mathbf{\Gamma}_t$ is uniquely determined by $\mathbf{\Gamma}_0$.

The particular correlation function of interest to free energy profile calculations involves averages over fast, irreversible trajectories, with a nonconservative flow obtained by augmenting the dynamics with a time-dependent pulling potential. On the basis of the Jarzynski identity [104], this strategy allows one to reconstruct free energy profiles along a given reaction coordinate $z(\mathbf{\Gamma}(t))$ from single-molecule pulling experiments [106] or from simulations. The pulling potential V is assumed to depend explicitly upon only the pulling coordinate z and the time t. This leads to a particular form of (8.46) that we shall employ in the present section:

$$C(t) \equiv \exp(-\beta A(z)) = \langle \delta(z - z(\mathbf{\Gamma}(t))) e^{-\beta(W_t - V(z(\mathbf{\Gamma}(t),t)))} \rangle, \qquad (8.48)$$

where A is the free energy profile to be obtained, W_t is the irreversible work performed by motion along the pulling coordinate z and where the average is over an ensemble of *trajectories* with initial points $\mathbf{\Gamma}$ canonically distributed on the time-dependent potential at $t = 0$. We note, however, that we can recast this average in the form of (8.47):

$$\exp(-\beta A(z)) = \langle Q_t(\mathbf{\Gamma}_0) \rangle_0 \qquad (8.49)$$

where $Q_t(\mathbf{\Gamma}_0)$ is the average of $\delta(z - z_t) \exp(-\beta(W_t - V_t))$ over all trajectories of length t initiated from the initial point $\mathbf{\Gamma}_0$.

The quantity W_t is the *irreversible* work done along a particular trajectory in time t, defined by $W_t \equiv \int_0^t d\tau \, \partial \mathscr{H}((\mathbf{\Gamma}(\tau), \tau)/\partial \tau$. Because the system's Hamiltonian evolves at a finite rate, the dynamical flow of the system no longer preserves an equilibrium distribution. Moreover, W_t is a functional of the trajectory, rather than a function of the end points.

Although the calculation of the correlation function in (8.48) does not require prior knowledge of the energy landscape, a significant computational burden still remains because in complex systems one expects a multitude of reaction paths to contribute to the average. Although low-work trajectories contribute most in the formula, they are rarely sampled. This issue is especially prevalent in the case of free energy profiles found from computational pulling simulations. In order to surmount the barriers typically found in such profiles, fast trajectories are usually far from reversible, and therefore require significant work. Consequently, the average in (8.48) converges slowly. In most instances, a smaller number of slow-pulling trajectories will provide a more accurate estimate of the potential of mean force [109]. However, in numerical analyses, the time required for a specific estimate varies linearly with the speed of the pulling simulation but not with the number of trajectories, because easy parallelization allows for the evaluation of additional trajectories without increasing the simulation time. Therefore, if massively parallel computers are used it may be more efficient to estimate a free energy surface from many fast trajectories rather than a few slower ones.

It therefore becomes apparent that enhanced sampling strategies are beneficial for surmounting the difficulties presented above. To exemplify possible approaches, we present a skewed momenta method, in which we calculate the quantities given in (8.46) [or (8.48)] as weighted averages over initial phase space distributions, in which the momenta are 'skewed' to bias the dynamics along certain directions. In the calculation of $C(t)$ in (8.46), we generate low-work trajectories that can surmount entropic and enthalpy barriers, and by doing so yield improved accuracy of potentials of mean force estimated from (8.48).

We start with some background on existing methods that alter the initial distributions in the 'reactant' basin, focusing in particular on the puddle jumping method of Tully and coworkers [69, 110], which is the inspiration for the skewed momenta method developed in the following section. We continue with a description of the skewed momenta method, as applied to (8.48), with numerical examples for each case. We end with a concluding discussion.

8.8.2 Puddle Jumping and Related Methods

In a recent development, Corcelli et al. [110] introduced a convenient bias function with general applicability that promises to accelerate the convergence of rate calculations in systems with large enthalpy barriers. They apply a *puddle* potential (used previously by the same group to enhance thermodynamic averaging [69]) that changes the potential energy surface from which the trajectories are initiated to become

$$V^*(\mathbf{x}) = \begin{cases} V(\mathbf{x}) & \text{if } V(\mathbf{x}) \geq V_{\text{pud}} \\ V_{\text{pud}} & \text{otherwise,} \end{cases} \tag{8.50}$$

and the correlation function in (8.46) can be written, without approximation

$$C(t) = \frac{\int d\mathbf{\Gamma}_0 \rho^*(\mathbf{x}_0)\rho(\mathbf{p}_0)w(\mathbf{x}_0)A[\mathbf{\Gamma}_t]}{\int d\mathbf{\Gamma}_0 \rho^*(\mathbf{x}_0)\rho(\mathbf{p}_0)w(\mathbf{x}_0)} = \frac{\langle A[\mathbf{\Gamma}_t]w(\mathbf{x}_0)\rangle^*}{\langle w(\mathbf{x}_0)\rangle^*}, \tag{8.51}$$

where $w = \exp[\beta(V^*(\mathbf{x}_0) - V(\mathbf{x}_0))]$, $\rho^*(\mathbf{x}_0)$ is the equilibrium spatial distribution on V^*, $\rho(\mathbf{p}_0)$ is the equilibrium distribution of momenta, and $\langle \cdots \rangle^*$ indicates an average over the equilibrium distribution corresponding to $V^*(\mathbf{x})$.

The puddle potential removes the deep energy minima that would ordinarily dominate the initial distribution; trajectories from these deep minima have little chance of crossing a large barrier into another important region (say, the *product* region of a conformational change reaction). This strategy bears a resemblance to the hyper-dynamics method of Voter [111, 112], (in which the bottoms of the potential wells are raised without affecting the barrier tops), and to the related accelerated dynamics method developed by Hamelberg et al. [113]. In the same category of approaches are the methods of Laio and Parrinello outlined in Sect. 4.7 [114], that of Huber and Kim [115], as well as Grubmüller's *conformational flooding* method [116]. In an even earlier reference, the strategy bears a resemblance, to

some extent, to that used by Carter et al. [21] to generate a constrained 'blue moon' ensemble from which one originates free trajectories to obtain rates from correlation functions.

Because the strategy of modifying the initial coordinate distribution proves to be effective in accelerating the convergence of time-correlation calculations, it is of interest to explore further modifications along the same lines. To this end, we will consider in detail in the next few sections a modification of the method of Corcelli et al. [110]. Instead of (or in addition to) applying a puddle potential, let us skew the momentum distribution along certain directions in conformational space (e.g., along the long axis of the right-panel distribution in Fig. 8.2). These directions are to be chosen to correspond to the local slow manifold, which is the conformational subspace in which the natural dynamics of the systems evolves more slowly than in the rest of the space. By increasing the probability to sample initial momenta that have large-magnitude components in the slow manifold, the subsequent relaxation dynamics is accelerated relative to that of the equilibrium distribution. The essence of the skewed momenta method is to extend the variance along a precalculated direction along which motion is encouraged, while discouraging motion in directions that might lead the trajectories away from the states of interest. A two-dimensional example of a spherical (in mass-weighted coordinates) and a skewed momentum distribution is shown in Fig. 8.2. Using the skewed distribution, passage both over enthalpy barriers and through entropic holes can be accentuated. In a procedure similar to the Corcelli et al. method, the trajectories are reweighted and the appropriate correlation function is eventually recovered.

Unlike the methods mentioned above, the skewed momenta method involves accentuating the dynamics only along particularly relevant directions. Because the
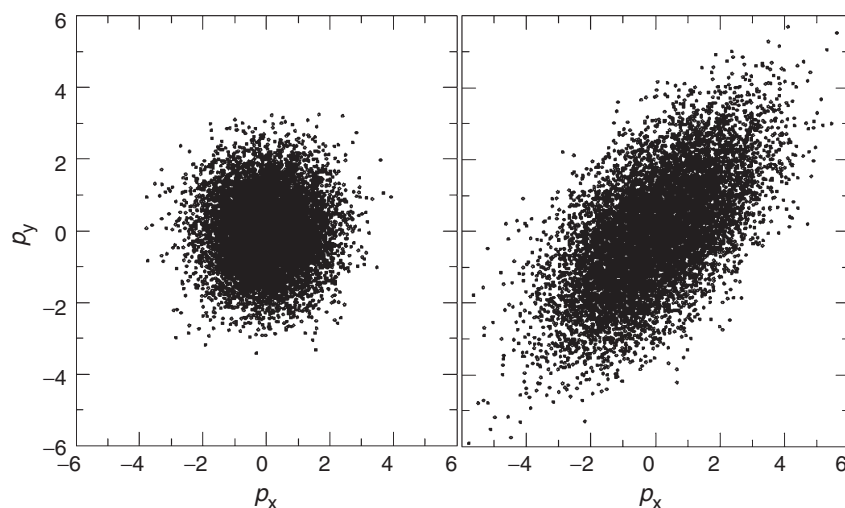

Fig. 8.2. Maxwell (left) and skewed momenta (right) distributions in two dimensions. If a slow direction is identified, the probability can be skewed along that direction such that it is more likely to *kick* the system along it; exact kinetics is recovered by reweighting

puddle jumping method is designed to change the *potential* energy (i.e., to modify the coordinate distribution), whereas the proposed method yields altered *kinetic* energy (i.e., a modified momentum distribution), the latter method can be trivially combined with the former to enable a more efficient exploration of the entire phase space dynamics.

Methods such as skewed momenta are expected to have an additional advantage in high-dimensional systems. Puddle jumping is efficient in such systems only if the 'puddle' can be selectively applied across a few pertinent degrees of freedom. In contrast, the skewed momenta method can be applied without modification to trajectories involving concerted changes to many degrees of freedom.

In addition to accelerating trajectories over enthalpy barriers, skewing momenta can also accelerate convergence in high-dimensional systems with entropic barriers: the momentum distribution can be chosen to encourage motion along a particular direction while discouraging motion in directions that might lead the trajectories away from the states of interest.

We will now turn our attention to the reconstruction of free energy profiles using the Jarzynski identity. This identity can be cast in terms of an equilibrium average, (8.49), as explained in Chap. 5. We can then bias the dynamics to follow the motion of the pulling potential, enhancing sampling of the low-work tail of the work distribution and thereby increasing the accuracy of the calculation.

8.8.3 The Skewed Momenta Method

As alluded to above, the method relies upon the identification of a $3N$-dimensional vector in configuration space, $\hat{\mathbf{e}}_s$, which points along a favored direction for the motion of the system. We then choose the initial momenta for the trajectory ensemble from a Gaussian distribution artificially extended in the direction of $\hat{\mathbf{e}}_s$, as illustrated in the right panel of Fig. 8.2. In the case of free energy reconstructions from (8.49), we wish to induce motion along a predefined pulling direction, and so $\hat{\mathbf{e}}_s$ can be found by inspection.

Because it is cumbersome to generate a Gaussian distribution oriented along an arbitrary axis in the natural coordinates of $\mathbf{p}$, we generate momenta in a rotated coordinate system $\mathbf{p}'$ in which $\hat{\mathbf{e}}_s$ lies along the p_1' axis and then rotate the coordinate axes to place the generated momenta in the original frame. To this end, we seek a rotation matrix R that transforms $\hat{\mathbf{e}}_s$ onto the p_1 axis; this matrix will transform momenta generated in the $\mathbf{p}'$ system to the natural coordinates of the system. For the detailed algorithm to calculate R we refer the reader to [117].

In the case of (8.49), the situation is once again simplified by prior knowledge of the pulling direction: $\hat{\mathbf{e}}_s$ can be taken without loss of generality to lie along one of the natural Cartesian axes of the system, so that the $\mathbf{p}$ and $\mathbf{p}'$ systems are equivalent, and no axis rotations need be performed.

Skewed Momenta and Reweighting

At equilibrium, the components of the momentum vector $\mathbf{p}$ are drawn from the distribution

$$\rho(\mathbf{p}) = C\exp(-\mathbf{p}^T\mathscr{A}\mathbf{p}), \tag{8.52}$$

where C is a normalization constant and $\mathscr{A}$ is the $3N \times 3N$ diagonal matrix $\mathscr{A} = 1/2\beta M^{-1}$, in which M is the mass matrix.

We choose instead from a biased distribution, ρ_B, defined in the $\mathbf{p}'$ coordinate system as

$$\rho_B(\mathbf{p}) = C'\exp(-\mathbf{p}'^T\mathscr{A}'\mathbf{p}') \tag{8.53}$$

where C' is another normalization constant and $\mathscr{A}'$ is a diagonal matrix with entries $\mathscr{A}'_i$ proportional to the variance of the Gaussian distribution along the ith axis of the rotated coordinate system. Since we wish to bias the dynamics along $\hat{\mathbf{e}}_s$ (which should then become the 'longest' direction of the skewed distribution), $\mathscr{A}'_1$ should be the longest axis, but in general there is no restriction on the diagonal entries of $\mathscr{A}'$ save that they be positive. The $\mathbf{p}'$ are related to $\mathbf{p}$ via $\mathbf{p} = R\mathbf{p}'$, where R is the rotation matrix.

We write the average in (8.47) in the exact factorized form

$$C(t) = \frac{\displaystyle\int \mathrm{d}\boldsymbol{\Gamma}_0 \rho(\mathbf{x}_0)\rho_B(\mathbf{p}_0)w(\mathbf{p}_0)Q_t(\boldsymbol{\Gamma}_0)}{\displaystyle\int \mathrm{d}\boldsymbol{\Gamma}_0 \rho(\mathbf{x}_0)\rho_B(\mathbf{p}_0)w(\mathbf{p}_0)}, \tag{8.54}$$

in which $\rho(\mathbf{x}_0)$ is the equilibrium spatial distribution, $\rho_B(\mathbf{p}_0)$ is defined in (8.53), and the weighting function is $w(\mathbf{p}_0) = \exp(\mathbf{p}_0'^T A'\mathbf{p}_0' - \mathbf{p}_0^T A\mathbf{p}_0)$. $Q_t(\boldsymbol{\Gamma}_0)$ is defined as before as the average of some observable $A[\boldsymbol{\Gamma}(t)]$ over all trajectories of length t initiated from the initial point $\boldsymbol{\Gamma}_0$. $C(t)$ can then be calculated as a weighted average of $Q_t(\boldsymbol{\Gamma}_0)$ over the biased momentum distribution

$$C(t) = \frac{\langle Q_t(\boldsymbol{\Gamma}_0)w(\mathbf{p})\rangle_B}{\langle w(\mathbf{p})\rangle_B}, \tag{8.55}$$

where the notation $\langle\cdots\rangle_B$ indicates that the momenta in the calculation are drawn from the biased distribution.

We emphasize that the average represented by $Q_t(\boldsymbol{\Gamma}_0)$ is introduced in the interest of theoretical development only. In numerical simulations, $C(t)$ is found by averaging over the observable $A[\boldsymbol{\Gamma}(t)]$ directly. That is

$$C(t) = \frac{\displaystyle\sum_i A_t^{(i)}w^{(i)}}{\displaystyle\sum_i w^{(i)}}, \tag{8.56}$$

where $A_t^{(i)}$ and $w^{(i)}$ are the values for the ith trajectory (with initial momenta drawn from the biased distribution) of the observable and the weighting factor, respectively.

A two-dimensional example of a spherical and a skewed momentum distribution is shown in Fig. 8.2; for this simple case, $\hat{\mathbf{e}}_s = \frac{1}{\sqrt{2}}\binom{1}{1}$, and the variance along $\hat{\mathbf{e}}_s$ has been extended.

Choosing $\mathscr{A}'$

To this point we have said nothing about the other entries in the matrix $\mathscr{A}'$, needed to calculate the rotation matrix, R. Although (8.55) is exact for any momentum distribution in the rotated reference frame, in practical implementations it will often be desirable to choose the entries of $\mathscr{A}'$ so that the momenta in the directions perpendicular to $\hat{\mathbf{e}}_s$ are similar to their equilibrium values, thereby minimizing their contribution to the scaling factor. However, in systems in which different degrees of freedom have different masses, the momentum-space shell of constant kinetic energy will be a high-dimensional ellipsoid, the axes of which may not align with the axes of the rotated reference frame. In such systems it is convenient to work in mass-weighted coordinates, in which the equi-energetic shells are spherical. That is, define

$$\pi_i \equiv \frac{p_i}{\sqrt{m_i}}, \tag{8.57}$$

where m_i is the mass of the ith degree of freedom and p_i is the unweighted momentum. In these coordinates, the equilibrium matrix entries are simply defined by the equipartition theorem, $\mathscr{A}_i = \beta/2$, and the equilibrium distribution can be reproduced in any rotated frame by choosing $\mathscr{A}_i' = \mathscr{A}_i = \beta/2$. The desired bias along $\hat{\mathbf{e}}_s$ can then be obtained simply by choosing $\mathscr{A}_1 = \alpha$, for $\alpha < \beta/2$, and momenta in the natural coordinates of the system recovered subsequently by inverting (8.57).

The Slow Manifold

The objective of the method presented here is to develop a momentum distribution that will bias path dynamics along the slow manifold, permitting the efficient calculation of kinetic properties of infrequent reactions.

Just like in the MEHMC method described in Sect. 8.7, we can identify the slow manifold from the time average of the momentum, e.g., by choosing a conformational direction $\hat{\mathbf{e}}_s = \mathbf{p}_0/|\mathbf{p}_0|$, where $\mathbf{p}_0$ is calculated as in (8.31).

Alternatively, $\hat{\mathbf{e}}_s$ can be found from either a normal-mode or a quasi-harmonic-mode decomposition [118] by solving an eigenvalue–eigenvector problem

$$\hat{\mathbf{e}}_s = \min_{\lambda}\{\hat{\mathbf{e}} \mid \mathbf{H}\hat{\mathbf{e}} = \lambda\hat{\mathbf{e}}\}, \tag{8.58}$$

and choosing the lowest-eigenvalue eigenvector obtained from diagonalizing $\mathbf{H}$, which is either the Hessian of the potential (in the case of the harmonic modes) or the inverse of the covariance matrix of atomic fluctuations (in the case of the quasi-harmonic modes). The normal-mode decomposition can be performed for a minimized structure in the reactant well, and $\hat{\mathbf{e}}_s$ aligned along a low-frequency mode (or a linear combination of several such slow modes). The quasi-harmonic calculation can be performed on the same trajectory that was used to generate the initial distribution of the starting conformations; again, a combination of slow quasi-harmonic modes can be used.

Using the momentum-averaging scheme, it has been shown that slow-mode directions are of promise for enhanced sampling and for the exploration of large conformational changes, and provide a hybrid MC scheme to obtain exact thermal equilibrium properties [119]. For the case of the alternative choice of $\hat{\mathbf{e}}_s$ in (8.58), normal-mode analyses provided considerable insight into the nature of collective motions in many proteins [120–124]. It has been demonstrated that, when the initial and the final structure of the system are available, the first few low-frequency modes are often sufficient to describe the large-scale conformational changes involved in going from one structure to the other [125, 126]. This strategy has worked well both for protein–DNA complexes [127] and for systems as large as the ribosome [128].

In cases when the slow manifold is higher than one-dimensional (which is likely to be the case for complex biomolecular conformational changes), the guiding vector $\hat{\mathbf{e}}_s$ is to be calculated for each initial configuration, or, if there is little variation in the slow direction for certain initial regions, a single $\hat{\mathbf{e}}_s$ can be applied to some or all of the initial points.

8.8.4 Application to the Jarzynski Identity

In this section we explore the use of the skewed momenta method for estimating the equilibrium free energy from fast pulling trajectories via the Jarzynski identity [104]. The end result will be that generating trajectories with skewed momenta improves the accuracy of the calculated free energy. As described in Chap. 5, Jarzynski's identity states that

$$\exp(-\beta\Delta A) = \langle\exp(-\beta W_t)\rangle, \tag{8.59}$$

where ΔA is the free energy change, W_t is the work performed on the system during each nonequilibrium trajectory of length t, and $\langle\cdots\rangle$ indicates an average over an infinite number of trajectories. Equation (8.49), which we presented in Sect. 8.8.1, represents a specific application of (8.59) to free energy reconstructions from pulling experiments. As already mentioned, even though Jarzynski's identity is asymptotically exact, it suffers from the problem that the trajectories which count most for its convergence, i.e., the low-work trajectories, are statistically rare. In computational simulations, which can only sample a finite number of trajectories, it is important to look for lower-work distributions.

With Jarzynski's identity in the form of (8.46), we can apply to it the skewed momenta method simply by setting $A[\mathbf{\Gamma}(t)] = \exp(-\beta W_t)$. However, we anticipate that the method will be most useful in the particular case of calculating free energy profiles from pulling experiments, for which Hummer and Szabo have provided a modified form of Jarzynski's expression [106].

The average defining the potential of mean force, (8.49), can be written as an average over a skewed distribution of initial momenta as described by (8.55). We can anticipate that skewed trajectories are associated with lower work, as the momenta can be biased so that important degrees of freedom tend to move in the same direction as the pulling potential. Specifically, the instantaneous contribution to the work of
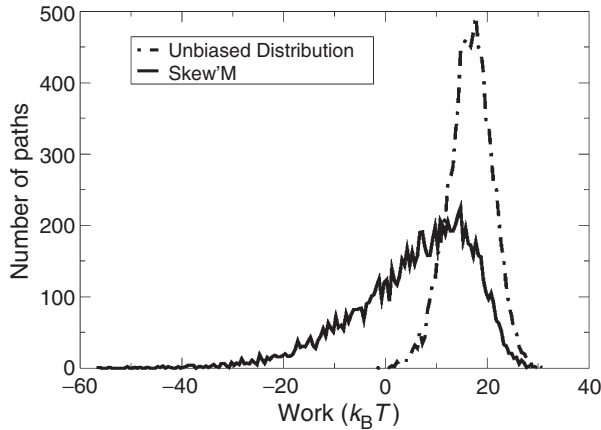
Fig. 8.3. Histogram of work values for Jarzynski's identity applied to the double-well potential, $V(x) = x^2(x-a)^2 + x$, with harmonic guide $V_{\text{pull}}(x,t) = k(x - vt)^2/2$, pulled with velocity v. Using skewed momenta, we can alter the work distribution to include more low-work trajectories. Langevin dynamics on $V_{\text{tot}}(x(t),t) = V(x(t)) + V_{\text{pull}}(x(t),t)$ with $k_{\text{B}}T = 1$, $k = 100$, was run with step size $\Delta t = 0.001$, and friction constant $\gamma = 0.2$ (in arbitrary units). We choose $v = 4$ and $a = 4$, so that the barrier height is many times $k_{\text{B}}T$ and the pulling speed far from reversible. Trajectories were run for a duration $t = 1000$. Work histograms for $10,000$ trajectories, for both equilibrium (Maxwell) initial momenta, with zero average and unit variance, and a skewed distribution with zero average and a variance of 16.0

a given trajectory is found from an infinitesimal movement of the time-dependent guiding potential at constant position and momentum, $dW = \frac{\partial V}{\partial t}(\mathbf{\Gamma}(t))dt$. The dW contribution will tend to be smaller, or even negative, when the system follows the motion of the guiding potential naturally rather than being 'pulled along.'

As an illustrative example of this approach, consider applying Jarzynski's identity to reconstruct a one-dimensional double-well energy profile of the form $V(x)$, assumed unknown and which is to be recovered by the method, from pulling trajectories with a harmonic guiding potential

$$V_{\text{pull}}(x,t) = \frac{k(x-vt)^2}{2} \tag{8.60}$$

where k is the spring constant and v the pulling velocity. By artificially increasing the variance of the momentum distribution, we can alter the work distribution to include more low-work trajectories. To accomplish this, one can run Langevin dynamics for the particle in the potential $V_{\text{tot}}(x(t),t) = V(x(t)) + V_{\text{pull}}(x(t),t)$. If the barrier height is many times $k_{\text{B}}T$ and the pulling speed is far from the reversible regime, convergence difficulties are expected in the regular application of the Jarzynski strategy, because the low-work values are rarely sampled. In Fig. 8.3 we display a typical histogram of work values for trajectories for both an equilibrium initial momentum

distribution, drawn from a Gaussian distribution with zero average and unit variance, and a biased distribution with zero average and an accentuated variance along the reactive coordinate.

Because an individual trajectory's contribution to the average in (8.59) depends exponentially on the work performed on the system during the trajectory, we expect that the increased sampling of low-work trajectories will improve the accuracy of the free energy calculation. As a test of the practical efficiency, we recreate the function $V(x)$ using (8.49) and (8.55) from a number of high-speed pulling simulations using a technique analogous to the weighted histogram method [106]

$$\exp(-\beta A_0(z)) = \frac{\sum_t \frac{\langle \delta[z - z(t)] \mathscr{W} \rangle}{\langle \mathscr{W} \rangle}}{\sum_t \frac{\exp(-\beta V(z, t))}{\langle \mathscr{W} \rangle}} \tag{8.61}$$

A detailed numerical implementation of this method is discussed in [106]. $\mathscr{W}$ is the statistical weight of a trajectory, and the averages are taken over the ensemble of trajectories. In the unbiased case, $\mathscr{W} = \exp(-\beta W_t)$, while in the biased case an additional factor must be included to account for the skewed momentum distribution: $\mathscr{W} = \exp(-\beta W_t) w(\mathbf{p})$. Such simulations can be shown to increase accuracy in the reconstruction using the skewed momenta method because of the increase in the likelihood of generating low work values. For such reconstructions and other applications, e.g., to estimate free energy barriers and rate constants, we refer the reader to [117].

In future work, it should be of interest to explore the combination of the skewed momenta method to alter the momentum distribution with novel methods that either apply periodic loading [129] or MC sampling of nonequilibrium trajectories from a work-weighted ensemble [130, 131]. A quantum analog of the Jarzynski method has recently been proposed as a means to describe dephasing of quantum coherences [132]. It would also be of interest to explore whether biasing the flow of adiabatic states in the corresponding master equation in a numerical realization of a nonequilibrium line shape measurement would yield a faster convergence of spectroscopic properties.

8.8.5 Discussion

By drawing momenta for molecular dynamics from a distribution that is artificially enhanced along an important degrees of freedom, we have shown that the 'skewed-momenta' method addresses a problem endemic to the reconstruction of free energy profiles from fast pulling experiments using Jarzynski's identity. Namely, although low-work trajectories have the largest statistical weight, they are rarely sampled, especially when the pulling is fast. The skewed momenta method generates more low-work trajectories by biasing the relevant degrees of freedom so that they tend to move with the pulling potential, thereby lowering the work done on the system and increasing the accuracy of the calculated potential of the mean force. This fact

is in accord with recent theoretical work analyzing the Jarzynski identity for an ideal gas. For this system, Lua and Grosberg [133] noted that the trajectories in the far tails of the Maxwell distribution are the key ones in determining the accuracy of the Jarzynski method when the system's volume is altered rapidly. It is true that an ideal gas is a crude 'flat landscape' approximation for a general system with a nonconstant potential energy landscape. However, it is important to realize that, for such a general system, the far tails of the Maxwell distribution are preferentially sampled in the skewed momenta method only along the slow degrees of freedom along which the potential is expected to be, at least locally, relatively flat (i.e., with low curvature).

The method is likely to be useful for the numerical calculation of other correlation functions of importance to complex molecules. An example is the orientation correlation functions of interest in NMR-derived dynamical estimates for proteins and nucleic acids [134]. Such correlations are difficult to converge numerically when multiple conformations separated by large free energy barriers contribute to their measurement.

Future improvements of the method presented here should take into account that the direction of the important manifold in general can change (i.e., the direction of importance can curve around conformational space). For the method to be effective in such instances, one would have to bias not only the initial momentum distribution, but, using a periodically updated $\hat{e}_s$, also the actual trajectories that lead to relaxation into the product well. In this case, one would also need to reweight the trajectory itself, not only the initial points. If the propagation uses Langevin dynamics, the formalism of stochastic path integrals [135] leads to the proper weight described by the exponential of an Onsager–Machlup action [136–138] that would have to be calculated along each trajectory according to the formula

$$S[\mathbf{\Gamma}(t)] = \int_{t_1}^{t_2} \mathrm{d}t (M\ddot{\mathbf{x}}(t) + \gamma\mathbf{p}(t) + \nabla V[\mathbf{x}(t)])^2. \qquad (8.62)$$

8.9 Quantum Free Energy Calculations

We end with a section exemplifying a method for enhanced sampling that is needed in a quantum context, i.e., when calculating quantum free energies through the quantum-classical isomorphism (as popularized by Chandler and Wolynes [139]). Although quantum free energies will be introduced later in Chap. 11, here we broaden the scope of the enhanced sampling using generalized distributions such as those in Sect. 8.4.2 by introducing a specialized method using generalized-ensemble path-integral methods.

There is considerable interest in the use of discretized path-integral simulations to calculate free energy differences or potentials of mean force using quantum statistical mechanics for many-body systems [140]. The reader has already become familiar with this approach to simulating with classical systems in Chap. 7. The theoretical basis of such methods is the Feynmann path-integral representation [141], from which is derived the isomorphism between the equilibrium canonical ensemble of a

quantum system and the canonical ensemble of a classical system of ring polymers of P beads, or pseudo-particles. Each such ring corresponds to one quantum particle. The classical system of ring polymers can be simulated with either MC [142] or molecular dynamics [143] methods. Examples of the applications of such simulations are the studies of the quantum mechanical contributions to the structure of water [144], electron localization in water clusters [145], and the reaction rates for intramolecular proton transfer in acetylacetone [146].

For a system at temperature $T = 1/k_B\beta$ in the potential $V(x)$ and having the Hamiltonian operator $H \equiv K + V = -(\hbar^2/2m)\nabla^2 + V(x)$, the elements of the density matrix operator $e^{-\beta H}$ in the canonical ensemble are defined, in the coordinate representation, as

$$\rho(x, x'; \beta) = \langle x|e^{-\beta\mathscr{H}}|x'\rangle. \tag{8.63}$$

These are Green's functions diffusing in what is interpreted as an imaginary time β, according to the Bloch equation, $\partial\rho/\partial\beta = \mathscr{H}\rho$ (a diffusion-type partial differential equation). These Green's functions satisfy the equation

$$\int \rho(x_1, x_1; \beta)\mathrm{d}x_1 = \int \ldots \int \rho(x_1, x_2; \beta/P) \ldots$$

$$\rho(x_{P-1}, x_P; \beta/P)\rho(x_P, x_1; \beta/P)\mathrm{d}x_1 \ldots \mathrm{d}x_P. \tag{8.64}$$

For large P, β/P is small and it is possible to find a good short-time approximation to the Green function ρ. This is usually done by employing the Trotter product formula for the exponentials of the noncommuting operators K and V

$$e^{-\beta\mathscr{H}} = \lim_{P\to\infty} \left(e^{-\beta K/P}e^{-\beta V/P}\right)^P. \tag{8.65}$$

In the so-called primitive representation of the discretized path-integral approach [141], the canonical partition function for finite P has the form

$$Q_P(\beta) \equiv \int \rho(x, x; \beta)\mathrm{d}x =$$

$$\left(\frac{mP}{2\pi\hbar^2\beta}\right)^{P/2} \int \ldots \int e^{-\beta\frac{mP}{2\hbar^2\beta^2}\sum_{i=1}^{P}(x_i - x_{i+1})^2 + \frac{1}{P}\sum_{i=1}^{P}V(x_i)}\mathrm{d}x_1 \ldots \mathrm{d}x_P.$$

$$\tag{8.66}$$

This is called 'primitive' in the sense that the short-time approximation, truncated after the first term, is in its crudest form. Nonprimitive schemes would be those that would improve this approximation, for instance by replacing the 'bare' potential $V(x)$ by an effective quantum potential (see [142, 149]).

With neglect of the quantum effects that arise from the exchange of identical particles [147], (8.66) gives the exact quantum partition function in the limit $P \to \infty$. For finite P, $Q_P(\beta)$ is the canonical partition function of a classical system composed of ring polymers. Each quantum particle corresponds to a ring polymer of P beads in which neighboring beads are connected by harmonic springs with force

constant $mP/2\hbar^2\beta^2$ and each bead is acted on by the interaction potential V/P. The formulation gives a good approximation if $\sqrt{\hbar^2\beta/mP}$ is small compared to a characteristic length of the system. By simulating a classical system of ring polymers for a large enough value of P, convergence to the quantum mechanical result can be achieved. Quantum thermodynamic averages can be calculated using appropriate estimators. For instance, the estimator $\mathscr{U}_P$ for the internal energy $U = \langle \mathscr{U}_P \rangle$ (the average is over the canonical distribution) can be found by inserting (8.66) into (8.67)

$$U = -\frac{\partial \ln Q}{\partial \beta} \tag{8.67}$$

to find

$$\mathscr{U}_P = \frac{P}{2\beta} - \frac{mP}{2\hbar^2\beta^2}\sum_{i=1}^{P}(x_i - x_{i+1})^2 + \frac{1}{P}\sum_{i=1}^{P}V(x_i). \tag{8.68}$$

The quantum delocalization arising from the Heisenberg uncertainty principle is equivalent, in the classical isomorphism picture, to the entropic part of the free energy of the system of ring polymers [148]. Just as a quantum-mechanical particle cannot be localized exactly at the bottom of a potential energy well because of the resulting infinite uncertainty in the momentum, in the classical isomorphic picture at finite T, entropy, as a number of accessible states, competes with the energetics that would favor the bottom minimum. For highly quantum systems, one thus expects that the entropy of the corresponding classical system plays an important role. However, in the simulation of the thermodynamics of classical systems, 'statistical' properties such as entropy are difficult to calculate. The problem is even more acute for systems (such as proteins) which suffer from 'broken ergodicity.' For such systems, the sampling of the configurations representative of the thermodynamical equilibrium takes too long from the simulation point of view. For such systems, it is desirable to use methods that will enhance sampling of the configurational space.

For systems exhibiting sizeable quantum effects, a large value of P is needed. However, this causes, in addition to an increase in the number of degrees of freedom for the polymer rings, very slow relaxation of the springs connecting the pseudo-particles. Thus, methods either to reduce the number of beads or to improve the sampling efficiency for a finite value of P are highly desirable. Several techniques have been devised for better convergence as a function of P. For systems with strongly repulsive interactions Barker [142] proposed and Pollock and Ceperley [149] implemented a short-time approximation scheme that replaces the potential V with a quantum potential calculated by a numerical matrix multiplication. This potential involves calculating a smoother version $V_{\text{eff}}(\beta)$ of V, a quantum effective potential. The short-time propagator $\hat{\rho}(\beta/P)$ can be thought of as a matrix. Multiplying this matrix n times (where $n = \log_2 P$) just up to time β yields the density matrix that is proportional to $\exp(-\beta V_{\text{eff}}(x;\beta))$, from which V_{eff} is found. It contains more of the quantum mechanics, so in this sense it is a quantum potential.

This scheme shows better convergence with increasing P for systems with hard-core repulsions. A reduction in the value of P that is required can also be achieved by renormalization techniques [139] or by choosing a higher-order short-time approximation [150]. Staging MC [151], in which one grows the polymer ring in stages, is another useful technique for improved sampling. Normal-mode decomposition has also been used for improving the sampling efficiency by having different sizes for the MC steps along the eigendirections depending on the frequency of the respective mode [152]. In addition, umbrella sampling techniques have been employed to reduce the number P of beads that are needed [153]. All of these methods are within the framework of the regular primitive algorithm given in (8.66). We present here a method that generalizes the distribution of the classical ring polymers. The method not only shows better convergence but can also be used as a framework for all other schemes (normal modes, staging, etc.) that improve the regular primitive algorithm.

Which methods improve quantum convergence and which have better sampling? It is not clear whether one can really separate these effects. For instance, one can have an improved method that is designed primarily to reduce P, but which also yields enhanced sampling. Thus, in the list of methods given above, we only state the primary reason for the apparent success of each of them.

In previous work on enhanced configurational sampling [154, 155] it was conjectured that a method based on the Tsallis generalization of the canonical ensemble [156] would be expected to have faster convergence with P, and an application was produced to prove this conjecture [34]. Here we present a pedagogical outline of the ideas behind this approach.

As we have seen in Sect. 8.4.2, in the Tsallis generalization of the canonical ensemble [31], the probability density that the system is at position x is

$$p_q(x) = (1 - (1 - q)\beta V(x))^{1/1-q} \tag{8.69}$$

which has the property that $\lim_{q \to 1} p_q(x) = \exp(-\beta V(x))$; i.e., Boltzmann statistics is recovered in the limit of $q = 1$. For values of $q > 1$, the generalized probability distributions $p_q(x)$ are more delocalized than the Boltzmann distribution at the same temperature. This feature has been used as the main ingredient for a set of successful methods to enhance the configurational sampling of classical systems suffering from broken ergodicity [33, 35]. Because of their power-law form (Lévy-like distributions are obtained naturally as stable distributions in the generalized formalism), the generalized distributions appear naturally in systems with fractal properties of the relevant space and time. In what follows, we present the application of this approach to quantum systems simulated via the discretized path-integral representation. Note that the structure of the chain of beads for highly quantum systems exhibits fractal scaling: the variance of a polymer chain of P beads equals P times the nearest-neighbor distance variance [157].

According to the central limit theorem, if one sums up random variables which are drawn from any (but the same for all variables) distribution (as long as this distribution has finite variance), then the sum is distributed according to a Gaussian. In this

sense the Gaussian distribution is the *attractor* of all finite-variance distributions; it is a stable distribution. If one relaxes the requirement that the variance is finite, then the stable distribution is called a Lévy distribution. The Cauchy–Lorentz distribution is an example. The Tsallis distributions have 'tails' which fit the Lévy distributions, so in this sense they are Lévy-like. Lévy distributions appear whenever there exists fractal scaling, i.e., multiplying the length by a constant factor yields the distribution that is equal to the distribution before multiplication times the constant factor to some power. Chandler has observed that the polymer of beads in path integrals has this property. This is the reason for speculating that the Tsallis distribution might be rather appropriate for these polymers. While this speculation seems very attractive, it has yet to be proven.

If we take

$$P = \frac{1}{q-1}, \tag{8.70}$$

the Tsallis probability density becomes

$$p_P(x) = \left(\frac{1}{1 + \beta V(x)/P} \right)^P \tag{8.71}$$

and the sequence $P = 1, 2, 3, \ldots, \infty$ is equivalent to $q = 2, \frac{3}{2}, \frac{4}{3}, \ldots, 1$. Now, instead of a small imaginary time step for the regular density matrix operator

$$\mathrm{e}^{-\beta \mathcal{H}/P} \simeq \mathrm{e}^{-\beta K/P} \mathrm{e}^{-\beta V/P} \tag{8.72}$$

we write

$$\mathrm{e}^{-\beta \mathcal{H}/P} \simeq \mathrm{e}^{-\beta K/P} \left(\frac{1}{1 + \beta V(x)/P} \right). \tag{8.73}$$

One can show that (8.73) is exact up to first order in $1/P$. To this end, start by expanding the operators in powers of $\tau = \beta/P$:

$$\mathrm{e}^{-\tau K} = 1 - \tau K + \frac{\tau^2}{2} K^2 - \cdots,$$

$$\mathrm{e}^{-\tau \bar{V}} = 1/(1 + \tau V) = 1 - \tau V + \tau^2 V^2 - \cdots$$

$$\mathrm{e}^{-\tau(K + \bar{V})} = 1 - \tau(K + V) + \frac{\tau^2}{2}((K + V)^2 + V^2).$$

Then observe that the product of the series expansions of $\mathrm{e}^{-\tau K}$ and $\mathrm{e}^{-\tau \bar{V}}$ differs from $\mathrm{e}^{-\tau(K+\bar{V})}$ by a correction of $\mathscr{O}(\tau^2)$.

Notice that the sequence limit $P \to \infty$ needed for convergence to quantum mechanics of the *regular* primitive algorithm (8.66) also yields the correct quantum mechanics for the *generalized* primitive algorithm which makes use of the propagator in (8.73) instead of that in (8.72). However, the key observation is that, while for infinite P the two approaches yield the same, theoretically exact, quantum

thermodynamics, for *finite* P there exists an important advantage in using the generalized kernel since it corresponds to a more delocalized distribution. This is very important because, for the simple case of a harmonic oscillator, it is known [148] that for all finite P the classical treatment in the regular primitive representation underestimates the delocalization of the particle. Faster convergence (with lower values of P) obtained using the generalized algorithm is due to the fact that the p_q, for $q > 1$, are more delocalized functions than the Boltzmann distribution.

By defining

$$\bar{V} = \frac{P}{\beta} \ln \left(1 + \frac{\beta}{P} V \right) \tag{8.74}$$

the generalized algorithm can be cast in a familiar form, in which the canonical partition function of the isomorphic classical system becomes

$$Q_P(\beta) \equiv \left(\frac{mP}{2\pi\hbar^2\beta} \right)^{P/2} \int \dots \int e^{-\beta W_P} dx_1 \dots dx_P. \tag{8.75}$$

where

$$W_P = \left(\frac{mP}{2\hbar^2\beta^2} \right) \sum_{i=1}^{P} (x_i - x_{i+1})^2 + \frac{1}{P} \sum_{i=1}^{P} \bar{V}(x_i). \tag{8.76}$$

As with the regular primitive path-integral algorithm, it is possible to use any MC and molecular dynamics method to calculate thermodynamical averages of quantum many-body systems by sampling the configuration space of the isomorphic classical ring polymers according to $\exp(-\beta W_P)$. Generalizations of path-integral molecular dynamics algorithms as well as centroid molecular dynamics [158] methods are straightforward; the later is particularly important as an approach to quantum dynamics.

8.10 Summary

In this chapter we have presented an eclectic set of enhanced sampling methods to address the problem of conformational sampling, which is one of several key factors for a successful free energy estimation. We have outlined the problem facing simulations of complex systems, that of *broken ergodicity*, and have introduced a metric, the ergodic measure, which can be used as a (necessary) criterion to assess whether sampling is thorough. We have then reviewed several established, as well as newly developed, methods for classical systems. Last, calculations of quantum free energy in the context of an isomorphism to classical systems has been addressed, stressing the utility of developing ad hoc enhanced sampling methods for this case as well.

While all enhanced sampling methods, reviewed here or not, are good strategies in themselves because they improve convergence in the calculation of equilibrium

averages, it is important to note that there is no 'best' method. Rather various methods may be good for various systems or reasons. We hope to have left the reader with the conviction that it is often the case that ad hoc modification of existing strategies, or newly devised ones, can bring about worthwhile improvements to aid the calculation of thermodynamical observables.

References

1. Eckmann, J.P.; Ruelle, D., Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.* **1985**, *57*, 617–656
2. Berne, B.J.; Straub, J.E., Novel methods of sampling phase space in the simulation of biological systems, *Curr. Opin. Struct. Biol.* **1997**, *7*, 181–189
3. Hodel, A.; Simonson, T.; Fox, R.O.; Brunger, A.T., Conformational substates and uncertainty in macromolecular free-energy calculations, *J. Phys. Chem.* **1993**, *97*, 3409–3417
4. Mountain, R.D.; Thirumalai, D., Measures of effective ergodic convergence in liquids, *J. Phys. Chem.* **1989**, *93*, 6975–6979
5. Mountain, R.D.; Thirumalai, D., Quantitative measure of efficiency of Monte-Carlo simulations, *Physica A* **1994**, *210*, 453–460
6. Straub, J.E.; Thirumalai, D., Exploring the energy landscape in proteins, *Proc. Natl Acad. Sci. USA* **1993**, *90*, 809–813
7. Straub, J.E.; Rashkin, A.; Thirumalai, D., Dynamics in rugged energy landscapes with applications to the S-peptide and Ribonuclease A, *J. Am. Chem. Soc.* **1994**, *116*, 2049
8. Siepmann, J.I.; Sprik, M., Folding of model heteropolymers by configurational-bias Monte Carlo, *Chem. Phy. Lett.* **1992**, *199*, 220
9. Rossky, P.J.; Doll, J.D.; Friedman, H.L., Brownian dynamics as smart Monte Carlo simulation, *J. Chem. Phys.* **1978**, *69*, 4628
10. Cao, J.; Berne, B.J., Monte Carlo methods for accelerating barrier crossing: anti-force-bias and variable step algorithms, *J. Chem. Phys.* **1990**, *92*, 1980
11. Frantz, D.D.; Freeman, D.L.; Doll, J.D., Reducing quasi-ergodic behavior in Monte Carlo simulation by J-walking: Applications to atomic clusters, *J. Chem. Phys.* **1990**, *93*, 2769
12. Tsai, C.J.; Jordan, K.D., Use of the histogram and jump-walking methods for overcoming slow barrier crossing behavior in Monte Carlo simulations: applications to the phase transitions in the $(Ar)_{13}$ and $(H_2O)_8$ clusters, *J. Chem. Phys.* **1993**, *99*, 6957
13. Marinari, E.; Parisi, G., Simulated tempering – a new Monte-Carlo scheme, *Europhys. Lett.* **1992**, *19*, 451–458
14. Geyer, C.J.; Thompson, E.A., Annealing Markov-chain Monte-Carlo with applications to ancestral inference, *J. Am. Stat. Assoc.* **1995**, *90*, 909–920
15. Hukushima, K.; Nemoto, K., Exchange Monte Carlo method and application to spin glass simulations, *J. Phys. Soc. Jpn.* **1996**, *65*, 1604–1608
16. Berg, B.A.; Neuhaus, T., Multicanonical algorithms for 1st order phase-transitions, *Phys. Lett. B* **1991**, *267*, 249–253
17. Berg, B.A.; Neuhaus, T., Multicanonical ensemble – a new approach to simulate 1st-order phase-transitions, *Phys. Rev. Lett.* **1992**, *68*, 9–12
18. Swendsen, R.H.; Wang, J.S., Nonuniversal critical-dynamics in Monte-Carlo simulations, *Phys. Rev. Lett.* **1987**, *58*, 86–88

19. Hansmann, U.H.E.; Okamoto, Y.; Eisenmenger, F., Molecular dynamics, Langevin and hybrid Monte Carlo simulations in a multicanonical ensemble, *Chem. Phys. Lett.* **1996**, *259*, 321–330

20. Torrie, G.M.; Valleau, J.P., Non-physical sampling distributions in Monte-Carlo free-energy estimation – umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199

21. Carter, E.A.; Ciccotti, G.; Haynes, J.T.; Kapral, R., Constrained reaction coordinate dynamics for the simulation of rare events, *Chem. Phys. Lett.* **1989**, *156*, 472–477

22. Darve, E.; Pohorille, A., Calculating free energies using average force, *J. Chem. Phys.* **2001**, *115*, 9169–9183

23. Torrie, G.M.; Valleau, J.P., Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the subcritical Lennard-Jones fluid, *Chem. Phys. Lett.* **1974**, *28*, 578–581

24. Lee, J., New Monte-Carlo algorithm – entropic sampling, *Phys. Rev. Lett.* **1993**, *71*, 211–214

25. Nakajima, N.; Nakamura, H.; Kidera, A., Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides, *J. Phys. Chem. B* **1997**, *101*, 817–824

26. Lyubartsev, A.P.; Martsinovski, A.A.; Shevkunov, S.V.; Vorontsovvelyaminov, P.N., New approach to Monte-Carlo calculation of the free-energy – method of expanded ensembles, *J. Chem. Phys.* **1992**, *96*, 1776–1783

27. Hesselbo, B.; Stinchcombe, R.B., Monte-Carlo simulation and global optimization without parameters, *Phys. Rev. Lett.* **1995**, *74*, 2151–2155

28. Bartels, C.; Karplus, M., Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations, *J. Comput. Chem.* **1997**, *18*, 1450–1462

29. Darve, E.; Wilson, M.A.; Pohorille, A., Calculating free energies using a scaled-force molecular dynamics algorithm, *Mol. Simul.* **2002**, *28*, 113–144

30. Kumar, S.; Bouzida, D.; Swendsen, R.H.; Kollman, P.A.; Rosenberg, J.M., The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J. Comput. Chem.* **1992**, *13*, 1011–1021

31. Tsallis, C., Possible generalization of Boltzmann–Gibbs statistics, *J. Stat. Phys.* **1988**, *52*, 479–487

32. Curado, E.M.F.; Tsallis, C., Generalized statistical mechanics: Connection with thermodynamics, *J. Phys. A: Math. Gen.* **1991**, *24*, L69

33. Andricioaei, I.; Straub, J.E., Generalized simulated annealing algorithms using Tsallis statistics: application to conformational optimization of a tetrapeptide, *Phys. Rev. E* **1996**, *53*, R3055–R3058

34. Andricioaei, I.; Straub, J.E.; Karplus, M., Simulation of quantum systems using path integrals in a generalized ensemble, *Chem. Phys. Lett.* **2001**, *346*, 274–282

35. Andricioaei, I.; Straub, J.E., On Monte Carlo and molecular dynamics methods inspired by Tsallis statistics: methodology, optimization, and application to atomic clusters, *J. Chem. Phys.* **1997**, *107*, 9117–9124

36. Bhattacharya, K.K.; Sethna, J.P., Multicanonical methods, molecular dynamics, and Monte Carlo methods: comparison for Lennard-Jones glasses, *Phys. Rev. E* **1998**, *57*, 2553–2562

37. Swendsen, R.H.; Wang, J.S., Replica Monte-Carlo simulation of spin-glasses, *Phys. Rev. Lett.* **1986**, *57*, 2607–2609

38. Manousiouthakis, V.I.; Deem, M.W., Strict detailed balance is unnecessary in Monte Carlo simulation, *J. Chem. Phys.* **1999**, *110*, 2753–2756

39. Kofke, D.A., On the acceptance probability of replica-exchange Monte Carlo trials, *J. Chem. Phys.* **2002**, *117*, 6911–6914

40. Predescu, C.; Predescu, M.; Ciobanu, C.V., On the efficiency of exchange in parallel tempering Monte Carlo simulations, *J. Phys. Chem. B* **2005**, *109*, 4189–4196

41. Schug, A.; Herges, T.; Wenzel, W., All-atom folding of the three-helix HIV accessory protein with an adaptive parallel tempering method, *Proteins-Struct. Funct. Bioinform.* **2004**, *57*, 792–798

42. Rathore, N.; Chopra, M.; de Pablo, J.J., Optimal allocation of replicas in parallel tempering simulations, *J. Chem. Phys.* **2005**, *122*

43. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding, *Chem. Phys. Lett.* **1999**, *314*, 141–151

44. Calvo, F., All-exchanges parallel tempering, *J. Chem. Phys.* **2005**, *123*

45. Calvo, F.; Neirotti, J.P.; Freeman, D.L.; Doll, J.D., Phase changes in 38-atom Lennard-Jones clusters. II. A parallel tempering study of equilibrium and dynamic properties in the molecular dynamics and microcanonical ensembles, *J. Chem. Phys.* **2000**, *112*, 10350–10357

46. Yan, Q.L.; de Pablo, J.J., Hyper-parallel tempering Monte Carlo: Application to the Lennard-Jones fluid and the restricted primitive model, *J. Chem. Phys.* **1999**, *111*, 9509–9516

47. Andricioaei, I.; Straub, J.E., On Monte Carlo and molecular dynamics methods inspired by Tsallis statistics: methodology, optimization, and application to atomic clusters, *J. Chem. Phys.* **1997**, *107*, 9117–9124

48. Whitfield, T.W.; Bu, L.; Straub, J.E., Generalized parallel sampling, *Physica A – Stat. Mech. Appl.* **2002**, *305*, 157–171

49. Jang, S.M.; Shin, S.; Pak, Y., Replica-exchange method using the generalized effective potential, *Phys. Rev. Lett.* **2003**, *91*

50. Liu, H.B.; Jordan, K.D., On the convergence of parallel tempering Monte Carlo simulations of LJ(38), *J. Phys. Chem. A* **2005**, *109*, 5203–5207

51. Sugita, Y.; Kitao, A.; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations, *J. Chem. Phys.* **2000**, *113*, 6042–6051

52. Fukunishi, H.; Watanabe, O.; Takada, S., On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction, *J. Chem. Phys.* **2002**, *116*, 9058–9067

53. Sugita, Y.; Okamoto, Y., Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape, *Chem. Phys. Lett.* **2000**, *329*, 261–270

54. Faller, R.; Yan, Q.L.; de Pablo, J.J., Multicanonical parallel tempering, *J. Chem. Phys.* **2002**, *116*, 5419–5423

55. Hansmann, U.H.E., Parallel tempering algorithm for conformational studies of biological molecules, *Chem. Phys. Lett.* **1997**, *281*, 140–150

56. Garcia, A.E.; Onuchic, J.N., Folding a protein in a computer: An atomic description of the folding/unfolding of protein A, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 13898–13903

57. Falcioni, M.; Deem, M.W., A biased Monte Carlo scheme for zeolite structure solution, *J. Chem. Phys.* **1999**, *110*, 1754–1766

58. Haliloglu, T.; Kolinski, A.; Skolnick, J., Use of residual dipolar couplings as restraints in ab initio protein structure prediction, *Biopolymers* **2003**, *70*, 548–562

59. Earl, D.J.; Deem, M.W., Parallel tempering: theory, applications, and new perspectives, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916

60. Frantz, D.D.; Freeman, D.L.; Doll, J.D., Reducing quasi-ergodic behavior in Monte-Carlo simulations by J-walking – applications to atomic clusters, *J. Chem. Phys.* **1990**, *93*, 2769–2784

61. Neirotti, J.P.; Calvo, F.; Freeman, D.L.; Doll, J.D., Phase changes in 38-atom Lennard-Jones clusters. I. A parallel tempering study in the canonical ensemble, *J. Chem. Phys.* **2000**, *112*, 10340–10349

62. Stolovitzky, G.; Berne, B.J., Catalytic tempering: A method for sampling rough energy landscapes by Monte Carlo, *Proc. Natl Acad. Sci. USA* **2000**, *97*, 11164–11169

63. Purisima, E.O.; Scheraga, H.A., An approach to the multiple-minima problem by relaxing dimensionality, *Proc. Natl Acad. Sci. USA* **1986**, *83*, 2782–2786

64. Faken, D.B.; Voter, A.F.; Freeman, D.L.; Doll, J.D., Dimensional strategies and the minimization problem: barrier-avoiding algorithms, *J. Phys. Chem. A* **1999**, *103*, 9521–9526

65. Stillinger, F.H.; Weber, T.A., Hidden structure in liquids, *Phys. Rev. A* **1982**, *25*, 978–989

66. Stillinger, F.H.; Weber, T.A., Packing structures and transitions in liquids and solids, *Science* **1984**, *225*, 983–989

67. Zhou, R.; Berne, B.J., Smart walking: A new method for Boltzmann sampling of protein conformations, *J. Chem. Phys.* **1997**, *107*, 9185

68. Li, Z.Q.; Scheraga, H.A., Monte-Carlo-minimization approach to the multiple-minima problem in protein folding, *Proc. Natl Acad. Sci. USA* **1987**, *84*, 6611–6615

69. Rahman, J.A.; Tully, J.C., Puddle-jumping: A flexible sampling algorithm for rare event systems, *Chem. Phys.* **2002**, *285*, 277–287

70. Bogdan, T.V.; Wales, D.J.; Calvo, F., Equilibrium thermodynamics from basin-sampling, *J. Chem. Phys.* **2006**, *124*

71. Nigra, P.; Freeman, D.L.; Doll, J.D., Combining smart darting with parallel tempering using Eckart space: Application to Lennard-Jones clusters, *J. Chem. Phys.* **2005**, *122*

72. Amadei, A.; Linssen, A.B.M.; Berendsen, H.J.C., Essential dynamics of proteins, *Proteins* **1993**, *17*, 412–425

73. Balsera, M.A.; Wriggers, W.; Oono, Y.; Schulten, K., Principal component analysis and long time protein dynamics, *J. Phys. Chem.* **1996**, *100*, 2567–2572

74. Duane, S.; Kennedy, A.D.; Pendleton, B.J.; Roweth, D., Hybrid Monte Carlo, *Phys. Lett. B* **1987**, *195*, 216–222

75. Mehlig, B.; Heermann, D.W.; Forrest, B.M., Hybrid Monte Carlo method for condensed-matter systems, *Phys. Rev. B* **1992**, *45*, 679–685

76. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, N.N.; Teller, A.H.; Teller, E., Equation of state calculations by fast computing machines, *J. Chem. Phys.* **1953**, *21*, 1087–1092

77. Miller, M.A.; Amon, L.M.; Reinhardt, W.P., Should one adjust the maximum step size in a Metropolis Monte Carlo simulation? *Chem. Phys. Lett.* **2000**, *331*, 278–284

78. Bouzida, D.; Kumar, S.; Swendsen, R.H., Efficient Monte Carlo methods for the computer simulation of biological systems, *Phys. Rev. A* **1992**, *45*, 8894–8901

79. Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C., Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes, *J. Comput. Phys.* **1977**, *23*, 327–341

80. Chun, H.M.; Padilla, C.E.; Chin, D.N.; Watanabe, M.; Karlov, V.I.; Alper, H.E.; Soosaar, K.; Blair, K.B.; Becker, O.M.; Caves, L.S.D.; Nagle, R.; Haney, D.N.; Farmer, B.L., MBO(N)D: A multibody method for long-time molecular dynamics simulations, *J. Comput. Chem.* **2000**, *21*, 159–184

81. Tuckerman, M.E.; Martyna, G.J.; Berne, B.J., Molecular-dynamics algorithm for condensed systems with multiple time scales, *J. Chem. Phys.* **1990**, *93*, 1287–1291

82. Elber, R.; Meller, J.; Olender, R., Stochastic path approach to compute atomically detailed trajectories: application to the folding of C peptide, *J. Phys. Chem. B* **1999**, *103*, 899–911

83. Elber, R.; Ghosh, A.; Cardenas, A., Long time dynamics of complex systems, *Acc. Chem. Res.* **2002**, *35*, 396–403

84. Nadler, W.; Schulten, K., Generalized moment expansion for Brownian relaxation processes, *J. Chem. Phys.* **1985**, *82*, 151–160

85. Kostov, K.S.; Freed, K.F., Mode coupling theory for calculating the memory functions of flexible chain molecules: influence on the long time dynamics of oligoglycines, *J. Chem. Phys.* **1997**, *106*, 771–783

86. Space, B.; Rabitz, H.; Askar, A., Long time scale molecular dynamics subspace integration method applied to anharmonic crystals and glasses, *J. Chem. Phys.* **1993**, *99*, 9070–9079

87. Dauber-Osguthorpe, P.; Maunder, C.M.; Osguthorpe, D.J., Molecular dynamics: Deciphering the data, *J. Comput. Aided Mol. Des.* **1996**, *10*, 177–185

88. Phillips, S.C.; Essex, J.W.; Edge, C.M., Digitally filtered molecular dynamics: the frequency specific control of molecular dynamics simulations, *J. Chem. Phys.* **2000**, *112*, 2586–2597

89. Elber, R.; Karplus, M., Enhanced sampling in molecular-dynamics – use of the time-dependent Hartree approximation for a simulation of carbon-monoxide diffusion through myoglobin, *J. Am. Chem. Soc.* **1990**, *112*, 9161–9175

90. Ulitsky, A.; Elber, R., Application of the locally enhanced sampling (LES) and a mean-field with a binary collision correction (CLES) to the simulation of Ar diffusion and NO recombination in myoglobin, *J. Phys. Chem.* **1994**, *98*, 1034–1043

91. Huber, T.; van Gunsteren, W.F., SWARM-MD: searching conformational space by cooperative molecular dynamics, *J. Phys. Chem. A* **1998**, *102*, 5937–5943

92. Simmerling, C.; Fox, T.; Kollman, P.A., Use of locally enhanced sampling in free energy calculations: Testing and application to the alpha → beta anomerization of glucose, *J. Am. Chem. Soc.* **1998**, *120*, 5771–5782

93. Piela, L.; Kostrowicki, J.; Scheraga, H.A., The multiple-minima problem in the conformational-analysis of molecules – deformation of the potential-energy hypersurface by the diffusion equation method, *J. Phys. Chem.* **1989**, *93*, 3339–3346

94. Liu, Z.H.; Berne, B.J., Method for accelerating chain folding and mixing, *J. Chem. Phys.* **1993**, *99*, 6071–6077

95. Whitfield, T.W.; Bu, L.; Straub, J.E., Generalized parallel sampling, *Physica A* **2002**, *305*, 157–171

96. Krivov, S.V.; Chekmarev, S.F.; Karplus, M., Potential energy surfaces and conformational transitions in biomolecules: A successive confinement approach applied to a solvated tetrapeptide, *Phys. Rev. Lett.* **2002**, *88*, 038101

97. Andricioaei, I.; Straub, J.E., On Monte Carlo and molecular dynamics methods inspired by Tsallis statistics: methodology, optimization, and application to atomic clusters, *J. Chem. Phys.* **1997**, *107*, 9117–9124

98. Hansmann, U.H.E.; Okamoto, Y., Generalized-ensemble Monte Carlo method for systems with rough energy landscapes, *Phys. Rev. E* **1997**, *56*, 2228–2233

99. Frantz, D.D.; Freeman, D.L.; Doll, J.D., Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking: applications to atomic clusters, *J. Chem. Phys.* **1990**, *93*, 2769–2784

100. Marinari, E.; Parisi, G., Simulated tempering – A new Monte Carlo scheme, *Europhys. Lett.* **1992**, *19*, 451–458

101. Hansmann, U.H.E., Parallel tempering algorithm for conformational studies of biological molecules, *Chem. Phys. Lett.* **1997**, *281*, 140–150

102. Hess, B., Similarities between principal components of protein dynamics and random diffusion, *Phys. Rev. E* **2000**, *62*, 8438–8448

103. Tuckerman, M.; Berne, B.J.; Martyna, G.J., Reversible multiple time scale molecular dynamics, *J. Chem. Phys.* **1992**, *97*, 1990–2001

104. Jarzynski, C., Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **1997**, *78*, 2690–2693

105. Jorgensen, W.L.; Ravimohan, C., Monte Carlo simulation of differences in free energies of hydration, *J. Chem. Phys.* **1985**, *83*, 3050–3054

106. Hummer, G.; Szabo, A., Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3659–3661

107. Dellago, C.; Bolhuis, P.G.; Csajka, F.S.; Chandler, D., Transition path sampling and the calculation of rate constants, *J. Chem. Phys.* **1998**, *108*, 1964–1977

108. Dellago, C.; Bolhuis, P.G.; Chandler, D., On the calculation of reaction rate constants in the transition path ensemble, *J. Chem. Phys.* **1999**, *110*, 6617–6625

109. Oberhofer, H.; Dellago, C.; Geissler, P.L., Biased sampling of nonequilibrium trajectories: can fast switching simulations outperform conventional free energy calculation methods?, *J. Phys. Chem. B* **2005**, *109*, 6902–6915

110. Corcelli, S.A.; Rahman, J.A.; Tully, J.C., Efficient thermal rate constant calculation for rare event systems, *J. Chem. Phys.* **2003**, *118*, 1085–1088

111. Voter, A.F., Hyperdynamics: Accelerated molecular dynamics of infrequent events, *Phys. Rev. Lett.* **1997**, *78*, 3908–3911

112. Voter, A.F., A method for accelerating the molecular dynamics simulation of infrequent events, *J. Chem. Phys.* **1997**, *106*, 4665–4677

113. Hamelberg, D.; Mongan, J.; McCammon, J.A., Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules, *J. Chem. Phys.* **2004**, *120*, 11919–11929

114. Laio, A.; Parrinello, M., Escaping free-energy minima, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 12562–12566

115. Huber, G.A.; Kim, S., Weighted-ensemble Brownian dynamics simulations for protein association reactions, *Biophys. J.* **1996**, *70*, 97–110

116. Grubmuller, H., Predicting slow structural transitions in macromolecular systems – conformational flooding, *Phys. Rev. E* **1995**, *52*, 2893–2906

117. MacFadyen, J.; Andricioaei, I., A skewed-momenta method to efficiently generate conformational-transition trajectories, *J. Chem. Phys.* **2005**, *123*, 074107

118. Brooks, B.R.; Janezic, D.; Karplus, M., Harmonic-analysis of large systems. 1. Methodology, *J. Comput. Chem.* **1995**, *16*, 1522–1542

119. Andricioaei, I.; Dinner, A.R.; Karplus, M., Self-guided enhanced sampling methods for thermodynamic averages, *J. Chem. Phys.* **2003**, *118*, 1074–1084

120. Go, N.; Noguti, T.; Nishikawa, T., Dynamics of a small globular protein in terms of low-frequency vibrational-modes, *Proc. Natl Acad. Sci. USA – Biol. Sci.* **1983**, *80*, 3696–3700

121. Levitt, M.; Sander, C.; Stern, P.S., Protein normal-mode dynamics – trypsin-inhibitor, crambin, ribonuclease and lysozyme, *J. Mol. Biol.* **1985**, *181*, 423–447

122. Brooks, B.; Karplus, M., Normal-modes for specific motions of macromolecules – application to the hinge-bending mode of lysozyme, *Proc. Natl Acad. Sci. USA* **1985**, *82*, 4995–4999

123. Ma, J.P.; Karplus, M., Ligand-induced conformational changes in ras p21: a normal mode and energy minimization analysis, *J. Mol. Biol.* **1997**, *274*, 114–131

124. Cui, Q.; Li, G.H.; Ma, J.P.; Karplus, M., A normal mode analysis of structural plasticity in the biomolecular motor F-1-ATPase, *J. Mol. Biol.* **2004**, *340*, 345–372

125. Tama, F.; Sanejouand, Y.H., Conformational change of proteins arising from normal mode calculations, *Protein Eng.* **2001**, *14*, 1–6

126. Krebs, W.G.; Alexandrov, V.; Wilson, C.A.; Echols, N.; Yu, H.Y.; Gerstein, M., Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic, *Proteins-Struct. Funct. Gene.* **2002**, *48*, 682–695

127. Delarue, M.; Sanejouand, Y.H., Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model, *J. Mol. Biol.* **2002**, *320*, 1011–1024

128. Tama, F.; Valle, M.; Frank, J.; Brooks, C.L., Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 9319–9323

129. Braun, O.; Hanke, A.; Seifert, U., Probing molecular free energy landscapes by periodic loading, *Phys. Rev. Lett.* **2004**, *93*

130. Sun, S.X., Equilibrium free energies from path sampling of nonequilibrium trajectories, *J. Chem. Phys.* **2003**, *118*, 5769–5775

131. Ytreberg, F.M.; Zuckerman, D.M., Single-ensemble nonequilibrium path-sampling estimates of free energy differences, *J. Chem. Phys.* **2004**, *120*, 10876–10879

132. Mukamel, S., Quantum extension of the Jarzynski relation: analogy with stochastic dephasing, *Phys. Rev. Lett.* **2003**, *90*

133. Lua, R.C.; Grosberg, A.Y., Practical applicability of the Jarzynski relation in statistical mechanics: a pedagogical example, *J. Phys. Chem. B* **2005**, *109*, 6805–6811

134. Palmer, A.G., NMR characterization of the dynamics of biomacromolecules, *Chem. Rev.* **2004**, *104*, 3623–3640

135. Kleinert, H., *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets*, (3rd edition), World Scientific, Singapore

136. Onsager, L.; Machlup, S., Fluctuations and irreversible processes, *Phys. Rev.* **1953**, *91*, 1505–1512

137. Elber, R.; Meller, J.; Olender, R., Stochastic path approach to compute atomically detailed trajectories: application to the folding of C peptide, *J. Phys. Chem. B* **1999**, *103*, 899–911

138. Zuckerman, D.M.; Woolf, T.B., Efficient dynamic importance sampling of rare events in one dimension, *Phys. Rev. E* **2001**, *6302*, 016702

139. Chandler, D.; Wolynes, P.G., Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids, *J. Chem. Phys.* **1981**, *74*, 4078–4095

140. Berne, B.J.; Thirumalai, D., On the simulation of quantum systems: path integral methods, *Annu. Rev. Phys. Chem.* **1986**, *37*, 401–424

141. Feynman, R.P.; Hibbs, A.R., *Quantum Mechanics and Path Integrals*, McGraw-Hill: New York, 1965

142. Barker, J.A., A quantum-statistical Monte Carlo method: Path integrals with boundary conditions, *J. Chem. Phys.* **1979**, *70*, 2914–2918

143. Parrinello, M.; Rahman, A., Study of an F center in molten KCl, *J. Chem. Phys.* **1980**, *80*, 860–867

144. Kuharski, R.A.; Rossky, P.J., Quantum mechanical contributions to the structure of liquid water, *Chem. Phys. Lett.* **1984**, *103*, 357–362

145. Thirumalai, D.; Wallqvist, A.; Berne, B.J., Path-integral Monte Carlo simulations of electron localization in water clusters, *J. Stat. Phys.* **1986**, *43*, 973–984

146. Hinsen, K.; Roux, B., Potential of mean force and reaction rates for proton transfer in acetylacetone, *J. Chem. Phys.* **1997**, *106*, 3567–3577

147. Ceperley, D.M., Path-integrals in the theory of condensed helium, *Rev. Mod. Phys.* **1995**, *67*, 279–355

148. Schweizer, K.S.; Stratt, R.M.; Chandler, D.; Wolynes, P.G., Convenient and accurate discretized path integral methods for equilibrium quantum mechanical calculations, *J. Chem. Phys.* **1981**, *75*, 1347–1364

149. Pollock, E.L.; Ceperley, D.M., Simulation of quantum many-body systems by path-integral methods, *Phys. Rev. B* **1984**, *30*, 2555–2568

150. Raedt, H.De; Raedt, B.De, Applications of the generalized Trotter formula, *Phys. Rev. A* **1983**, *28*, 3575–3580

151. Sprik, M.; Klein, M.L.; Chandler, D., Staging: A sampling technique for the Monte Carlo evaluation of path integrals, *Phys. Rev. B* **1985**, *31*, 4234–4244

152. Herman, M.F.; Bruskin, E.J.; Berne, B.J., On path integral Monte Carlo simulations, *J. Chem. Phys.* **1982**, *76*, 5150–5155

153. Friesner, R.A.; Levy, R.M., An optimized harmonic reference system for the evaluation of discretized path integrals, *J. Chem. Phys.* **1984**, *80*, 4488–4495

154. Straub, J.E.; Andricioaei, I., Computational methods inspired by Tsallis statistics: Monte Carlo and molecular dynamics algorithms for the simulation of classical and quantum systems, *Braz. J. Phys.* **1999**, *29*, 179–186

155. Andricioaei, I.; Straub, J.E., Computational methods for the simulation of classical and quantum many body systems sprung from the nonextensive thermostatistics. In *Nonextensive Statistical Mechanics and Its Application*, Abe, S.; Okamoto, Y., Eds., Lecture Notes in Physics. Springer: Berlin, Heidelberg, New York, **2001**, ch. IV, pp. 195–235

156. Tsallis, C., Possible generalization of Boltzmann–Gibbs statistics, *J. Stat. Phys.* **1988**, *52*, 479–487

157. Chandler, D. Quantum processes in liquids. In *Liquids, Freezing and Glass Transition*, Levesque, D.; Hansen, J.; Zinn-Justin, J., Eds. Elsevier: New York, **1990**, pp. 195–285

158. Cao, J.; Voth, G.A., The formulation of quantum statistical mechanics based on Feynman path centroid density, *J. Chem. Phys.* **1994**, *100*, 5093–5105

9

Potential Distribution Methods and Free Energy Models of Molecular Solutions

Lawrence R. Pratt and Dilip Asthagiri

9.1 Introduction

Computing thermodynamic properties is the most important validation of simulations of solutions and biophysical materials. The potential distribution theorem (PDT) presents a partition function to be evaluated for the *excess* chemical potential of a molecular component which is part of a general thermodynamic system. The excess chemical potential of a component α is that part of the chemical potential of Gibbs which would vanish if the intermolecular interactions were to vanish. Therefore, it is just the part of that chemical potential that is interesting for consideration of a complex solution from a molecular basis. Since the excess chemical potential is measurable, it also serves the purpose of validating molecular simulations.

In this chapter, we discuss, exemplify, and thus support the assertions that the potential distribution theorem provides:

1. A *general* basis for the theory of solutions
2. A *practical* basis for the calculation of solution thermodynamic properties
3. A useful tool for the development of physically motivated *approximate* models of solution thermodynamics, particularly in view of quasichemical extensions

The last of these assertions deserves emphasis here in advance of more-technical developments. In the era of computing capabilities that are widely available and steadily advancing, the lack of a revealing theoretical model often means that simulation results are not as informative as they might be. An example is the theory of classic hydrophobic effects. Decades of correct simulation of aqueous solutions of hydrophobic solutes produced imperceptible progress in our physical understanding of these systems [1]. But when cogent theoretical concepts based upon the potential distribution and quasichemical theories were recognized [2–5], unanticipated conclusions could be identified [6–10], conclusions that could then be tested, refined, and consolidated.

Potential distribution methods are conventionally called *test particle* methods. Because the assertions above outline a general and basic position for the potential distribution theorem, it is appropriate that the discussion below states the potential

distribution theorem in a general way. This will seem somewhat unconventional. The generality we seek will emphasize cases of nonrigid solute molecules, and cases for which a sufficiently accurate force field describing solute–solvent interactions is not available. The latter case is one for which a chemical description of solute–solvent interactions is required as a practical matter. The ability to treat chemical interactions is an important pay-off of our generality here.

We achieve this generality on the basis of two devices introduced below. The first is the concept and notation of a *conditional mean,* average, or expected value. The condition is often the conformation of a distinguished solute molecule. This permits a proper and economical description of nonrigid molecules by consideration of rigid cases. The second device is a *rule of averages* that permits translation of PDT results into results for physical average values. These two devices have a close correspondence to the classic statistical tools of *stratification* and *importance weighting* [11]. These devices are what makes the PDT a generally practical approach, and it is remarkable that this approach can be viewed with such beautiful economy.

These devices – and the general picture that results – are discussed expansively elsewhere [10]. Here, we emphasize brevity in discussing the general results in order to leave space for consideration of some detailed examples.

9.1.1 Example: Zn^{2+}(aq) and Metal Binding of Zn Fingers

Understanding the role of metals in biological processes is a frontier area in biophysics. Metals are widely distributed in the body and play key roles in many processes, usually in association with metallo-proteins. These metallo-proteins, thought to comprise nearly a third of the proteome, carry out tasks as diverse as gene regulation, metal homeostasis, respiration, and metabolism. Understanding the functioning of metallo-proteins at a molecular level using atomically detailed simulations is the desideratum, but extant force fields are of limited utility in dealing realistically with metals in aqueous and biological systems. This is due largely to the chemical intricacies of the interactions involving the metal. The quasichemical approach discussed in Sect. 9.3 provides guidance for future studies of metallo-proteins.

Take the case of zinc-finger proteins (Fig. 9.1), which are important in gene regulation. The metal ion – Zn^{2+} – is necessary in maintaining the folded state. Why is zinc selected for this task? How is this selection achieved? For example, Fe^{2+}, a cation comparable to Zn^{2+}, is redox active; i.e., readily exists in different oxidation states. This ability to accept or release an electron proximal to genetic material makes Fe^{2+} undesirable. Indeed, nature utilizes a tetrahedral coordination of the metal – see Fig. 9.1 – so that the protein binds Zn^{2+} more favorably than Fe^{2+} by about 8 kcal mol^{-1}. This energy difference is large on the thermal scale. This is even more intriguing because coarse descriptors, such as the size of the ion and the net charge, would suggest only modest differences between Fe^{2+} and Zn^{2+}.

This difference between Fe^{2+} and Zn^{2+} is eventually related to the presence of unfilled d-orbitals in Fe^{2+}, whereas Zn^{2+} is a closed-shell ion. The organization of these d-orbitals is sensitive to the environment. In bulk water, Fe^{2+}(aq) is coordinated by six water molecules, which leads to the splitting of the d-orbitals into
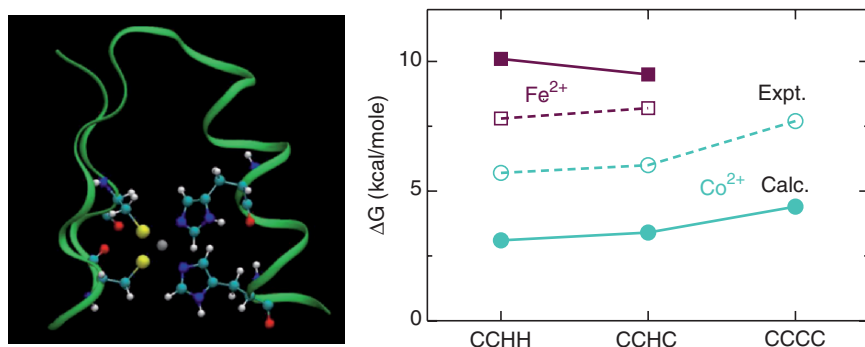
Fig. 9.1. Left panel: A model zinc finger obtained using the second domain of the transcription factor IIIA. The zinc ion (*gray sphere*) is coordinated tetrahedrally by two histidine (H) and two cysteine (C) residues. Right panel: Results showing the free energy change for displacing Zn^{2+} by other comparable ions Fe^{2+} and Co^{2+} from different binding motifs CCHH, CCHC, and CCCC, respectively
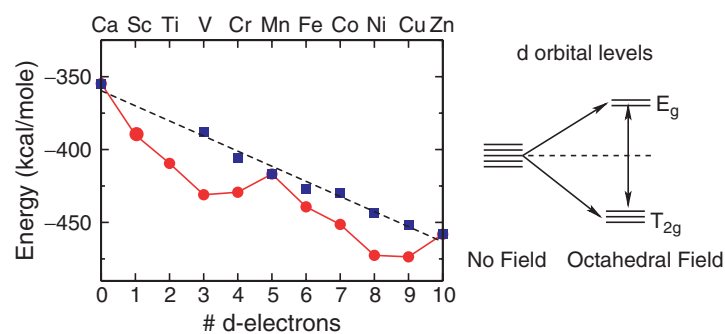

Fig. 9.2. Hydration free energies of dication transition metal ions in water, calculated as described in Sect. 9.3.3, p. 339; see also [12]. *Filled circles*: the actual hydration free energy computed using the quasichemical approach. *Filled squares*: the expected trend when the ligand field stabilization energy is removed from the hydration free energy. The stabilization energy was inferred from spectroscopic experiments. Note that removing the ligand field stabilization reveals the linear decrease along the period. The panel on the right illustrates the splitting of the metal d-orbitals due to the six water ligands arranged octahedrally around the cation

two groups, as shown in Fig. 9.2. The three T_{2g} orbitals are lowered in energy, and the two E_g orbitals are elevated above the original degenerate level. The *size* of the splitting depends on the ligand. In water, the splitting is small, and the orbitals are filled in a high-spin pattern. In a tetrahedral field, as occurs in the zinc-finger protein, the splitting pattern is *inverted*. Considering again how electrons organize in this new set of orbitals, the stabilization is less than in the octahedral field. This is the feature that governs why Zn^{2+} preferentially binds the zinc-finger rather than either Fe^{2+} or Co^{2+}.

In Fig. 9.2 we present results of a first-of-a-kind study of the hydration of the first-transition-row metals within the quasichemical framework. The biphasic behavior of the actual hydration free energy is consistent with features inferred experimentally. Removing the ligand field effects reveals the linear decrease [12]. The results shown in Fig. 9.2 are largely outside the purview of extant simulation techniques, but are treated simply in the quasichemical framework developed below.

9.2 Background Notation and Discussion of the Potential Distribution Theorem

Here we establish notation that is integral to this topic in the course of discussion of basic features of the potential distribution theorem (PDT).

9.2.1 Some Thermodynamic Notation

The PDT focuses on the chemical potentials μ_α that compose the Gibbs free energy

$$G\left(T, p, \mathbf{n}\right) = \sum_\alpha n_\alpha \mu_\alpha \,, \tag{9.1}$$

of a fluid solution. Here T is the thermodynamic temperature, p is the pressure, and $\mathbf{n} = \{n_1, \ldots, n_\alpha, \ldots\}$ are the particle numbers of molecules of each type.

For our problems here these chemical potentials are cast as

$$\mu_\alpha = k_{\mathrm{B}} T \ln \left[\frac{\rho_\alpha \Lambda_\alpha^3}{q_\alpha^{\mathrm{int}}} \right] + \mu_\alpha^{\mathrm{ex}} \,. \tag{9.2}$$

The first term on the right is the formula for the chemical potential of component α at density $\rho_\alpha = n_\alpha/V$ in an ideal gas, as would be the case if interactions between molecules were negligible. k_{B} is Boltzmann's constant, and V is the volume of the solution. The other parameters in that ideal contribution are properties of the isolated molecule of type α, and depend on the thermodynamic state only through T. Specifically, V/Λ_α^3 is the translational contribution to the partition function of single α molecule at temperature T in a volume V

$$\frac{1}{\Lambda_\alpha} = \frac{1}{h} \int \exp \left[-\frac{1}{k_{\mathrm{B}} T} \left(\frac{p^2}{2m_\alpha} \right) \right] \mathrm{d}p \,, \tag{9.3}$$

with h Planck's constant, and m_α the mass of the molecule. q_α^{int} is the contribution to that partition function due to all other degrees of freedom of that molecule. Thus, we could follow the standard practice of textbooks on statistical thermodynamics, writing (9.2) as

$$\mu_\alpha = k_{\mathrm{B}} T \ln \left[\frac{n_\alpha}{Q\left(n_\alpha = 1, T, V\right)} \right] + \mu_\alpha^{\mathrm{ex}} \,, \tag{9.4}$$

where $Q\left(n_\alpha = 1, T, V\right) = V q_\alpha^{\text{int}}/\Lambda_\alpha^3$ is the canonical partition function for the circumstances $\left(n_\alpha = 1, T, V\right)$. It is helpful to recognize the natural multiplicative contribution of V to $Q\left(n_\alpha = 1, T, V\right)$, so (9.2) is preferred.

The PDT then asserts that

$$\boxed{\mathrm{e}^{-\beta\mu_\alpha^{\text{ex}}} = \left\langle\!\left\langle \mathrm{e}^{-\beta\Delta U_\alpha} \right\rangle\!\right\rangle_0}.\tag{9.5}$$

Here $\beta^{-1} = k_{\text{B}} T$, ΔU_α is the binding potential energy of a molecule of type α to the solution of interest, and the brackets $\left\langle\!\left\langle \ldots \right\rangle\!\right\rangle_0$ mean the average over the thermal motion of the solution and the distinguished molecule, uncoupled. Dissecting, analyzing, and exemplifying this simple formula will be a principal task of what follows. Note that it is a defined sum over cases of the Boltzmann factor – with temperature in a customary position – of a defined energy. Thus, this formula should be seen as analogous to conventional partition functions.

The potential distribution theorem has been around for a long time [13–17], but not as long as the edifice of Gibbsian statistical mechanics where traditional partition functions were first encountered. We refer to other sources [10] for detailed derivations of this PDT, suitably general for the present purposes.

Our point of view is that the evaluation of the partition function (9.5) can be done by using any available tool, specifically including computer simulation. If that computer simulation evaluated the mechanical pressure, or if it simulated a system under conditions of specified pressure, then μ_α^{ex} would have been determined at a known value of p. With temperature, composition, and volume also known, (9.2) and (9.1) permit the construction of the full thermodynamic potential. This establishes our first assertion that the potential distribution theorem provides a basis for the general theory of solutions.

The remaining two assertions are to be established more inductively, and we reveal facts and include examples supporting these assertions in the following discussions.

9.2.2 Some Statistical Notation

Notation to describe general results for possibly complicated molecular components can be tricky. The notation exploited here is satisfactorily detailed, yet not burdensome. Specifically, we strive to cast important results in terms of coordinate-independent averages to permit generality and transparency.

Carrying out a simulation, which we view as the readiest source of data, does require coordinate choices. Thus, we do need some notation for coordinates, and we use $\mathscr{R}^n$ generically to denote the configuration of a molecule of n atoms, including translational, orientational, and conformational positioning; see Fig. 9.3. This notation suggests that Cartesian coordinates of each atom would be satisfactory, in principle, but does not require any specific choice.

Specification of the configuration of a complex molecule leads next to an essential element in our notation: conditional averages [18, 19]. The joint probability $P(A, B)$ of events A and B may be expressed as $P(A, B) = P(A|B)P(B)$ where
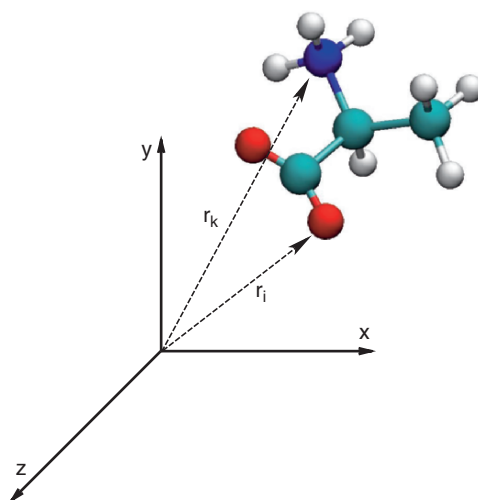
Fig. 9.3. Illustration of the definition of conformational coordinates, and the rotation $\mathscr{R}^n = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$

$P(B)$ is the marginal distribution of B and $P(A|B)$ is the distribution of A conditional on B, provided that $P(B) \neq 0$. The expectation of A conditional on B is $\langle A|B \rangle$, the expectation of A evaluated with the distribution $P(A|B)$ for specified B. In many texts [19], that object is denoted as $E(A|B)$ but the bracket notation for 'average' is firmly establish in the present subject, so we follow that precedent.

Our statement of the PDT (9.5) specifically considers two independent systems: first, a *distinguished* molecule of the type of interest and, second, the solution of interest. Our general expression of the PDT (9.5) can then be cast as

$$\mathrm{e}^{-\beta \mu_\alpha^{\mathrm{ex}}} = V^{-1} \int \left\langle \mathrm{e}^{-\beta \Delta U_\alpha} | \mathscr{R}^n \right\rangle_0 s_\alpha^{(0)}(\mathscr{R}^n) \mathrm{d}\left(\mathscr{R}^n\right) \ . \tag{9.6}$$

$V^{-1} s_\alpha^{(0)}(\mathscr{R}^n)$ is the normalized thermal distribution of configurations of the distinguished molecule in isolation [10], i.e., the required marginal distribution. The remaining set of brackets here indicates the average over solvent coordinates. The second set of brackets are not written on the right here because the averaging over solute coordinates is explicitly written out. This last formula is

$$\mathrm{e}^{-\beta \mu_\alpha^{\mathrm{ex}}} = V^{-1} \int \mathrm{e}^{-\beta \mu_\alpha^{\mathrm{ex}}(\mathscr{R}^n)} s_\alpha^{(0)}(\mathscr{R}^n) \mathrm{d}\left(\mathscr{R}^n\right) \ , \tag{9.7}$$

with the natural identification

$$\boxed{\mathrm{e}^{-\beta \mu_\alpha^{\mathrm{ex}}(\mathscr{R}^n)} \equiv \left\langle \mathrm{e}^{-\beta \Delta U_\alpha} | \mathscr{R}^n \right\rangle_0} \ . \tag{9.8}$$

The average indicated on the right is the average over the thermal motion of the solution with the solute positioned at $\mathscr{R}^n$, with no coupling between these subsystems.

Notice that if we assign number densities of solute molecules, $\rho_\alpha(\mathscr{R}^n)$, in conformation $\mathscr{R}^n$, according to

$$\rho_\alpha(\mathscr{R}^n) = s_\alpha^{(0)}(\mathscr{R}^n) \left(\frac{\mathrm{e}^{\beta\mu_\alpha} q_\alpha^{\mathrm{int}}}{\Lambda_\alpha{}^3} \right) \left\langle \mathrm{e}^{-\beta\Delta U_\alpha} | \mathscr{R}^n \right\rangle_0 \; , \tag{9.9}$$

then (9.5) is obtained from (9.9) by integrating over all solute conformations, recognizing that

$$n_\alpha = \int \rho_\alpha(\mathscr{R}^n) \mathrm{d}\left(\mathscr{R}^n \right) \; . \tag{9.10}$$

That integrating out produces the desired interaction contribution μ_α^{ex} to the chemical potential. Eliminating the quantities in parentheses produces the interesting form [20]

$$\beta\mu_\alpha^{\mathrm{ex}} = \left\langle \ln \left[\frac{\rho_\alpha(\mathscr{R}^n)}{\rho_\alpha s_\alpha^{(0)}(\mathscr{R}^n)} \right] \right\rangle + \left\langle \beta\mu_\alpha^{\mathrm{ex}}(\mathscr{R}^n) \right\rangle \; . \tag{9.11}$$

The first term on the right suggests an entropic contribution associated with a shift of conformational probability due to the solution environment. In fact, the form (9.11) holds for each $\mathscr{R}^n$ without the brackets. This then shows that

$$\mathrm{e}^{\beta\mu_\alpha^{\mathrm{ex}}} = n_\alpha{}^{-1} \int \mathrm{e}^{\beta\mu_\alpha^{\mathrm{ex}}(\mathscr{R}^n)} \rho_\alpha(\mathscr{R}^n) \mathrm{d}\left(\mathscr{R}^n \right) \; , \tag{9.12}$$

a relation pleasingly symmetrical to (9.7). This connects to the inverse formula (9.15) that comes up later.

9.2.3 Observations on the PDT

Physical Generality of Molecular Interactions is Permitted

The PDT partition function formula (9.5) does not require that ΔU_α adopt a specifically simplified form such as additive contributions over pairs of molecules involved. ΔU_α is simply

$$\Delta U_\alpha = U_{N+\alpha} - U_N - U_\alpha \; , \tag{9.13}$$

where U_N and U_α are the mechanical potential energies of the two independent subsystems, the solution and distinguished molecule, respectively, and $U_{N+\alpha}$ is the mechanical potential energy of the joint system. Thus, systems with N-body interaction forms, for example, with polarizability and induction effects, are naturally treated in this development. For example, ab initio molecular dynamics calculations involve non-pair-decomposable interactions.

Rule of Averages and the Inverse Formula

To evaluate physical averages involving the two subsystems considered jointly, supply the Boltzmann factor of the coupling energy in a numerator and denominator according to

$$\langle F_\alpha \rangle = \frac{\langle\langle e^{-\beta \Delta U_\alpha} F_\alpha \rangle\rangle_0}{\langle\langle e^{-\beta \Delta U_\alpha} \rangle\rangle_0} \ . \tag{9.14}$$

Taking the case $F_\alpha = e^{\beta \Delta U_\alpha}$ leads to a nontrivial formula for the inverse of (9.5)

$$\boxed{\langle e^{\beta \Delta U_\alpha} \rangle = e^{\beta \mu_\alpha^{\mathrm{ex}}}} \ . \tag{9.15}$$

The brackets on the left of (9.15) indicate the *fully coupled* thermal averages, involving the *actual* interactions between the solution and the distinguished molecule of the joint system, specifically with the distinguished solute present. Equation (9.12) made a preliminary presentation of this result.

To assist in considering the averages in (9.5) and (9.15), introduce the probability distribution functions $\mathscr{P}_\alpha^{(0)}(\varepsilon) = \langle\langle \delta(\varepsilon - \Delta U_\alpha) \rangle\rangle_0$, and $\mathscr{P}_\alpha(\varepsilon) = \langle \delta(\varepsilon - \Delta U_\alpha) \rangle$. For example, then

$$e^{-\beta \mu_\alpha^{\mathrm{ex}}} = \langle\langle e^{-\beta \Delta U_\alpha} \rangle\rangle_0 = \int \mathscr{P}_\alpha^{(0)}(\varepsilon) e^{-\beta \varepsilon} \mathrm{d}\varepsilon \ , \tag{9.16}$$

and, with (9.14)

$$\mathscr{P}_\alpha(\varepsilon) = e^{-\beta(\varepsilon - \mu_\alpha^{\mathrm{ex}})} \mathscr{P}_\alpha^{(0)}(\varepsilon) \ . \tag{9.17}$$

This casts the PDT partition function as a normalizing denominator, as is customary for a partition function.

The determination of the distribution $\mathscr{P}_\alpha^{(0)}(\varepsilon)$, or $\mathscr{P}_\alpha(\varepsilon)$, to sufficient accuracy in the required ε range is a common practical difficulty. Since the averaging that defines $\mathscr{P}_\alpha(\varepsilon)$ concentrates on thermally optimal interactions ΔU_α, sampling of configurations with less-than-favorable interactions can be sketchy or nonexistent. Those unfavorable interactions are typically important to the integrated value

$$e^{\beta \mu_\alpha^{\mathrm{ex}}} = \int \mathscr{P}_\alpha(\varepsilon) e^{\beta \varepsilon} \mathrm{d}\varepsilon \ , \tag{9.18}$$

because of the exponential weight in the integrand. This is an issue that we take up again in the example of Sect. 9.4. In such a case, coming at the problem from an inverse direction, also utilizing (9.17), can be advantageous.

As an attractive statement of such a two-sided approach, consider cutting off the integral of (9.18) at a value $\varepsilon = \bar{\varepsilon}$ above which the determination of $\mathscr{P}_\alpha(\varepsilon)$ is not sufficiently accurate. For $\varepsilon > \bar{\varepsilon}$, consider exploiting $\mathscr{P}_\alpha^{(0)}(\varepsilon)$ and (9.17) to characterize $\mathscr{P}_\alpha(\varepsilon)$ in that high-interaction-energy regime. A simple calculation then shows that

$$\mu_\alpha^{\mathrm{ex}} = k_{\mathrm{B}} T \ln \int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha(\varepsilon) e^{\beta \varepsilon} \mathrm{d}\varepsilon - k_{\mathrm{B}} T \ln \int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha^{(0)}(\varepsilon) \mathrm{d}\varepsilon \ . \tag{9.19}$$

The mean value of the interaction potential energy should provide some guidance on the value of the first of the terms on the right; it helps that those interaction energies will have a lower bound. The second term then primarily addresses entropic contributions to μ_α^{ex}; that integral accumulates the weight of the favorable configurations, well-bound to the solute, that the solvent host offers the solute without coercion.

The result (9.19) should not depend on the cut-off parameter $\bar{\varepsilon}$, but a pragmatic choice is necessary for a practical calculation. It is clear that the practical success of such an approach depends on the availability of a parameter $\bar{\varepsilon}$ that cuts off both distributions to a measurable extent. This formula therefore expresses the basic idea of histogram overlap procedures [21, 22]; see Fig. 9.5, p. 345, for an example.

Dependence of the PDT Partition Function on the Simulation Ensemble

Occasionally alternative expressions of the PDT (9.5) have been proposed [23–25]. These alternatives arise from consideration of statistical thermodynamic manipulations associated with a particular ensemble, and the distinguishing features of those alternative formulae are relics of the particular ensemble considered. On the other hand, relics specific to an ensemble are not evident in the PDT formula (9.5). These alternative formulae should give the same result in the thermodynamic limit.

As an example [24, 25] consider calculation of $\left\langle\left\langle V\mathrm{e}^{-\beta\Delta U_\alpha}\right\rangle\right\rangle_0 / \left\langle\left\langle V\right\rangle\right\rangle_0$ in an isothermal–isobaric ensemble, for which the volume V fluctuates. In this formula V is a relic of the ensemble considered. A calculation using the rule of averages (9.14) then leads to [10]

$$
\frac{\left\langle\left\langle V\mathrm{e}^{-\beta\Delta U_\alpha}\right\rangle\right\rangle_0}{\left\langle\left\langle V\right\rangle\right\rangle_0} = \left\langle\left\langle\mathrm{e}^{-\beta\Delta U_\alpha}\right\rangle\right\rangle_0 \frac{\left\langle V\right\rangle}{\left\langle\left\langle V\right\rangle\right\rangle_0}
$$

$$
\sim \left\langle\left\langle\mathrm{e}^{-\beta\Delta U_\alpha}\right\rangle\right\rangle_0 \left(1 + \frac{1}{\left\langle V\right\rangle}\left[\frac{\partial\left\langle V\right\rangle}{\partial n_\alpha}\right]_{T,p,n_{\gamma\neq\alpha}}\right)
$$

$$
= \left\langle\left\langle\mathrm{e}^{-\beta\Delta U_\alpha}\right\rangle\right\rangle_0 \left(1 + \frac{1}{\left\langle V\right\rangle}\left[\frac{\partial\mu_\alpha}{\partial p}\right]_{T,\mathbf{n}}\right) . \quad (9.20)
$$

The correction displayed is negligible relative to 1, in the macroscopic limit. The *independence* in the thermodynamic limit of the PDT on a choice of simulation ensemble used for statistical evaluation is a difference from the partition functions encountered in Gibbsian statistical thermodynamics.

Size Consistency and the Thermodynamic Limit

The chemical potentials sought are intensive properties of the system, in the usual thermodynamic language [26]. Furthermore, ΔU_α is a quantity of molecular order of magnitude. Specifically, the ΔU_α defined by (9.13) should be system-size independent for typical configurations of thermodynamically large systems. Because of

that, the probability distribution functions of (9.17) should be independent of system size for large systems. This aspect facilitates the development of physical models for those distribution functions.

The General Computational Tricks Work also for the PDT

The general principles for estimating free energies and of high-dimensional integrals [11] typically also apply to the estimation of the PDT partition function (9.5). These principles include importance weighting and stratification, which can lead to thermodynamic integration methods. These topics are treated extensively elsewhere in the earlier chapters of this book, and we limit ourselves here to a couple of specific points. It is worth noting examples [27–33], not discussed further, of coordinated theoretical studies of realistic cases using these general tools.

The PDT approach enables the precise assessment of the differing consequences of intermolecular interactions of different types. Separate the interaction ΔU_α into two contributions $\Delta \tilde{U}_\alpha + \Phi_\alpha$. If Φ_α were not present, we would have

$$\mathrm{e}^{-\beta \tilde{\mu}_\alpha^{\mathrm{ex}}} = \left\langle\!\!\left\langle \mathrm{e}^{-\beta \Delta \tilde{U}_\alpha} \right\rangle\!\!\right\rangle_0 . \tag{9.21}$$

The tilde over $\tilde{\mu}_\alpha^{\mathrm{ex}}$ indicates that this is the interaction contribution to the chemical potential of the solute when $\Phi_\alpha = 0$. The properties of the solution alone are unchanged.

With the result (9.21) available, consider the remainder

$$\mathrm{e}^{-\beta(\mu_\alpha^{\mathrm{ex}} - \tilde{\mu}_\alpha^{\mathrm{ex}})} = \frac{\left\langle\!\!\left\langle \mathrm{e}^{-\beta \Delta U_\alpha} \right\rangle\!\!\right\rangle_0}{\left\langle\!\!\left\langle \mathrm{e}^{-\beta \Delta \tilde{U}_\alpha} \right\rangle\!\!\right\rangle_0} . \tag{9.22}$$

Noting, $\mathrm{e}^{-\beta \Delta U_\alpha} = \mathrm{e}^{-\beta \Delta \tilde{U}_\alpha} \times \mathrm{e}^{-\beta \Phi_\alpha}$, choose $F_\alpha = \mathrm{e}^{-\beta \Phi_\alpha}$, and then use the rule of averages to find

$$\mathrm{e}^{-\beta(\mu_\alpha^{\mathrm{ex}} - \tilde{\mu}_\alpha^{\mathrm{ex}})} = \left\langle \mathrm{e}^{-\beta \Phi_\alpha} \right\rangle_{\mathrm{r}} . \tag{9.23}$$

This takes the conventional form of standard thermodynamic perturbation theory, but with the decisive feature that interactions with only one molecule need be manipulated. Here $\langle \ldots \rangle_{\mathrm{r}}$ indicates averaging for the case that the solution contains a distinguished molecule which interacts with the rest of the system on the basis of the function $\Delta \tilde{U}_\alpha$, i.e., the subscript 'r' identifies an average for the *reference system*. Notice that a normalization factor for the *intra*molecular distribution cancels between the numerator and denominator of (9.22).

As an example, consider assessment of classic electrostatic solute–solvent interactions associated with solute partial charges in force-field models. The contribution of electrostatic interactions is then isolated as

$$\mathrm{e}^{-\beta(\mu_\alpha^{\mathrm{ex}} - \tilde{\mu}_\alpha^{\mathrm{ex}})} = \int \tilde{\mathscr{P}}_\alpha(\varepsilon) \mathrm{e}^{-\beta \varepsilon} \mathrm{d}\varepsilon , \tag{9.24}$$

where $\tilde{\mathscr{P}}_\alpha(\varepsilon) \equiv \langle\delta\left(\varepsilon - \Phi_\alpha\right)\rangle_\mathrm{r}$, and $\Phi_\alpha = \Delta U_\alpha - \Delta\tilde{U}_\alpha$ is the electrostatic contribution to the solute–solvent interactions. The reference interactions would typically include van der Waals interactions, and the perturbation corresponds to transforming an uncharged molecule to a charged molecule.

Modeling of $\tilde{\mathscr{P}}_\alpha(\varepsilon)$ can be motivated by a simple thermodynamic model for this electrostatic contribution. The *Born model* [34] for the hydration free energy of a spherical ion of radius R_α with a charge q_α at its center is

$$\mu_\alpha^\mathrm{ex} \approx \tilde{\mu}_\alpha^\mathrm{ex} - \frac{q_\alpha^2}{2R_\alpha}\left(\frac{\epsilon - 1}{\epsilon}\right). \qquad (9.25)$$

The dielectric constant of the external medium is ϵ. The electrostatic contribution is proportional to $q_\alpha{}^2$. To isolate this behavior, change variables as $\Phi_\alpha = q_\alpha\varphi_\alpha$, and consider the Gaussian model

$$\tilde{\mathscr{P}}_\alpha(\varepsilon|\mathscr{R}^n) \approx \frac{1}{\sqrt{2\pi\left\langle\delta\varphi_\alpha{}^2|\mathscr{R}^n\right\rangle_\mathrm{r}}}\exp\left[-\frac{1}{2}\frac{\left(\varepsilon - \langle\varphi_\alpha|\mathscr{R}^n\rangle_\mathrm{r}\right)^2}{\left\langle\delta\varphi_\alpha{}^2|\mathscr{R}^n\right\rangle_\mathrm{r}}\right], \qquad (9.26)$$

adopting the notation of conditional means to treat the conformational status in the general molecular solute case. φ_α is the electrostatic potential exerted by the solution on the distinguished solute. Consideration of the quantity $\langle\varphi_\alpha|\mathscr{R}^n\rangle_\mathrm{r}$ requires some conceptual subtlety. This is intended to be the electrostatic potential of the solution *induced* by reference interactions between the solute and the solution. Any contribution to the electrostatic potential that exists in the absence of those reference interactions, i.e., the *potential of the phase*, is trickier, and we defer discussion of those issues to another forum [10]. Here, we find

$$\mu_\alpha^\mathrm{ex}(\mathscr{R}^n) \approx \tilde{\mu}_\alpha^\mathrm{ex}(\mathscr{R}^n) + q_\alpha\langle\varphi_\alpha|\mathscr{R}^n\rangle_\mathrm{r} - \frac{\beta q_\alpha{}^2}{2}\left\langle\delta\varphi_\alpha{}^2|\mathscr{R}^n\right\rangle_\mathrm{r}. \qquad (9.27)$$

The thermodynamic chemical potential is then obtained by averaging the Boltzmann factor of this conditional result using the *isolated solute* distribution function $s_\alpha^{(0)}(\mathscr{R}^n)$. Notice that the fluctuation contribution necessarily lowers the calculated free energy.

This analysis of the consequences of interactions of different types is an example of the general technique of importance weighting, discussed in Chap. 3. An operational view is that the additional factor $\mathrm{e}^{\beta\Delta\tilde{U}_\alpha}$ serves to broaden the sampling. Following this idea, we might consider another configurational function $1/\Omega$ that helpfully broadens the sampling and write

$$\mathrm{e}^{-\beta\mu_\alpha^\mathrm{ex}(\mathscr{R}^n)} = \frac{\left\langle\mathrm{e}^{-\beta\Delta U_\alpha}\Omega|\mathscr{R}^n\right\rangle_{1/\Omega}}{\langle\Omega|\mathscr{R}^n\rangle_{1/\Omega}}. \qquad (9.28)$$

The sampling distribution indicated for $\langle\ldots|\mathscr{R}^n\rangle_{1/\Omega}$ is proportional to

$$\mathscr{P}_B(\mathscr{N})/\Omega(\mathscr{N}).$$

The denominator of (9.28) would correspond to the factor $e^{\beta \tilde{\mu_\alpha}^{\text{ex}}(\mathscr{R}^n)}$ of (9.22). In contrast to the view of (9.22), here it is not assumed that $\langle \Omega | \mathscr{R}^n \rangle_\Omega$ is separately known. Thus bias or variance-reduction issues would involve both the numerator and the denominator of (9.28). Compare the discussion of this topic in Chap. 6.

$1/\Omega$ is called an *importance function* or sometimes an *umbrella function* [35]. The latter name arises from the view that $1/\Omega$ broadens the sampling to cover relevant cases more effectively. Since $1/\Omega$ is involved as an unnormalized probability, it should not change sign, and it should not be zero throughout regions where the unmodified distribution and integrand are nonzero.

As an example of importance-weighting ideas, consider the situation that the actual interest is in hydration free energies of a distinct conformational states of a complex solute. Is there a good reference system to use to get comparative thermodynamic properties for all conformers? There is a theoretical answer that is analogous to the Hebb training rule of neural networks [36, 37], and generalizes a procedure of [21]

$$\frac{1}{\Omega} = \sum_{\text{conformations}} e^{-\beta \Delta U_c}. \tag{9.29}$$

The sum is over interesting conformations to be exploited in this way, and the summand is the Boltzmann factor of the solute–solvent interaction when the solute adopts a specific conformational structure. When this umbrella function is used to get the free energy for the solute in a specific conformation, it should resemble at least one conformer in the sum. So the umbrella should cover every conformation in the family, and this is literally the point of the original *umbrella sampling:* $1/\Omega$ "should cover simultaneously the regions of configuration space relevant to two or more physical systems" [35]. Jointly matching several members of the family will help too. The penalty is just the sum over the family. The hydration free energy of the reference system, that is the denominator of (9.28), is not required for the evaluation of free energy *differences* between conformations.

Thermodynamic Integration and Stratification

To organize the description of interactions of a specified type, it is often helpful to introduce an ordering parameter λ in

$$\mu_\alpha^{\text{ex}}(\mathscr{R}^n) = \tilde{\mu}_\alpha^{\text{ex}}(\mathscr{R}^n) - k_B T \ln \left\langle e^{-\beta \lambda \Phi_\alpha} | \mathscr{R}^n \right\rangle_{\text{r}} \tag{9.30}$$

with the intention that $\lambda = 1$ provides the physical interactions of interest. λ might be viewed as a perturbative parameter in cases where it appears naturally as a gauge of the strength of solute–solvent interactions. The derivative

$$\frac{\partial \mu_\alpha^{\text{ex}}(\mathscr{R}^n)}{\partial \lambda} = \frac{\left\langle \Phi_\alpha e^{-\beta \lambda \Phi_\alpha} | \mathscr{R}^n \right\rangle_{\text{r}}}{\left\langle e^{-\beta \lambda \Phi_\alpha} | \mathscr{R}^n \right\rangle_{\text{r}}} = \left\langle \Phi_\alpha | \mathscr{R}^n \right\rangle_\lambda \tag{9.31}$$

is then a straightforward consequence that uses a rule of averages in a natural way. The final average indicates that the solute–solvent interactions are $\Delta \tilde{U}_\alpha + \lambda \Phi_\alpha$, i.e., the reference system plus the perturbation coupled to the extent λ. The use of λ to

scale partial atomic charges on the distinguished solute is an example. On this basis, mere quadrature provides an evaluation of

$$\mu_\alpha^{\mathrm{ex}}(\mathscr{R}^n) = \tilde{\mu}_\alpha^{\mathrm{ex}}(\mathscr{R}^n) + \int_0^1 \langle \Phi_\alpha | \mathscr{R}^n \rangle_\lambda \, \mathrm{d}\lambda \, . \tag{9.32}$$

From (9.27), we see that this approach will work nicely if the variance is always small; Taylor's theorem with remainder tells us that the error of the first-derivative – mean-field – contribution is proportional to the second derivative evaluated at an intermediate λ. That second derivative can be identified with the variance as in (9.27). If that variance is never large, then this approach should be particularly effective. For further discussion, see Chap. 4 on thermodynamic integration, and Chap. 6 on error analysis in free energy calculations.

Evaluation of the integrand of (9.32) at several intermediate values of λ to perform the quadrature amounts to a divide-and-conquer method that is an example of stratification. The advantage of stratification can be considered generally [11, Sect. 5.3], [38, Sect. 4.5], [39, Sect. 7.8], or for specific cases [10]. The idea is that statistical uncertainties are mitigated by a nonstatistical subdivision of the problem, solution of the subdivided problems, and then recomposition of the whole. Stratified calculations such as thermodynamic integration are typically decisive, enabling maneuvers [10]. Such a computational strategy can be embarrassingly parallel.

We can return to the issue of analysis of electrostatic contributions to hydration free energies of ions to give an example of stratification [40]. Suppose that we recognize a partitioning of the statistical possibilities so that the distribution $\tilde{\mathscr{P}}_\alpha(\varepsilon|\mathscr{R}^n)$ in (9.24) can be expressed as a linear combination of contributions from different strata corresponding to configurational substates of the system. For the case of electrostatic interactions in aqueous solutions the strata, or substates, might be distinct configurations defined by different hydrogen-bonding possibilities for the solute and solvent molecules. Indexing those substates by s, we then analyze the joint probability distribution of φ_α and s, $\tilde{\mathscr{P}}_\alpha(\varepsilon, s|\mathscr{R}^n)$, assuming that the marginal distributions $p_\alpha(s|\mathscr{R}^n) = \int \tilde{\mathscr{P}}_\alpha(\varepsilon, s|\mathscr{R}^n)\mathrm{d}\varepsilon$ can be obtained from simulation calculations, and further assuming that the conditional probability distributions $\tilde{\mathscr{P}}_\alpha(\varepsilon|s, \mathscr{R}^n)$ are Gaussian, again with parameters obtained from simulation data. Then

$$\tilde{\mathscr{P}}_\alpha(\varepsilon|\mathscr{R}^n) = \sum_s \tilde{\mathscr{P}}_\alpha(\varepsilon|s, \mathscr{R}^n) p_\alpha(s|\mathscr{R}^n) \tag{9.33}$$

is the generally valid total probability formula. Any structural parameter s that is considered to be significant could be exploited here.

Test Particle Techniques

As noted in the Introduction, the PDT is widely recognized with the moniker *test particle method*. This name reflects a view of how calculations of $\langle\langle \mathrm{e}^{-\beta \Delta U_\alpha} \rangle\rangle_0$ might be tried: solute conformations are sampled, solvent configurations are sampled, and then the two systems are superposed; the energy change is calculated, and

the Boltzmann factor for that the energy change is scored. Such an approach can be practical when the test solute interrogates the shortest length scale of the solvent, as when the solute is literally a particle. In fact, in Sect. 9.4, we give an example of such a calculation.

Direct test particle methods are expected to be inefficient, compared to other possibilities, for molecular systems described with moderate realism. Successful placements of a test particle may be complicated, and placements with favorable Boltzmann factor scores may be rare. Fortunately, the tools noted above are generally available to design more-specific approaches for realistic cases.

Nevertheless, direct test particle calculations have been of great conceptual importance, particularly in cases where there is a consensus on the relevance of simplified model solutes [2–4, 6, 9, 10, 41–45]. The related particle insertion techniques are used for simulating phase equilibria, as discussed in Chap. 10.

9.3 Quasichemical Theory

Quasichemical theory (QCT) is an important general extension of the PDT formula; expansive discussions are available elsewhere [10, 46–48]. As with the PDT, we focus on a distinguished molecule of the species of interest. We then define an *inner-shell* region for association of the solution components with the distinguished molecule. This inner shell is a region proximal to the distinguished molecule where the interactions with other solution components are particularly important, intricate, or strong. Consider an ion such as Be^{2+} in water as an example. The interactions with contacting water molecules are fundamentally chemical in character, and thus are expected to be strong and complicated.
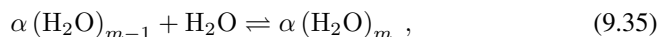
For simplicity of notation, we will here discuss the circumstance that the distinguished molecule is present at the lowest concentration. The occupants of the inner shell will be of one type only, solvent of type denoted by "w"; "w" = H_2O for example. Then our discussion can be more economical, though the ideas do have broader relevance. For a distinguished molecule of type α, we will encode the definition of the inner shell by an indicator function $b_\alpha(k)$ which is one when the kth solvent molecule occupies the defined inner shell, and zero otherwise. Then the PDT formula can be recast as

$$ e^{-\beta\mu_\alpha^{\mathrm{ex}}} = \left(1 + \sum_{m \geq 1} K_m \rho_{\mathrm{w}}{}^m\right) \times \left\langle\!\!\left\langle e^{-\beta\Delta U_\alpha} \prod_k [1 - b_\alpha(k)]\right\rangle\!\!\right\rangle_0 . \quad (9.34)$$

The right-most term is similar to the familiar PDT formula except that the indicator function combinations forbid binding of solution molecules to the defined inner shell. That last factor is recognized as the Boltzmann factor of the hydration free energy that would result if inner-shell binding were prohibited. The K_m are recognizable ratios of equilibrium concentrations – *equilibrium constants* – that are discussed

further below, and ρ_w is the bulk number density of the solvent. The leading factor on the right of (9.34) is a sum over cases for binding of solvent to the defined inner shell. Only a finite number of terms contribute to this sum. Notice that this expresses a stratification of the statistical problem. Here the variable identifying the strata is the number m of solvent occupants of the inner shell.

The K_m are descriptors of the aggregation reactions

$$\alpha \left(H_2O\right)_{m-1} + H_2O \rightleftharpoons \alpha \left(H_2O\right)_m \ , \tag{9.35}$$

specifically

$$\frac{K_m}{K_{m-1}} = \frac{\rho_{\alpha w_m}}{\rho_{\alpha w_{m-1}} \rho_w}$$

$$= \frac{\left(\dfrac{q^{int}_{\alpha w_m}}{\Lambda_{\alpha w_m}{}^3}\right) e^{-\beta \mu^{ex}_{\alpha w_m}}}{\left(\dfrac{q^{int}_{\alpha w_{m-1}}}{\Lambda_{\alpha w_{m-1}}{}^3}\right) e^{-\beta \mu^{ex}_{\alpha w_{m-1}}} \left(\dfrac{q^{int}_w}{\Lambda_w{}^3}\right) e^{-\beta \mu^{ex}_w}} \ , \tag{9.36}$$

with $K_0 = 1$. Together with the identification

$$e^{-\beta \mu^{ex}_{\alpha w_m}} = \left\langle\!\!\left\langle e^{-\beta \Delta U_{\alpha w_m}} \prod_{k \in \mathbf{m}} b_\alpha \left(k\right) \prod_{j \notin \mathbf{m}} \left[1 - b_\alpha \left(j\right)\right] \right\rangle\!\!\right\rangle_0 \ , \tag{9.37}$$

this adapts the concepts of the PDT to the cluster species αw_m viewed as conventional components of the solution. Further definition, and a fuller discussion of how this arises, can be found in [10]. The appearance of the language of chemical thermodynamics suggests an essentially thermodynamical derivation of the QCT (9.34); that is developed next as a worked exercise.

9.3.1 Cluster-Variation Exercise Sketched

Here we sketch a heuristic derivation of the quasichemical formula (9.34). Consider a solution of species α and w. The Gibbs free energy is

$$G = \mu_w n_w + \mu_\alpha n_\alpha \ . \tag{9.38}$$

If we can identify complexes αw_m that may form, we might wish to express G as

$$G = \hat{\mu}_w \hat{n}_w + \sum_{m \geq 1} \mu_{\alpha w_m} n_{\alpha w_m} \ . \tag{9.39}$$

$\hat{n}_w$ is intended to be the number of *uncomplexed* water molecules, and (9.39) requests the chemical potential $\hat{\mu}_w$ for those uncomplexed water molecules. The motivation for (9.39) is that species αw_m might be available to be purchased and scooped from a jar, so treating it as another chemical component seems natural. A conceptual hitch

is that (9.39) seems to have more composition variables than (9.38) does. To address this hitch, note that

$$n_\alpha = \sum_{m \geq 0} n_{\alpha w_m} \,, \tag{9.40a}$$

$$n_w = \hat{n}_w + \sum_{m \geq 1} m \, n_{\alpha w_m} \,. \tag{9.40b}$$

Substitution of (9.40) into (9.38) permits a comparison with (9.39). Such a comparison helps, but does not solve the problem that (9.39) seems to have superfluous variables.

This conceptual hitch is addressed by adjusting the values of the superfluous variables, subject to the constraints of (9.40), to make G stationary – minimal. Accounting for the constraints (9.40) by the standard procedure of Lagrange's undetermined multipliers [49] yields:

$$(\hat{\mu}_w - \lambda_w) \, \delta \hat{n}_w + \sum_{m \geq 1} (\mu_{\alpha w_m} - \lambda_\alpha - m\lambda_w) \, \delta n_{\alpha w_m} = 0 \,, \tag{9.41}$$

where λ_w and λ_α are the necessary Lagrange multipliers, and the indicated composition variations are now unconstrained. The conclusion is that:

$$\hat{\mu}_w = \lambda_w \,, \tag{9.42a}$$

$$\mu_{\alpha w_m} = \lambda_\alpha + m\lambda_w \,. \tag{9.42b}$$

With these specifications, the variation of G is

$$\delta G = \lambda_w \delta \hat{n}_w + \sum_{m \geq 1} (\lambda_\alpha + m\lambda_w) \, \delta n_{\alpha w_m} \,, \tag{9.43}$$

$$= \lambda_w \delta n_w + \lambda_\alpha \delta n_\alpha \,, \tag{9.44}$$

using (9.40). Now comparison with (9.38) leads to the identification $\lambda_w = \mu_w$, $\lambda_\alpha = \mu_\alpha$, and then further $\hat{\mu}_w = \mu_w$.

The relation $\mu_\alpha = \mu_{\alpha w_m} - m\mu_w$ leads to the quasichemical approach. This relation can be put into the form

$$\beta \mu_\alpha^{ex} = \beta \mu_{\alpha w_{m=0}}^{ex} + \ln x_m - \ln K_m \rho_w{}^m \,, \tag{9.45}$$

using the definitions (9.2) and (9.36), and the notation

$$x_n = \frac{\rho_{\alpha w_m}}{\rho_\alpha} = \frac{\rho_{\alpha w_m}}{\sum\limits_{m \geq 0} \rho_{\alpha w_m}} = \frac{K_m \rho_w{}^m}{1 + \sum\limits_{m \geq 1} K_m \rho_w{}^m} \,. \tag{9.46}$$

Then, finally

$$\beta \mu_\alpha^{ex} = \beta \mu_{\alpha w_{m=0}}^{ex} - \ln \left(1 + \sum_{m \geq 1} K_m \rho_w{}^m \right) \tag{9.47}$$

which is the first important message of (9.34).

The ideal factor $K_m^{(0)}(T)$ can be extracted from the result (9.36)

$$K_m = K_m^{(0)}(T) \frac{\left\langle\left\langle e^{-\beta \Delta U_{\alpha w m}}\right\rangle\right\rangle_0}{\left\langle\left\langle e^{-\beta \Delta U_{\alpha w m=0}}\right\rangle\right\rangle_0 \left(\left\langle\left\langle e^{-\beta \Delta U_{w}}\right\rangle\right\rangle_0\right)^m} \ . \tag{9.48}$$

In the cases that are the target of the present development, the evaluation of $K_m^{(0)}(T)$ can be a challenging calculation in its own right.

9.3.2 Results of Clustering Analyses

The development leading to (9.47) is extremely broad, and the notation of (9.48) is correspondingly broad. An important question is just what is meant precisely by notations such as $\left\langle\left\langle e^{-\beta \Delta U_{\alpha w m}}\right\rangle\right\rangle_0$ and specifically $\left\langle\left\langle e^{-\beta \Delta U_{\alpha w m=0}}\right\rangle\right\rangle_0$. A development that uses more specific notation [10] clarifies that

$$\left\langle\left\langle e^{-\beta \Delta U_{\alpha w m=0}}\right\rangle\right\rangle_0 = \left\langle\left\langle e^{-\beta \Delta U_{\alpha}} \prod_k \left[1 - b_\alpha(k)\right]\right\rangle\right\rangle_0 \ . \tag{9.49}$$

The quantities averaged on the right here are the Boltzmann factor for the binding energy, as is usual for the PDT, but multiplied by the indicator function for the event that there are zero occupants of the defined inner shell.

Notice that the quantity of (9.49) is the right-most factor of (9.34), and also appears in the denominators of the equilibrium ratios (9.48). Thus, that factor in (9.34) cancels precisely the same factor in the denominator of contributions from (9.48). This observation gives some perspective on how the basic quasichemical formula (9.34) is built, but does not address the leading term in (9.34) of "1", which does not have such a denominator. But in multiplying out of (9.34) we see that the term (9.49) is the correct contribution to the PDT formula for the case $m = 0$. Again, this observation serves to explain how the equilibrium ratios and the outer-shell contribution come together to formulate the general result.

If we simplify the sum of (9.34) to include only the maximum term, the $\hat{m}$th term, we find the formula

$$\mu_\alpha^{\text{ex}} \approx -k_{\text{B}} T \ln\left(K_{\hat{m}}^{(0)} \rho_{\text{w}}^{\hat{m}}\right) + \left(\mu_{\alpha w \hat{m}}^{\text{ex}} - \hat{m} \mu_{\text{w}}^{\text{ex}}\right) \tag{9.50}$$

in which the outer-shell contribution does not appear explicitly because it has been cancelled precisely.

9.3.3 Primitive Quasichemical Approximation

Equation (9.50) is the primitive quasichemical approximation that was used to obtain the results of Figs. 9.1 and 9.2. *Primitive* emphasizes that the equilibrium constants are obtained with initial neglect of the effects of the outer-shell material, as (9.50)

suggests. It remains to establish practical methods of calculation for the various quantities that appear there, and we here provide some of the specifics that form the basis of the results in Figs. 9.1 and 9.2.

$K_m^{(0)}$ is readily calculated by considering the inner-shell clustering reactions

$$\mathrm{M}^{q+} + m\mathrm{H_2O} \rightleftharpoons \mathrm{M(H_2O)}_m{}^{q+},$$
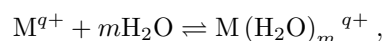
where M^{q+} is the metal ion, and utilizing the Gaussian [50] suite of programs, as an example. First the metal–water clusters were geometry optimized using the B3LYP hybrid density functional [51] and the 6-31+G(d,p) basis set. Frequency calculations confirmed a true minimum, and also yielded zero point corrections to the energy and vibrational contributions to the entropy change. The final electronic energies are calculated with the larger 6-311+G(2d,p) basis set. For the open-shell transition metals, the unrestricted formalism was used. The transition-metal ions and their water complexes were modeled in the high-spin state, since water typically leads to only a small splitting in the d-orbital energies. Further technical details can be found in [12].

The second necessary ingredient in the primitive quasichemical formulation is the excess chemical potential of the metal–water clusters and of water by itself. These quantities $\mu_{\alpha \mathrm{w}_m}^{\mathrm{ex}} - m\mu_{\mathrm{w}}^{\mathrm{ex}}$ can typically be obtained from widely available computational packages for molecular simulation [52]. In hydration problems where electrostatic interactions dominate, dielectric models of those hydration free energies are usually satisfactory. The combination $\mu_{\alpha \mathrm{w}_m}^{\mathrm{ex}} - m\mu_{\mathrm{w}}^{\mathrm{ex}}$ is typically insensitive to computational approximations because the water molecules coat the surface of the $\alpha \mathrm{w}_m$ complex, and computational errors can compensate between the bound and free ligands.

The final ingredient that enters the calculation is the density factor ρ_{w}. This is the actual density of water appropriate to the thermodynamic state intended in the calculation. For the usual case of 1 atm. pressure and 298 K, this is $1\,\mathrm{g\,cm^{-3}}$. The reference density in the electronic structure calculations is $\rho^\circ = 1\,\mathrm{atm}/RT$, however. Hence to account properly for the entropic cost of sequestering water in the metal–water complexes, the free energies should be adjusted by $-mRT \ln (\rho_{\mathrm{H_2O}}/\rho^\circ) = -mRT \ln (1354)$. With these inputs the excess chemical potential is readily composed as per (9.50), provided the optimal value of $\hat{m}$ is known. This is found by composing the excess chemical potential for different assumed m values and identifying the most stable case. For the dication transition metals studied, this is found to be six, consistent with experiment [12].

The above procedure readily yields Fig. 9.2. For estimating the excess chemical potential in the protein, again we need to know the ligation state of the metal ion. This is well known for the zinc-finger case. So we followed the above procedure, deciding what clusters to study quantum mechanically, and then composing the free energies as above.

For the metal ions that have been tried [12], this procedure works remarkably well in determining hydration free energies. There has been little [47] detailed attempt to determine hydration entropies and volumes in this way, and those properties are not

expected to be accurately represented in this model. For those properties it is probably necessary to consider seriously the description of packing problems in dense solutions. Though the necessary ideas seem to be in place [10], serious experience is nonexistent.

For anions in water – with the exception of HO^-(aq) [53–55] – this primitive quasichemical model is complicated for a different reason. In anion cases, determining $\hat{m}$ and the geometry for the clusters on the basis of gas-phase calculations can lead to problematic results. In contrast to the hydrated metal ions, in anion cases the hydrogen bonding of inner-shell water molecules to the outer-shell material is sometimes decisive in establishing the most probable inner-shell structure. A signature of this problem is effective H-bonding between inner-shell water molecules – *intra-cluster H-bonding* – in gas-phase calculations. This is not observed in the hydrated metal ion and the HO^-(aq) cases. The approach for treating those problematic anion cases is to treat the outer-shell effects on the basis of molecular-field developments as discussed below.

9.3.4 Molecular-Field Approximation $K_m \approx K_m^{(0)}[\varphi]$

Let us return to our basic formula (9.34). Notice that lots of the detail of ΔU_α expressed in the outer-shell factor (9.49) will have no effect because the indicator function $\prod_k [1 - b_\alpha(k)]$ multiplies by zero in the inner-shell region. This suggests that analysis of this contribution has a chance of producing results of general utility. The extreme case is the one in which ΔU_α is short-ranged to the extent that it is nonzero only in the inner-shell region. To analyze this, let us consider the case for which $\Delta U_\alpha = 0$ everywhere. In that case, (9.34) becomes

$$1 = \left(1 + \sum_{m \geq 1} \tilde{K}_m \rho_{\mathrm{w}}{}^m\right) \times \left\langle\!\!\left\langle \prod_k [1 - b_\alpha(k)] \right\rangle\!\!\right\rangle_0 . \qquad (9.51)$$

We adopt the tilde-notation $\tilde{K}_m$ as a reminder that these equilibrium ratios are the ones appropriate to the $\Delta U_\alpha = 0$ case considered. Notice then that the indicator function $\prod_k [1 - b_\alpha(k)]$ corresponds to the Boltzmann factor for the solute–solvent interactions of hard-core type, i.e., interactions that are infinitely unfavorable for molecular overlap, but otherwise zero. Thus [56]

$$\left\langle\!\!\left\langle \prod_k [1 - b_\alpha(k)] \right\rangle\!\!\right\rangle_0 = \mathrm{e}^{-\beta \mu_{\mathrm{HC}}^{\mathrm{ex}}} = \frac{1}{1 + \displaystyle\sum_{m \geq 1} \tilde{K}_m \rho_{\mathrm{w}}{}^m} . \qquad (9.52)$$

The identifier "HC" means that this is the solvation free energy for the case of a hard-core solute with the excluded-volume region established by the inner-shell definition.

In applying this approach to the equation of state of the hard-sphere fluid [57], it was found that the molecular-field approximation

$$\tilde{K}_m \approx \tilde{K}_m^{(0)}\left[\varphi\right] \; , \tag{9.53a}$$

$$\frac{\delta}{\delta \beta \varphi(\mathbf{r})} \ln \left[1 + \sum_{m \geq 1} \tilde{K}_m^{(0)}\left[\varphi\right] \rho_{\mathrm{w}}{}^m \right] = \rho_{\mathrm{w}} \; . \tag{9.53b}$$

was a suitable physical approximation, producing an equation of state for that system as accurate as the most accurate previous theory. The motivation for this approximation scheme is that the integrals $\tilde{K}_m^{(0)}\left[\varphi\right]$ are numerically accessible, the molecular field φ expresses the effects of the outer-shell material on each m-cluster, and the prescription for determination of that field – the second member of (9.53) – requires that the predicted density pattern implied within the defined inner shell be the known uniform density.

This approach is suggestive of the Hartree approximation of atomic and molecular physics. The outer-shell interactions are important, but complicated because of the correlations involved when they are considered directly. The suggested response to this difficulty is to treat these effects as uncorrelated – as a product contribution to the distribution – but with the product factors optimized by (9.53) to be consistent with the basic data.

The first point from this development and example is that, although the quasichemical approach is directed towards treating strong attractive – chemical – interactions at short range, it can describe traditional packing problems accurately. The second point is that this molecular-field idea permits us to go beyond the *primitive* quality noted above of the primitive quasichemical approximation, and specifically to account approximately for the influence of the outer-shell material on the equilibrium ratios K_m required by the general theory. This might help with cases of delicate structures noted above with anion hydrates.

Equation (9.53) for the desired molecular field is nonlinear, typically solved iteratively. For this molecular-field approach to become practical, an alternative to this nonlinear iterative calculation is required. A natural idea is that a useful approximation to this molecular field might be extracted from simulations with available generic force fields. Then with a satisfactory molecular field in hand, the more ambitious quasichemical evaluation of the free energy can be addressed, presumably treating the actual binding interactions with chemical methods specifically. This is work currently in progress.

How the latter application of QCT can be formulated has been discussed in some detail [10]. That discussion nearly closes a logical circle: PDT → QCT → *pdt*. The final *pdt* is, however, approximate, as is natural when utilizing a molecular-field description of the influence of the outer-shell material. Specifically,

$$\mu_\alpha^{\mathrm{ex}}\left(\mathscr{R}^n\right) \approx -k_{\mathrm{B}} T \ln \left\langle \mathrm{e}^{-\beta \Delta U_\alpha} \,\middle|\, \mathscr{R}^n \right\rangle_0 \left[GC : \varphi\right] \; . \tag{9.54}$$

Here the notation $[GC : \varphi]$ indicates that the system to be treated is *only* the inner-shell volume, and the material enclosed is described by an ensemble of fluctuating composition – as with the *grand canonical* ensemble – under the influence of the molecular-field φ. With longer-ranged interactions, a correction for those

interactions would also be required, much as a tail correction for traditional computer simulations [58, 59]. This emphasizes that these approximations are so firmly grounded that implementations are reincarnated as simulations, but with the system size drastically reduced.

9.4 Example

9.4.1 $\mu_\alpha^{\text{ex}} = k_{\text{B}}T \ln \int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha(\varepsilon) \, e^{\beta\varepsilon} d\varepsilon - k_{\text{B}}T \ln \int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha^{(0)}(\varepsilon) \, d\varepsilon$

Here we present and discuss an example calculation to make some of the concepts discussed above more definite. We treat a model for methane (CH_4) solute at infinite dilution in liquid under conventional conditions. This model would be of interest to conceptual issues of hydrophobic effects, and general hydration effects in molecular biosciences [1, 9], but the specific calculation here serves only as an illustration of these methods. An important element of this method is that nothing depends restrictively on the representation of the mechanical potential energy function. In contrast, the problem of methane dissolved in liquid water would typically be treated from the perspective of the van der Waals model of liquids, adopting a reference system characterized by the pairwise-additive repulsive forces between the methane and water molecules, and then correcting for methane–water molecule attractive interactions. In the present circumstance this should be satisfactory in fact. Nevertheless, the question frequently arises whether the attractive interactions substantially affect the statistical problems [60–62], and the present methods avoid such a limitation.

The example we consider is based upon the beautiful (9.17), which relates the probability distribution function for the interaction energy of the solute with the rest of the system in the fully coupled and the uncoupled cases, $\mathscr{P}_\alpha(\varepsilon)$ and $\mathscr{P}_\alpha^{(0)}(\varepsilon)$, respectively. This relation leads to the identity (9.19) which motivates an interesting view of the hydration free energy. The initial contribution on the right of (9.19) seems to be principally a coupling interaction energy between a CH_4 molecule and the liquid water matrix. The next term suggests a packing contribution even though we have not troubled to separate a repulsive force of interaction on which a packing theory might be based.

By seeking a satisfactory value of $\bar{\varepsilon}$ in a low range, (9.19) attempts to isolate a mean interaction potential energy as a contribution to the free energy that might be weakly dependent on the thermodynamic state. In the case of model interactions that accurately conform to a van der Waals equation of state [64], $\mathscr{P}_\alpha(\varepsilon)$ would be essentially δ-distributed. The distribution associated with the packing contribution is expected to be extremely broad, but only the cumulative probability is required, specifically without an exponential weight. Thus, evaluation of that packing contribution could be accomplished with insertion trials [2, 41] as are applicable to hard-core excluded-volume interactions.
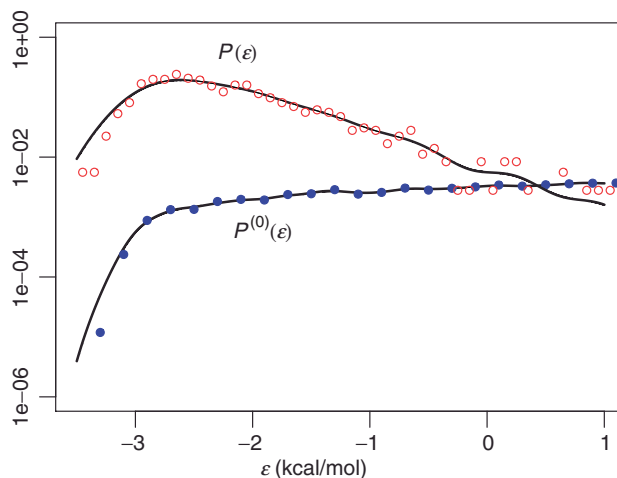
Fig. 9.4. $\mathscr{P}_\alpha(\varepsilon)$ and $\mathscr{P}_\alpha^{(0)}(\varepsilon)$ as a function of the binding energy. The simulations treated 216 water molecules, utilizing the SPC/E water model, and the Lennard-Jones parameters for methane were from [63]. The number density for both the systems is fixed at $0.03333\,\text{Å}^{-3}$, and $T = 298\,\text{K}$ established by velocity rescaling. These calculations used the NAMD program (www.ks.uiuc.edu/namd). After equilibration, the production run comprised 200 ps in the case of the pure water simulation and 500 ps in the case of the methane–water system. Configurations were saved every 0.5 ps for analysis

To exemplify this interesting formula, we performed two simulations, one of pure water and another of a single methane molecule in water, as shown in Fig. 9.4 which also shows the distribution functions $\mathscr{P}_\alpha(\varepsilon)$ and $\mathscr{P}_\alpha^{(0)}(\varepsilon)$ obtained from the simulation records. $\mathscr{P}_\alpha$ is the more tightly confined distribution. $\mathscr{P}_\alpha^{(0)}$, on the other hand, is a broader distribution. This suggests again how this latter contribution is supplying information on entropic effects attendant with the hydration process. Both data sets find about the same lower limit on the binding energy.

Figure 9.5 shows the standard histogram overlap analysis following (9.17), and helps identify an overlap region for $\bar{\varepsilon}$ of (9.19). The computed excess chemical potential is in excellent agreement with the earlier result of [63]. The inferred temperature agrees with the specified simulation temperature. Equation (9.17) is best applied in the region $-2.5\,\text{kcal mol}^{-1} < \bar{\varepsilon} < -0.5\,\text{kcal mol}^{-1}$.

The application of (9.19) is shown in Fig. 9.6. Note that the packing contribution, the second term on the right-hand side of (9.19), is always positive. Here it is of sufficient magnitude to make the net hydration free energy positive, consistent with the folklore of hydrophobic effects [9]. It is a remarkable point that truncation of the integral (9.18) over the range of observations of Fig. 9.4 would produce a result of the wrong sign, leading to qualitatively erroneous conclusions.

The upper, 'packing,' curve of Fig. 9.6 will decrease to zero for large $\bar{\varepsilon}$; the lower, 'chemistry,' curve will increase, cross the 'packing' curve, and approach the desired
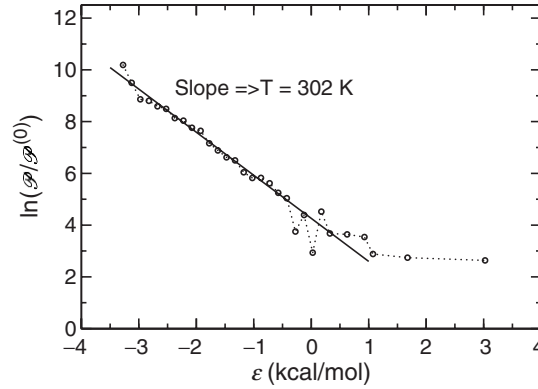
Fig. 9.5. Plot of $\ln[\mathscr{P}_\alpha(\varepsilon)/\mathscr{P}_\alpha^{(0)}(\varepsilon)]$. The region between $-2.5\,\text{kcal mol}^{-1}$ and $-0.5\,\text{kcal mol}^{-1}$ is satisfactorily linear, and (9.17) yields $T \approx 302$ K, in good agreement with the simulation temperature of 298 K. The intercept gives $\mu^{\text{ex}} = 2.5\,\text{kcal mol}^{-1}$ in agreement with a value of $2.5\,\text{kcal mol}^{-1}$ for the SPC water model [63]
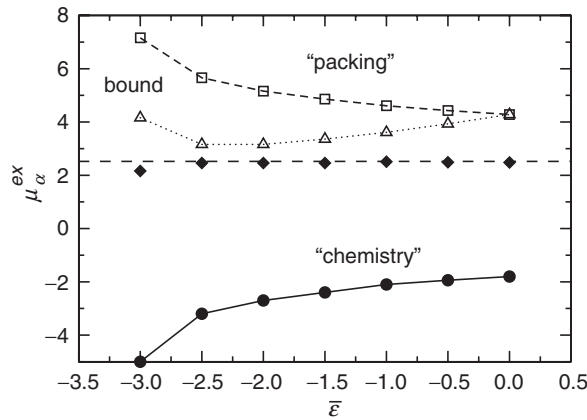

Fig. 9.6. Analysis of μ^{ex} following (9.19), p. 330. The *packing* curve is the rightmost term of (9.19), the *chemistry* curve is the preceding term, the diamonds are the sum as given by (9.19), and the *horizontal dashed line* is the hydration free energy of Fig. 9.5. μ^{ex} is insensitive to $\bar{\varepsilon}$. Here the packing contribution raises the hydration free energy to a substantial positive value consistent with the low solubility of CH_4 in water. The upper bound of (9.56) is given by the *open triangles*. The minimum value of this upper bound is attained near the thermal mean binding energy, about $-2.2\,\text{kcal mol}^{-1}$

μ_α^{ex} asymptotically from below. The desired weight $\int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha^{(0)}(\varepsilon)\,\text{d}\varepsilon$ in the low-energy wing conforms to the bound [65]

$$\int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha^{(0)}(\varepsilon)\,\text{d}\varepsilon \leq \text{e}^{\beta(\bar{\varepsilon} - \mu_\alpha^{\text{ex}})}, \tag{9.55}$$

which is analogous the Tchebycheff inequality [66]. (Note that this gives a trivial, but true, result when $\bar{\varepsilon} \geq \mu_\alpha^{\mathrm{ex}}$.) This relation yields an *upper* bound on the free energy

$$\mu_\alpha^{\mathrm{ex}} \leq \bar{\varepsilon} - k_{\mathrm{B}}T \ln \int_{-\infty}^{\bar{\varepsilon}} \mathscr{P}_\alpha^{(0)}\left(\varepsilon\right) \mathrm{d}\varepsilon \; . \tag{9.56}$$

Choosing $\bar{\varepsilon} = \langle \varepsilon \rangle$ gives the interesting result

$$\mu_\alpha^{\mathrm{ex}} \leq \langle \varepsilon \rangle - k_{\mathrm{B}}T \ln \int_{-\infty}^{\langle \varepsilon \rangle} \mathscr{P}_\alpha^{(0)}\left(\varepsilon\right) \mathrm{d}\varepsilon \; . \tag{9.57}$$

The evidence from Fig. 9.6 suggests that this can be accurate enough to be useful; that might be important in cases where the interactions are less simple than those assumed here. We reemphasize that these relations do not depend on particularly simple forms of the interactions. Of course

$$\langle \varepsilon \rangle \leq \mu_\alpha^{\mathrm{ex}} \; , \tag{9.58}$$

also. Equations (9.58) and (9.57) thus bracket the desired free energy.

9.4.2 Physical Discussion and Speculation on Hydrophobic Effects

The present example does support some speculation on an outstanding puzzle for our conceptualization of hydrophobic effects. It is known that the sum of the standard hydration entropies of $K^+(aq)$ and $Cl^-(aq)$ is about *twice* the standard hydration entropy of Ar(aq) [67]. The case of methanol as the solvent is qualitatively different. If hydrophobic effects are conceptualized on the basis of hydration entropies and specific hydration structures then this is paradoxical: according to the measured entropies $K^+(aq)$ + $Cl^-(aq)$ is about as hydrophobic as Ar(aq) + Ar(aq), but the hydration structures neighboring $K^+(aq)$, $Cl^-(aq)$, and Ar(aq) should be qualitatively different. This paradox has not been given a satisfactory explanation.

We can initiate this discussion with the following straw-man argument: viewed from the perspective of (9.27), p. 333, evidently the contribution of the leading – zeroth-order – term is most of the net entropy of hydration; and the entropy contributions of the succeeding terms must be comparatively small. This is not an explanation because it does not explain why those succeeding contributions make small contributions to the entropy despite the fact that the interactions involved with those terms have large effects on the hydration structure. Therefore, this straw-man argument serves only to sharpen the paradox. We pass over the question of why a perturbative treatment should be satisfactory when those interactions have a large effect on the hydration structure.

The interactions that have to be considered in the three cases here are also different from each other. Nevertheless, some features of the graphs of Figs. 9.4 and 9.6 are likely to be generally observed. $\mathscr{P}_\alpha^{(0)}(\varepsilon)$ will populate higher energies than $\mathscr{P}(\varepsilon)$, and $\mathscr{P}_\alpha^{(0)}(\varepsilon)$ will be extremely broad relative to $\mathscr{P}(\varepsilon)$. $\mathscr{P}(\varepsilon)$ will be much more

concentrated, but – because the interactions are so different in the cases considered – with significantly different locations. We speculate that for the cases considered here a satisfactory low value of $\bar{\varepsilon}$ can be identified so that the 'chemistry' contribution might be less sensitive to temperature T because of the exceptionally slow decrease of the density of liquid water with increasing T along the vapor saturation curve [7]; then we further speculate that the T dependence of the 'packing' contribution at that value of $\bar{\varepsilon}$ might be somewhat generic. In fact, the 'packing' contribution only assesses a fraction of favorable cases, and so might be insensitive to details of the different interactions. This speculation then amounts to the suggestion that:

$$\mu_\alpha^{\mathrm{ex}} \approx \langle \Delta U_\alpha \rangle - k_{\mathrm{B}} T \ln \int_{-\infty}^{\mu_\alpha^{\mathrm{ex}}} \mathscr{P}_\alpha^{(0)} (\varepsilon) \, \mathrm{d}\varepsilon \; , \qquad (9.59)$$

with the interesting temperature sensitivity in the right-most term. It is interesting to note also that $\langle \Delta U_\alpha \rangle$ probably will exhibit parabolic dependence on the ionic charge [68], as does the approximation (9.27). Nevertheless, refinement of this speculation will require detailed studies for the cases of interest here.

9.5 Conclusions

Return now to the assertions of the Introduction. The explanation of assertion (1) was pointed out previously. Assertion (2), that the PDT is a practical computational tool, was the subject of Sect. 9.2.3. See especially the discussion *"the general computational tricks work for the PDT also,"* emphasizing the general statistical methods of stratification and importance weighting, and their correspondence to the natural theoretical analyses of the PDT partition function.

The final assertion (3) is that the PDT is a natural vehicle for the organization and comprehension of data, either experimental or simulation, i.e., for the development of physically motivated models. This is an evolving conclusion to be supported by scientific induction. The discussion above supports this view in several places, including dielectric models of hydration involving electrostatic interactions (9.27), p. 333, chemical problems (9.50), and packing problems (9.52). Reference [10] discusses the further examples of the van der Waals model of liquids and liquid mixtures, the Debye–Hückel model of electrolyte solutions, the Flory–Huggins model of polymer solutions, and the information theory approaches that underlie recent progress in understanding classic hydrophobic phenomena [7, 9].

The reason that the PDT is an effective tool for the generation of physical models is that it treats an intensive thermodynamic property, and the distribution functions involved are simpler in the thermodynamic limit than if this were not the case [10]. An extended family of modeling tools then applies directly. The quasichemical approach is a 'general example.' It amounts to stratification of the statistical problem

according to the number of neighbors of a solution molecule of interest. The cluster-variation exercise sketched in Sect. 9.3.1 gives a new derivation of the principal features of that quasichemical approach.

References

1. Pratt, L. R., Molecular theory of hydrophobic effects: "She is too mean to have her name repeated.", *Annu. Rev. Phys. Chem.* **2002**, *53*, 409–436
2. Pohorille, A.; Pratt, L. R., Cavities in molecular liquids and the theory of hydrophobic solubilities, *J. Am. Chem. Annu.* **1990**, *112*, 5066–5074
3. Pratt, L. R.; Pohorille, A., Theory of hydrophobicity: transient cavities in molecular liquids, *Proc. Natl Acad. Sci. USA* **1992**, *89*, 2995–2999
4. Hummer, G.; Garde, S.; García, A. E.; Pohorille, A.; Pratt, L. R., An information theory model of hydrophobic interactions, *Proc. Natl Acad. Sci. USA* **1996**, *93*, 8951–8955
5. Gomez, M. A.; Pratt, L. R.; Hummer, G.; Garde, S., Molecular realism in default models for information theories of hydrophobic effects, *J. Phys. Chem. B* **1999**, *103*, 3520–3523
6. Garde, S.; Hummer, G.; García, A. E.; Paulaitis, M. E.; Pratt, L. R., Origin of entropy convergence in hydrophobic hydration and protein folding, *Phys. Rev. Lett.* **1996**, *77*, 4966–4968
7. Pratt, L. R.; Pohorille, A., Hydrophobic effects and modeling of biophysical aqueous solution interfaces, *Chem. Rev.* **2002**, *102*, 2671–2691
8. Ashbaugh, H. S.; Asthagiri, D.; Pratt, L. R.; Rempe, S. B., Hydration of krypton and consideration of clathrate models of hydrophobic effects from the perspective of quasi-chemical theory, *Biophys. Chem.* **2003**, *105*, 323–338
9. Ashbaugh, H. S.; Pratt, L. R., Colloquium: Scaled particle theory and the length scales of hydrophobicity, *Rev. Mod. Phys.* **2006**, *78*, 159–178
10. Beck, T. L.; Paulaitis, M. E.; Pratt, L. R., *The Potential Distribution Theorem and Models of Molecular Solutions*, Cambridge University Press: Cambridge, 2006
11. Hammersley, J. M.; Handscomb, D. C., *Monte Carlo Methods*, Chapman and Hall: London, 1964
12. Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B., Hydration structure and free energy of biomolecularly specific aqueous dications, including Zn^{2+} and first transition row metals, *J. Am. Chem. Soc.* **2004**, *126*, 1285–1289
13. Kirkwood, J. G.; Poirier, J. C., The statistical mechanical basis of the Debye–Hückel theory of strong electrolytes, *J. Phys. Chem.* **1954**, *86*, 591–596
14. Widom, B., Some topics in the theory of fluids, *J. Chem. Phys.* **1963**, *39*, 2808–2812
15. Jackson, J. L.; Klein, L. S., Potential distribution method in equilibrium statistical mechanics, *Phys. Fluids* **1964**, *7*, 228–231
16. Widom, B., Potential-distribution theory and the statistical mechanics of fluids, *J. Phys. Chem.* **1982**, *86*, 869–872
17. Stell, G. Mayer-Montroll equations (and some variants) through history for fun and profit. in *The Wonderful World of Stochastics A Tribute to Elliot W. Montroll*, Shlesinger, M. F.; Weiss, G. H., Eds., vol. XII, *Studies in Statistical Mechanics*. Elsevier: New York, **1985**, pp. 127–156
18. Lebowitz, J. L.; Percus, J. K.; Verlet, L., Ensemble dependence of fluctuations with application to machine computations, *Phys. Rev.* **1967**, *153*, 250
19. Resnick, S. I., *A Probability Path*, Birkhäuser: New York, 2001

20. Imai, T.; Hirata, F., Partial molar volume and compressibility of a molecule with internal degrees of freedom., *J. Chem. Phys.* **2003**, *119*

21. Bennett, C. H., Efficient estimation of free-energy differences from Monte Carlo data, *J. Comp. Phys.* **1976**, *22*, 245–268

22. Ciccotti, G.; Frenkel, D.; McDonald, I. R., *Simulation of Liquids and Solids. Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics*, North-Holland: Amsterdam, 1987

23. Frenkel, D. Free-energy computation and first-order phase transitions. in *International School of Physics 'Enrico Fermi'*, vol. XCVII. Soc. Italiana di Fisica: Bologna, **1986**, pp. 151–188

24. Shing, K. S.; Chung, S. T., Computer-simulation methods for the calculation of solubility in supercritical extraction systems, *J. Phys. Chem.* **1987**, *91*, 1674–1681

25. Smith, P. E., Computer simulation of cosolvent effects on hydrophobic hydration, *J. Phys. Chem. B* **1999**, *103*, 525–534

26. Callen, H. B., *Thermodynamics*, [2nd edition]. See Chapter 5

27. Wood, R. H.; Yezdimer, E. M.; Sakane, S.; Barriocanal, J. A.; Doren, D. J., Free energies of solvation with quantum mechanical interaction energies from classical mechanical simulations, *J. Chem. Phys.* **1999**, *110*, 1329–37

28. Sakane, S.; Yezdimer, E. M.; Liu, W. B.; Barriocanal, J. A.; Doren, D. J.; Wood, R. H., Exploring the ab initio/classical free energy perturbation method: the hydration free energy of water, *J. Chem. Phys.* **2000**, *113*, 2583–93

29. Wood, R. H.; Liu, W. B.; Doren, D. J., Rapid calculation of the structures of solutions with ab initio interaction potentials., *J. Phys. Chem. A* **2002**, *106*, 6689–6693

30. Liu, W. B.; Sakane, S.; Wood, R. H.; Doren, D. J., The hydration free energy of aqueous Na^+ and Cl^- at high temperatures predicted by ab initio/classical free energy perturbation: 973 K with 0.535 g/cm^3 and 573 K with 0.725 g/cm^3, *J. Phys. Chem. A* **2002**, *106*, 1409–1418

31. Sakane, S.; Liu, W. B.; Doren, D. J.; Shock, E. L.; Wood, R. H., Prediction of the Gibbs energies and an improved equation of state for water at extreme conditions from ab initio energies with classical simulations, *Geochim. Cosmochim. Acta* **2001**, *65*, 4067–4075

32. Liu, W. B.; Wood, R. H.; Doren, D. J., Hydration free energy and potential of mean force for a model of the sodium chloride ion pair in supercritical water with ab initio solute–solvent interactions, *J. Chem. Phys.* **2003**, *118*, 2837–2844

33. Liu, W. B.; Wood, R. H.; Doren, D. J., Density and temperature dependences of hydration free energy of Na^+ and Cl^- at supercritical conditions predicted by a initio/classical free energy perturbation, *J. Phys. Chem. B* **2003**, *107*, 9505–9513

34. Pettitt, B. M., A Perspective on "Volume and heat of hydration of ions" - Born M. (1920) Z Phys. 1 : 45, *Theor. Chem. Acc.* **2000**, *103*, 171–172

35. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199

36. Hertz, J.; Krogh, A.; Palmer, R. G., *Introduction to the Theory of Neural Computation*, Addison-Wesley: Redwood City, CA, 1991

37. Plishke, M.; Bergerson, B., *Equilibrium Statistical Physics*, World Scientific: Singapore, 1994

38. Kalos, M. H.; Whitlock, P. A., *Monte Carlo Methods, Volume I: Basics*, Wiley–Interscience: New York, 1986

39. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical Recipes in Fortran 77*, [2nd edition]

40. Hummer, G.; Pratt, L. R.; García, A. E., Multistate Gaussian model for electrostatic solvation free energies, *J. Am. Chem. Soc.* **1997**, *119*, 8523–8527

41. Pratt, L. R.; Pohorille, A. in *Proceedings of the EBSA 1992 International Workshop on Water–Biomolecule Interactions*, Palma, M. U.; Palma-Vittorelli, M. B.; Parak, F., Eds. Societá Italiana de Fisica: Bologna, **1993**, pp. 261–268

42. Pohorille, A.; Wilson, M. A., Molecular structure of aqueous interfaces., *Theochem* **1993**, *284*, 271–98

43. Pohorille, A., Transient cavities in liquids and the nature of the hydrophobic effect, *Pol. J. Chem.* **1998**, *72*, 1680–1690

44. Pratt, L. R., *Hydrophobic effects* Wiley: Chichester, **1998**, pp. 1286–1294

45. Pohorille, A.; Wilson, M. A., Excess chemical potential of small solutes across water-membrane and water–hexane interfaces, *J. Chem. Phys.* **1996**, *104*, 3760–3773

46. Pratt, L. R.; LaViolette, R. A., Quasi-chemical theories of associated liquids, *Mol. Phys.* **1998**, *94*, 909–915

47. Pratt, L. R.; Rempe, S. B. Quasi-chemical theory and implicit solvent models for simulations. in *Simulation and Theory of Electrostatic Interactions in Solution. Computational Chemistry, Biophysics, and Aqueous Solutions*, Pratt, L. R.; Hummer, G., Eds., vol. 492, *AIP Conference Proceedings*. American Institute of Physics, Melville: New York, **1999**, pp. 172–201

48. Paulaitis, M. E.; Pratt, L. R., Hydration theory for molecular biophysics, *Adv. Prot. Chem.* **2002**, *62*, 283–310

49. Mathews, J.; Walker, R. L., *Mathematical Methods of Physics*, Benjamin: New York, 1964

50. Frisch, M. J. *et al.* Gaussian 98 (Revision A.2), 1998, Gaussian, Inc.: Pittsburgh PA

51. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.* **1993**, *98*, 5648

52. Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S., Absolute hydration free energies of ions, ion–water clusters, and quasi-chemical theory, *J. Chem. Phys.* **2003**, *119*, 2702–2708

53. Asthagiri, D.; Pratt, L. R.; Kress, J. D.; Gomez, M. A., The hydration state of HO^-(aq), *Chem. Phys. Lett.* **2003**, *380*, 530–535

54. Asthagiri, D.; Pratt, L. R.; Kress, J. D.; Gomez, M. A., HO^-(aq) hydration and mobility, *Proc. Natl Acad. Sci. USA* **2004**, *101*, 7229–7233

55. Asthagiri, D.; Pratt, L. R.; Kress, J. D., Ab initio molecular dynamics and quasichemical study of H^+(aq)., *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6704–6708

56. Pratt, L. R.; LaViolette, R. A.; Gomez, M. A.; Gentile, M. E., Quasi-chemical theory for the statistical thermodynamics of the hard-sphere fluid, *J. Phys. Chem. B* **2001**, *105*, 11662–11668

57. Pratt, L. R.; Ashbaugh, H. S., Self-consistent molecular field theory for packing in classical liquids, *Phys. Rev. E* **2003**, *68*, 021505

58. Allen, M .P.; Tildesley, D. J., *Computer Simulation of Liquids*, Oxford Science: Oxford, 1987

59. Frenkel, D.; Smit, B., *Understanding Molecular Simulation. From Algorithms to Applications*, [2nd edition]

60. Gallicchio, E.; Kubo, M. M.; Levy, R. M., Enthalpy–entropy and cavity decomposition of alkane hydration free energies: numerical results and implications for theories of hydrophobic solvation, *J. Phys. Chem. B* **2000**, *104*, 6271–6285

61. Grossman, J. C.; Schwegler, E.; Galli, G., Quantum and classical molecular dynamics simulations of hydrophobic hydration structure around small solutes, *J. Phys. Chem. B* **2004**, *108*, 15865–15872

62. Raschke, T. M.; Levitt, M., Detailed hydration maps of benzene and cyclohexane reveal distinct water structures, *J. Phys. Chem. B* **2004**, *108*, 13492–13500

63. Hummer, G.; Pratt, L.R.; García, A.E., Free energy of ionic hydration, *J. Phys. Chem.* **1996**, *100*, 1206–1215
64. Lebowitz, J. L.; Waisman, E. M., Statistical-mechanics of simple fluids: beyond van der Waals, *Phys. Today* **1980**, *33*, 24–30
65. Jarzynski, C., Microscopic analysis of Clausius–Duhem processes, *J. Stat. Phys.* **1999**, *96*, 415–427
66. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill: Singapore, 1991
67. Friedman, H. L.; Krishnan, C. V. Thermodynamics of ion hydration. in *Water A Comprehensive Treatise*, Franks, F., Ed., vol. 3. Plenum: New York, **1973**, pp. 1–118
68. Stell, G. Fluids with long-range forces. in *Statistical Mechanics. Part A: Equilibrium Techniques*, Berne, B. J., Ed. Plenum: New York, **1977**, pp. 47–84

10

Methods for Examining Phase Equilibria

M. Scott Shell and Athanassios Z. Panagiotopoulos

10.1 Introduction

This chapter focuses on methods to obtain free energies and phase equilibria of classical, i.e., nonquantum fluids. One distinguishing issue with these systems relative to lattice models of fluids is of course the existence of continuous degrees of freedom and their associated continuous macroscopic parameters. This expands configuration space relative to the lattice systems, but also allows greater flexibility in propagating the system, including the possible application of hybrid Monte Carlo/molecular dynamics methods. Nonetheless, there are a number of subtleties associated with both the formalism and implementation of off-lattice phase-equilibria calculations, and it is these issues which form the focus of this chapter. In particular, we have endeavored to review a collection of algorithms which have an established record of effectiveness with fluids; these methods do not always correspond to the most efficient approaches one would employ for systems reviewed elsewhere in this book.

The main emphasis of the chapter is on Monte Carlo (MC) rather than molecular dynamics (MD) methods. While some of the algorithms described here have analogs using MD, MC is a more natural and flexible framework for performing the calculations of free energies and phase equilibria presented. MC methods easily incorporate 'unphysical' moves, such as particle insertions and deletions, or cutting and regrowing segments of molecules. Furthermore, while MD can be formulated in various ensembles including those of non-natural state probabilities, as in the multicanonical approach, both the conceptual framework and implementation in these situations is often significantly more straightforward in an MC setting. We therefore feel that, in an introductory text such as this, the reader is best served if we focus on MC algorithms and provide references to related MD methods where appropriate.

In any MC simulation, three ingredients form the basis of the way in which properties are extracted from the model system: the prescribed microstate probabilities: the ensemble of interest, the set of random moves which propagate the system according to these probabilities, and the *estimators* which extract the appropriate property averages. In this discussion, we will not be concerned with the intermediate

of these; we will largely ignore the issue of good move design. This is not to trivialize that aspect, as effective move schemes can often mean the difference between feasible and impractically long simulations, but it is because the free energy calculations themselves are more suitably addressed by the other two elements. We therefore assume that the reader is familiar with conventional moves such as single-particle displacement, volume scaling, and particle addition/deletion.

We remind the reader that a particular ensemble and a collection of moves completely specifies the acceptance criteria, as a consequence of detailed balance. In general mathematical form, these two components are described by a microstate probability scheme, the probabilities associated with each configuration in the ensemble, denoted here by $\wp(\mathbf{q})$ where $\mathbf{q}$ are the positions of the N particles, and by and the transition proposal probabilities, $t_{\text{prop}}(\mathbf{q}_A \to \mathbf{q}_B)$, which give the probability that a particular move will be attempted between two configurations A and B. By applying the detailed balance condition, one ensures that the limiting distribution of the simulation is $\wp(\mathbf{q})$. This condition provides an expression for the probabilities with which proposed random moves from A to B must be accepted, $\mathscr{P}_{\text{acc}}(\mathbf{q}_A \to \mathbf{q}_B)$. The general relationship is

$$\frac{\mathscr{P}_{\text{acc}}(\mathbf{q}_A \to \mathbf{q}_B)}{\mathscr{P}_{\text{acc}}(\mathbf{q}_B \to \mathbf{q}_A)} = \frac{t_{\text{prop}}(\mathbf{q}_B \to \mathbf{q}_A)}{t_{\text{prop}}(\mathbf{q}_A \to \mathbf{q}_B)} \frac{\wp(\mathbf{q}_B)}{\wp(\mathbf{q}_A)}. \tag{10.1}$$

This expression is most frequently satisfied using the Metropolis criterion

$$\mathscr{P}_{\text{acc}}(\mathbf{q}_A \to \mathbf{q}_B) = \min\left\{1, \frac{t_{\text{prop}}(\mathbf{q}_B \to \mathbf{q}_A)}{t_{\text{prop}}(\mathbf{q}_A \to \mathbf{q}_B)} \frac{\wp(\mathbf{q}_B)}{\wp(\mathbf{q}_A)}\right\}. \tag{10.2}$$

A more detailed discussion of the subtleties in formulating correct acceptance criteria can be found in [1]. For the purpose of this chapter, we will focus on ensembles in general rather than acceptance criteria specifically, with the understanding that once the configurational probabilities are fixed the criteria follow directly. With this in mind, we will sometimes present the microstate probability scheme without discussing the associated acceptance criteria.

A number of textbooks and review articles are available which provide background and more-general simulation techniques for fluids, beyond the calculations of the present chapter. In particular, the book by Frenkel and Smit [1] has comprehensive coverage of molecular simulation methods for fluids, with some emphasis on algorithms for phase-equilibrium calculations. General review articles on simulation methods and their applications – e.g., [2–6] – are also available. Sections 10.2 and 10.3 of the present chapter were adapted from [6]. The present chapter also reviews examples of the recently developed flat-histogram approaches described in Chap. 3 when applied to phase equilibria.

This chapter presents a collection of algorithms varied in application and generality, and is designed to provide a diverse toolkit for dealing with free energies and phase equilibria of fluids. We discuss both methods which aim to calculate the free energy or entropy directly, as well as those which characterize these functions indirectly, as in Gibbs ensemble calculations. Our discussion begins with more-specific

methods, such as the test particle method for the chemical potential, proceeds to more-general phase-equilibrium techniques, and finally ends with the widely applicable flat-histogram approaches. Our goal is to illustrate the general principles of these calculations, rather than be an exhaustive reference. Indeed, it is our belief that much remains to be explored with many of the methods, and thus a 'big-picture' formalism better serves the reader in applying them to novel problems.

10.2 Calculating the Chemical Potential

10.2.1 Widom Test Particle Method

Most free energy and phase-equilibrium calculations by simulation up to the late 1980s were performed with the Widom 'test particle' method [7]. The method is still appealing in its simplicity and generality – for example, it can be applied directly to MD calculations without disturbing the time evolution of a system. The potential distribution theorem on which the test particle method is based as well as its applications are discussed in Chap. 9.

The method is based on the following expression for constant-*NVT* simulations (modified expressions are available in other ensembles [1])

$$\exp(-\beta\mu) = \langle\exp(-\beta U_{\text{test}})\rangle/\rho, \tag{10.3}$$

where μ is the chemical potential of a component in a system (to within a temperature-dependent constant that does not affect phase-equilibrium calculations), U_{test} is the energy experienced by a 'test' particle of that component placed in a random position in the simulation cell, and ρ is the molar density. Most attempted insertions of test particles result in overlap with existing particles in the fluid, with U_{test} large and positive. These insertions do not contribute significantly to the ensemble average. For simple single-site intermolecular potentials, such as the Lennard-Jones potential, sampling using the Widom test particle method can be performed throughout the fluid state. Sampling fails for ordered solid phases. For multisegment molecules, the Widom method can be combined with configurational bias techniques [1]. Even if a molecule is not multisegmented but is not spherically symmetric, one must perform averaging over orientations and possibly conformation. This is discussed in Chap. 9.

10.2.2 *NPT* + Test Particle Method

The NPT + test particle method [8, 9] aims to determine phase coexistence points based on calculations of the chemical potentials for a number of state points. A phase coexistence point is determined at the intersection of the vapor and liquid branches of the chemical potential versus pressure diagram. The Widom test particle method [7] of the previous paragraph or any other suitable method [10] can be used to obtain the chemical potentials. Corrections to the chemical potential of the liquid and vapor phases can be made, using standard thermodynamic relationships, for deviations

between the pressure at which the calculation were made and the actual coexistence pressure. Extrapolations with respect to temperature are also possible [11].

In contrast to the Gibbs ensemble discussed later in this chapter, a number of simulations are required per coexistence point, but the number can be quite small, especially for vapor–liquid equilibrium calculations away from the critical point. For example, for a one-component system near the triple point, the density of the dense liquid can be obtained from a single NPT simulation at zero pressure. The chemical potential of the liquid, in turn, determines the density of the (near-ideal) vapor phase so that only one simulation is required. The method has been extended to mixtures [12, 13]. Significantly lower statistical uncertainties were obtained in [13] compared to earlier Gibbs ensemble calculations of the same Lennard-Jones binary mixtures, but the NPT + test particle method calculations were based on longer simulations.

The NPT + test particle method shares many characteristics with the histogram reweighting methods discussed in Chap. 3. In particular, histogram reweighting methods also yield the chemical potentials and pressures of the coexisting phase from a series of simulations. The corrections to the chemical potentials for changes in pressure [9] and temperature [11] are similar to the concept of reweighting the combined histograms from grand canonical simulations to new densities and temperatures.

Spyriouni et al. [14, 15] have presented a powerful method (called 'SPECS') for calculations of polymer phase behavior related to the NPT + test particle method. The method of Spyriouni et al. targets the calculation of the phase behavior of long-chain systems for which the test particle method for the calculation of chemical potentials fails. For sufficiently long chains, even configurational bias sampling methods become impractical. For binary mixtures of a low-molecular-weight solvent (species 1) and a polymer (species 2), two parallel simulations are performed in the (μ_1, N_2, P, T) ensemble at conditions near the expected coexistence curve. The chemical potential of component 2 is determined through the 'chain increment' technique [16]. Iterative calculations at corrected values of the chemical potential of the solvent are performed until the chemical potential of the polymer in the two phases is equal. For the special case of a dilute solutions, estimates of the chemical potentials of the solvent and polymer for compositions different from the original simulation conditions can be made using standard thermodynamic relations and the number of required iterations is significantly reduced.

10.3 Ensemble-Based Free Energies and Equilibria

10.3.1 Gibbs Ensemble

The Gibbs Ensemble MC simulation methodology [17–19] enables direct simulations of phase equilibria in fluids. A schematic diagram of the technique is shown in Fig. 10.1. Let us consider a macroscopic system with two phases coexisting at equilibrium. Gibbs ensemble simulations are performed in two separate microscopic regions, each within periodic boundary conditions (denoted by the dashed lines in Fig. 10.1). The thermodynamic requirements for phase coexistence are that each
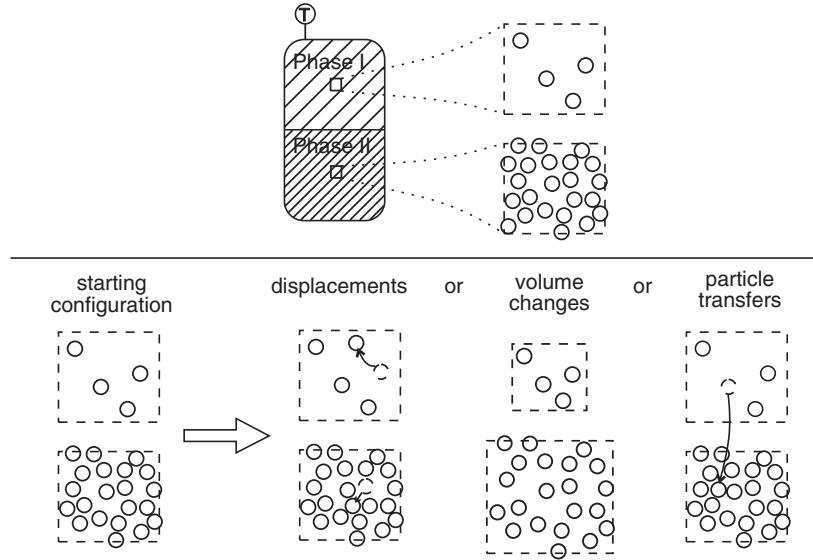
Fig. 10.1. Schematic diagram of the Gibbs ensemble MC simulation methodology. Reprinted with permission from [6]. ©2000 IOP Publishing Ltd

region should be in internal equilibrium, and that temperature, pressure and the chemical potentials of all components should be the same in the two regions. System temperature in MC simulations is specified in advance. The remaining three conditions are satisfied by performing three types of MC moves, displacements of particles within each region (to satisfy internal equilibrium), fluctuations in the volume of the two regions (to satisfy equality of pressures) and transfers of particles between regions (to satisfy equality of chemical potentials of all components).

The acceptance criteria for the Gibbs ensemble were originally derived from fluctuation theory [17]. An approximation was implicitly made in the derivation that resulted in a difference in the acceptance criterion for particle transfers proportional to $1/N$ relative to the exact expressions given subsequently [18]. A full development of the statistical mechanics of the ensemble was given by Smit et al. [19] and Smit and Frenkel [20], which we follow here. A one-component system at constant temperature T, total volume V, and total number of particles N is divided into two regions, with volumes V_{I} and $V_{\mathrm{II}} = V - V_{\mathrm{I}}$, and number of particles N_{I} and $N_{\mathrm{II}} = N - N_{\mathrm{I}}$. The partition function, Q_{NVT} is

$$Q_{NVT} = \frac{1}{\Lambda^{3N} N!} \sum_{N_{\mathrm{I}}=0}^{N} \binom{N}{N_{\mathrm{I}}} \int_0^V \mathrm{d}V_{\mathrm{I}} V_{\mathrm{I}}^{N_{\mathrm{I}}} V_{\mathrm{II}}^{N_{\mathrm{II}}} \int \mathrm{d}s_{\mathrm{I}}^{N_{\mathrm{I}}} \exp[-\beta U_{\mathrm{I}}(N_{\mathrm{I}})]$$

$$\times \int \mathrm{d}s_{\mathrm{II}}^{N_{\mathrm{II}}} \exp[-\beta U_{\mathrm{II}}(N_{\mathrm{II}})], \tag{10.4}$$

where s_I and s_{II} are the scaled coordinates of the particles in the two regions and $U(N)$ is the total intermolecular potential of interaction of N particles. Equation (10.4) represents an ensemble with probability density, $\wp(N_I, V_I; N, V, T)$ proportional to

$$\wp(N_I, V_I; N, V, T) \propto \frac{N!}{N_I! N_{II}!} \exp\left[N_I \ln V_I + N_{II} \ln V_{II}\right. \\ \left. -\beta U_I(N_I) - \beta \, U_{II}(N_{II})\right]. \tag{10.5}$$

Smit et al. [19] used the partition function given by (10.4) and a free energy minimization procedure to show that, for a system with a first-order phase transition, the two regions in a Gibbs ensemble simulation are expected to reach the correct equilibrium densities.

The acceptance criteria for the three types of moves can be obtained immediately from (10.4). For a displacement step internal to one of the regions, the probability of acceptance is the same as for conventional constant$-NVT$ simulations

$$\wp_{move} = \min[1, \exp(-\beta \Delta U)]. \tag{10.6}$$

where ΔU is the configurational energy change resulting from the displacement. For a volume change step during which the volume of region I is increased by ΔV with a corresponding decrease of the volume of region II,

$$\wp_{volume} = \min\left[1, \exp\left(-\beta \Delta U_I - \beta \Delta U_{II} + N_I \ln V_I + \frac{\Delta V}{V_I}\right.\right. \\ \left.\left. + N_{II} \ln V_{II} - \frac{\Delta V}{V_{II}}\right)\right]. \tag{10.7}$$

Equation (10.7) implies that sampling is performed uniformly in the volume itself. The acceptance criterion for particle transfers, written here for transfer from region II to region I is

$$\wp_{transfer} = \min\left[1, \frac{N_{II} V_I}{(N_I + 1) V_{II}} \exp\left(-\beta \Delta U_I - \beta \Delta U_{II}\right)\right]. \tag{10.8}$$

Equation (10.8) can be readily generalized to multicomponent systems. The only difference is that the number of particles of species j in each of the two regions, $N_{I,j}$ and $N_{II,j}$ replace N_I and N_{II}, respectively. In simulations of multicomponent systems dilute in one component, it is possible that the number of particles of a species in one of the two regions becomes zero after a successful transfer out of that region. Equation (10.8) in this case is taken to imply that the probability of transfer out of an empty region is zero.

To this point, the acceptance rules have been defined for a simulation, in which the total number of molecules in the system, temperature and volume are constant. For pure component systems, the phase rule requires that only one intensive variable (in this case the system temperature) can be independently specified when two phases

coexist. The vapor pressure is obtained from the simulation. In contrast, for multi-component systems pressure can be specified in advance, with the total system being considered at constant *NPT*. The probability density for this case, $\wp(N_\mathrm{I}, V_\mathrm{I}; N, P, T)$ is proportional to

$$\wp(N_\mathrm{I}, V_\mathrm{I}; N, P, T) \propto \frac{N!}{N_\mathrm{I}! N_\mathrm{II}!} \exp \Big[N_\mathrm{I} \ln V_\mathrm{I} + N_\mathrm{II} \ln V_\mathrm{II}$$
$$-\beta U_\mathrm{I}(N_\mathrm{I}) - \beta U_\mathrm{II}(N_\mathrm{II}) - \beta P(V_\mathrm{I} + V_\mathrm{II}) \Big]. \quad (10.9)$$

and the only change necessary in the algorithm is that the volume changes in the two regions are now made independently. The acceptance criterion for a volume change step in which the volume of region I is changed by ΔV, while the other region remains unchanged is then:

$$\wp_\mathrm{volume} = \min \left[1, \exp \left(-\beta \Delta U_\mathrm{I} + N_\mathrm{I} \ln V_\mathrm{I} + \frac{\Delta V}{V_\mathrm{I}} - \beta P \Delta V \right) \right]. \quad (10.10)$$

An interesting extension of the original methodology was proposed by Lopes and Tildesley to allow the study of more than two phases at equilibrium [21]. The extension is based on setting up a simulation with as many boxes as the maximum number of phases expected to be present. Kristof and Liszi [22, 23] have proposed an implementation of the Gibbs ensemble in which the total enthalpy, pressure and number of particles in the total system are kept constant. Molecular dynamics versions of the Gibbs ensemble algorithm are also available [24–26].

The physical reason for the ability of the Gibbs ensemble to converge to a state that contains phases at their equilibrium density in the corresponding boxes, rather than a mixture of the two phases in each box, is the free energy cost of creating and maintaining an interface. Near critical points, Gibbs ensemble simulations become unstable because the free energy penalty for creating an interface becomes small. A better approach to dealing with systems near critical points is provided by the histogram reweighting methods described in Chap. 3. The finite-size critical behavior of the Gibbs ensemble has been examined by Bruce [27], Mon and Binder [28], and Panagiotopoulos [29]. The 'standard' procedure for obtaining critical points from Gibbs ensemble simulations is to fit subcritical coexistence data to universal scaling laws. This approach has a weak theoretical foundation, since the universal scaling laws are only guaranteed to be valid in the immediate vicinity of the critical point, where simulations give the wrong (classical) behavior due to the truncation of the correlation length at the edge of the simulation box. In many cases, however, the resulting critical points are in fair agreement with more-accurate results obtained from finite-size scaling methods.

In summary, the Gibbs ensemble MC methodology provides a direct and efficient route to the phase coexistence properties of fluids, for calculations of moderate accuracy. The method has become a standard tool for the simulation community, as evidenced by the large number of applications using the method. Histogram reweighting techniques (Chap. 3) have the potential for higher accuracy, especially if

equilibria for a large number of state points are to be determined. Histogram methods are also inherently better at determining critical points. In its original form, the Gibbs ensemble method is not practical for multisegment or strongly interacting systems, but development of configurational bias sampling methods [30, 31] has overcome this limitation.

10.3.2 Gibbs–Duhem Integration

Most methods for the determination of phase equilibria by simulation rely on particle insertions to equilibrate or determine the chemical potentials of the components. Methods that rely on insertions experience severe difficulties for dense or highly structured phases. If a point on the coexistence curve is known (e.g., from Gibbs ensemble simulations), the remarkable method of Kofke [32, 33] enables the calculation of a complete phase diagram from a series of constant-pressure, NPT, simulations that do not involve any transfers of particles. For one-component systems, the method is based on integration of the Clausius–Clapeyron equation over temperature,

$$\left(\frac{\mathrm{d}P}{\mathrm{d}\beta}\right)_{\mathrm{sat}} = -\frac{\Delta H}{\beta \Delta V}, \tag{10.11}$$

where sat indicates that the equation holds on the saturation line, and ΔH is the difference in enthalpy between the two coexisting phases. The right-hand side of (10.11) involves only 'mechanical' quantities that can be simply determined in the course of a standard MC or MD simulation. From the known point on the coexistence curve, a change in temperature is chosen, and the saturation pressure at the new temperature is predicted from (10.11). Two independent simulations for the corresponding phases are performed at the new temperature, with gradual changes of the pressure as the simulations proceed to take into account the enthalpies and densities at the new temperature as they are being calculated.

Questions related to the propagation of errors and numerical stability of the method have been addressed in [33, 34]. Errors in initial conditions resulting from uncertainties in the coexistence densities can propagate and increase with distance from the starting point when the integration path is toward the critical point [34]. Near critical points, the method suffers from instability of a different nature. Because of the small free energy barrier for conversion of one phase into the other, even if the coexistence pressure is set properly, the identity of each phase is hard to maintain and large fluctuations in density are likely. The solution to this last problem is to borrow an idea from the Gibbs ensemble and couple the volume changes of the two regions [33]. Extensions of the method to calculations of three-phase coexistence lines are presented in [35] and to multicomponent systems in [34]. Unfortunately, for multicomponent systems the Gibbs–Duhem integration method cannot avoid particle transfers; however, it avoids transfers for one component, typically the one that is the hardest to transfer. The method and its applications have been reviewed in [36].

In some cases, in particular for lattice and polymeric systems, volume change moves may be hard to perform, but particle insertions and deletions may be relatively

easy, especially when using configurational bias methods. Escobedo and de Pablo [37, 38] proposed a modification of the Gibbs–Duhem approach that is based on the expression:

$$\left(\frac{\mathrm{d}(\beta\mu)}{\mathrm{d}\beta}\right)_{\mathrm{sat}} = -\frac{\Delta(\rho u)}{\Delta\rho},\qquad(10.12)$$

where ρ is the density $(= N/V)$ and u is the energy per particle. This method has been applied to continuous-phase polymeric systems in [37] and to lattice models in [39].

The Gibbs–Duhem integration method excels in calculations of solid–fluid coexistence [40, 41], for which other methods described in this chapter are not applicable. The calculation of the free energies of solids is typically performed via thermodynamic integration along an appropriate singularity-free coordinate, as described elsewhere in this book and in [1]. An extension of the Gibbs–Duhem method, in which it is assumed that the initial free-energy difference between the two phases is known in advance, rather than requiring it to be zero, has been proposed by Meijer and El Azhar [42]. The procedure has been used in [42] to determine the coexistence lines of a hard-core Yukawa model for charge-stabilized colloids.

The Gibbs–Duhem integration method represents a successful combination of numerical methods and molecular simulations. Taking this concept even further, Mehta and Kofke [43] proposed a 'pseudo-grand canonical ensemble' method, in which a system maintains a constant number of particles and temperature, but has a fluctuating volume to ensure that, at the final density, the imposed value of the chemical potential is matched. The formalism still requires that estimates of the chemical potential be made during the simulation. The main advantage of the approach over more traditional grand canonical ensemble methods is that it provides additional flexibility with respect to the method to be used for determination of the chemical potential. For example, the 'chain increment' method [16] for chain molecules, which cannot be combined with grand canonical simulations, can be used for the chemical potential evaluations in a pseudo-grand canonical simulation (as in [14]).

The same 'pseudo-ensemble' concept has been used by Camp and Allen [44] to obtain a 'pseudo-Gibbs' method in which particle transfers are substituted by volume fluctuations of the two phases. The volume fluctuations are unrelated to the ones required for pressure equality (10.7) but are instead designed to correct imbalances in the chemical potentials of some of the components detected, for example, by test particle insertions.

While the main driving force in [43, 44] was to avoid direct particle transfers, Escobedo and de Pablo [38] designed a 'pseudo-NPT' method to avoid direct volume fluctuations which may be inefficient for polymeric systems, especially on lattices. Escobedo [45] extended the concept for bubble-point and dew-point calculations in a 'pseudo-Gibbs' method and proposed extensions of the Gibbs–Duhem integration techniques for tracing coexistence lines in multicomponent systems [46].

10.3.3 Phase Equilibria in the Grand Canonical Ensemble

A grand-canonical Monte Carlo (GCMC) simulation for a one-component system is performed as follows. The simulation cell has a fixed volume V, and periodic

boundary conditions are applied. The inverse temperature, $\beta = 1/k_{\mathrm{B}}T$ and the chemical potential, μ, are specified as input parameters to the simulation. In this ensemble, histogram reweighting requires collection of data for the joint probability $\wp(N, U)$ of the occurrence of N particles in the simulation cell with total configurational energy in the vicinity of U. This probability distribution function follows the grand-canonical partition function

$$\wp(N, U) = \frac{\Lambda^{-3N}\Omega(N, V, U)\exp(-\beta U + \beta\mu N)}{\Xi(\mu, V, \beta)}, \tag{10.13}$$

where $\Xi(\mu, V, \beta)$ is the grand partition function. We do not know Ω or Ξ at this stage, but Ξ is a constant for a run at given conditions. Since the left-hand side of (10.13) can be easily measured in a simulation, an estimate for Ω and its corresponding thermodynamic function, the entropy $\mathscr{S}(N, V, U)$, can be obtained by a simple transformation of (10.13)

$$\mathscr{S}(N, V, U) = \ln \Omega(N, V, U) = \ln \wp(N, U) + \beta U - \beta\mu N + 3N \ln \Lambda + \mathrm{C}, \tag{10.14}$$

where C is a run-specific constant. As in the example of Chap. 3, using the measured distribution in energy, (10.14) is meaningful only over the range of particle number and energies covered in a simulation. If two runs at different chemical potentials and temperatures have a region of overlap in the space of (N, U) sampled, then the entropy functions can be 'merged' by requiring that the entropy functions are identical in the region of overlap. To illustrate this concept, we make a one-dimensional projection of (10.13) to obtain

$$\wp(N) = \frac{Q(N, V, \beta)\exp(\beta\mu N)}{\Xi(\mu, V, \beta)}. \tag{10.15}$$

Histograms for two runs at different chemical potentials are presented in Fig. 10.2. There is a range of N over which the two runs overlap. In Fig. 10.3 we show the function $\ln \wp(N) - \beta\mu N$ for the data of Fig. 10.2. Rearranging (10.15) and taking the logarithm, we see that this function is related to the Helmholtz free energy

$$\beta A(N, V, \beta) = -\ln Q(N, V, \beta) = \ln \wp(N) - \beta\mu N + \ln \Xi(\mu, V, \beta). \tag{10.16}$$

The raw curves for μ_1 and μ_2 as well as a 'composite' curve formed by shifting data for the two runs by the amount indicated by the arrows are shown in Fig. 10.3 . The combined curve provides information over the combined range of particle numbers, N, covered by the two runs. Note that by keeping one-dimensional histograms for N we are restricted to combining runs of the same temperature, while the more general form (10.14) allows combination of runs at different temperatures.

As before, our simulation data are subject to statistical (sampling) uncertainties, which are particularly pronounced near the extremes of particle numbers and energies visited during a run. When data from multiple runs are combined, as shown in Fig. 10.3, the question arises of how to determine the optimal amount by which to shift the raw data in order to obtain a global free energy function. As reviewed in
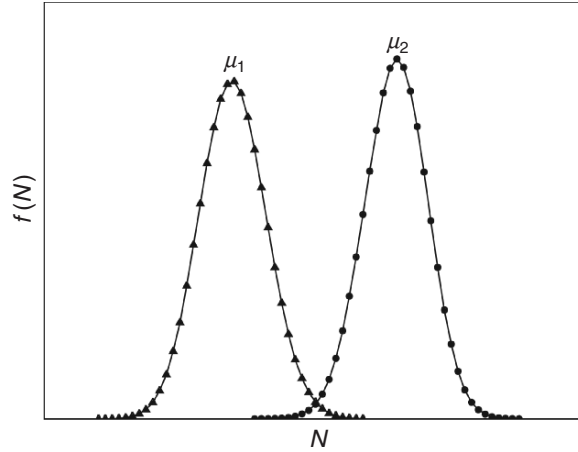
Fig. 10.2. Schematic diagram of the probability $f(N)$ of occurrence of N particles for two GCMC runs of a pure component system at the same volume V and temperature T, but different chemical potentials, μ_1 and μ_2, respectively. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd
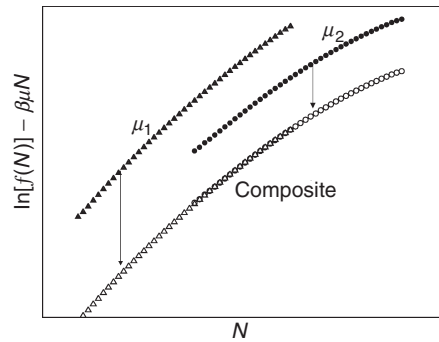

Fig. 10.3. The function $\ln[f(N)] - \beta\mu N$ for the data of Fig. 10.2. The figure shows the raw curves for μ_1 and μ_2 as well as a 'composite' curve formed by shifting the data by the amount indicated by the *arrows*. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd

Chap. 3, Ferrenberg and Swendsen [47] provided a solution to this problem by minimizing the differences between predicted and observed histograms. Since a complete derivation of this procedure is available in [1], we will only present the result here. In this approach, one considers a collection of histogram data from multiple simulations at differing conditions, denoted by $i = 1, 2, \ldots, R$. The Ferrenberg–Swendsen approach then determines the optimal macrostate probabilities when data from all of these runs is patched together and reweighted to another condition. The composite probability, $\wp(N, U; \mu, \beta)$, of observing N particles and potential energy U, if one

takes into account all runs and assumes that they have the same statistical efficiency, is

$$\widetilde{\wp}^*(N, U; \mu, \beta) = \frac{\displaystyle\sum_{i=1}^{R} f_i(N, U)\exp[-\beta U + \beta \mu N]}{\displaystyle\sum_{i=1}^{R} K_i \exp[-\beta_i U + \beta_i \mu_i N - C_i]} \tag{10.17}$$

where $\widetilde{\wp}^*$ is the un-normalized probability prediction, f is the histogram of counts, and K_i is the total number of observations ($K_i = \sum_{N,U} f_i(N, U)$) for run i. The constants C_i (also known as 'weights') are obtained by iteration from the relationship

$$\exp(C_i) = \sum_N \sum_U \widetilde{\wp}^*(N, U; \mu_i, \beta_i). \tag{10.18}$$

Given an initial guess for the set of weights C_i, (10.17) and (10.18) can be iterated until convergence. When many histograms are to be combined, this convergence of the Ferrenberg–Swendsen weights can take a long time. Once this has been achieved, however, all thermodynamic quantities for the system over the range of densities and energies covered by the histograms can be obtained. For example, the mean configurational energy is

$$\langle U \rangle (\mu, \beta) = \sum_N \sum_U \widetilde{\wp}(N, U; \mu, \beta)U, \tag{10.19}$$

and the mean density $\rho(\mu, \beta)$ is

$$\langle \rho \rangle (\mu, \beta) = \frac{1}{V} \sum_N \sum_U \widetilde{\wp}(N, U; \mu, \beta)N. \tag{10.20}$$

In both of these equations, we have used a summation rather than an integral over the potential energy for both clarity and the connection with the actual, discretized measurements in a simulation.

In the canonical example, we could estimate the free energy difference between two runs by examining the overlap in their probability distributions. Similarly, in the grand canonical ensemble, we can estimate the pressure difference between the two runs. If the conditions for run 1 are (μ_1, V, β_1) and for run 2 (μ_2, V, β_2), then

$$C_2 - C_1 = \ln\frac{\Xi(\mu_2, V, \beta_2)}{\Xi(\mu_1, V, \beta_1)} = \beta_2 P_2 V - \beta_1 P_1 V, \tag{10.21}$$

where P is the pressure, since $\ln \Xi = \beta PV$. We can use (10.21) to obtain the absolute value of the pressure for one of the two runs, provided that the absolute pressure can be estimated for the other run. Typically, this is done by performing simulations for low-density states for which the system follows the ideal-gas equation of state, $PV = Nk_BT$.

Up to this point, we assumed that a system exists in a one-phase region over the range of densities and energies sampled. If a phase transition exists then *in principle*,

states on either side of the phase transition should be sampled, resulting in histograms with multiple peaks. This is illustrated in Fig. 10.4, in which actual simulation data (from a single run) are plotted for a simple cubic lattice homopolymer system [48] at a slightly subcritical temperature. There are two states sampled by the run, one at low and one at high particle numbers, corresponding to the gas and liquid states. The conditions for phase coexistence are equality of temperature, chemical potential and pressure; the first two are satisfied by construction in the grand canonical ensemble. From (10.21), the integral under the un-normalized probability distribution function with respect to N and U is proportional to the pressure. In the case of two distinct phases, the integrals should be calculated separately under the liquid and gas peaks. The condition of equality of pressures can be satisfied by reweighing the data until this condition is met, i.e., until the integrated 'volumes' under the two peaks are equal. In Sect. 10.3.4, we discuss how near-critical histogram data can be used to obtain precise estimates of the critical parameters for a transition.

In the absence of phase transitions or at temperatures near a critical point, the values of all observable quantities (such as the histograms of energy and density) are independent of initial conditions, since free energy barriers for transitions between states are small or nonexistent and the system readily progresses through its configuration space. However, at lower temperatures, free energy barriers for nucleation of new phases become increasingly larger. At these conditions, we will find that the states sampled at a given temperature and chemical potential depend on the initial conditions, a phenomenon known as hysteresis. This is illustrated
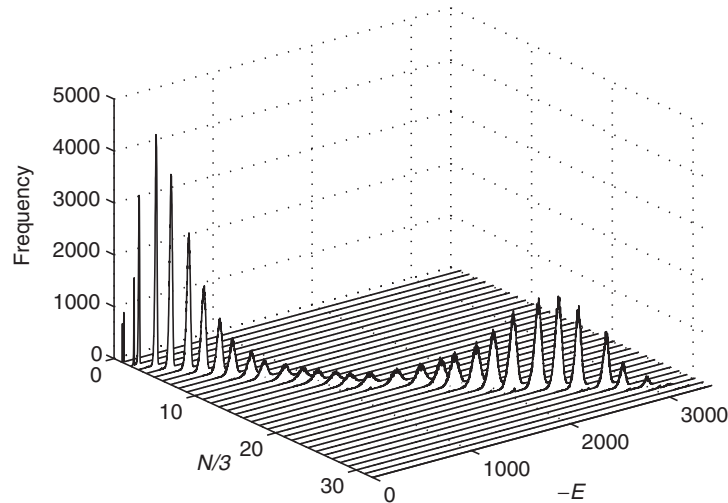

Fig. 10.4. Frequency of observation of states versus energy, E, and number of particles, N, for a homopolymer of chain length $r = 8$ and coordination number $z = 6$ on a $10 \times 10 \times 10$ simple cubic lattice. Conditions, following the notation of [48] are $T^* = 11.5, \mu^* = -60.4$. In order to reduce clutter, data are plotted only for every third particle. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd

schematically in Fig. 10.5. For a supercritical isotherm, $T > T_c$, the mean value of the density is a continuous function of the chemical potential, and the same value is obtained for given conditions, irrespective of the starting configuration. By contrast for a subcritical isotherm, when we start runs at low-density state, we will observe a discontinuous 'jump' to a state of higher density at some value of the chemical potential. The exact location of the jump depends on the initial state and the specific mix of MC moves used to change the configurations of the system. When simulations are started in a high-density state, the system remains on the high-density branch of the isotherm until some value of the chemical potential that is lower than the chemical potential of the jump from low- to high-density states.

The histogram reweighting method can be applied to systems with large free-energy barriers for transitions between states, provided that care is taken to link all states of interest via reversible paths. One possibility is to utilize umbrella or multicanonical sampling techniques [49, 50] to artificially enhance the frequency with which the intermediate density region is sampled in a simulation [51]. These methods are discussed in detail in Chap. 3. Essentially, multicanonical and umbrella sampling require as input an estimate of the free energy in the intermediate density region, which has to be obtained by trial and error. In addition, a significant fraction of simulation time is spent sampling nonphysical configurations of intermediate density. An alternative approach is to link states by providing connections through a supercritical path, in a process analogous to thermodynamic integration [1]. This approach is illustrated schematically in Fig. 10.6. The filled square represents the critical point for a transition, and open squares linked by dashed lines represent tie lines. Ellipses represent the range of particle numbers and energies sampled by a single simulation. A near-critical simulation samples states on both sides of the coexistence curve, while subcritical simulations are likely to be trapped in
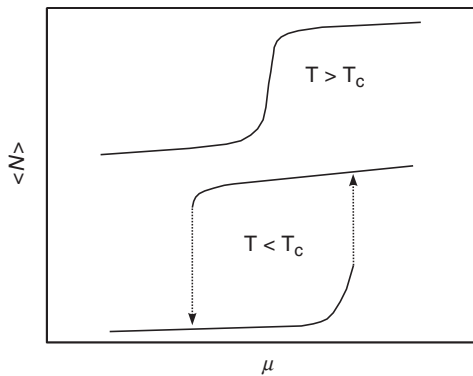

Fig. 10.5. Schematic diagram of the mean number of particles, $\langle N \rangle$, versus chemical potential, μ for a subcritical and a supercritical isotherm of a one-component fluid. The curve for the supercritical isotherm has been shifted up for clarity. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd
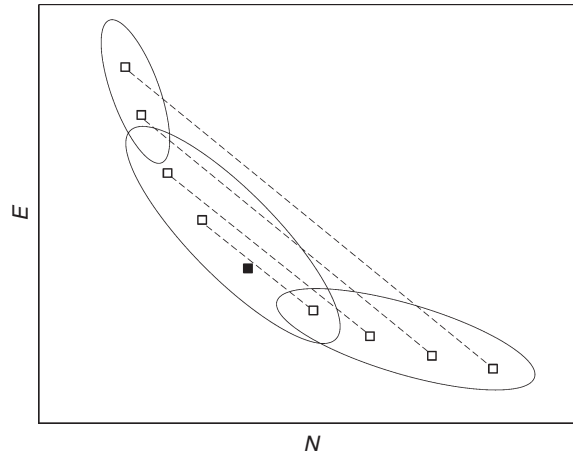
Fig. 10.6. Schematic diagram of the energy, E, versus the number of particles, N, for a one-component fluid with a phase transition. Squares linked by dashed lines are coexisting phases joined by tie lines and the filled square indicates the critical point of the transition. Ellipses represent the range of particle numbers and energies sampled during different GCMC runs. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd

(possibly metastable) states on either side. However, as long as there is a continuous path linking all states of interest, the free energies and pressures can be calculated correctly, and an accurate phase envelope can be obtained.

An example of the application of histogram reweighting for determining the phase behavior of a homopolymer model on the simple cubic lattice is illustrated in Fig. 10.7. The phase behavior and critical properties of the model for a range of chain lengths have been studied in [48]. The system in this example is for chain length $r = 8$ and coordination number $z = 6$. In this example, we first performed a simulation at reduced temperature $T^* = 11.5$ and chemical potential $\mu^* = -60.4$, for which the raw histogram data are shown in Fig. 10.4. The resulting average volume fraction for the run is indicated in Fig. 10.7 by the filled circle at $T^* = 11.5$. The range of volume fractions sampled during the simulation is indicated in Fig. 10.7 by the arrows originating at the run point. Because this run is near the critical point, a very broad range of particle numbers and thus volume fractions is sampled during this single run. The histogram from this run was then reweighted to lower temperatures and a preliminary phase diagram was obtained. The estimated coexistence chemical potential at $T^* = 9$ was used as the input to a new simulation, which sampled states near the saturated liquid line. The same procedure was repeated, now with combined histograms from the first two runs, to obtain an estimate of the coexistence chemical potential at $T^* = 7$. A new simulation was performed to sample the properties of the liquid at that temperature. The final result of these three calculations was the phase coexistence lines shown by the thick continuous lines on Fig. 10.7.
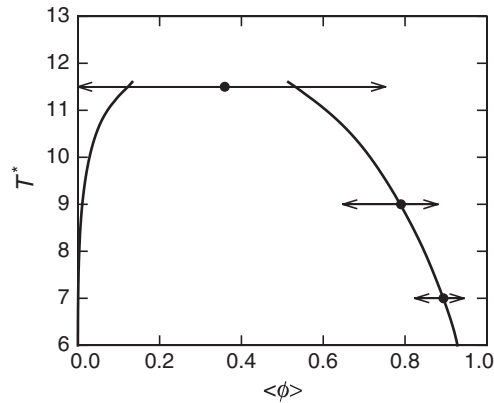
Fig. 10.7. Phase diagram for a homopolymer of chain length $r = 8$ on a $10 \times 10 \times 10$ simple cubic lattice of coordination number $z = 6$. Filled circles give the reduced temperature, T^* and mean volume fraction, $\langle \phi \rangle$ of the three runs performed. Arrows from the run points indicate the range of densities sampled for each simulation. The thick continuous line is the estimated phase coexistence curve. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd

Two general observations can be made in relation to this example. First, we should point out that the histogram reweighting method works much faster for smaller system sizes. As system size increases, relative fluctuations in the number of particles and energy for a single run at specified conditions decrease as the $1/2$ power of the system volume V. This implies that more simulations are required to obtain overlapping histograms that cover the range of energies and densities of interest. Moreover, the number of MC moves required to sample properties increases approximately linearly with system size in order to keep the number of moves per particle constant. The computational cost of each MC move is proportional to system size for pairwise additive long-range interactions and independent of system size for short-range interactions. The net effect is that total computational effort required to obtain a phase diagram at a given accuracy scales as the 1.5–2.5 power of system volume, respectively for short- and long-range interactions. Fortunately, away from critical points, the effect of system size on the location of the coexistence curves for first-order transitions is typically small. In this example, calculations on a 15^3 system result in phase coexistence lines practically indistinguishable from the ones shown in Fig. 10.7. The mean absolute relative differences for the coexistence densities between the small and large systems are 0.1% for the liquid and 1% for the (much lower density) gas, well within the width of the coexistence lines in Fig. 10.7.

A second observation relates to calculations near critical points. The coexistence lines in Fig. 10.7 do not extend above a temperature of $T^* = 11.6$ because above that temperature significant overlap exists between the liquid and vapor peaks of the histograms. This overlap renders calculations of the liquid and gas densities imprecise.

Larger system sizes suffer less from this effect and can be used to obtain coexistence densities near critical points.

10.3.4 Advanced Approaches

The histogram reweighting methodology for multicomponent systems [52–54] closely follows the one-component version described above. The probability distribution function for observing N_1 particles of component 1 and N_2 particles of component 2 with configurational energy in the vicinity of E for a GCMC simulation at imposed chemical potentials μ_1 and μ_2, respectively, at inverse temperature β in a box of volume V is

$$\wp(N_1, N_2, U) = \frac{\Lambda_1^{-3N_1} \Lambda_2^{-3N_2} \Omega(N_1, N_2, V, U)\exp(-\beta U + \beta \mu_1 N_1 + \beta \mu_2 N_2)}{\Xi(\mu_1, \mu_2, V, \beta)}.$$

(10.22)

Equations (10.14–10.21) can be similarly extended to multicomponent systems.

The main complication in the case of multicomponent systems relative to the one-component case is that the dimensionality of the histograms is one plus the number of components, thus making their machine storage and manipulation somewhat more challenging. For example, in the case of one-component systems, it is possible to store the histograms directly as two-dimensional arrays. The memory requirements for storing three-dimensional arrays for a two-component system makes it impractical to do so. Instead, lists of observations of particle numbers and energies are periodically stored on disk. It is important to select the frequency of sampling of the histogram information so that only independent configurations are sampled. This implies that sampling is less frequently at high densities for which the acceptance ratio of the insertion and removal steps is lower. Sampling essentially independent configurations also enforces the condition of equal statistical efficiency underlying the Ferrenberg–Swendsen histogram combination (10.17) and (10.18).

Recent advances in the determination of critical parameters for fluids lacking special symmetries have been based on the concept of mixed-field finite-size scaling and have been reviewed in detail by Wilding [55]. As a critical point is approached, the correlation length ξ grows without bound and eventually exceeds the linear system size L of the simulation box. Singularities and discontinuities that characterize critical behavior in the thermodynamic limit are smeared out and shifted in finite systems. The infinite-volume critical point of a system can, however, be extracted by examining the size dependence of thermodynamic observables, through finite-size scaling theory [56–58]. The finite-size scaling approach proposed by Bruce and Wilding [59, 60] accounts for the lack of symmetry between coexisting phases in most continuous-space fluids. For one component systems, the ordering operator, M, is proportional to a linear combination of the number of particles N and total configurational energy U

$$M \propto N - sU,$$

(10.23)

where s is the field mixing parameter. For multicomponent systems, an extra field mixing parameter appears for each added component – for example binary systems

$$M \propto N_1 - sU - qN_2, \tag{10.24}$$

where q is the field mixing parameter for the number of particles of component 2.

General finite-size scaling arguments predict that the normalized probability distribution for the ordering operator M at criticality, $\wp(M)$, has a universal form. The order parameter distribution for the three-dimensional Ising universality class is shown in Fig. 10.8 as a continuous line. Also shown in Fig. 10.8 are data for a homopolymer of chain length $r = 200$ on a $50 \times 50 \times 50$ simple cubic lattice of coordination number $z = 26$ [48]. The data were obtained by histogram reweighting methods, by adjusting the chemical potential, temperature and field mixing parameter s so as to obtain the best possible fit to the universal distribution. The nonuniversal constant A and the critical value of the ordering operator M_c were chosen so that the data have zero mean and unit variance. Due to finite-size corrections to scaling, the apparent critical temperature, $T_c(L)$, and density, $\rho_c(L)$, deviate from their infinite-system values, $T_c(\infty)$ and $\rho_c(\infty)$. They are expected to follow the scaling relationships with respect to the simulated system size, L

$$
\begin{aligned}
T_c(L) - T_c(\infty) &\propto L^{-(\theta+1)/\nu} \\
\rho_c(L) - \rho_c(\infty) &\propto L^{-(1-\alpha)/\nu},
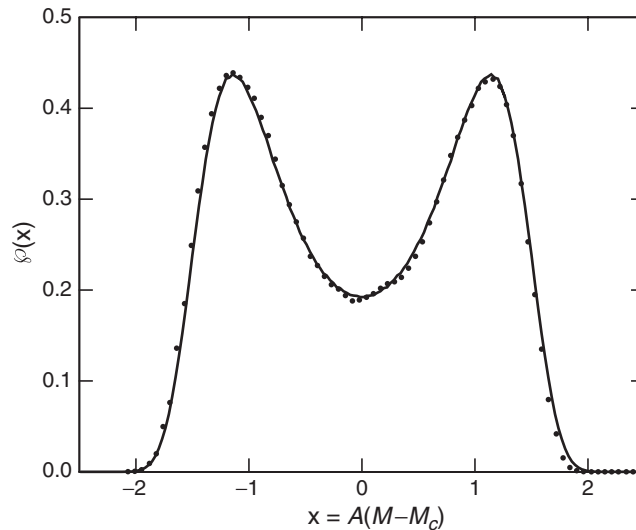\end{aligned}
\tag{10.25}
$$

Fig. 10.8. The ordering operator distribution for the three-dimensional Ising universality class (continuous line – data are courtesy of N.B. Wilding). Points are for a homopolymer of chain length $r = 200$ on a $50 \times 50 \times 50$ simple cubic lattice of coordination number $z = 26$ [48]. The nonuniversal constant A and the critical value of the ordering operator M_c were chosen so that the data have zero mean and unit variance. Reprinted by permission from [6]. ©2000 IOP Publishing Ltd

where θ, ν and α are, respectively, the correction-to-scaling exponent, the correlation length exponent and the exponent associated with the heat capacity divergence. For the three-dimensional Ising universality class, the approximate values of these exponents are [62, 63] $(\theta, \nu, \alpha) \approx (0.54, 0.629, 0.11)$. Figure 10.9 demonstrates these scaling relationships for the critical temperature and density of the square-well fluid of range $\lambda = 3$ [61].

Finally in this section, we would like to mention briefly two methods that are related to histogram reweighting. Thermodynamic scaling techniques proposed by Valleau [64] are based on calculations in the NPT, rather than the grand canonical (μVT) ensemble and provide information about the free energy over a range of volumes, rather than a range of particle numbers. Thermodynamic scaling techniques can also be designed to cover a range of Hamiltonians (potential models) in the Gibbs [65] or grand canonical [66] ensembles. In their Hamiltonian scaling form, the methods are particularly useful for optimizing parameters in intermolecular potential models to reproduce experimental data such as the coexisting densities and vapor pressures. Thermodynamic and Hamiltonian scaling methods require estimates for the free energy of the system as a function of conditions, so that the system can
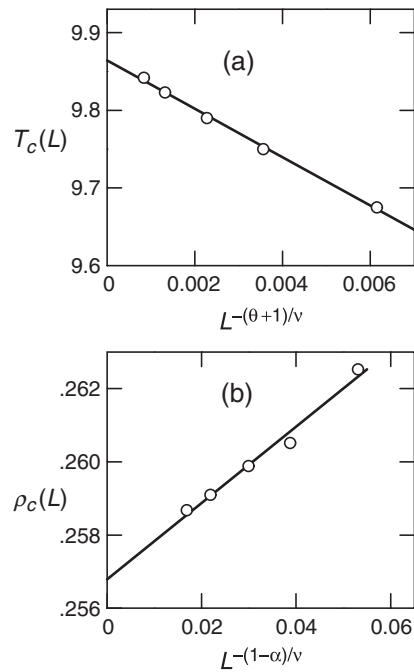

Fig. 10.9. Critical temperature (**a**) and density (**b**) scaling with linear system size for the square-fluid of range $\lambda = 3$. *Solid lines* represent a least-squares fit to the points. Reprinted by permission from [61]. ©1999 American Institute of Physics

be forced to sample the range of states of interest with roughly uniform probability, as for umbrella sampling MC [49].

10.4 Selected Applications of Flat Histogram Methods

The flat-histogram algorithms described in Chap. 3 [67–73] provide a straightforward basis for performing phase equilibrium calculations. Many of these methods are designed to determine an underlying thermodynamic potential, such as an entropy or free energy, over a very wide range of macrostates. Once such a simulation is complete, the calculated potential can be used directly to determine state probabilities and ensemble averages through a reweighing-like procedure. For phase equilibrium calculations, one then seeks to find conditions at which the calculated macrostate probability distribution function is bimodal – for instance, the temperature at a given pressure where the probability distribution of densities exhibits two peaks.

In the sections below, we describe several studies in which flat-histogram methods were used to examine phase equilibria in model systems. The discussion assumes the reader is familiar with this general family of techniques and the theory behind them, so it may be useful to consult the material in Chap. 3 for background reference. Although the examples provided here entail specific studies, their general form and the principles behind them serve as useful templates for using flat-histogram methods in novel phase equilibria calculations.

10.4.1 Liquid–Vapor Equilibria using the Wang–Landau Algorithm

Liquid–vapor phase behavior is important to numerous technologies, and fortunately the phase envelope of model systems is readily studied using flat-histogram methods. By performing a single density of states simulation to measure the energy and density dependence of $\mathscr{S}$, a subsequent reweighing procedure can be used to identify state conditions which result in a bimodal macroscopic probability distribution – the signature feature of first-order phase transitions. The Wang–Landau algorithm is well-suited for this task, and was first applied by Yan et al. [74] and Shell et al. [75] to the frequently studied Lennard-Jones system.

In order to map the thermodynamic properties of this system, the Gibbs phase rule implies that two parameters are necessary to specify the thermodynamic state. Consequently, in setting up the density of states simulation, we desire to achieve a flat-histogram in both energy and density. The density component is frequently addressed by allowing the number of particles N to fluctuate via addition and deletion moves, in which case the subject of calculation is $\mathscr{S}(N,U)$. That is, in the Wang–Landau scheme, a dimensionless entropy whose dependent variables are the number of particles and potential energy will be dynamically modified, while the simulation volume remains fixed. Examining the equations of Chap. 3, the microscopic sampling scheme for this simulation is

$$\wp(\mathbf{q}, N) \propto \frac{1}{N!} \exp[-\mathscr{S}(N, U)]. \tag{10.26}$$

The acceptance criteria for particle insertion and deletion moves are determined from the detailed balance condition applied to these probabilities. The final expressions are

$$\mathscr{P}_{\text{acc}}(N \rightarrow N+1) = \min\left\{1, \exp[\mathscr{S}(N, U_\text{o}) - \mathscr{S}(N+1, U_\text{n}) + \ln V - \ln(N+1)]\right\}$$
(10.27a)

$$\mathscr{P}_{\text{acc}}(N \rightarrow N-1) = \min\left\{1, \exp[\mathscr{S}(N, U_\text{o}) - \mathscr{S}(N-1, U_\text{n}) + \ln N - \ln V]\right\},$$
(10.27b)

where the subscripts on U indicate their correspondence with the original and new states. The corresponding acceptance criterion for particle displacement moves, which leads to fluctuations in the potential energy at fixed density is

$$\mathscr{P}_{\text{acc}}(\text{o} \rightarrow \text{n}) = \min\left\{1, \exp[\mathscr{S}(N, U_\text{o}) - \mathscr{S}(N, U_\text{n})]\right\}.$$
(10.28)

Recognizing that $\mathscr{S} = \mathscr{S}_{\text{ex}} + N\ln V - \ln N!$ from Chap. 1, all three acceptance criteria can be expressed as

$$\mathscr{P}_{\text{acc}}(\text{o} \rightarrow \text{n}) = \min\left\{1, \exp[\mathscr{S}_{\text{ex}}(N_\text{o}, U_\text{o}) - \mathscr{S}_{\text{ex}}(N_\text{n}, U_\text{n})]\right\}.$$
(10.29)

We can, therefore, let $\mathscr{S}_{\text{ex}}$ be the subject of our calculations (which we approximate via an array in the computer). Post-simulation, we desire to examine the joint probability distribution $\wp(N, U)$ at normal thermodynamic conditions. The reweighting ensemble which is appropriate to fluctuations in N and U is the grand-canonical ensemble; consequently, we must specify a chemical potential and temperature to determine $\wp$. Assuming $\mathscr{S}_{\text{ex}}$ has converged upon the true function $\ln \Omega_{\text{ex}}$, the state probabilities are given by

$$\wp(N, U; \mu, T) = c \times \Lambda^{-3N} \exp[\mathscr{S}(N, U) - \beta U + \beta\mu N]$$
$$\wp(N, U; \mu, T) = c \times \exp[\mathscr{S}(N, U) - \beta U + \beta\mu' N],$$
(10.30)

where c is a normalization constant, calculated so that the probabilities sum to unity. In the second line, we have absorbed the de Broglie wavelength into the chemical potential, notated by μ', which simply serves to shift its zero. Phase coexistence is determined by a procedure identical to that described in Sect. 10.3.3. For a given subcritical T, μ is adjusted until $\wp$ exhibits two peaks of equal volume (integrated over N and U). The corresponding equilibrium densities and energies are then given by

$$\langle a \rangle_{\text{liq}} = \frac{\displaystyle\sum_{U} \sum_{N > N_{\text{mid}}} a(N, U)\wp(N, U; \mu, T)}{\displaystyle\sum_{U} \sum_{N > N_{\text{mid}}} \wp(V, U; \mu, T)}$$

$$\langle a \rangle_{\text{gas}} = \frac{\displaystyle\sum_{U} \sum_{N < N_{\text{mid}}} a(N, U)\wp(N, U; \mu, T)}{\displaystyle\sum_{U} \sum_{N < N_{\text{mid}}} \wp(V, U; \mu, T)}$$
(10.31)

where N_{mid} is the particle number at the valley between the two peaks and a is either N or U.

The density dependence of the entropy can also be studied by introducing fluctuations in volume rather than particle number. Typically the particle number approach is favored; the computational demands of volume scaling moves scale faster with system size than do addition and deletion moves. Nevertheless, the Wang–Landau approach provides a means for studying volume fluctuations as well. In this case, the excess entropy is determined as a function of volume and potential energy for fixed particle number; one, therefore, calculates $\mathscr{S}(V, U)$. Here the microstate probabilities follow:

$$\wp(\mathbf{q}, V) \propto \exp[-\mathscr{S}(V, U)]. \tag{10.32}$$

The acceptance criterion for increasing or decreasing the volume by a random, linear amount follows as

$$\mathscr{P}_{\text{acc}}(V_{\text{o}} \rightarrow V_{\text{n}}) = \min\left\{1, \exp[\mathscr{S}(V_{\text{o}}, U_{\text{o}}) - \mathscr{S}(V_{\text{n}}, U_{\text{n}}) + N \ln(V_{\text{n}}/V_{\text{o}})]\right\}, \tag{10.33}$$

which also conforms to (10.29) when expressed using the excess entropy and substituting volume for particle number. In the case where changes are made to the logarithm of the volume, rather than the volume itself, the N in this expression becomes $N + 1$. Finally, the appropriate reweighting probabilities are those of the isothermal–isobaric ensemble:

$$\wp(V, U; \mu, T) = c \times \exp[\mathscr{S}(V, U) - \beta U - \beta P V]. \tag{10.34}$$

To determine phase equilibria in this case, P is adjusted at fixed T to achieve a bimodal distribution, in exactly the same manner as before.

The results from the Wang–Landau study of the Lennard-Jones fluid in [75] are shown in Figs. 10.10 and 10.11, in which the 'Gibbs surface' for the excess entropy as a function of U and either N or V are plotted. In these simulations, the entropy was determined using the usual Wang–Landau approach, in conjunction with random particle displacements and changes in density through either isotropic volume dilations or particle additions and deletions. Therefore, the relevant flat-histogram and results are functions of two variables. In each case, the modification factor was initialized at $g = 1$ and decreased by a factor of two at each subsequent stage. A stage was deemed complete once each (N, U) or (N, V) pair was visited at least 20 times, and the simulation was terminated for $g < 10^{-5}$. In order to determine the accessible energy range at each density, a lower-bound potential energy border was first determined from short NVT runs at the lowest temperature of interest; states beyond this border did not count towards the flat-histogram requirement for stage completion, and proposed moves to this region were rejected.

Figures 10.10 and 10.11 contain the complete phase behavior of the Lennard-Jones system over the range of energies and densities indicated, but do not immediately reveal any important features. More revealingly, using the reweighting procedures just described, the authors generated the liquid–vapor phase envelope. Figure 10.12 contains these results, which are in good agreement with both literature values and those results generated from the grand canonical histogram reweighing techniques. Although the authors did not observe any speed advantage over the latter approach, the results of the Wang–Landau method contain, in principle, all relevant
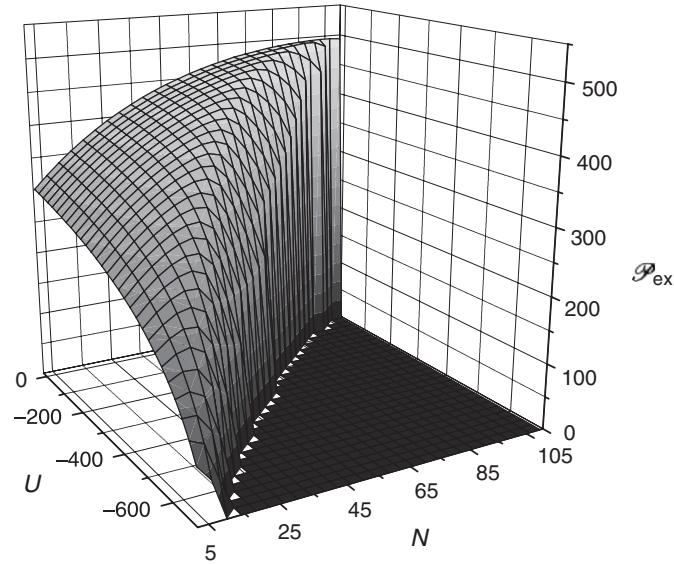
Fig. 10.10. Calculated $\mathscr{S}_{ex}(N, U)$ from a Wang–Landau simulation for the Lennard-Jones fluid at $V = 125$. The potential energy has been discretized into 1,000 bins and is expressed in Lennard-Jones units. Reprinted figure with permission from [75]. ©2002 by the American Physical Society
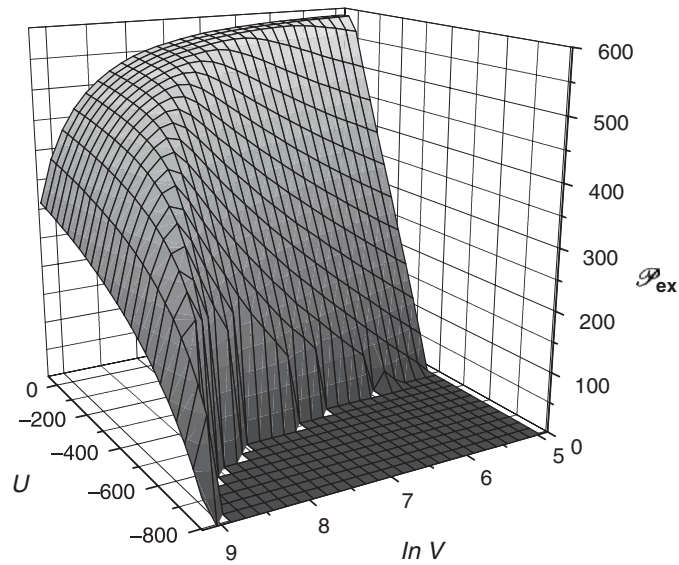

Fig. 10.11. Calculated $\mathscr{S}_{ex}(V, U)$ from a Wang–Landau simulation for the Lennard-Jones fluid at $N = 128$. The potential energy and volume have been discretized into 500 and 200 bins, respectively, and are expressed in Lennard-Jones units. Reprinted figure with permission from [75]. ©2002 by the American Physical Society
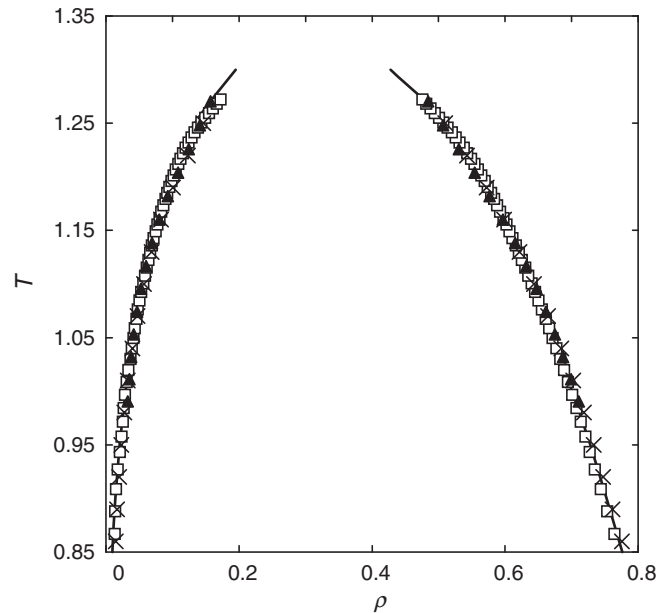
Fig. 10.12. Vapor–liquid phase behavior for the Lennard-Jones fluid. Solid triangles and hollow squares indicate the results of the particle addition/deletion and volume scaling variants of the flat-histogram simulation using the Wang–Landau algorithm. Crosses are from a histogram reweighting study based on grand-canonical measurements at seven state points. The solid line is from Lotfi, et al. [76]. Reprinted figure with permission from [75]. ©2002 by the American Physical Society

phase transitions within the state space studied and therefore are potentially very useful for mapping out the complete phase behavior of new systems. This contrasts with the GCMC histogram reweighting technique, in which a good estimate of the state conditions for equilibrium must be determined beforehand.

10.4.2 Prewetting Transitions in Confined Fluids using Transition Matrix Methods

Fluids confined to extremely small regions and in the presence of surfaces often exhibit a wealth of interesting phase transitions. Understanding the origins and behavior of these transitions, especially relative to the bulk fluid, is vital to technologies which operate at the nanoscale. A model system, that of argon confined between a solid carbon dioxide surface and a hard wall, has been studied by Errington using an efficient transition matrix approach [77]. In the model, argon–argon interactions are modeled by the Lennard-Jones interaction and the argon-surface energies are given by a 9–3 potential [78]. Periodic boundary conditions are applied in the directions parallel to the adsorbing and hard surfaces.

Errington used the grand-canonical ensemble with a transition matrix to estimate the weights $\eta(N)$ needed to achieve a flat-histogram in particle number at a given temperature. This is identical to the example discussed in Chap. 3, where the grand-canonical equations are given for microstate probabilities and acceptance criteria for particle addition and deletion. For a selected temperature, a starting chemical potential is chosen and the initial weights are set to zero. The simulation is run for some time before the weights are updated. The relevant transition probability equation for this initial purely grand-canonical scenario is

$$\wp(N_1)T(N_1 \to N_2) = \wp(N_2)T(N_2 \to N_1) \tag{10.35}$$

where all quantities are implicitly a function of the temperature, chemical potential, and zero weights. In order to achieve a flat-histogram in N, the weights must be such that $\eta(N) \sim \ln \wp(N)$. Using (10.35), this estimate can be determined directly from the transition probabilities. One first sets $\eta(N_{\min}) = 0$, and then evaluates each neighbor in series through

$$\eta(N+1) = \eta(N) + \ln T(N+1 \to N) - \ln T(N \to N+1). \tag{10.36}$$

Estimates of the transition probabilities are constructed from values of $\mathscr{P}_{\text{acc}}$ collected during the simulation. The $C(N_1, N_2)$ matrix is updated at each particle addition or deletion during the run: $\mathscr{P}_{\text{acc}}$ and $(1-\mathscr{P}_{\text{acc}})$ are added to the appropriate off-diagonal and diagonal entries, respectively. Due to the nature of these moves, the C matrix always remains tridiagonal. Estimates for the transition probabilities then follow directly. In his simulations, Errington periodically updated the weighting function with these estimates, in long intervals similar to multicanonical updating. One important distinction, however, is that he did not re-zero transition matrix in between these updates. Instead, the matrix continued to be updated as if the weights were zero. That is, he calculated two $\mathscr{P}_{\text{acc}}$: one for the actual acceptance criterion which included the weights and one for the transition matrix data which did not. In this manner, all of the data in the transition matrix were consistent and able to be retained. See Chap. 3 for a graphical schematic of this procedure.

Errington performed individual transition matrix simulations for temperatures in the range $T = 0.6 - 0.9$. Each simulation was run for several hundred million steps, with weight updates every million. For the lowest temperatures, the transition matrix algorithm was implemented in a parallel scheme, with each processor assigned to a window of particle numbers and occasional configuration swaps between adjacent windows. The final weights were calculated to very high accuracy, and were used directly to reweigh state probabilities to alternate chemical potentials at the simulation temperatures. The relevant equation is

$$\wp(U; \mu, T) \propto \exp\left[\beta(\mu - \mu_0)N + \eta(N)\right]. \tag{10.37}$$

By tuning μ to obtain a bimodal particle number distribution, it was possible to locate the point of coexistence between saturated thin and thick films on the adsorbing surface. The prewetting coexistence curve was then obtained through the application of this procedure at each simulation temperature.

A significant prediction of this study was the estimated wetting temperature, essentially the low-temperature bound for the prewetting phase coexistence. The wetting temperature is defined as the point at which the saturated bulk liquid–vapor (μ_{b}) and prewetting (μ_{pw}) chemical potentials become equal. Both chemical potentials were determined by reweighting η to obtain a bimodal distribution in $\wp(N)$ in the relevant density range. Through a modest extrapolation in temperature, Errington predicted that $\mu_{\text{b}} = \mu_{\text{pw}}$ occurred at $T = 0.598(5)$. The use of transition matrix estimators in this study was essential to the precise determination of this temperature. The high accuracy of the transition matrix predictions enabled simulations at lower temperatures than previous work, which yielded results much closer to the actual wetting temperature.

10.4.3 Isomerization Transition in $(\text{NaF})_4$ using the Wang–Landau Algorithm

One of the most familiar scenarios for studying the free energy along predetermined coordinates is that of chemical reactions. Calvo has used the Wang–Landau algorithm to elicit the entropy along the reaction path of $(\text{NaF})_4$ isomerization between cubic and octagonal ring conformations [79]. He employed a system of eight Na^+ and F^- ions interacting via Coulombic and Born–Mayer repulsive forces whose parameters are derived from ab initio calculations. Through trial and error, Calvo selected the principal moment of inertia as the appropriate reaction coordinate, given by

$$\lambda(\mathbf{q}) = \sum_i |\mathbf{q}_i - \mathbf{q}_G|^2, \tag{10.38}$$

where $\mathbf{q}_G$ is the coordinate of the center of mass of the system. The system was studied in the microcanonical ensemble and a bias was introduced to create a flat-histogram in the order parameter λ. Following the equations of Chaps. 1 and 3, the overall partition function Ω_{tot} can be written as

$$\Omega_{\text{tot}}(N, V, E) \propto \int_{V^N} [E - U(\mathbf{q})]^{3N/2-1} \, \theta \, [E - U(\mathbf{q})] \, \mathrm{d}\mathbf{q}$$

$$= \int_\lambda \left\{ \int_{V^N} [E - U(\mathbf{q})]^{3N/2-1} \, \theta \, [E - U(\mathbf{q})] \, \delta[\hat{\lambda}(\mathbf{q}) - \lambda] \mathrm{d}\mathbf{q} \right\} \mathrm{d}\lambda$$

$$= \int \Omega_{\text{tot},\lambda}(N, V, E, \lambda) \mathrm{d}\lambda, \tag{10.39}$$

where the N- and V-dependent factor in the first line has been omitted for clarity and the 'coordinate' density of states defined in the last line, $\Omega_{\text{tot},\lambda}$, includes the kinetic contribution from the microcanonical ensemble. The entropic analog of the coordinate microcanonical partition function is denoted by $\mathscr{S}_\lambda = \ln \Omega_{\text{tot},\lambda}$. The final line in (10.39) indicates that the probability of a particular value of λ for a given E is proportional to $\exp[\mathscr{S}(N, V, E, \lambda)]$. In order to enforce a flat histogram, therefore, the inverse factor must be inserted in the microstate sampling scheme.

The microscopic probabilities are microcanonical (see Chap. 1) with an additional biasing factor

$$\wp(\mathbf{q}; E) \propto [E - U(\mathbf{q})]^{3N/2-1} \, \theta \, [E - U(\mathbf{q})] \exp[-\mathscr{S}_\lambda(\lambda)], \qquad (10.40)$$

where the dependence of $\mathscr{S}_\lambda$ on N, V and E has been made implicit. Since the correspondence is with the microcanonical ensemble, the total energy E must be specified in the simulation, rather than the temperature. The acceptance criterion for single particle displacement moves based on (10.40) is

$$\mathscr{P}_{\text{acc}}(o \to n) = \min \left[1, \left(\frac{E - U_{\text{n}}}{E - U_{\text{o}}} \right)^{3N/2-1} \theta(E - U_{\text{n}}) \exp[\mathscr{S}_\lambda(\lambda_{\text{o}}) - \mathscr{S}_\lambda(\lambda_{\text{n}})] \right].$$
$$(10.41)$$

Employing this sampling scheme, Calvo used the Wang–Landau technique to determine $\mathscr{S}_\lambda$ over a range spanning values which correspond to octagonal and cubic configurations. He then repeated the calculations for several values of the total energy. The entropy was discretized into 1,000 bins, and the calculations required approximately 24 million steps for each run. Once determined, the entropy gave directly the microcanonical probabilities for λ

$$\wp(\lambda; E) \propto \exp[\mathscr{S}_\lambda(\lambda)], \qquad (10.42)$$

where $\mathscr{S}_\lambda$ implicitly corresponds to the Wang–Landau results from the run at energy E. Calvo showed that the favored isomer changed with total energy, and that fluctuations in λ were greater in the cubic configuration than in the octagonal one. It was also noted that the converged $\mathscr{S}_\lambda$ could be used in principal to determine transition rates. This is analogous to examining the free energy profile along a reaction coordinate, and using its value at the transition state to predict kinetic rate coefficients.

10.4.4 Other Applications

Numerous other off-lattice flat-histogram studies exist, and for the reader's reference, we briefly mention a handful of examples here. The multicanonical algorithm has often been used to study phase coexistence; Wilding has used the technique to investigate model fluids [80] and provides an instructive overview [81]. First-order transitions in confined mixtures [82] and polymer chains [83–87] have also been examined. Other applications have included performance characterization for the equilibration of glasses [88], ab initio study of water dimers [89], and the entropy of hard-sphere crystals [90]. Faller et al. [91] have also developed a replica-exchange multicanonical algorithm for the study of fluid phase transitions. The multicanonical algorithm also has a substantial history in the simulation of proteins, and has been reviewed in [92–94]. For example, the folding pathway of the beta-hairpin structure has been examined by Dinner et al. [95]. Higo et al. [96] used multicanonical sampling along various order parameters to enhance configurational sampling, and

Sayano et al. [97] generated the potential of mean force for interactions between amino acids and base pairs. Okamoto and coworkers [98, 99] have also developed replica-exchange multicanonical methods for proteins. Numerous additional biological examples can be found among the references within [94].

The WL algorithm has also been applied to a wealth of problems. It has been used to determine the bead-size dependence of the entropy in freely jointed hard-sphere chains [100], to measure the chemical potential in dense or low-temperature fluids [101], to evaluate the potential of mean force of a confined particle in a liquid crystal [102], and to study a binary Lennard-Jones glass at low temperatures [103]. The method has also been used to determine surface tension in fluids [104] and, somewhat related, the condensation behavior of a supersaturated vapor into liquid droplets [105]. The Wang–Landau approach has even been used to solve general constrained optimization problems [106]. In biology, the Wang–Landau algorithm has been applied to several continuum protein models. Rathore et al. [107, 108] studied the folding behavior of short united-atom peptide sequences, and estimated a folding temperature through subsequent reweighing of the calculated entropy. They have also adapted the method to study mechanically induced unfolding of proteins, in which the free energy along a 'stretching' coordinate is determined [109]. Similarly, Liang has used a derivative of the WL method to locate the ground states of two- and three-dimensional 'AB' models of proteins [110].

Transition matrix estimators have received less attention than the multicanonical and Wang–Landau methods, but have been applied to a small collection of informative examples. Smith and Bruce [111, 112] applied the transition probability approach to the determination of solid–solid phase coexistence in a square-well model of colloids. Errington and coworkers [113, 114] have also used the method to determine liquid–vapor and solid–liquid [115] equilibria in the Lennard-Jones system. Transition matrices have also been used to generate high-quality data for the evaluation of surface tension [114, 116] and for the estimation of order parameter weights in 'phase-switch' simulations [117].

10.5 Summary: Comparison of Methods

A natural question that arises at this point is that of selection among the many possible methods that are available to obtain the free energy and phase behavior of a system of interest. In most cases there are tradeoffs between accuracy and precision on one hand and programming and computational effort required. In order to facilitate the selection a method appropriate to a given problem, we present a brief summary of capabilities and limitations below for each of the approaches we have described in this chapter. In reviewing the differences between the algorithms, we find it beneficial to evaluate each along five principal characteristics: generality (ability to determine various free energies and applicability to a wide class of systems), calculation type (whether yields free energy values directly or indirectly, e.g., via phase behavior), extent of implementation (programming effort required and whether the method is standalone or can serve as a supplementary

measurement), computational efficiency (overhead required), and accuracy (the extent to which desired quantities can be determined to high precision for a given simulation time):

– **Widom test-particle method**. Provides the chemical potential in various ensembles. Relatively easy to implement and can be used as an additional measurement in standard MC ensembles (and also MD). Computational overhead is small. Yields good accuracy in simple systems, although less reliable in very dense or complex systems (i.e., chain molecules).

– **Gibbs ensemble**. Good for obtaining a few points for subcritical phase coexistence between phases of moderate densities; does not provide free energies directly. Primarily used to study fluid (disordered) phases. Is a standalone approach, and requires modest programming and computational effort to set up and equilibrate the multiple simulation boxes. Provides accurate coexistence points at intermediate temperatures below the critical point but with sufficient thermal mobility to equilibrate.

– **Gibbs–Duhem integration**. Provides a way to trace a coexistence curve using results from a series of individual MC simulations; does not provide free energies directly. Is the only method widely used to determine solid–fluid coexistence, and is also good for parametric studies of the effect of variation in potential parameters. Implementation is straightforward, relying directly on conventional ensemble algorithms; several variants exist to accommodate different MC moves. Can yield reasonable accuracy if: (a) an initial point on the coexistence curve is known, and (b) enough simulations are performed to reduce propagation of error.

– **Histogram-reweighting methods**. Used to extract and combine underlying thermodynamic estimates from multiple conventional simulations, which can then be reweighted to arbitrary state conditions. Requires small programming and computational effort in the tabulation of histograms; somewhat more effort is necessary for post-run data processing. Number of simulations required depends on the calculation; for liquid–vapor coexistence, a small number is usually sufficient, although some effort is required in selecting their conditions. More accurate, especially near critical points.

– **Multicanonical approach**. Allows recursive estimation of thermodynamic potentials based on histograms, simultaneously facilitating flat-histogram-type sampling of the system. Very general standalone simulation method, but implementation requires the introduction of weights into the sampling scheme, as well as a schedule for their update based on measured histograms. Single runs can require significant computational time to refine the weights and collect final data. Mostly useful for aiding equilibration where conventional ensembles have difficulty; can also be implemented in parallel. Values for entropies or free energies themselves are typically not highly accurate, although calculated weights and final histogram data can be accurately reweighted.

– **Wang–Landau algorithm**. Permits direct determination of entropies and free energies using a flat-histogram ensemble outright and performing dynamic update of the state weights. Useful for mapping out broad regions of thermodynamic space (as opposed to specific loci) and for extended ensemble or

reaction-coordinate studies. Highly general approach but implementation is more intricate, requiring the introduction of weights, their monitoring, and an update scheme. Computational expense can be significant if large ranges of thermodynamic conditions are to be explored; can be implemented in parallel to improve real-time convergence. Also requires the determination of several parameters affecting the rate of calculation and its final accuracy. For common implementations, yields modest to high accuracy and can be useful for rapidly generating rough estimates of free energies.

– **Transition matrix estimators**. Provide a method for calculating arbitrary thermodynamic potentials based on the statistics of transitions between macrostates. Can be used as an additional free energy measurement in conventional simulations, or as a method for updating weights in a multicanonical or Wang–Landau type approach. Implementation requires additional bookkeeping of move transitions, which can become memory-intensive for some applications. Can be parallelized easily. Generally speaking, are an extremely efficient use of simulation data (do not discard information) and provide highly accurate free energy estimates.

References

1. Frenkel, D.; Smit, B., *Understanding Molecular Simulation*, (2nd edition) Academic: San Diego, 2002
2. Binder, K., Applications of Monte Carlo methods to statistical physics, *Rep. Prog. Phys.* **1997**, *60*, 487–559
3. Deem, M. W., Recent contributions of statistical mechanics in chemical engineering, *AIChE J.* **1998**, *44*, 2569–2596
4. Siepmann, J. I., Monte Carlo methods for simulating phase equilibria of complex fluids, *Monte Carlo Methods Chem. Phys.* **1999**, *105*, 443–460
5. de Pablo, J. J.; Yan, Q.; Escobedo, F. A., Simulation of phase transitions in fluids, *Ann. Rev. Phys. Chem.* **1999**, *50*, 377–412
6. Panagiotopoulos, A. Z., Monte Carlo methods for phase equilibria of fluids, *J. Phys.: Condens. Matt.* **2000**, *12*, R25–R52
7. Widom, B., Some topics in the theory of fluids, *J. Chem. Phys.* **1963**, *39*, 2808-2817
8. Möller, D.; Fischer, J., Vapour liquid equilibrium of a pure fluid from test particle method in combination with npt molecular dynamics simulations, *Mol. Phys.* **1990**, *69*, 463–473
9. Lotfi, A.; Vrabec, J.; Fischer, J., Vapour liquid equilibria of the Lennard-Jones fluid from the NPT plus test particle method, *Mol. Phys.* **1992**, *76*, 1319–1333
10. Kofke, D. A; Cummings, P. T., Quantitative comparison and optimization of methods for evaluating the chemical potential by molecular simulation, *Mol. Phys.* **1997**, *92*, 973–996
11. Boda, D.; Liszi, J.; Szalai, I., An extension of the NPT plus test particle method for the determination of the vapour–liquid equilibria of pure fluids, *Chem. Phys. Lett.* **1995**, *235*, 140–145
12. Vrabec, J.; Fischer, J., Vapour liquid equilibria of mixtures from the NPT plus test particle method, *Mol. Phys.* **1995**, *85*, 781–792

13. Vrabec, J.; Lotfi, A.; Fischer, J., Vapour liquid equilibria of Lennard-Jones model mixtures from the NPT plus test particle method, *Fluid Phase Equil.* **1995**, *112*, 173–197

14. Spyriouni, T.; Economou, I. G.; Theodorou, D. N., Vapour liquid equilibria of Lennard-Jones model mixtures from the NPT plus test particle method, *Phys. Rev. Lett.* **1998**, *80*, 4466

15. Spyriouni, T.; Economou, I. G.; Theodorou, D. N., Phase equilibria of mixtures containing chain molecules predicted through a novel simulation scheme, *Macromolecules* **1998**, *31*, 1430–1431

16. Kumar, S. K.; Szleifer, I.; Panagiotopoulos, A. Z., Molecular simulation of the pure n-hexadecane vapor–liquid equilibria at elevated temperature, *Phys. Rev. Lett.* **1991**, *66*, 2935

17. Panagiotopoulos, A. Z., Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.* **1987**, *61*, 813–826

18. Panagiotopoulos, A. Z.; Quirke, N.; Stapleton, M.; Tildesley, D. J., Phase equilibria by simulation in the Gibbs ensemble: alternative derivation, generalization and application to mixture and membrane equilibria, *Mol. Phys.* **1988**, *63*, 527–545

19. Smit, B.; Smedt, P. De; Frenkel, D., Computer simulations in the Gibbs ensemble, *Mol. Phys.* **1989**, *68*, 931–950

20. Smit, B.; Frenkel, D., Calculation of the chemical potential in the Gibbs ensemble, *Mol. Phys.* **1989**, *68*, 951–958

21. Lopes, J. N. C.; Tildesley, D. J., Multiphase equilibria using the Gibbs ensemble Monte Carlo method, *Mol. Phys.* **1997**, *92*, 187–196

22. Kristof, T.; Liszi, J., Application of a new Gibbs ensemble Monte Carlo method to site–site interaction model fluids, *Mol. Phys.* **1997**, *90*, 1031–1034

23. Kristof, T.; Liszi, J., Application of a new Gibbs ensemble Monte Carlo method to site–site interaction model fluids, *Mol. Phys.* **1998**, *94*, 519

24. Baranyai, A.; Cummings, P. T., On the molecular dynamics algorithm for Gibbs ensemble simulation, *Mol. Simul.* **1996**, *17*, 21–25

25. Kotelyanskii, M. J.; Hentschke, R., Gibbs-ensemble molecular dynamics: liquid–gas equlibria for Lennard-Jones spheres and n-hexane, *Mol. Simul.* **1996**, *17*, 95–112

26. Hentschke, R.; Bast, T.; Aydt, E.; Kotelyanskii, M., Gibbs-ensemble molecular dynamics: a new method for simulations involving particle exchange, *J. Mol. Model.* **1996**, *2*, 319–326

27. Bruce, A. D., Finite-size critical behavior in the Gibbs ensemble, *Phys. Rev. E* **1997**, *55*, 2315–2320

28. Mon, K. K.; Binder, K., Finite size effects for the simulation of phase coexistence in the Gibbs ensemble near the critical point, *J. Chem. Phys.* **1992**, *96*, 6989–6995

29. Panagiotopoulos, A. Z., Exact calculations of fluid-phase equilibria by Monte Carlo simulation in a new statistical ensemble, *Int. J. Thermophys.* **1989**, *10*, 447–457

30. Laso, M.; de Pablo, J. J.; Suter, U. W., Simulation of phase equilibria for chain molecules, *J. Chem. Phys.* **1992**, *97*, 2817–2819

31. Mooij, G. C. A. M.; Frenkel, D.; Smit, B., Direct simulation of phase equilibria of chain molecules, *J. Phys.: Condens. Matter* **1992**, *4*, L255–L259

32. Kofke, D. A., Gibbs–Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation, *Mol. Phys.* **1993**, *78*, 1331–1336

33. Kofke, D. A., Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line, *J. Chem. Phys.* **1993**, *98*, 4149–4162

34. Mehta, M.; Kofke, D. A., Coexistence diagrams of mixtures by molecular simulation, *Chem. Eng. Sci.* **1994**, *49*, 2633–2645

35. Agrawal, R.; Mehta, M.; Kofke, D. A., Efficient evaluation of three-phase coexistence lines, *Int. J. Thermophys.* **1994**, *15*, 1073–1083

36. Kofke, D. A., Semigrand canonical Monte Carlo simulation; integration along coexistence lines, *Adv. Chem. Phys.* **1999**, *105*, 405–442

37. Escobedo, F. A.; de Pablo, J. J., Monte Carlo simulation of athermal mesogenic chains: pure systems, mixtures, and constrained environments, *J. Chem. Phys.* **1997**, *106*, 9858–9868

38. Escobedo, F. A.; de Pablo, J. J., Pseudo-ensemble simulations and Gibbs–Duhem integrations for polymers, *J. Chem. Phys.* **1997**, *106*, 2911–2923

39. Escobedo, F. A.; de Pablo, J. J., Gibbs–Duhem integrations in lattice systems, *Europhys. Lett.* **1997**, *40*, 111–116

40. Agrawal, R.; Kofke, D. A., Solid–fluid coexistence for inverse-power potentials, *Phys. Rev. Lett.* **1995**, *74*, 122–125

41. Bolhuis, P.; Frenkel, D., Tracing the phase-boundaries of hard spherocylinders, *J. Chem. Phys.* **1997**, *106*, 666–687

42. Meijer, E. J.; Azhar, F. El, Novel procedure to determine coexistence lines by computer simulation. application to hard-core Yukawa model for charge-stabilized colloids, *J. Chem. Phys.* **1997**, *106*, 4678–4683

43. Mehta, M.; Kofke, D. A., Molecular simulation in a pseudo grand-canonical ensemble, *Mol. Phys.* **1995**, *86*, 139–147

44. Camp, P. J.; Allen, M. P., Phase coexistence in a pseudo-Gibbs ensemble, *Mol. Phys.* **1996**, *88*, 1459–1469

45. Escobedo, F. A., Novel pseudoensembles for simulation of multicomponent phase equilibria, *J. Chem. Phys.* **1998**, *108*, 8761–8772

46. Escobedo, F. A., Tracing coexistence lines in multicomponent fluid mixtures by molecular simulation, *J. Chem. Phys.* **1999**, *110*, 11999–12010

47. Ferrenberg, A. M.; Swendsen, R. H., Optimized Monte Carlo data analysis, *Phys. Rev. Lett.* **1989**, *63*, 1195–1198

48. Panagiotopoulos, A. Z.; Wong, V.; Floriano, M. A., Phase equilibria of lattice polymers from histogram reweighting Monte Carlo simulations, *Macromolecules* **1998**, *31*, 912–918

49. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free-energy estimation – umbrella sampling, *J. Comp. Phys.* **1977**, *23*, 187–199

50. Berg, B. A.; Neuhaus, T., Multicanonical ensemble: a new approach to simulate first-order phase transitions, *Phys. Rev. Lett.* **1992**, *68*, 9–12

51. Wilding, N. B., Critical-point and coexistence-curve properties of the Lennard-Jones fluid: a finite-size scaling study, *Phys. Rev. E* **1995**, *52*, 602–611

52. Wilding, N. B., Critical end point behavior in a binary fluid mixture, *Phys. Rev. E* **1997**, *55*, 6624–6631

53. Wilding, N. B.; Schmid, F.; Nielaba, P., Liquid–vapor phase behavior of a symmetrical binary fluid mixture, *Phys. Rev. E* **1998**, *58*, 2201–2212

54. Potoff, J. J.; Panagiotopoulos, A. Z., Critical point and phase behavior of the pure fluid and a Lennard-Jones mixture, *J. Chem. Phys.* **1998**, *109*, 10914–10920

55. Wilding, N. B., Simulation studies of fluid critical behaviour, *J. Phys.: Condens. Matter* **1997**, *9*, 585–612

56. Ferdinand, A. E.; Fisher, M. E., Bounded and inhomogeneous Ising models: I. specific-heat anomaly of a finite lattice, *Phys. Rev.* **1969**, *185*, 832–846

57. Fisher, M. E.; Barber, M. N., Scaling theory for finite-size effects in the critical region, *Phys. Rev. Lett.* **1972**, *28*, 1516–1519

58. Privman, V., (Ed.), *Finite Size Scaling and Numerical Simulation of Statistical Mechanical Systems*, World Scientific, Singapore, 1990

59. Bruce, A. D.; Wilding, N. B., Scaling fields and universality of the liquid–gas critical point, *Phys. Rev. Lett.* **1992**, *68*, 193–196

60. Wilding, N. B.; Bruce, A. D., Density fluctuations and field mixing in the critical fluid, *J. Phys.: Condens. Matter* **1992**, *4*, 3087–3108

61. Orkoulas, G.; Panagiotopoulos, A. Z., Phase behavior of the restricted primitive model and square-well fluids from Monte Carlo simulations in the grand canonical ensemble, *J. Chem. Phys.* **1999**, *110*, 1581–1590

62. Sengers, J. V.; Levelt-Sengers, J. N. H., Thermodynamic behavior of fluids near the critical point, *Ann. Rev. Phys. Chem.* **1986**, *37*, 187–222

63. Liu, A. J.; Fisher, M. E., The three-dimensional ising model revisited numerically, *A. Physica* **1989**, *156*, 35–76

64. Valleau, J. P., Thermodynamic scaling methods in Monte Carlo and their application to phase equilibria, *Monte Carlo Methods Chem. Phys.* **1999**, *105*, 369–404

65. Kiyohara, K.; Spyriouni, T.; Gubbins, K. E.; Panagiotopoulos, A. Z., Thermodynamic scaling Gibbs ensemble Monte Carlo: a new method for determination of phase coexistence properties of fluid, *Mol. Phys.* **1996**, *89*, 965–974.

66. Errington, J. R.; Panagiotopoulos, A. Z., Phase equilibria of the modified Buckingham exponential-6 potential from Hamiltonian scaling grand canonical Monte Carlo, *J. Chem. Phys.* **1998**, *109*, 1093–1100

67. Berg, B. A.; Celik, T., New approach to spin-glass simulations, *Phys. Rev. Lett.* **1992**, *69*, 2292–2295

68. Lee, J., New Monte Carlo algorithm: entropic sampling, *Phys. Rev. Lett.* **1993**, *71*, 211–214

69. Smith, G. R.; Bruce, A. D., A study of the multicanonical Monte Carlo method, *J. Phys. A* **1995**, *28*, 6623–6643

70. Fitzgerald, M.; Picard, R. R.; Silver, R. N., Monte Carlo transition dynamics and variance reduction, *J. Stat. Phys.* **1999**, *98*, 321–345

71. Wang, F.; Landau, D. P., Efficient, multiple-range random walk algorithm to calculate the density of states, *Phys. Rev. Lett.* **2001**, *86*, 2050–2053

72. Wang, F.; Landau, D. P., Determining the density of states for classical statistical models: a random walk algorithm to produce a flat histogram, *Phys. Rev. E* **2001**, *64*, 056101

73. Wang, J. S.; Swendsen, R. H., Transition matrix Monte Carlo method, *J. Stat. Phys.* **2001**, *106*, 245–285

74. Yan, Q.; Faller, R.; de Pablo, J. J., Density-of-states Monte Carlo method for simulation of fluids, *J. Chem. Phys.* **2002**, *116*, 8745–8749

75. Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z., Generalization of the Wang–Landau method for off-lattice simulations, *Phys. Rev. E* **2002**, *90*, 056703

76. Lotfi, A.; Vrabec, J.; Fischer, J., Vapour liquid equilibria of the Lennard-Jones fluid from the NPT plus test particle method, *Mol. Phys.* **1992**, *76*, 1319–1333

77. Errington, J. R., Prewetting transitions for a model argon on solid carbon dioxide system, *Langmuir* **2004**, *20*, 3798–3804

78. Ebner, C.; Saam, W. F., New phase-transition phenomena in thin argon films, *Phys. Rev. Lett* **1977**, *38*, 1486–1489

79. Calvo, F., Sampling along reaction coordinates with the Wang–Landau method, *Mol. Phys.* **2002**, *100*, 3421–3427

80. Wilding, N. B., Critical-point and coexistence-curve properties of the Lennard-Jones fluid: a finite-size scaling study, *Phys. Rev. E* **1995**, *52*, 602–611

81. Wilding, N. B., Computer simulation of fluid phase transitions, *Am. J. Phys* **2001**, *69*, 1147–1155

82. Gelb, L. D.; Gubbins, K. E., Studies of binary liquid mixtures in cylindrical pores: phase separation, wetting and finite-size effects from Monte Carlo simulations, *Physica A* **1997**, *244*, 112–123

83. Noguchi, H.; Yoshikawa, K., First-order phase transition in a stiff polymer chain, *Chem. Phys. Lett.* **1997**, *278*, 184–188

84. Noguchi, H.; Yoshikawa, K., Morphological variation in a collapsed single homopolymer chain, *J. Chem. Phys.* **1998**, *109*, 5070–5077

85. Liang, H. J.; Chen, H. N., First-order transition of a homopolymer chain with Lennard-Jones potential, *J. Chem. Phys.* **2000**, *113*, 4469–4471

86. Wang, Y. U.; Chen, H. N.; Liang, H. J., Monte Carlo simulation on thermodynamic properties of a heteropolymer chain, *J. Chem. Phys.* **2001**, *115*, 3951–3956

87. Huang, L.; He, X. H.; Wang, Y. Y.; Chen, H. N.; Liang, H. J., Phase transitions of short chains at ultralow temperature, *J. Chem. Phys.* **2003**, *119*, 2432–2438

88. Bhattacharya, K. K.; Sethna, J. P., Multicanonical methods, molecular dynamics, and Monte Carlo methods: comparison for Lennard-Jones glasses, *Phys. Rev. E* **1998**, *57*, 2553–2562

89. Bandyopadhyay, P.; Ten-No, S.; Iwata, S., Ab initio Monte Carlo simulation using multicanonical algorithm: temperature dependence of the average structure of water dimer, *Mol. Phys.* **1999**, *96*, 349–358

90. Mau, S. C.; Huse, D. A., Stacking entropy of hard-sphere crystals, *Phys. Rev. E* **1999**, *59*, 4396–4401

91. Faller, R.; Yan, Q. L.; de Pablo, J. J., Multicanonical parallel tempering, *J. Chem. Phys.* **2002**, *116*, 5419–5423

92. Hansmann, U. H. E.; Okamoto, Y., New Monte Carlo algorithms for protein folding, *Curr. Opin. Struct. Biol.* **1999**, *9*, 177–183

93. Hansmann, U. H. E., Computer simulation of biological macromolecules in generalized ensembles, *Int. J. Mod. Phys. C* **1999**, *10*, 1521–1530

94. Mitsutake, A.; Sugita, Y.; Okamoto, Y., Generalized-ensemble algorithms for molecular simulations of biopolymers, *Biopolymers* **2001**, *60*, 96–123

95. Dinner, A. R.; Lazaridis, T.; Karplus, M., Understanding beta-hairpin formation, *Proc. Natl Acad. Sci. USA* **1999**, *96*, 9068–9073

96. Higo, J.; Nakajima, N.; Shirai, H.; Kidera, A.; Nakamura, H., Two-component multi-canonical Monte Carlo method for effective conformation sampling, *J. Comput. Chem.* **1997**, *18*, 2086–2092

97. Sayano, K.; Kono, H.; Gromiha, M. M.; Sarai, A., Multicanonical Monte Carlo calculation of the free-energy map of the base–amino acid interaction, *J. Comput. Chem.* **2000**, *21*, 954–962

98. Sugita, Y.; Kitao, A.; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations, *J. Chem. Phys.* **2000**, *113*, 6042–6051

99. Mitsutake, A.; Okamoto, Y., Replica-exchange simulated tempering method for simulations of frustrated systems, *Chem. Phys. Lett.* **2000**, *332*, 131–138

100. Vorontsov-Velyaminov, P. N.; Yolkov, N. A.; Yurchenko, A. A., Entropic sampling of simple polymer models within Wang–Landau algorithm, *J. Phys. A* **2004**, *37*, 1573–1588

101. Vlugt, T. J. H., Measurement of chemical potentials of systems with strong excluded volume interactions by computing the density of states, *Mol. Phys.* **2002**, *100*, 2763–2771

102. Kim, E. B.; Faller, R.; Yan, Q.; Abbot, N. L.; de Pablo, J. J., Potential of mean force between a spherical particle suspended in a nematic liquid crystal and a substrate, *J. Chem. Phys.* **2002**, *117*, 7781–7787

103. Faller, R.; de Pablo, J. J., Density of states of a binary Lennard-Jones glass, *J. Chem. Phys.* **2003**, *119*, 4405–4408

104. Jain, T. S.; de Pablo, J. J., Calculation of interfacial tension from density of states, *J. Chem. Phys.* **2003**, *118*, 4226–4229

105. MacDowell, L. G.; Virnau, P.; Müller, M.; Binder, K., The evaporation/condensation transition of liquid droplets, *J. Chem. Phys.* **2004**, *120*, 5293–5308

106. de Menezes, M. A.; Lima, A. R., Using entropy-based methods to study general constrained parameter optimization problems, *Physica A* **2003**, *323*, 428–434

107. Rathore, N.; IV, T. A. Knotts; de Pablo, J. J., Density of states simulations of proteins, *J. Chem. Phys.* **2002**, *118*, 4285–4290

108. Rathore, N.; IV, T. A. Knotts; de Pablo, J. J., Configurational temperature density of states simulations of proteins, *Biophys. J.* **2003**, *85*, 3963

109. Rathore, N.; Yan, Q.; de Pablo, J. J., Molecular simulation of the reversible mechanical unfolding of proteins, *J. Chem. Phys.* **2004**, *120*, 5781–5788

110. Liang, F., Annealing contour Monte Carlo algorithm for structure optimization in an off-lattice protein model, *J. Chem. Phys.* **2004**, *120*, 6756–6763

111. Smith, G. R.; Bruce, A. D., Multicanonical Monte Carlo study of a structural phase transition, *Europhys. Lett.* **1996**, *34*, 91–96

112. Smith, G. R.; Bruce, A. D., Multicanonical Monte Carlo study of solid–solid phase coexistence in a model colloid, *Phys. Rev. E* **1996**, *53*, 6530–6543

113. Errington, J. R., Direct calculation of liquid–vapor phase equilibria from transition matrix Monte Carlo simulation, *J. Chem. Phys.* **2003**, *118*, 9915–9925

114. Singh, J. K.; Kofke, D. A.; Errington, J. R., Surface tension and vapor–liquid phase coexistence of the square-well fluid, *J. Chem. Phys.* **2003**, *119*, 3405–3412

115. Errington, J. R., Solid–liquid phase coexistence of the Lennard-Jones system through phase-switch Monte Carlo simulation, *J. Chem. Phys.* **2004**, *120*, 3130–3141

116. Errington, J. R., Evaluating surface tension using grand-canonical transition-matrix Monte Carlo simulation and finite-size scaling, *Phys. Rev. E* **2003**, *67*, 012102

117. Bruce, A. D.; Jackson, A. N.; Ackland, G. J.; Wilding, N. B., Lattice-switch Monte Carlo method, *Phys. Rev. E* **2000**, *61*, 906–919

11

Quantum Contributions to Free Energy Changes in Fluids

Thomas L. Beck

11.1 Introduction

For the majority of problems in molecular solutions, a classical description is sufficient. In nearly all of the applications discussed so far in this book, that classical limit was assumed in computer simulation calculations of free energy changes. A principal focus is on biological macromolecules such as proteins, and, neglecting any solvent effects for now, those large molecules do not display appreciable quantum effects on their conformational states. Of course internal modes such as vibrations are highly quantized, but those modes can typically be treated with rigid constraints or simple classical models due to their small amplitudes of motion. Whatever small quantum effects are present are likely incorporated during empirical fitting of the classical potentials. When should we worry about quantum effects in free energy calculations?

Addressing this question is the subject of this chapter. The emphasis of the chapter is on physical arguments and relatively simple derivations to give a start to somebody interested in checking whether quantum effects should be considered and to provide a set of tools to compute those effects if initial modeling suggests that they are important. For two reasons, the approach taken is that of the path integral formulation of quantum mechanics [1, 2]. First, this method gives an intuitive physical picture that adapts naturally to the classical limit: variational approaches and resulting approximate theories emerge quite easily. Second, the path integral formulation of equilibrium statistical mechanics is well suited for simulation. The analogy to classical polymer simulations is direct [3], meaning that the quantum simulation merely requires extending the number of degrees of freedom and employing either Monte Carlo or molecular dynamics procedures to generate the equilibrium distributions. For this reason, the majority of finite-temperature condensed-phase quantum simulations of large systems have utilized path integral methods due to the favorable scaling compared with other quantum methods.

Let us first consider physical systems, in which quantum effects might be important, in order of decreasing effect. The prototype quantum liquid is liquid helium with its well-known exotic properties. This liquid requires a full quantum

treatment and has been examined using path integral simulation [4]. At very low temperatures, particle exchange effects are important in addition to quantum effects due to wave packet spreading.[1] A second highly quantum example is the solvated electron, which has received a lot of attention from the simulation community [5–7]. Aside from the interaction of the single electron with atomic cores in the solution modeled with a pseudopotential, exchange effects are not important. Liquid hydrogen is another example with significant quantum effects [8–10]. Liquids which exhibit quantum effects of smaller magnitude are methane [11, 12], neon [9, 12–16], and even nitrogen [17]. These moderately quantum systems have been treated successfully with path integral methods and approximations which will be derived in this chapter. Path integral methods have also been applied to compute numerically exact gas-phase partition functions for small molecules [18, 19], to predict cluster free energies of formation [20], and to examine quantum effects on melting temperatures of small clusters [20]. So there has been some work on quantum free energy calculations [21–24], but much less than for the classical counterpart.

Most of the above systems are prototype 'simple' liquids which require handling of quantum effects. They are foundational in establishing computational methods and approximations, yet they are not of great interest to biological modelers. But liquid water exhibits several interesting quantum effects which are likely to be of some importance for biological molecules in solution. The temperature of maximum density of H_2O is $4°C$, while for D_2O it is $7°C$ higher. The heat of vaporization is 9% larger for D_2O under normal conditions [25]. The solubilities of simple solutes [26] and biomolecules [27] are measurably different in the two liquids, and changes from one solvent to the other can affect protein stability [28]. The triple temperature of H_2O is $4\,K$ lower than that for D_2O, but the critical temperature for H_2O is $3\,K$ higher. And the static dielectric constant for H_2O is higher than that for D_2O by about $\Delta\epsilon \approx 0.35$ [29]. Most of these effects are consistent with a picture of H_2O as a slightly more disordered liquid, which at first glance could be rationalized as a small-mass quantum effect [29]. Since protein conformational states can be quite sensitive to the solvent environment, these effects are large enough to stimulate investigation. Quantum effects can also be expected to contribute significantly to acid–base reactions in water; the structure and dynamics of the proton in water is a subject of intensive current research [30]. In addition, quantum effects on reaction rates in enzymes have received some attention [31–33]; those models typically employ transition state theory (TST) which requires the activation free energy as input. An added complication in studies of enzyme proton transfer reactions is the chemical rearrangements which occur during proton motion. The redistribution of the electron density makes treatment with empirical force fields questionable. To address the electronic structure issues in these large systems, a great deal of work has been directed at treating the local environment of the active site with quantum mechanical methods, while the more distant regions are modeled with molecular mechanics (QM/MM methods) [33]. Thus these systems are doubly challenging: quantum mechanics is required both for

---

[1] Indicated by a thermal de Broglie wavelength which is a substantial fraction of the average intermolecular separation.

the local electronic structure changes and for the zero-point/tunneling effects on the proton nuclear dynamics.

It is interesting that quantum effects in water are not fully understood, and are the topic of current research. This is due to a subtle interplay between several factors influencing water structure and thermodynamics. Classical force fields have been empirically fit to reproduce a range of structural and thermodynamic experimental data [34]. So these force fields implicitly incorporate whatever quantum effects might be present. Also, some force fields do not include intramolecular flexibility or polarizability factors which may have some importance for the complicated condensed-phase water behavior. Yet when these 'improvements' are incorporated, they often lead to worse agreement with experiment. All of these factors lead to a morass of effects which are hard to disentangle. As a first step in water simulations, if the interest is in quantum effects computed using a realistic empirical potential, the force field fitting should be based on a quantum simulation [34].

An even more challenging step has been taken, namely to employ path integral methods for the nuclear degrees of freedom on top of an ab initio simulation [35].[2] As long as a density functional approximation capable of accurately describing water is used for the simulations [36–39], this approach should yield a good description of the liquid. The calculations are enormously time consuming though; the referenced simulation included 64 water molecules with a total simulation 'time' on the order of 15 ps. An issue with this approach is that the path integral description is asked to represent simultaneously both the large intramolecular and small intermolecular[3] quantum effects on the same footing, creating statistical sampling and perhaps convergence problems [39].

In this chapter, it will be argued that the relatively minor intermolecular quantum effects in a liquid like water are amenable to approximation, and the more quantized intramolecular degrees of freedom can be handled simply say with a normal-mode, harmonic-oscillator picture or even a rigid-molecule representation [39, 40]. Then calculations of free energy changes require only a classical simulation on a modified potential surface.[4] The focus will be on free energies, which provide stringent physical tests of force fields or ab initio methods. The excess chemical potential is the natural variable to study these free energy calculations since it reflects the local environment of a molecule in the solution. The theoretical description of the excess chemical potential stems from the potential distribution theorem (PDT) [29, 41, 42]. The PDT is discussed in detail in Chap. 9 of this book. We will see that this description provides a nice route to present the exact quantum results with path integral methods, and variational approximations fall out in a few easy steps. The Gibbs free energy is easy to construct once the chemical potentials are obtained: $G = \sum_\alpha \mu_\alpha n_\alpha$, where

---

[2] In an ab initio simulation, the electronic structure problem is solved for each nuclear configuration, and forces are computed using the Hellmann–Feynman theorem.

[3] All degrees of freedom except the intramolecular vibrations.

[4] That classical calculation may be a density functional theory (DFT) ab initio simulation. An ab initio treatment may be important to handle charge redistribution effects in the condensed phase.

α labels the chemical species in the solution, μ_α is the chemical potential for the species α, and n_α is the total number of α molecules.

11.2 Historical Backdrop

A first step toward quantum mechanical approximations for free energy calculations was made by Wigner and Kirkwood. A clear derivation of their method is given by Landau and Lifshitz [43]. They employ a plane-wave expansion to compute approximate canonical partition functions which then generate free energy models. The method produces an expansion of the free energy in powers of $\hbar$. Here we just quote several of the results of their derivation.

To second order in $\hbar$ the Helmholtz free energy is

$$A \approx A_{\mathrm{cl}} + \frac{\beta^2 \hbar^2}{24} \sum_i \frac{1}{m_i} \left\langle (\nabla_i U)^2 \right\rangle_{\mathrm{cl}}, \tag{11.1}$$

where A_{cl} is the classical free energy, m_i is the mass of particle i, ∇_i specifies differentiation with respect to the Cartesian coordinates of one of the particles, U is the total system potential energy, and the averaging utilizes classical mechanics. We will see below that several approximation steps are necessary to get to this point, but it follows from the formula that the quantum correction to the classical free energy is positive and related to the average squared forces on the particles. Landau and Lifshitz argue that, under conditions to which the above expansion applies, particle exchange effects appear at third order in $\hbar$.

By integrating out the coordinate dependence, they also obtain an approximate quantum-corrected formula for the momentum distribution which leads to the definition of an effective temperature. That is, the approximate distribution is still Gaussian in the momenta, but with an increased temperature for each particle

$$T_{i,\mathrm{eff}} \approx T + \frac{\beta^2 \hbar^2}{12 k_{\mathrm{B}} m_i} \left\langle (\nabla_i U)^2 \right\rangle_{\mathrm{cl}}, \tag{11.2}$$

which may differ for particles with different masses. This effective temperature suggests that quantum effects can possibly be modeled by simply increasing the temperature in a way related to each particle's mass and the mean square forces on it during its classical motion.

Alternatively, if the momenta are integrated out, an effective distribution of the coordinates is obtained. In this distribution the Boltzmann factor is sampled with the classical potential energy replaced by a quantum effective potential

$$U_{\mathrm{eff}} \approx U - \frac{\beta^2 \hbar^2}{24} \sum_i \frac{1}{m_i} (\nabla_i U)^2 + \frac{\beta \hbar^2}{12} \sum_i \frac{1}{m_i} \nabla_i^2 U. \tag{11.3}$$

This potential is similar to one derived by Stratt [44] using different techniques. It is problematic though for simulations of molecular fluids due to the middle term on the right side [12]. The square of the repulsive component of the force leads to nonphysical behavior in the effective potential for small separations.

11.3 The Potential Distribution Theorem

A statistical mechanical expression for the chemical potential was derived by Widom in 1963 [41]. His formula is called the potential distribution theorem. He employed the canonical ensemble, but the PDT can also be derived using a grand canonical approach [29] (see also Chap. 9). The latter derivation leads to a *building up* picture, in which the molecule of interest is successively solvated by increasing numbers of surrounding solvent molecules. The resulting expression for the excess chemical potential has the appearance of a partition function formula, but the information required for its evaluation involves only the local neighborhood of the solute. The locality feature distinguishes this picture from the traditional partition function developments. The standard tools of statistical mechanics can be utilized, and often derivations take a particulary simple form when cast in the language of the PDT.

The PDT expression for the chemical potential of a classical molecular species is

$$\beta\mu_\alpha = \ln\left[\frac{\rho_\alpha\Lambda_\alpha^3}{q_\alpha^{\text{int}}}\right] - \ln\left\langle\!\left\langle e^{-\beta\Delta U_\alpha}\right\rangle\!\right\rangle_0. \tag{11.4}$$

The chemical potential μ_α on the left is the full chemical potential including ideal and excess parts. In this chapter we will scale the chemical potentials by β and often refer to this unitless quantity as the chemical potential. $\beta\mu_\alpha$ yields the absolute activity. The first term on the right is the ideal-gas chemical potential, where ρ_α is the number density, Λ_α is the de Broglie wavelength, and q_α^{int} is the internal (neglecting translations) partition function for a single molecule without interactions with any other molecules.

The second term on the right is the excess chemical potential resulting from interactions of the distinguished molecule with the surrounding solution. The double brackets indicate averaging over two separate thermal motions, one for the molecule of interest and one for the rest of the system, but with no interactions between the two. Then we take the average of the Boltzmann factor of the interaction energy with the molecule superposed on the solution. This is an exact expression for the excess chemical potential assuming a classical description is valid.

At equilibrium, the chemical potential for a given molecular species is constant throughout the system. The two terms on the right-hand side of (11.4) can vary in space, however, so as to add up to a constant. In an inhomogeneous system, the number density and excess chemical potential adjust so as to yield the same constant chemical potential. Due to the local nature of the excess chemical potential, it is reasonable to define an excess chemical potential at a single point in space and/or for a single molecular conformation [29]. That excess chemical potential then determines

the local density of the chemical species of interest, once the total chemical potential is specified. This key feature of the PDT permits examination of free energy changes along a specified pathway. An example would be examining free energy profiles along a reaction path in an enzyme [31, 32].

For book-keeping purposes, it is often easiest to perform the averages in the canonical ensemble. Then the double average for the excess chemical potential can be viewed as follows. The overall distribution to be sampled consists of the Boltzmann factors which contain the potential energies for the N-molecule solution and the internal potential energy for the distinguished molecule. The normalizing factor in the denominator is an integral of this distribution over all coordinates of the solution and the molecule, neglecting the translational degrees of freedom of the molecule.

A useful view of averages can be expressed via the PDT. The configurational average of the quantity F is

$$\langle F \rangle = \frac{\langle\langle \mathrm{e}^{-\beta \Delta U_\alpha} F \rangle\rangle_0}{\langle\langle \mathrm{e}^{-\beta \Delta U_\alpha} \rangle\rangle_0}. \tag{11.5}$$

The normalization integrals for the averages in the numerator and denominator cancel each other, leaving the traditional expression for the thermal average of F with the distinguished molecule present in the solution. This expression for the average will prove helpful several times below. The PDT is discussed extensively in Chap. 9, and in [29].

11.4 Fourier Path Integrals

Feynman, building on earlier ideas of Dirac, developed a formulation of quantum mechanics alternative to the older Heisenberg and Schrödinger pictures [1]. This view expresses the transition amplitude to go from points x to x' in a time t as the superposition of $\exp(iS/\hbar)$ phase factors for an infinite number of paths linking the two points, where S is the action evaluated along the path. When this formulation for real-time evolution is carried over to equilibrium statistical mechanics, the physical time t is replaced by the imaginary time $i\tau$, where τ ranges from 0 to $\beta\hbar$. This converts the action S to an energy function in a space with an increased number of degrees of freedom specifying the path coordinates for each particle [2]. For all of the systems of interest here, particle exchange effects are negligible. Then, the paths are cyclic, starting and ending at the same place. This means that we are assuming Boltzmann statistics. The assumption of Boltzmann statistics is the only approximation required to derive the quantum PDT. Then, the configuration of the whole system consists of the location of the path origin for each particle and an added set of variables to describe the cyclic paths.

Path integrals can be expressed directly in Cartesian coordinates [1, 2] or can be transformed to Fourier variables [1, 2, 20, 45]. A Fourier path integral method will be used here [20]. The major reason for doing this is that length scales are directly built

into the representation. The Cartesian and Fourier path integrals generate the same physical information, and their mathematical relationship and equivalence have been studied by Coalson [46]. As we will see below, going to a Fourier representation does not involve a Fourier transform operation, but is simply a coordinate change; all of the calculations are still carried out in real space (not **k**-space).[5] At first appearance, the mathematics appears daunting in path integral calculations. But after some exposure to the expressions and practice using them to derive useful results, things begin to look very similar to the more usual classical Boltzmann statistical mechanics. The main purposes in introducing path integrals here are to permit the derivation of a quantum version of the PDT and then to obtain useful approximations based on the exact expressions. Those approximations are discussed in the beginning of Sect. 11.7. After all the path integral gyrations, the Feynman–Hibbs effective potential to model quantum effects turns out to be remarkably simple, namely a Gaussian smear of the potential where the width depends on the temperature and mass.

In what follows, a one-dimensional picture will mainly be employed to avoid a notational proliferation. We will also consider a single *atomic* solute. The methods discussed are easily generalized to three dimensions and many particles; the $x, y,$ and z coordinates of the cyclic path for each atom are represented by a set of Fourier coefficients.

In equilibrium statistical mechanics involving quantum effects, we need to know the density matrix in order to calculate averages of the quantities of interest. This density matrix is the quantum analog of the classical Boltzmann factor. It can be obtained by solving a differential equation very similar to the time-dependent Schrödinger equation

$$\frac{\partial \rho}{\partial \beta} = -\mathscr{H}\rho, \tag{11.6}$$

where ρ is the density matrix and H is the system's Hamiltonian operator. Notice that, when you insert the Hamiltonian operator on the right, the equation appears quite similar to a diffusion equation. There are many possibilities for solving (11.6) for ρ; here we present the path integral representation of the solution.

In coordinate space, the diagonal elements of the canonical density matrix in the Fourier path integral representation are given by [20]

$$\rho(x, x; \beta) = \frac{1}{N!} J(\beta) \int \prod_k \mathrm{d}a_k \exp\left(-\sum_k \frac{a_k^2}{2\sigma_k^2}\right) \exp\left(-\beta \int_0^1 U[x(\tau)]\mathrm{d}\tau\right), \tag{11.7}$$

where the a_k variables are the Fourier coefficients, and, for one particle in one dimension

$$J(\beta) = \left(\frac{m}{2\pi\beta\hbar^2}\right)^{1/2} \prod_k \frac{1}{\sqrt{2\pi\sigma_k^2}}, \tag{11.8}$$

---

[5] This coordinate change is analogous to changing integration variables x, y, z to spherical polar coordinates r, θ, ϕ.

and

$$\sigma_k^2 = \frac{2\beta\hbar^2}{m(\pi k)^2}. \tag{11.9}$$

$J(\beta)$ is the Jacobian for the transformation from Cartesian to Fourier variables; the Fourier coefficients $\{a_k\}$ used to specify a given path (see below) have the same units as the positional coordinates in the Cartesian path integral. The spreads σ_k in the Gaussian factors in the kinetic-energy piece are inversely related to the square roots of the temperature and mass; also, the spreads decrease with increasing Fourier index k. After choosing a position x which is the origin of a given path, we integrate over all the Fourier variables $\{a_k\}$ to obtain the density matrix element $\rho(x, x; \beta)$. The integral of the potential appearing in the last term involves the fictitious scaled 'time' variable τ which ranges from 0 to 1 and expresses the motion along the path. The integrand in (11.7) is called a functional; instead of connecting one variable to another as in a regular function, we need to specify the entire path to return a value for the integrand. We will discuss in more detail the two terms of the integrand below. The integral of (11.7) over the coordinate x yields the canonical partition function Q.

A path is specified by a set of Fourier coefficients $\{a_k\}$

$$x(\tau) = x + \sum_k a_k \sin k\pi\tau. \tag{11.10}$$

The location x is the origin of the path, and the Fourier coefficients generate deviations away from that origin. In practical calculations the expansion is truncated at a finite value of $k = k_{\max}$, and the integral of the potential along the path is performed at a discrete set of points chosen fine enough to capture the physical features. It is clear that the high-k modes describe short-wavelength oscillations about the path structure determined by the longer-wavelength, low-k modes. If there are many particles moving in three dimensions, the Jacobian in (11.8) contains a product over all the masses (with the exponent 1/2 replaced by 3/2) and over all the Fourier coefficients specifying the paths.

A good example problem is the one-dimensional harmonic oscillator, where $U(x) = m\omega^2 x^2/2$. This case can be solved analytically, and the calculation is fairly lengthy. Here we just quote the results, and encourage the reader to give the derivation a try. To outline the steps, first compute the τ integral of the potential, and then do all of the Gaussian integrals involving the kinetic energy. The result is

$$\rho(x, x; \beta) = C \exp[-S(x, x; \beta)], \tag{11.11}$$

where

$$S(x, x; \beta) = \frac{1}{2} m\omega^2 \beta x^2 - \frac{mx^2}{\beta\hbar^2} \sum_k \left(\frac{\beta\hbar\omega}{k\pi}\right)^2 \frac{[1 - (-1)^k]^2}{[1 + (k\pi/\beta\hbar\omega)^2]}. \tag{11.12}$$

Several of the system parameters have been collected in C, which is independent of x. Notice that this result depends on $k_{\max}$. The analytical result in the $k_{\max} \to \infty$ limit, obtained by other means in Feynman's book [2], is

$$\rho(x, x; \beta) = C' \exp\left(-\frac{m\omega}{\hbar} x^2 \tanh f\right) , \tag{11.13}$$

where C' is another constant and $f = \beta\hbar\omega/2$. We can analyze (11.13) in the high- and low-temperature limits. At high temperature, $f \to 0$ and we can expand the exponentials appearing in the $\tanh f$ term. The result is just the classical distribution for the harmonic oscillator at a given temperature

$$\rho(x, x; \beta) = C' \exp\left(-\frac{1}{2} m\omega^2 x^2\right). \tag{11.14}$$

At the other extreme of low temperature, $\tanh \to 1$ and $\rho(x, x; \beta)$ just becomes the square of the ground-state wave function

$$\rho(x, x; \beta) = C' \exp\left(-\frac{m\omega}{\hbar} x^2\right). \tag{11.15}$$

What does this example illustrate? First, at high temperatures we know that the paths shrink due to the decrease in the σ_k values with increasing temperature. Eventually, the paths shrink to points, and that is the classical limit 11.14. At the other extreme of low temperature, the paths are more extended since the σ_k values become large, but the potential confines the paths to be distributed in a way that reflects the ground-state wave function. The approximation methods discussed in this chapter are valid at temperatures where the paths have shrunk to small, but not point-like, sizes.

Given the density matrix, the average of a coordinate-dependent function $F(x)$ is then

$$< F >= \frac{\int \mathrm{d}x \rho(x, x; \beta) F(x)}{\int \mathrm{d}x \rho(x, x; \beta)}. \tag{11.16}$$

This shows that the probability density for the averaging involves a Boltzmann factor which contains a kinetic-energy part $\left(\sum_k a_k^2/2\sigma_k^2\right)$ and the average interaction potential along the path $x(\tau)$. The kinetic-energy piece creates an energetic cost when the paths become too extended and/or 'kinked,' indicated by large values of the Fourier coefficients. The interaction potential is computed only between points with a given τ value and then integrated over τ to get the average; a clear discussion of these points is given in [13]. This discussion shows that obtaining a quantum average using path integrals is virtually the same as a classical average, except we have to sample in an enlarged space of a_k variables besides the usual positional coordinates x.

Imagine that we are using the Metropolis Monte Carlo method to calculate an average. For each trial x and set of Fourier coefficients $\{a_k\}$ during the sampling, we first generate a path and then calculate the kinetic and potential energetic pieces appearing in (11.7). If this energy in our enlarged space decreases, we accept the trial move; if it increases, we accept it with a probability proportional to the exponential of the difference of the energy factor between the current and trial configurations in the usual Metropolis way.

If we are interested in the off-diagonal elements of the density matrix, that is $\rho(x, x'; \beta)$, the paths are no longer cyclic but begin at x and terminate at x'. Then the Fourier representation of the path is

$$x(\tau) = x + (x - x')\tau + \sum_k a_k \sin k\pi\tau. \tag{11.17}$$

This picture suggests a trajectory passing from one point to another in a total time $\tau = 1$. For fixed x and x', we can then imagine sampling those trajectories with a Boltzmann-like weight from an integrand similar to that in (11.7).[6] Although the details differ, the transition path sampling method for locating transition states in high-dimensional problems is close in spirit to the path integrals discussed above: both model a diffusion-like process with fixed endpoints by sampling over a distribution of paths or 'trajectories.' This provides an alternative view of boundary conditions for dynamical processes. Instead of producing an ensemble of trajectories with initial positions and momenta, we generate paths which link two points in space in a given time interval. See Chap. 7 and the early pioneering work by Pratt [47].

11.5 The Quantum Potential Distribution Theorem

As mentioned above, there are multiple ways to derive the PDT for the chemical potential. Here we utilize the older method in the canonical ensemble which says that $\beta\mu_\alpha$ is just minus the logarithm of the ratio of two partition functions, one for the system with the distinguished atom or molecule present, and the other for the system with no solute. Using (11.7) we obtain [9, 48, 49]

$$\beta\mu_\alpha = \ln \rho_\alpha \Lambda_\alpha^3 - \ln \left\langle \left\langle e^{-\beta \int_0^1 \Delta U_\alpha[x(\tau)]d\tau} \right\rangle_{a_k} \right\rangle_0. \tag{11.18}$$

First notice the close similarity to the classical formula for a molecular solute (11.4). The ideal part of the chemical potential does not contain the internal partition function q_α^{int} since we are considering an atomic solute here. The inner Gaussian average in (11.18) comes from the normalizing factors for the Fourier coefficients in (11.7). That average pertains to an average over the Boltzmann weight containing the kinetic energy factor alone

$$\langle \ldots \rangle_{a_k} = \frac{\int \prod_k da_k \, (\ldots) \, e^{-\sum_k \frac{a_k^2}{2\sigma_k^2}}}{\int \prod_k da_k e^{-\sum_k \frac{a_k^2}{2\sigma_k^2}}}, \tag{11.19}$$

that is, a Gaussian average with no interactions between the solute and surrounding solvent. The outer average is over the thermal motions of the solvent decoupled from the solute. In place of the interaction energy ΔU_α in (11.4), we have the average interaction energy along the instantaneous solute path configuration generated during the inner-Gaussian sampling; (11.18) has been used directly to obtain the excess chemical potential for hydrogen [9] and neon [9, 16] liquids.

---

[6] The energetic factors are slightly more complicated with differing endpoints x and x'.

We now use a trick to partition this exact expression for the chemical potential into classical and quantum correction parts [29]. To do this we multiply and divide inside the logarithm of the excess term by the classical average

$$\beta\mu_\alpha = \ln\rho_\alpha\Lambda_\alpha^3 - \ln\left\langle e^{-\beta\Delta U_\alpha(\bar{x})}\right\rangle_0 - \ln\frac{\left\langle\left\langle e^{-\beta\int_0^1\Delta U_\alpha[x(\tau)]\mathrm{d}\tau}\right\rangle_{a_k}\right\rangle_0}{\left\langle\left\langle e^{-\beta\Delta U_\alpha(\bar{x})}\right\rangle_{a_k}\right\rangle_0} . \quad (11.20)$$

The Fourier coefficient average in the denominator of the last term is added to make the numerator and denominator symmetrical. It has no effect on the classical average. The classical factor $\Delta U_\alpha(\bar{x})$ signifies that the potential is evaluated at the centroid of the path

$$\bar{x} = \int_0^1 x(\tau)\mathrm{d}\tau = x + \frac{1}{\pi}\sum_k\frac{a_k}{k}\left[1 - (-1)^k\right] . \quad (11.21)$$

This is helpful in deriving approximations later. Intuitively it makes sense since we would like to evaluate fluctuations about the 'center of mass' of the path.

In the numerator of the last term, we then multiply and divide by $\exp(-\beta\Delta U_\alpha(\bar{x}))$, which means adding and subtracting $-\beta\Delta U_\alpha(\bar{x})$ inside the exponent. This puts the term in the form of (11.5), where F is everything left over besides the $\exp(-\beta\Delta U_\alpha(\bar{x}))$ factor. We can finally write the quantum PDT as

$$\beta\mu_\alpha = \ln\rho_\alpha\Lambda_\alpha^3 - \ln\left\langle e^{-\beta\Delta U_\alpha(\bar{x})}\right\rangle_0$$
$$- \ln\left\langle\left\langle e^{-\beta\int_0^1[\Delta U_\alpha[x(\tau)] - \Delta U_\alpha(\bar{x})]\mathrm{d}\tau}\right\rangle_{a_k}\right\rangle_{\mathrm{cl}} . \quad (11.22)$$

The first two terms on the right are the classical chemical potential and the last term is an *exact* quantum correction. The averaging in that last term is over the Gaussian kinetic energy piece and the "cl" subscript on the outer average now says that the *classical* solute is included during the calculation; the average is over the classical reference system. This partitioning is fruitful in deriving approximations for the quantum correction to the excess chemical potential. The inclusion of the classical solute during the averaging process makes the calculation less noisy than the brute-force approach suggested by (11.18). In the case of (11.18), the entire cyclic path is randomly 'inserted' into the liquid, which can lead to frequent substantial overlaps with the solvent atoms; this results in a noisy averaging process with most terms close to zero and rare favorable insertions into available cavities. Equation (11.22) reorganizes the calculation into a classical part and a correction term in which the classical solute already exists in the fluid. The averaging in the last term of (11.22) focuses on differences between the interaction energy of the classical point particle and the quantum particle along the cyclic path. Thus, the sampling is expected to be less noisy.

11.6 The Variational Approach to Approximations

The calculation of the quantum correction to the excess chemical potential is of the form

$$-\ln\langle e^{-\beta f}\rangle_{a_k}. \tag{11.23}$$

We will consider expression (11.23) as a schematic for the inner a_k averaging in (11.22). Let us now explore an approximation for this expression. First, expand the exponential and take the averages of the terms in increasing powers of β. We will consider terms here up to second order in β. Then, expand the logarithm and we get

$$-\ln\langle e^{-\beta f}\rangle_{a_k} \approx \beta\langle f\rangle_{a_k} - \frac{\beta^2}{2}\left[\langle f^2\rangle_{a_k} - \langle f\rangle_{a_k}^2\right] + \cdots \tag{11.24}$$

These are the first two terms in a cumulant expansion [50]. We note here that the convergence of cumulant expansions is a subtle issue. Generally, if the statistics are nearly Gaussian, the cumulant expansion yields a good approximation. If the statistical distribution is not Gaussian, however, the cumulant expansion diverges with the inclusion of higher-order terms. See [29] and references therein for more discussion of this point.

An inequality then proves useful [2]

$$\left\langle e^{-\beta f}\right\rangle_{a_k} \geq e^{-\beta\langle f\rangle_{a_k}} . \tag{11.25}$$

This inequality is called the Gibbs–Bogoliubov–Feynman bound [51], and it can be obtained as the instantaneous-switching limit of Jarzynski's identity relating non-equilibrium trajectories to free energy changes.[7] It says that if we just retain the first-order term in β, the approximated quantum correction lies above the exact quantum result. The second-order term is always negative, which also gives an indication that the exact result lies below the approximate one. In addition, the classical free energy lies below the exact quantum result [45, 52, 53]. Thus, the exact quantum free energy is bounded above and below by values that can be obtained using classical mechanics for the sampling. This useful point does not seem to have been exploited much in computations of free energies.

11.7 The Feynman–Hibbs Variational Method

The physical principle underlying the following approximations is that relatively weak quantum effects are reflected in narrow Gaussian distributions in (11.19). The small widths are due to the large mass and high temperature. This means that the potential does not vary much over the length scales sampled by the kinetic energy distribution. Additionally, the higher-k Fourier modes become successively narrower. This length-scale argument was invoked in the development of the partial averaging method by Doll and coworkers [20]. In this method, the equality in (11.25)

---

[7] See discussion of Jarzynski's identity in Chap. 5.

is employed to obtain an effective potential along the path specified by the long-wavelength modes. That effective potential involves a Gaussian smear of the potential with a width determined by the largest chosen $k = k_{\max}$. The effective potential allows a much smaller required $k_{\max}$ in many-body simulations. When the averaging process in the partial averaging method is continued through *all* of the Fourier variables, the Feynman–Hibbs effective potential [2] is obtained.

We just summarize the steps of the derivation here and refer the reader to [29] where the procedure is outlined in an exercise. We first make approximation (11.25) for the inner average over the a_k variables in (11.22). We are then interested in the average of the difference between the interaction energy of the solute along the path and that evaluated at the centroid. But we note that no point along the path is special, and we choose the point x for our sampling point; $\bar{x}$ remains fixed. Since the variable x is generated by a linear combination of the a_k variables, the multidimensional Gaussian can be collapsed into a single-Gaussian integral after some algebra [50]

$$\Delta\Delta U_\alpha^{\text{eff}}(\bar{x}) = \frac{\int \mathrm{d}y\, [\Delta U_\alpha(\bar{x} - y) - \Delta U_\alpha(\bar{x})]\, \mathrm{e}^{-\frac{6my^2}{\beta\hbar^2}}}{\int \mathrm{d}y\, \mathrm{e}^{-\frac{6my^2}{\beta\hbar^2}}}. \tag{11.26}$$

Here, $\Delta\Delta U_\alpha^{\text{eff}}(\bar{x})$ is an effective potential representing the quantum deviations from the classical interaction potential. This effective potential is variational since the only approximation made so far is from the inequality (11.25).

Now if we assume that the deviations are small, and expand $\Delta U_\alpha(\bar{x} - y)$ about the point $\bar{x}$ to second order, we get

$$\boxed{\Delta\Delta U_\alpha^{\text{eff}}(\bar{x}) \approx \frac{\beta\hbar^2}{24m} \Delta U_\alpha''(\bar{x})}, \tag{11.27}$$

where $\Delta U_\alpha''(\bar{x})$ is the second derivative of the potential evaluated at $\bar{x}$. We will call this the quadratic Feynman–Hibbs (QFH) correction. Kleinert [45] shows that this quadratic expansion is also variational; all approximations derived beyond this one may remove this property. For a pairwise interaction, the QFH correction reads

$$\Delta\Delta U_\alpha^{\text{eff}}(\bar{x}) \approx \frac{\beta\hbar^2}{24\mu} \left[\Delta U_\alpha''(r) + 2\frac{\Delta U_\alpha'(r)}{r} \right], \tag{11.28}$$

where μ is the reduced mass, and the second ($\Delta U''$) and first ($\Delta U'$) derivatives are taken with respect to the scalar distance between the atoms.

The QFH potential approximately captures two key quantum effects. When an atom is near a potential minimum, the curvature is positive and thus so is the QFH correction; this models the zero-point effect. On the other hand, near potential maxima the curvature is negative, and the QFH potential models tunneling.

The Feynman–Hibbs and QFH potentials have been used extensively in simulations examining quantum effects in atomic and molecular fluids [12, 15, 25]. We note here that the centroid molecular dynamics method [54, 55] is related and is intermediate between a full path integral simulation and the Feynman–Hibbs approximation;

the averaged forces during the classical propagation are determined using a path integral simulation. It has been shown that the QFH potential and the centroid approach yield similar results in water simulations [25].

We can view obtaining the QFH correction to the excess chemical potential in two ways. If we simply insert (11.27) back into (11.22), this suggests that we first compute the classical excess chemical potential and then insert the classical solute into the system and evaluate

$$
-\ln \left\langle e^{-\frac{\beta^2 \hbar^2}{24m} \Delta U''_\alpha(\bar{x})} \right\rangle_{\mathrm{cl}}
\tag{11.29}
$$

for the quantum correction. Alternatively, we can recombine the classical and quantum correction terms in (11.22) and calculate the excess chemical potential as

$$
\beta \mu_\alpha^{\mathrm{ex,QFH}} = -\ln \left\langle e^{-\beta[\Delta U_\alpha(\bar{x}) + \frac{\beta \hbar^2}{24m} \Delta U''_\alpha(\bar{x})]} \right\rangle_0 .
\tag{11.30}
$$

If we are interested in the excess chemical potential change for mutating mass m_{A} into mass m_{B}, we obtain

$$
\begin{aligned}
\beta \Delta \mu_\alpha^{\mathrm{ex,QFH}}(m_{\mathrm{A}} \to m_{\mathrm{B}}) &= -\ln \frac{\left\langle e^{-\beta[\Delta U_\alpha(\bar{x}) + \frac{\beta \hbar^2}{24m_{\mathrm{B}}} \Delta U''_\alpha(\bar{x})]} \right\rangle_0}{\left\langle e^{-\beta[\Delta U_\alpha(\bar{x}) + \frac{\beta \hbar^2}{24m_{\mathrm{A}}} \Delta U''_\alpha(\bar{x})]} \right\rangle_0} \\
&= -\ln \left\langle e^{-\frac{\beta^2 \hbar^2}{24} \Delta\left(\frac{1}{m}\right) \Delta U''_\alpha(\bar{x})} \right\rangle_{m_{\mathrm{A}}} ,
\end{aligned}
\tag{11.31}
$$

where $\Delta(1/m) = 1/m_{\mathrm{B}} - 1/m_{\mathrm{A}}$ and again we have used (11.5). The calculation is performed with the mass m_{A} particle included, interacting with the rest of the system with the QFH potential.

Feynman and Kleinert derived a method which is a significant improvement over the Feynman–Hibbs variational approach. A detailed discussion of this method is given in [45]. The method focuses on a local harmonic oscillator reference system rather than on performing the Gaussian integrals directly as done above. The derived effective potential goes to the classical potential at high temperatures, but in addition gives a remarkably good estimate of the ground-state energy at low temperatures. The approximated free energy is an upper bound, just like in the Feynman–Hibbs method, but a better approximation is obtained at lower temperatures. It does involve increased complexity in obtaining an optimal local harmonic frequency. Moreover, the applications considered here are mainly in the higher-temperature regime discussed in [45].

Another issue to mention concerns the sampling of the solvent degrees of freedom. The cl subscript on the outside average of the last term of (11.22) refers to treating the *solute* classically. If the solvent is expected to display minor

quantum effects, then it too can be modeled with a QFH potential; the solute should be treated classically in the cl-labeled averaging though. If the interaction potential happens to be a pairwise potential, then a very simple form for the QFH potential results [12, 15, 25], as shown in (11.28).

11.8 A Worked Example

We have covered a lot of ground starting from an exact quantum PDT and deriving a physically based Feynman–Hibbs effective potential designed to approximately include quantum effects during a classical calculation. The path integral methods used to derive the Feynman–Hibbs potential are involved, but the result (11.26) is simple: take a Gaussian smear of the potential centered at the classical point $\bar{x}$. Here we stop and consider an example, the harmonic oscillator, which illustrates some of the results discussed above. The harmonic oscillator partition function, and thus the Helmholtz free energy, is easy to obtain analytically. The exact free energy is:

$$A_{\mathrm{QM}} = \frac{1}{\beta} \ln \left[\sinh \left(\frac{\hbar\omega\beta}{2} \right) \right], \tag{11.32}$$

where ω is the harmonic frequency. The high-temperature or classical limit of this expression is

$$A_{\mathrm{CM}} = \frac{1}{\beta} \ln(\hbar\omega\beta). \tag{11.33}$$
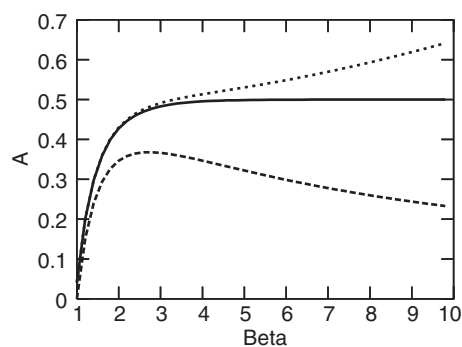
Fig. 11.1. The Helmholtz free energy as a function of β for the three free energy models of the harmonic oscillator. Here we have set $\hbar = \omega = 1$. The exact result is the *solid line*, the Feynman–Hibbs free energy is the *upper dashed line*, and the classical free energy is the *lower dashed line*. The classical and Feynman–Hibbs potentials bound the exact free energy, and the Feynman–Hibbs free energy becomes inaccurate as the quantum system drops into the ground state at low temperature

Using path integral methods, Feynman [2] showed that the Feynman–Hibbs form for the harmonic oscillator free energy is

$$A_{\mathrm{FH}} = \frac{1}{\beta}\left[\ln(\hbar\omega\beta) + \frac{\hbar^2\beta^2\omega^2}{24}\right]. \tag{11.34}$$

This approximation can also be easily obtained from the expression (11.29). Based on our discussion above, we expect $A_{\mathrm{CM}} < A_{\mathrm{QM}} < A_{\mathrm{FH}}$. Figure 11.1 illustrates these bounds. The Feynman–Hibbs free energy provides a very good representation at moderate temperatures, at which some excited states are populated. See Kleinert's book [45] for applications to anharmonic systems.

We will finish the example with a couple of points which will prove useful later in the chapter. If we insert the harmonic oscillator potential into the temperature correction formula (11.2), we get

$$\Delta T = T_{i,\mathrm{eff}} - T \approx \frac{1}{12}(\beta\hbar\omega)^2 T \tag{11.35}$$

for the correction. At room temperature, $\beta^{-1} \approx 208\,\mathrm{cm}^{-1}$ in typical units for vibrational frequencies, so the correction is $25(\tilde{\nu}/208)^2$. If we assume a frequency of $294\,\mathrm{cm}^{-1}$, we get a temperature correction of $50\,\mathrm{K}$. By examining the density of states for liquid water [56], the choice of a few hundred cm^{-1} as a characteristic frequency is not unreasonable. The high density of states in that frequency range is due to the hindered rotations of the water molecules in the liquid. As we will see below, the quantum effects on the structural properties of liquid water are roughly equivalent to a $50\,\mathrm{K}$ temperature rise in the classical liquid.

Second, there has been a discussion of whether it is better to treat the water molecule as rigid or flexible during simulations of the fluid [40]. One argument has been that, since the molecules are largely in their ground vibrational states at room temperature, it might be better to treat them as rigid. But this assumption seems somewhat questionable when the root-mean-square proton fluctuations $(\sqrt{<x^2>})$ are calculated in the classical and quantum (ground-state) limits. Let us assume a harmonic oscillator with a vibrational frequency of roughly $1,500\,\mathrm{cm}^{-1}$, a mass of 2 proton masses, and a temperature of $300\,\mathrm{K}$. These parameters lead to classical and quantum predictions of the rms fluctuations of 0.04 and 0.075 Å, respectively. So the notion that the molecule is effectively more rigid in the quantum system may not be physically correct.

11.9 Wigner–Kirkwood Approximations

Wigner–Kirkwood related expansions follow by taking the approximations further. We assume (11.30) as a starting point and linearize the exponential of the correction term for the potential

$$-\ln \left\langle \mathrm{e}^{-\beta \left[\Delta U_\alpha(\bar{x}) + \frac{\beta \hbar^2}{24m} \Delta U_\alpha''(\bar{x}) \right]} \right\rangle_0$$

$$\approx -\ln \frac{\int \mathrm{d}x^N \mathrm{e}^{-\beta U_N} \mathrm{e}^{-\beta \Delta U_\alpha(\bar{x})} \left(1 - \frac{\beta^2 \hbar^2}{24m} \Delta U_\alpha''(\bar{x}) \right)}{\int \mathrm{d}x^N \mathrm{e}^{-\beta U_N} \mathrm{e}^{-\beta \Delta U_\alpha(\bar{x})}}$$

$$-\ln \frac{\int \mathrm{d}x^N \mathrm{e}^{-\beta U_N} \mathrm{e}^{-\beta \Delta U_\alpha(\bar{x})}}{\int \mathrm{d}x^N \mathrm{e}^{-\beta U_N}}, \tag{11.36}$$

assuming the total potential for the solvent U_N is a classical potential for now.

The second term on the right of (11.36) is just the classical excess chemical potential. The first term is

$$-\ln \left[1 - \frac{\beta^2 \hbar^2}{24m} \langle \Delta U_\alpha''(\bar{x}) \rangle_{\mathrm{cl}} \right], \tag{11.37}$$

which can also be obtained by expanding the exponential in (11.29). An effective potential closely related to (11.37) was derived by Stratt [44] and has been examined in molecular simulations [12]. Assuming the second term inside the logarithm in (11.37) is small and expanding the logarithm, we get

$$\beta \mu_\alpha^{\mathrm{ex,WK}} \approx \beta \mu_\alpha^{\mathrm{ex,cl}} + \frac{\beta^2 \hbar^2}{24m} \langle \Delta U_\alpha''(\bar{x}) \rangle_{\mathrm{cl}}. \tag{11.38}$$

Now consider the classical average of the second derivative appearing in (11.38). This average can be integrated by parts if we assume a very large system where boundary effects are negligible [43]

$$\boxed{\beta \mu_\alpha^{\mathrm{ex,WK}} \approx \beta \mu_\alpha^{\mathrm{ex,cl}} + \frac{\beta^3 \hbar^2}{24m} \langle (\Delta U_\alpha'(\bar{x}))^2 \rangle_{\mathrm{cl}}}. \tag{11.39}$$

This is of the form of the correction to the free energy in (11.1). Extensions for rigid molecules are given in [57].

By performing the integration by parts on only a portion of the correction, we can also say

$$\boxed{\beta \mu_\alpha^{\mathrm{ex,WK}} \approx \beta \mu_\alpha^{\mathrm{ex,cl}} - \frac{\beta^3 \hbar^2}{24m} \langle (\Delta U_\alpha'(\bar{x}))^2 \rangle_{\mathrm{cl}} + \frac{\beta^2 \hbar^2}{12m} \langle \Delta U_\alpha''(\bar{x}) \rangle_{\mathrm{cl}}}. \tag{11.40}$$

If we now re-express (11.40) as the expansion of a logarithm, and then re-exponentiate the terms inside the classical average, we obtain (11.3) as an effective potential. Even though we have retained the dependence on the coordinate $\bar{x}$ to make the connection to an effective potential for a given system configuration, the averaging process implied by the integration by parts means that the solute particle is free to move throughout the system volume. In this sense, the Wigner–Kirkwood effective potential is an averaged potential which is designed to reproduce system properties

globally. The Feynman–Hibbs potential, on the other hand, results from averaging only over the quantum degrees of freedom, and retains the well-defined local nature inherent in the PDT; this is a distinct advantage of the Feynman–Hibbs approach when modeling inhomogeneous systems.

At this point it might be helpful to summarize what has been done so far in terms of effective potentials. To obtain the QFH correction, we started with an exact path integral expression and obtained the effective potential by making a first-order cumulant expansion of the Boltzmann factor and analytically performing all of the Gaussian kinetic energy integrals. Once the first-order cumulant approximation is made, the rest of the derivation is exact up to (11.26). A second-order expansion of the potential then leads to the QFH approximation.

Once the QFH formula for the excess chemical potential is linearized in (11.37), the logarithmic expression can be expanded to first order and all or part of the classical-average term can be integrated by parts to yield the Wigner–Kirkwood correction to the free energy. Then if (11.40) is reorganized, computation of the chemical potential can be viewed as a classical average with a modified interaction potential of the same form as (11.3).

How do the two effective potentials compare? In [12], the effective potentials corresponding to (11.27), (11.37), and (11.40) are plotted for neon at its triple-point temperature. All three effective potentials mimic zero-point motion by raising the potential near the minimum relative to the classical Lennard-Jones form. The potentials differ widely at small separations, however. The QFH approximation is more repulsive than the classical potential at small r values, while the logarithmic form of the Wigner–Kirkwood effective potential equation (11.37) is less repulsive than the classical potential. Also, this logarithmic form of the Wigner–Kirkwood potential matches up quite well at small r with path integral calculations for the atomic pair [13]. The effective potential equation (11.3) derived from (11.40) exhibits a non-physical negative divergence at small r due to the negative sign in front of the square of the gradient of the interaction potential.

At first glance, it would appear that the QFH approximation is a better one since it is only a few steps removed from the exact path integral results, and the Wigner–Kirkwood formulas are obtained only after several subsequent approximations. However, the Wigner–Kirkwood pair potential is closer to a path-integral-derived effective potential at small distances. There does not appear to be a conclusive comparison of the two effective potentials for all of the thermodynamic and structural properties of fluids, although there has been significant work in this direction [11, 15]. Sese concludes that the QFH approximation performs better (than the Wigner–Kirkwood form) in computing thermodynamic properties. He does note that QFH-derived pressures deviate from the exact results at low temperatures and/or high densities. This is probably due to the enhanced repulsive character of the potential. Also, he showed that a Gaussian deconvolution of the center-of-mass QFH radial distribution function leads to much better agreement with experiment [15]. Previously he had found that Wigner–Kirkwood models gave better agreement with experimental structural data [17]. The reader is referred to more-extensive discussion of the relative merits of the effective potentials in these original papers.

11.10 The PDT and Thermodynamic Integration for Exact Quantum Free Energy Changes

Say you have performed a classical calculation to determine the excess chemical potential from the first two terms on the right side of (11.22) followed by another classical calculation to obtain an estimate of the quantum correction from the expression (11.29), and the estimated correction is large. This suggests that a full quantum treatment is necessary. In this section, we derive the appropriate formulas for changes in the excess chemical potential due to mutating masses. If the original mass is very large, which corresponds to the classical limit, the derived expressions yield the quantum correction.

Consider a problem, in which we are interested in the mass-dependent partitioning of a solute between an ideal gas phase and a condensed phase. The ratio of the densities in the two phases or the partition coefficient for species α is then

$$K = e^{-\beta \mu_\alpha^{\text{ex}}}, \tag{11.41}$$

where μ_α^{ex} is the excess chemical potential in the condensed phase. Thus to calculate the ratio of the partition coefficients for two isotopes, we need only consider the difference in the excess chemical potentials, and that will be our focus here.

Let us first examine changes in the quantum correction of (11.22) due to a change in mass. For that change we obtain

$$-\ln \frac{\left\langle \left\langle e^{-\beta \int_0^1 [\Delta U_\alpha[x(\tau)] - \Delta U_\alpha(\bar{x})] d\tau} \right\rangle_{a_k, \text{B}} \right\rangle_{\text{cl}}}{\left\langle \left\langle e^{-\beta \int_0^1 [\Delta U_\alpha[x(\tau)] - \Delta U_\alpha(\bar{x})] d\tau} \right\rangle_{a_k, \text{A}} \right\rangle_{\text{cl}}}. \tag{11.42}$$

We will use (11.5) yet again, but we should be careful to note that the normalization integrals for the numerator and denominator are slightly different. Taking care of those terms, we get

$$\boxed{\beta \Delta \mu_\alpha^{\text{ex}}(m_\text{A} \to m_\text{B}) = -\frac{3k_{\max}}{2} \ln \frac{m_\text{B}}{m_\text{A}} - \ln \left\langle e^{-\frac{\pi^2 \Delta m}{4\beta \hbar^2} \sum_k a_k^2 k^2} \right\rangle_{m_\text{A}}}, \tag{11.43}$$

where $\Delta m = m_\text{B} - m_\text{A}$, and the 3 in the first term on the right-hand side comes from assuming the particle moves in three dimensions. In that case, the sum in the exponent of the second term should be over all the solute Fourier variables for the three dimensions. In the averaging of the second term it is assumed that the mass m_A particle is included in the system. If the mass change is not too large, this expression should suffice. But if there is a significant mass change and the quantum effects are large, the statistical averaging will be noisy. In fact, this is a somewhat disturbing limit since, as we let the initial mass m_A approach infinity, we get the difference of two infinite terms, which should in the end yield a well-defined and finite result. We will see later that calculating the excess chemical potential in this limit is not so bad

as it first appears. For practical purposes, we can break up the mutation process into a sequence of smaller mass changes, using (11.43) for each step. See Chap. 2 for a discussion of perturbation theory.

Alternatively, the mass change can be enacted by thermodynamic integration. It is shown in [29] that thermodynamic integration possesses excellent scaling properties as long as the free energy changes smoothly with the scaling parameter. This provides a good reason for its workhorse status in free energy calculations. We will assume that the mass m_B is smaller than the mass m_A with an eye toward the transition from classical to quantum limits. Take the second term on the right-hand side of (11.43) and replace Δm with $(1 - \lambda)m_B$. We will consider the transition from $\lambda = 1$ to $\lambda = \lambda_f$, where $m_A = \lambda_f m_B$. This creates the λ-dependent function $F(\lambda)$:

$$F(\lambda) = -\ln \left\langle e^{-\frac{\pi^2(1-\lambda)m_B}{4\beta\hbar^2}\sum_k a_k^2 k^2} \right\rangle_{m_A}. \tag{11.44}$$

Then we can integrate the derivative of this function to get our desired result

$$F(\lambda_f) - F(1) = \int_1^{\lambda_f} \frac{\partial F}{\partial \lambda} d\lambda, \tag{11.45}$$

since $F(\lambda_f)$ is what we seek and $F(1) = 0$. Once we calculate the derivative and assemble the averages, we get for the quantum correction

$$\boxed{\beta\Delta\mu_\alpha^{ex}(m_A \to m_B) = -\frac{3k_{max}}{2}\ln\frac{m_B}{m_A} - \frac{\pi^2 m_B}{4\beta\hbar^2}\int_1^{\lambda_f}\left\langle\sum_k a_k^2 k^2\right\rangle_\lambda d\lambda.} \tag{11.46}$$

The λ-dependent averaging inside the integral involves a path integral simulation with the particle of interest involved but with a mass of $m_A + (1 - \lambda)m_B$. This formula has been used to estimate isotope effects on solubilities of hydrogen and deuterium in model anharmonic solids in [48]. A similar expression was derived for Cartesian path integrals by Runge and Chester [21]. We are considering the transition from a large mass to a small mass, so $\lambda_f > 1$ and the second term is thus always negative due to the minus sign in front. But the net result for the quantum correction is positive [53], so the first term must exceed the second in magnitude. That first term is always positive for $m_A \gg m_B$. Equation (11.46) appears similar to the T method of kinetic energy estimation [20]. It has been noted that the standard deviation of the kinetic energy in the T method increases as the square root of k_{max} [20]. The growth of the standard deviation is not surprising due to the k^2 factor. Predescu and Doll [58], building on fundamental ideas from Brownian motion theory, have proposed a reweighting scheme for the Fourier method which yields better-convergent estimators, and Mielke and Truhlar [59] have compared various Fourier-based estimators in free energy calculations.

Physically, as we go to larger masses during the λ integration, the widths of the Gaussians in the kinetic-energy piece of the sampling function become very narrow. This means that the distributions of the a_k are essentially Gaussian due to no influence from the potential; on the scale of the very small particle fluctuations, the

potential does not vary. If we assume that the kinetic and potential pieces have decoupled in this way, then we can perform the Gaussian integrals in the $\langle \ldots \rangle_\lambda$ averages analytically, and indeed we get exactly the opposite of the first term in (11.46). This suggests that we perform the λ integration as given in (11.46), successively increasing m_A until the process converges to a stable value; we have then reached the classical limit.

We can view the classical$\rightarrow$quantum transition another way, namely through the potential rather than the kinetic energy. If we substitute $(1 - \lambda)\Delta U_\alpha(\bar{x}) + \lambda \Delta U_\alpha[x(\tau)]$ for $\Delta U_\alpha[x(\tau)]$ in the quantum correction term of (11.22), and then follow through with a thermodynamic integration procedure, we obtain

$$\beta[\mu_\alpha^{\text{ex,QM}} - \mu_\alpha^{\text{ex,CM}}] = \beta \int_0^1 \left\langle \int_0^1 [\Delta U_\alpha[x(\tau)] - \Delta U_\alpha(\bar{x})]\mathrm{d}\tau \right\rangle_\lambda \mathrm{d}\lambda, \quad (11.47)$$

where now the path integral λ-dependent averaging uses the mixed potential given above. The $\lambda = 0$ limit is the classical limit, and $\lambda = 1$ generates the full quantum correction. This method presumably does not suffer from the k_{max}-dependent statistical sampling issue in (11.46), but we are not aware of a direct comparison of the two views of the classical$\rightarrow$quantum transition. The approach of (11.47) was used by Morales and Singer [14] in calculations of free energies for liquid neon.

11.11 Assessment and Applications

As we have seen, the PDT gives a compact means to derive quantum corrections to the classical chemical potential for an atomic or molecular solute. By making a variational approximation, the Feynman–Hibbs effective potential emerges directly from the exact path integral expression. If we expand the potential about the centroid to second order, the QFH approximation results, and the Wigner–Kirkwood approximations are obtained from further approximations. The classical excess chemical potential and the QFH quantum approximation give lower and upper bounds, respectively, to the exact quantum result. Thus, using purely classical simulation, we can begin to get a handle on the importance of quantum effects for free energies of a fluid. We will see below that, if the estimated quantum correction to the chemical potential is roughly 15% of the classical value or larger, then we should probably consider using a full path integral treatment. It is recommended to use the QFH potential (11.27) for the estimate of the quantum correction of thermodynamic properties since it is easy to incorporate during a classical simulation and is closest to the exact quantum result as we proceed down the 'approximation chain' – see Sect. 11.9. For most of the problems of interest to molecular modelers, the quantum effects are expected to be relatively weak, so these approximate methods are likely to give good estimates.

In this section, we will discuss some examples from the literature, in which the approximation methods derived in this chapter have been used. In several cases, the approximations have been compared with more-accurate path integral simulations to assess their validity. This is not meant as a full review; rather, several case studies have been chosen to illustrate the tools we have developed. We will first look at simpler examples and then discuss water models and applications in enzyme kinetics.

11.11.1 Foundational Examples

As noted above, Wang et al. [9] utilized the exact form of the quantum PDT in studies of liquid *para*-hydrogen and neon. For *para*-hydrogen, they found large differences between the free energies with classical or quantum models at low temperatures; at 30 K, the quantum result is roughly $-250\,\mathrm{kJ\,kg^{-1}}$, while the classical prediction is $-750\,\mathrm{kJ\,kg^{-1}}$. The quantum and classical results do not converge until roughly 120 K. The deviations for liquid neon are much less, on the order of $15\,\mathrm{kJ\,kg^{-1}}$ at 30 K, and the classical and quantum calculations converge around 50 K. The quantum results for both cases agree well with experimental data. These simulations confirm that hydrogen is a highly quantum liquid, while neon exhibits small but non-negligible quantum effects at low temperatures. Ortiz and Lopez [16] calculated adsorption isotherms for a neon monolayer using the quantum PDT, and observed an appreciable shift of the isotherms due to quantum effects at low temperatures. Morales and Singer [14] computed quantum corrections to the classical Helmholtz free energy for liquid neon at the triple point and observed a 5% change with inclusion of quantum effects. The Gibbs free energy, on the other hand, changes by 15% due to large changes in the ratio of the pressure to the density. They also tested a Wigner–Kirkwood model and found a 10% overestimation of the quantum correction obtained from an expansion to second order in $\hbar$. By adding a fourth-order term, the Wigner–Kirkwood error is reduced to about 5%. The free energy computed with path integral simulations agrees with experiment to within 3% of the measured value.

Sese [15] presented a more thorough examination of approximate quantum models for neon, namely the Feynman–Hibbs and QFH models. Free energies were calculated using the PDT. The Gibbs free energy from the models is always higher than the value computed using path integral methods, and the classical free energy is below the path integral result. This is expected considering the variational bounds discussed above. Four state points were examined. For state point 2, with a reduced temperature of 0.9517 and reduced density of 0.7246, the calculated Gibbs free energy differs from the path integral results by 2% and 3% for the Feynman–Hibbs and QFH models, respectively. The classical prediction differs from the path integral value by 8%, and the path integral result is 1% in error relative to experiment. The agreement between the Feynman–Hibbs models and path integral simulations is not as good at a lower-temperature and higher-density state point where the classical and quantum Gibbs free energies differ by 14%. Sese defined a dimensionless parameter $\Lambda^* = (h^2/2\pi m k_\mathrm{B} T \sigma^2)^{1/2}$, which gives the ratio of the quantum spread to the Lennard-Jones size parameter σ. When the reduced density is less than 0.89, the reduced temperature greater than 0.60, and $\Lambda^* \leq 0.28$, accurate thermodynamics can be obtained with the Feynman–Hibbs models.

Tchouar et al. [12] also utilized the QFH potential in simulations of neon, methane, and gaseous helium at low temperatures. They did not compute free energies, but obtained excellent agreement between the QFH and path integral calculations for the average total energies of the systems. They found similar conditions of validity to those given by Sese [15]. They also found much better agreement with experiment for diffusion constants and shear viscosity coefficients when the classi-

cal potentials were substituted with the QFH form, thus indicating these effective potentials may also be useful for dynamical quantities.

Summary

– The Feynman–Hibbs and QFH models perform quite well in free energy calculations as long as the quantum corrections are modest. The conditions for validity of the approximations are given above.

11.11.2 Force Field Models of Water

Moving on to quantum effects in water, we will first examine force field models of water, and then discuss recent ab initio simulation results. Early works utilizing path integral methods to study quantum effects in water include those by Kuharski and Rossky [60] and Wallqvist and Berne [61]. Kuharski and Rossky used a rigid (ST2) model of water, whereas Wallqvist and Berne examined a flexible model. Both observed a destructuring of the fluid with the inclusion of quantum effects. Kuharski and Rossky estimated the quantum correction to the free energy with a Wigner–Kirkwood model, obtaining a value of $0.68\,\text{kcal mol}^{-1}$ for H_2O. The major contributor to this correction is the librational component, not translations, which comprise less than 10% of the total correction. The excess chemical potential of water is $-6.1\,\text{kcal mol}^{-1}$ [36], so the estimated quantum correction is roughly 10% of the total. By differentiating the free energy with respect to temperature, they also estimated the quantum correction to the averaged interaction energy as $1.24\,\text{kcal mol}^{-1}$. The experimental total binding energy is $-9.92\,\text{kcal mol}^{-1}$ [56]. The larger magnitude for the energetic portion of the quantum correction implies a positive entropic contribution. Using path integral methods, their estimate of the energetic change was $0.82\,\text{kcal mol}^{-1}$. Reduced quantum corrections were observed for D_2O.

In more-recent path integral studies, Stern and Berne [56] examined a force field for flexible water obtained from ab initio calculations. Their classical simulation produced a total binding energy of $-11.34\,\text{kcal mol}^{-1}$, while their path integral result was $-9.8\,\text{kcal mol}^{-1}$, very close to the experimental value. The estimate of the quantum correction to the binding energy of $1.5\,\text{kcal mol}^{-1}$ is slightly larger than previous estimates. A simple harmonic model predicts a correction of $1.7\,\text{kcal mol}^{-1}$. In their flexible model, convergence of the energy of an isolated monomer is not obtained until more than 30 beads are included in the Cartesian path integral simulations.

Mahoney and Jorgensen [34] studied quantum, flexibility, and polarizability effects on water. They also developed a modified TIP5P(PIMC) rigid water potential, where the parametrization was based on path integral rather than classical simulations. Serious attention was paid to reproducing the temperature of maximum density. A Cartesian discretization with five beads was found sufficient to obtain converged results for rigid water models. For the TIP5P(PIMC) model, the authors observed an average intermolecular energy of $-9.94\,\text{kcal mol}^{-1}$, close to the $-9.92\,\text{kcal mol}^{-1}$ experimental value. Also, the predicted heat of vaporization is

$10.53\,\mathrm{kcal\,mol^{-1}}$ versus the experimental value of $10.51\,\mathrm{kcal\,mol^{-1}}$. The quantum correction to the interaction energy for the TIP5P model is $1.6\,\mathrm{kcal\,mol^{-1}}$, consistent with the results of Stern and Berne [56]. Mahoney and Jorgensen argue that real water is better approximated by rigid water models than by classical flexible models. We note here that there has also been significant recent effort devoted to developing classical force field models which accurately reproduce spectroscopic data for water clusters [62].

These path integral simulations serve as benchmarks for approximate models. Guillot and Guissani [25] employed the QFH approximation in extensive simulations of structure and dynamics in water and D_2O. A modified flexible force field was used. The QFH model predicts a heat of vaporization of $9.42\,\mathrm{kcal\,mol^{-1}}$ (corrected for zero-point differences) for water compared with path integral predictions of $9.79\,\mathrm{kcal\,mol^{-1}}$ (ST2 model) and $10.84\,\mathrm{kcal\,mol^{-1}}$ (SPC/E model), and an experimental value of $9.66\,\mathrm{kcal\,mol^{-1}}$. The quantum correction to the heat of vaporization predicted with the QFH model is $1.23\,\mathrm{kcal\,mol^{-1}}$ versus the path integral predictions of 1.18 and $1.27\,\mathrm{kcal\,mol^{-1}}$. Therefore, it appears the QFH model gives an excellent estimate of this thermodynamic quantity. The quantum contribution to the heat of vaporization is roughly 12% for water, a significant value. Guillot and Guissani noted an interesting discrepancy between the simulations and experimental results: the partial molar volume of D_2O is *greater* than that for water, and the QFH model does not reproduce this property. As for atomic fluids [15], for a given volume the pressure increases with inclusion of quantum effects, and a higher pressure is thus observed for water compared with D_2O in the QFH simulation. This may be related to the enhanced repulsive character of the QFH effective potential discussed above in Sect. 11.9. The temperature of maximum density shift is reproduced reasonably well with the QFH model. The authors also examined extensively the structural and dynamical properties of water. Consistent with many other simulations, the radial distribution functions soften with the inclusion of quantum effects, and diffusion is enhanced. The ratio of diffusion constants for water and D_2O increases markedly as the temperature is decreased; this quantity is also accurately reproduced by the QFH model. This study suggests that the QFH approximate model can accurately predict thermodynamic properties of water. The authors also compared their results with those from a more costly centroid molecular dynamics simulation of water [54], and found excellent agreement between the two methods.

Summary

- Inclusion of quantum effects leads to a destructuring of water.
- The quantum contribution to the excess chemical potential is roughly 10% of the total.
- The librational component is the major contributor to the quantum correction.
- The interaction-energy contribution to the quantum correction is larger in magnitude than the free energy contribution, suggesting the entropic part is positive.
- Modified force field models based on path integral simulations yield excellent agreement with experiment for thermodynamic properties.

- Flexible force field models require 30 or more beads in Cartesian path integral descriptions to obtain converged intramolecular energies. Rigid water models require only about five beads due to the weaker intermolecular quantum effects.
- Rigid water models appear to represent real water better than flexible classical models.
- The QFH effective potential gives good agreement with path integral results for the thermodynamic properties of water.
- Diffusion constants are enhanced with the approximate inclusion of quantum effects. Changes in the ratio of diffusion constants for water and D_2O with decreasing temperature are accurately reproduced with the QFH model. This ratio computed with the QFH model agrees well with the centroid molecular dynamics result at room temperature. Fully quantum path integral dynamical simulations of diffusion in liquid water are not presently possible.

11.11.3 Ab Initio Water

Finally, we mention recent ab initio simulations of liquid water. This promising area for fundamental studies of water thermodynamics, structure, reactivity, and dynamics is in active development. Calculations using different DFT functionals and different computational methods have resulted in quite different properties for water [36], and we make no attempt to assess these differences here. Generally, if the nuclei are propagated with classical mechanics, an overstructured liquid is observed, and the diffusion constant is much smaller than the experimental value. These differences could be due to the neglect of quantum effects on the proton motions, deficiencies in the DFT functionals, or lack of convergence in the calculations. Proton quantum effects during ab initio simulation have been included in the recent study of Chen et al. [35], but there is some question as to convergence of the simulation with the path integral discretization number P [39]. Schwegler et al. [39] found that the experimental radial distribution functions could be accurately reproduced by increasing the temperature in classical simulations by roughly 50 K for simulations using a rigid water structure and 100 K for a flexible water model.[8] They suggested that this temperature increase is consistent with path integral results using the TIP5P(PIMC) model discussed above, and thus quantum effects on proton motion are significant. The correspondence of quantum effects with a 50 K temperature rise seems quite large, but if one estimates the change in excess chemical potential from $\Delta\mu^{\mathrm{ex}} = -s_i \Delta T$, where s_i is the partial molar entropy of classical water, a 50 K temperature rise yields a chemical potential change of roughly $1 \, \mathrm{kcal \, mol^{-1}}$, consistent with previous estimates from path integral studies [60]. And, as discussed in Sect. 11.8, a reasonable estimate of the temperature correction for liquid water using (11.2) is also 50 K. This finding is further supported by recent centroid molecular dynamics simulations [64].

Allesch et al. [40] also carried out ab initio simulations of rigid water molecules and observed a decrease in water structure relative to a flexible water simulation.

---

[8] A recent ab initio study [63] has also shown that the melting point with neglect of quantum effects is elevated to nearly 400 K.

It was argued that, since the water is predominantly in the ground vibrational state at room temperature, it is not surprising that rigid models seem to more accurately mimic water properties. Calculations of the rms fluctuations in classical or quantum ground-state harmonic models call this argument into question, however; see Sect. 11.8. It is also possible that the underlying DFT functionals contain deficiencies which are difficult to disentangle from other contributions to the structure and thermodynamics [36, 37, 62]. Two of the most commonly used DFT functionals (PBE and BLYP) lead to very similar and over-structured radial distribution functions [65], while a modification of the PBE functional (rPBE) produces a less-structured fluid [36, 37].

The only estimate of the free energy of water determined from ab initio simulation is [36]. This ground-breaking work utilized data from ab initio simulation in conjunction with quasichemical theory to estimate the excess chemical potential of water in water. The quasichemical treatment partitions the problem into inner-sphere chemical effects and outer-sphere packing, electrostatic, and van der Waals effects [29]. Information theory, using occupancy statistics, yields estimates of the chemical and packing contributions. The theory can also be checked variationally by altering the inner-sphere radius. One key conclusion from these calculations is that the inner-sphere chemical effects nearly balance the outer-sphere packing effects. The estimates of the independent contributions to the chemical potential lead to helpful insights into the various contributions, and the estimated final value of $-5.1\,\mathrm{kcal\,mol^{-1}}$ is quite close to the experimental value of $-6.1\,\mathrm{kcal\,mol^{-1}}$. When a molecular-level simulation result is substituted for a simple dielectric model for the outer-sphere electrostatics, the prediction changes to $-7.5\,\mathrm{kcal\,mol^{-1}}$. With a quantum correction of roughly $0.7\,\mathrm{kcal\,mol^{-1}}$, the final prediction would be about $-6.8\,\mathrm{kcal\,mol^{-1}}$, within $k_\mathrm{B}T$ of the experimental result. The DFT functional (rPBE) used in this study is perhaps not of such a high quality to expect such remarkable agreement, but it is noteworthy that such a close estimate is possible. Quantum effects can also be expected for the proton in water, and steps toward modeling those effects are discussed in [30] and [29]. There is a lot of interest in extending ideas developed in fundamental studies of the proton in water to examine energetics and pathways for proton motion through membrane proteins [66–70].

Summary

– Ab initio simulations of water using classical propagation generally lead to an overstructured liquid compared with experiment.
– Deviations from experiment could be due to the neglect of quantum effects, overestimation of the flexibility of water in the liquid, deficiencies in the DFT functionals, or lack of convergence in the computational methods.
– Treating water as a rigid molecule in ab initio simulations leads to a destructuring of the fluid relative to classical flexible water.
– The radial distribution functions in ab initio simulations agree with experiment if the temperature is raised by roughly 50 K, consistent with results from the

TIP5P(PIMC) force field model. This implies that quantum effects are nonnegligible.

– A quasichemical theory, using ab initio simulation (rPBE functional) to generate data for the computation of the various contributions to the free energy, yields an estimate for the excess chemical potential of water very close to the experimental value.

11.11.4 Enzyme Kinetics and Proton Transport

Most of the discussion in this chapter has concerned equilibrium quantities, free energies in particular. How might the methods discussed here apply to kinetic phenomena? An area with significant current interest is the utilization of quantum methods to study the kinetics of enzyme reactions [31, 32]. The underlying theory typically employed is TST. In TST it is assumed that there is a local equilibrium along the path linking the reactants and the transition state. The central quantities in TST are the free energy of activation and the transmission coefficient. The methods discussed in this chapter are directly applicable to computing the free energies of activation in complex systems. The PDT allows a local definition of the free energy which can be followed along the progression of the reaction. Thus, in principle the free energy profile can be calculated from reactants to the transition state, or rather the ensemble of transition states.

Extensive theoretical work has been directed at understanding nuclear quantum effects on enzyme kinetics. These quantum effects enter in two ways. First, there can be a change in the system zero-point energy between the reactants and the transition state. Second, there can be tunneling effects which show up in the transmission coefficient and in the activation free energy. The Feynman–Hibbs models at least approximately account for both of these effects: the effective potential is raised in regions where the force constant is positive (potential minima) and lowered near transition states (potential maxima). If the quantum effects are large, then path integral methods can be implemented to compute more-accurate free energies along the reaction coordinate.

As an example, consider an early calculation of isotope effects on enzyme kinetics by Hwang and Warshel [31]. This study examines isotope effects on the catalytic reaction of carbonic anhydrase. The expected rate-limiting step is a proton transfer reaction from a zinc-bound water molecule to a neighboring water. The TST expression for the rate constant k is

$$k = F \frac{k_B T}{h} \exp(-\beta \Delta g^{\ddagger}) , \qquad (11.48)$$

where F is the transmission factor and $\Delta g^{\ddagger}$ is the free energy of activation. As discussed above, the free energy of activation is calculated as a difference between the reactant and transition states, and can be viewed as the difference of local excess chemical potentials with the particle centroid situated at those two locations. The major contributor to the quantum effects in proton transfer comes through the Boltzmann factor of the free energy change. The next step in the method of Hwang and

Warshel is to utilize a formula identical to (11.22) in this chapter to compute the free energy change. They employed an empirical valence bond (EVB, below) approach to approximately model electronic effects, and the calculations included the full experimental structure of carbonic anhydrase. An H/D isotope effect of 3.9 ± 1.0 was obtained in the calculation, which compared favorably with the experimental value of 3.8. This benchmark calculation gives optimism that quantum effects on free energies can be realistically modeled for complex biochemical systems.

Finally, enzyme reactions involve the making and breaking of chemical bonds. Another important physical problem involving chemical changes is proton transport in water or through membrane proteins. Classical force fields are questionable for these systems;[9] a second level of quantum treatment is then suggested in treating the electron density changes accompanying chemical transformations. Several approaches have been taken in modeling bond making and breaking: (1) empirical valence bond (EVB), (2) quantum mechanics/molecular mechanics (QM/MM), and (3) ab initio simulation methods. These three strategies are listed in order of expected increasing chemical accuracy (and thus computational cost). Calculating free energies requires substantial statistical averaging. The QM/MM and ab initio methods, especially the latter, are quite computationally costly even considering modern parallel architectures. For example, a typical ab initio molecular dynamics study at the present time simulates tens of molecules for tens of ps [71]. Thus, the ab initio simulation methods are not a practical option for modeling proton transport across a membrane channel.

In the EVB method pioneered by Warshel et al. [31] and extended by others [67], an empirical Hamiltonian is diagonalized in a basis of N states, which typically correspond to particular molecular structures in the condensed phase. The potential energy is given by the lowest eigenvalue of the Hamiltonian matrix. It is common in simulations of the proton in water to use around 10 basis states. The off-diagonal terms are parameterized to agree with higher-level calculations on structures for which classical force fields fail. The instantaneous force is computed using the Hellmann–Feynman theorem. Since each state corresponds to a distinct structure, the relative free energies for various isomers (such as the Eigen and Zundel structures) can be obtained from their relative populations during the simulation [67]. These techniques have been used to study proton transport through model hydrophobic channels [67] and more-realistic models of channels [68, 72]. An interesting observation in [67] is that the proton conduction speeds up considerably as the pore radius decreases to create a single-file chain of waters.

The basic idea of the QM/MM methods [32, 33] is to partition the system into an inner quantum zone, in which the interesting chemistry happens, and an outer classical force field region. While this division is physically sensible, it can be quite tricky to handle the boundary between the two domains. If there is no covalent bond linking the QM and MM regions, the partitioning is simpler. If the boundary cuts through chemical bonds, however, the partitioning is more difficult. Several approaches have

---

[9] Classical force fields [69] have been used to model proton transport, but their accuracy has been questioned [68].

been explored to handle the boundary. These have included link-atom and frozen-orbital approaches. It is important to allow the two subsystems to respond to each other via coupling terms in the Hamiltonian. The computational overhead can be only on the order of a factor of two greater than the cost for the electronic structure calculation on the QM region itself; typically that region may contain tens of atoms, so it is quite feasible to expect simulation studies using the QM/MM approach to begin to address problems such as proton motion through proteins in the near future. The field is rapidly developing and cannot be properly reviewed here, but the reader is referred to the above-referenced articles for technical details.

The QM/MM and ab initio methodologies have just begun to be applied to challenging problems involving ion channels [73] and proton motion through them [74]. Reference [73] utilizes Hartree–Fock and DFT calculations on the KcsA channel to illustrate that classical force fields can fail to include polarization effects properly due to the interaction of ions with the protein, and protein residues with each other. Reference [74] employs a QM/MM technique developed in conjunction with Car–Parrinello ab initio simulations [75] to model proton and hydroxide ion motion in aquaporins. Due to the large system size, the time scale for these simulations was relatively short (10 ps), but the influences of key residues and macrodipoles on the short time motions of the ions could be examined.

We can expect to see future research directed at QM/MM and ab initio simulation methods to handle these electronic structure effects coupled with path integral or approximate quantum free energy methods to treat nuclear quantum effects. These topics are broadly reviewed in [32]. Nuclear quantum effects for the proton in water have already received some attention [30, 76, 77]. Utilizing the various methods briefly described above (and other related approaches), free energy calculations have been performed for a wide range of problems involving proton motion [30, 67–69, 71, 72, 78–80].

Summary

– Methods similar to those discussed in this chapter have been applied to determine free energies of activation in enzyme kinetics and quantum effects on proton transport. They hold promise to be coupled with QM/MM and ab initio simulations to compute accurate estimates of nulcear quantum effects on rate constants in TST and proton transport rates through membranes.

11.12 Summary

The list of fluids which exhibit important quantum effects is not large. Getting back to the original question of this chapter, it is clear that for liquids like helium and hydrogen, a full quantum treatment is necessary. Liquids such as neon and water, however, show modest quantum effects which can be modeled with approximate free energy methods. The quantum correction to the free energy of water is roughly 10%

– this magnitude is large enough to warrant inclusion in calculations of free energies. As discussed in Sect. 11.1, there are experimentally observed isotope effects on the solubilities of small nonpolar molecules and biomolecules, and protein stability.

The quasichemical calculations discussed above are a first major step toward determining the relative magnitudes of the various factors contributing to the excess chemical potential of water. The factors include packing effects, chemical contributions from local interactions, electrostatics, van der Waals interactions, molecular flexibility, electronic polarization, and quantum effects. Ab initio simulation methods, although computationally challenging, remove many of the uncertainties inherent in empirical force fields. It appears that rigid water models better reproduce experimental structural properties than flexible models, but the origin of this observation is not entirely clear [40]. An alternative to the rigid model would be to compute an intramolecular potential of mean force from the ground-state vibrational wave function of water and include this potential during classical propagation; this approach is consistent with the fact that it is easy to generate the exact intramolecular partition function q_α^{int}. Adding quantum effects for flexible water with path integral calculations requires handling two very different energy-scale quantum effects on the same footing. But the *inter*molecular quantum effects are modest, and in this chapter we have discussed evidence that the QFH approach can handle those effects quite well. Thus, it would be interesting to see free energy computations performed using a combination of ab initio simulation along with the approximate quantum models. Establishing quantitative conclusions concerning the factors contributing to the excess chemical potential of water is a major challenge for molecular fluid free energy calculations. An even bigger challenge is extending the quantum mechanical methods discussed here to problems as complex as biomolecule solvation and enzyme kinetics.

At a practical level, what is the current status of methods for studying quantum effects on condensed-phase free energies? If the quantum effects are relatively large, path integral methods are required. These techniques are mature, and the convergence of the calculations with increasing numbers of quantum degrees of freedom can easily be monitored. As long as an underlying classical interaction potential is employed, the additional computational cost is directly proportional to the number of variables needed to describe the paths. Thus, systems with hundreds of atoms can be handled on single workstations. The combination of path integral simulation with ab initio DFT methods is extremely challenging, however, so systems with only tens of water molecules can be modeled for tens of ps. This is a frontier methods development problem for the computer simulation of liquids. Progress will involve both the further development of linear scaling algorithms for ab initio DFT and increased computer power. Improvements in DFT potentials should proceed in parallel with the development of more-efficient numerical methods.

The effective potentials described in this chapter, on the other hand, are suited to relatively weak intermolecular quantum effects and require only a slight additional computational overhead – more terms in the potential – relative to routine classical simulations. Therefore, systems with tens of thousands of atoms can readily be modeled. This makes possible the large-scale simulation of biomolecule solvation with

the inclusion of quantum effects. If the intermolecular effective potentials ride atop a classical[10] ab initio DFT simulation the overall cost should be comparable to the purely classical DFT modeling, but that approach has not been worked out yet.

To reiterate, a main obstacle to overcome is a useful partitioning of quantum effects into intra- and intermolecular contributions during the ab initio simulation of molecular fluids with minor quantum effects. Ab initio simulation of a liquid like water is necessary to treat the complex charge redistribution effects and perhaps chemical reactions which may occur in the condensed phase. And quantum effects cannot be entirely neglected since they have a significant magnitude. Therefore, development of new computational methods for this partitioning should open the door to the quantitative modeling of aqueous solutions and their interactions with biomolecules.

References

1. Feynman, R. P.; Hibbs, A. R., *Quantum Mechanics and Path Integrals*, McGraw-Hill: New York, 1965
2. Feynman, R. P., *Statistical Mechanics*, Benjamin/Cummings: London, 1972
3. Chandler, D.; Wolynes, P. G., Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids, *J. Chem. Phys.* **1981**, *74*, 4078–4095
4. Ceperley, D. M., Path integrals in the theory of condensed helium, *Rev. Mod. Phys.* **1995**, *67*, 279–355
5. Parrinello, M.; Rahman, A., Study of an F center in molten KCl, *J. Chem. Phys.* **1984**, *80*, 860–867
6. Laria, D.; Chandler, D., Comparative study of theory and simulation calculations for excess electrons in simple fluids, *J. Chem. Phys.* **1987**, *87*, 4088–4092
7. Marchi, M.; Sprik, M.; Klein, M. L., Calculation of the free energy of electron solvation in liquid ammonia using a path integral quantum Monte Carlo simulation, *J. Phys. Chem.* **1988**, *92*, 3625–3629
8. Wang, Q.; Johnson, J. K.; Broughton, J. Q., Thermodynamic properties and phase equilibrium of fluid hydrogen from path integral simulations, *Mol. Phys.* **1996**, *89*, 1105–1119
9. Wang, Q.; Johnson, J. K.; Broughton, J. Q., Path integral grand canonical Monte Carlo, *J. Chem. Phys.* **1997**, *107*, 5108–5117
10. Poulsen, J. A.; Nyman, G.; Rossky, P. J., Quantum diffusion in liquid para-hydrogen: an application of the Feynman–Kleinert linearized path integral approximation, *J. Phys. Chem. B* **2004**, *108*, 19799–19808
11. Sese, L. M., A quantum Monte Carlo study of liquid Lennard-Jones methane, path-integral and effective potentials, *Mol. Phys.* **1992**, *76*, 1335–1346
12. Tchouar, N.; Ould-Kaddur, F.; Levesque, D., Computation of the properties of liquid neon, methane, and gas helium at low temperature by the Feynman–Hibbs approach, *J. Chem. Phys.* **2004**, *121*, 7326–7331
13. Thirumalai, D.; Hall, R. W.; Berne, B. J., A path integral Monte Carlo study of liquid neon and the quantum effective pair potential, *J. Chem. Phys.* **1984**, *81*, 2523–2527

---

[10] Here classical refers to the nuclear propagation.

14. Morales, J. J.; Singer, K., Path integral simulation of the free energy of Lennard-Jones neon, *Mol. Phys.* **1991**, *73*, 873–880

15. Sese, L. M., Feynman–Hibbs quantum effective potentials for Monte Carlo simulations of liquid neon, *Mol. Phys.* **1993**, *78*, 1167–1177

16. Ortiz, V.; Lopez, G. E., Fourier path integral Monte Carlo study of a two-dimensional model quantum monolayer, *Mol. Phys.* **2002**, *100*, 1003–1009

17. Sese, L. M., Path integral and effective potential Monte Carlo simulations of liquid nitrogen, hard-sphere and Lennard-Jones potentials, *Mol. Phys.* **1991**, *74*, 177–189

18. Miller, T. F., III; Clary, D. C., Torsional path integral Monte Carlo method for calculating the absolute quantum free energy of large molecules, *J. Chem. Phys.* **2003**, *119*, 68–76

19. Srinivisan, J.; Volobuev, Y. L.; Mielke, S. L.; Truhlar, D. G., Parallel Fourier path-integral Monte Carlo calculations of absolute free energies and chemical equilibria, *Comput. Phys. Commun.* **2000**, *128*, 446–464

20. Doll, J. D.; Beck, T. L.; Freeman, D. L., Equilibrium and dynamical Fourier path integral methods, *Adv. Chem. Phys.* **1990**, *78*, 61–127

21. Runge, K. J.; Chester, G. V., Solid–fluid phase transition of quantum hard spheres at finite temperature, *Phys. Rev. B* **1988**, *38*, 135–162

22. Barrat, J.-L.; Loubeyre, P.; Klein, M. L., Isotopic shift in the melting curve of helium: a path integral Monte Carlo study, *J. Chem. Phys.* **1989**, *90*, 5644–5650

23. Li, D.; Voth, G. A., A path integral Einstein model for characterizing the equilibrium states of low temperature solids, *J. Chem. Phys.* **1992**, *96*, 5340–5353

24. Liu, A.; Beck, T. L., Determination of excess Gibbs free energy of quantum mixtures by MC path integral simulations, *Mol. Phys.* **1995**, *86*, 225–233

25. Guillot, B.; Guissani, Y., Quantum effects in simulated water by the Feynman–Hibbs approach, *J. Chem. Phys.* **1998**, *108*, 10162–10174

26. Ben-Naim, A.; Marcus, Y., Solvation thermodynamics of nonionic solutes, *J. Chem. Phys.* **1984**, *81*, 2016–2027

27. Gripon, C.; Legrand, L.; Rosenman, I.; Vidal, O.; Robert, M. C.; Boue, F., Lysozyme solubility in H_2O and D_2O solutions: a simple relationship, *J. Cryst. Growth* **1997**, *177*, 238–247

28. Bonnete, F.; Madern, D.; Zaccai, G., Stability against denaturation mechanisms in halophilic malate dehydrogenase "adapt" to solvent connditions, *J. Mol. Biol.* **1994**, *244*, 436–447

29. Beck, T. L.; Paulaitis, M. E.; Pratt, L. R., *The Potential Distribution Theorem and Models of Molecular Solutions*, Cambridge University Press: New York, 2006

30. Lobaugh, J.; Voth, G. A., The quantum dynamics of an excess proton in water, *J. Chem. Phys.* **1996**, *104*, 2056–2069

31. Hwang, J.-K.; Warshel, A., How important are quantum mechanical nuclear motions in enzyme catalysis, *J. Am. Chem. Soc.* **1996**, *118*, 11745–11751

32. Gao, J.; Truhlar, D. G., Quantum mechanical methods for enzyme kinetics, *Ann. Rev. Phys. Chem.* **2002**, *53*, 467–505

33. Friesner, R. A.; Guallar, V., Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis, *Ann. Rev. Phys. Chem.* **2005**, *56*, 389–427

34. Mahoney, M. W.; Jorgensen, W. L., Quantum, intramolecular flexibility, and polarizability effects on the reproduction of the density anomaly of liquid water by simple potential functions, *J. Chem. Phys.* **2001**, *115*, 10758–10768

35. Chen, B.; Ivanov, I.; Klein, M. L.; Parrinello, M., Hydrogen bonding in water, *Phys. Rev. Lett.* **2003**, *91*, 215503

36. Asthagiri, D.; Pratt, L. R.; Kress, J. D., Free energy of liquid water on the basis of quasichemical theory and ab initio molecular dynamics, *Phys. Rev. E* **2003**, *68*, 041505

37. Fernandez-Serra, M. V.; Ferlat, G.; Artacho, E., Two exchange-correlation functionals compared for first-principles liquid water, *Los Alamos Eprint archive: cond-mat/0407724*, 2004

38. Kuo, I.-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M., Liquid water from first principles: investigation of different sampling approaches, *J. Phys. Chem. B* **2004**, *108*, 12990–12998

39. Schwegler, E.; Grossman, J. C.; Gygi, F.; Galli, G., Towards an assessment of the accuracy of density functional theory for first principles simulations of water II, *J. Chem. Phys.* **2004**, *121*, 5400–5409

40. Allesch, M.; Schwegler, E.; Gygi, F.; Galli, G., A first principles simuation of rigid water, *J. Chem. Phys.* **2004**, *120*, 5192

41. Widom, B., Some topics in the theory of fluids, *J. Chem. Phys.* **1963**, *39*, 2808–2812

42. Widom, B., Potential-distribution theory and the statistical mechanics of fluids, *J. Phys. Chem.* **1982**, *86*, 869–872

43. Landau, L. D.; Lifshitz, E. M., *Statistical Physics*, (3rd edition, part 1), 1980

44. Stratt, R. M., Semiclassical statistical mechanics of fluids: nonperturbative incorporation of quantum effects in classical many body models, *J. Chem. Phys.* **1979**, *70*, 3630–3638

45. Kleinert, H., *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, World Scientific: Singapore, 1995

46. Coalson, R. D., On the connection between Fourier coefficient and discretized Cartesian path integration, *J. Chem. Phys.* **1986**, *85*, 926

47. Pratt, L. R., A statistical method for identifying transition states in high dimensional problems, *J. Chem. Phys.* **1986**, *85*, 5045–5048

48. Beck, T. L., Quantum path integral extension of Widom's test particle method for chemical potentials with application to isotope effects on hydrogen solubilities in model solids, *J. Chem. Phys.* **1992**, *96*, 7175–7177

49. Beck, T. L.; Marchioro, T. L., The quantum potential distribution theorem, in *Path Integrals from meV to MeV: Tutzing 1992*, Grabert, H.; Inomata, A.; Schulman, L.; Weiss, U., Eds., World Scientific: Singapore, **1993**, pp. 238–243

50. van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*, Elsevier: New York, 1992

51. Jarzynski, C., Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **1997**, *78*, 2690–2693

52. Roepstorff, G., *Path Integral Approach to Quantum Physics*, Springer: Berlin, Heidelberg, New York, 1994

53. Predescu, C., The partial averaging method, *J. Math. Phys.* **2003**, *44*, 1226–1239

54. Lobaugh, J.; Voth, G. A., A quantum model for water: equilibrium and dynamical properties, *J. Chem. Phys.* **1997**, *106*, 2400–2410

55. de la Pena, L. Hernandez; Kusalik, P. G., Quantum effects in light and heavy water: a rigid-body centroid molecular dynamics study, *J. Chem. Phys.* **2004**, *121*, 5992–6002

56. Stern, H. A.; Berne, B. J., Quantum effects in liquid water: path-integral simulations of a flexible and polarizable ab initio model, *J. Chem. Phys.* **2001**, *115*, 7622

57. Gray, C. G.; Gubbins, K. E., *Thoery of Molecular Fluids. Volume 1: Fundamentals*, Oxford University Press: Oxford, 1984

58. Predescu, C.; Doll, J. D., Optimal series representations for numerical path integral simulations, *J. Chem. Phys.* **2003**, *117*, 7448–7463

59. Mielke, S. L.; Truhlar, D. G., A new Fourier path integral method, a more general scheme for extrapolation, and comparison of eight path integral methods for the quantum mechanical calculation of free energies, *J. Chem. Phys.* **2001**, *114*, 621–630

60. Kuharski, R. A.; Rossky, P. J., A quantum mechanical study of structure in liquid H_2O and D_2O, *J. Chem. Phys.* **1985**, *82*, 5164–5177

61. Wallqvist, A.; Berne, B. J., Path-integral simulation of pure water, *Chem. Phys. Lett.* **1985**, *117*, 214

62. Goldman, N.; Leforestier, C.; Saykally, R. J., A 'first principles' potential energy surface for liquid water from VRT spectroscopy of water clusters, *Philos. Trans. R. Soc. A* **2005**, 1–16. doi:10.1098/rsta.2004.1504

63. Sit, P.; Marzari, N., Static and dynamical properties of heavy water at ambient conditions from first-principles molecular dynamics, *Los Alamos Eprint Server* **2005**. cond-mat/0504146

64. de la Pena, L. Hernandez; Kusalik, P. G., Temperature dependence of quantum effects in liquid water, *J. Am. Chem. Soc.* **2005**, *127*, 5246–5251

65. Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G., Towards an assessment of the accuracy of density functional theory for first principles simulations of water, *J. Chem. Phys.* **2004**, *120*, 300–311

66. Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K., Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin, *Biophys. J.* **2002**, *83*, 3097–3112

67. Brewer, M. L.; Schmitt, U. W.; Voth, G. A., The formation and dynamics of proton wires in channel environments, *Biophys. J.* **2001**, *80*, 1691–1702

68. Wu, Y.; Voth, G. A., A computer simulation study of the hydrated proton in a synthetic proton shannel, *Biophys. J.* **2003**, *85*, 864–875

69. Chakrabarti, N.; Roux, B.; Pomes, R., Structural determinants of proton blockage in aquaporins, *J. Mol. Biol.* **2004**, *343*, 493–510

70. Yin, J.; Kuang, Z.; Mahankali, U.; Beck, T. L., Ion transit pathways and gating in ClC chloride channels, *Proteins: Struct. Funct. Bioinform.* **2004**, *57*, 414–421

71. Asthagiri, D.; Pratt, L. R.; Kress, J. D., Ab initio molecular dynamics and quasichemical study of H^+(aq), *Proc. Natl Acad. Sci.* **2005**, *102*, 6704–6708

72. Burykin, A.; Warshel, A., What really prevents proton transport through aquaporin? Charge self-energy versus proton wire proposals, *Biophys. J.* **2003**, *85*, 3696–3706

73. Bliznyuk, A. A.; Rendell, A. P., Electronic effects in biomolecular simulations: investigations of the KcsA potassium ion channel, *J. Phys. Chem. B* **2004**, *108*, 13866–13873

74. Jensen, M. O.; Rothlisberger, U.; Rovira, C., Hydroxide and proton migration in aquaporins, *Biophys. J.* **2005**, *89*, 1744–1759

75. Laio, A.; VandeVondele, J.; Rothlisberger, U., A Hamiltonian electrostatic coupling scheme for hybrid Car–Parrinello molecular dynamics simulations, *J. Chem. Phys.* **2002**, *116*, 6941–6947

76. Pomes, R.; Roux, B., Theoretical study of H^+ translocation along a model proton wire, *J. Phys. Chem.* **1996**, *100*, 2519–2527

77. Pomes, R.; Roux, B., Free energy profiles for H^+ conduction along hydrogen-bonded chains of water molecules, *Biophys. J.* **1998**, *75*, 33–40

78. Zahn, D.; Brickmann, J., Quantum-classical simulation of proton transport via a phospholipid bilayer, *Phys. Chem. Chem. Phys.* **2001**, *3*, 848–852

79. deGroot, B. L.; Frigato, T.; Helms, V.; Grubmuller, J., The mechanism of proton exclusion in the aquaporin-1 water channel, *J. Mol. Biol.* **2003**, *333*, 279–293

80. Pomes, R.; Roux, B., Molecular mechanism of H^+ conduction in the single-file water chain of the gramicidin channel, *Biophys. J.* **2002**, *82*, 2304–2316

12

Free Energy Calculations: Approximate Methods for Biological Macromolecules

Thomas Simonson

12.1 Introduction

In this chapter we present the most important simplified free energy methods in use today and the main biological problems that have motivated their development. The area in which these methods are perhaps the most valuable is the study of molecular recognition between biological molecules, such as an enzyme and a substrate or inhibitor. Noncovalent association between biomolecules is a key element of the biochemistry and information flow in living systems. Many competing effects can contribute to receptor–ligand binding [1]: changes in rotational, translational, conformational, and vibrational entropy of the partners, entropy changes associated with solvent ordering around hydrophobic or charged groups, solute conformational strain, changes in electrostatic and van der Waals interactions within and between the partners and the solvent, counterion reorganization. Experimental studies often combine structure determination methods with point mutagenesis and thermodynamic measurements to obtain information on the binding [2]. However, there are considerable difficulties in the experimental analysis of longer-range electrostatic contributions, the cooperativity between amino acid residues of a protein, or disordered solvent, for example. Such effects can be determined using rigorous free energy simulations, described in the earlier chapters of this book. They can also be incorporated, at different levels of accuracy, into simplified free energy methods.

The basic principle of a receptor–ligand binding analysis by free energy simulations is explained in Fig. 12.1 [3]. Most applications focus on binding free energy differences between a series of ligands or protein mutants. For a review of the relevant statistical thermodynamics see [1]. Like experiments, the rigorous free energy simulation method requires a reversible (or near-reversible [4, 5]) path between the initial and final states. The Helmholtz free energy change along the horizontal legs in Fig. 12.1 can be obtained, for example, by a thermodynamic integration [6] (see also Chap. 4)

$$\Delta A(0 \to 1) = \int_0^1 \frac{\mathrm{d}A}{\mathrm{d}\lambda}\,\mathrm{d}\lambda = \int_0^1 \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda. \qquad (12.1)$$
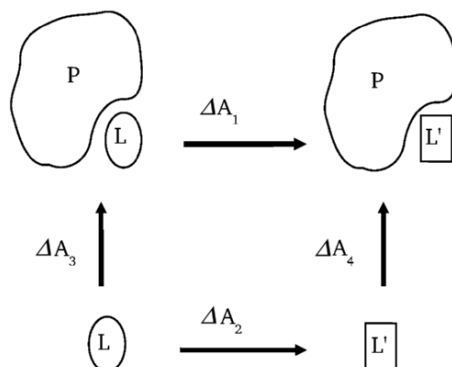
Fig. 12.1. Thermodynamic cycle for ligand binding. Solutes L and L' in solution (*below*) and bound to the receptor P (*above*). Vertical legs correspond to the binding reactions. Horizontal legs correspond to the alchemical transformation of L into L'. The binding free energy difference can be obtained from either route: $\Delta\Delta A = \Delta A_4 - \Delta A_3 = \Delta A_1 - \Delta A_2$

U represents the energy function, which depends on λ, and the brackets represent an average over the ensemble corresponding to $U(\lambda)$. At either endpoint, $\lambda = 0$ or 1, the energy function is that of the native or mutant state; intermediate values correspond to 'alchemical' states (see Chap. 2). λ is referred to as a coupling parameter. The double free energy difference $\Delta\Delta A$ (Fig. 12.1) can also be obtained from the vertical legs of the cycle, which correspond to 'chemical,' i.e., binding reactions, but the simulations are usually more difficult and costly [7–9]. For either choice of pathway, molecular dynamics simulations are usually employed. We refer to this approach as the MDFE (molecular dynamics free energy) method, bearing in mind that Monte Carlo is sometimes used in place of MD.

The rigorous MDFE approach has two main drawbacks. First, it is a complex technique (as illustrated by this book), so that significant expertise is needed, especially for applications to proteins. Second, the need to simulate intermediate points along a complete pathway between any two systems of interest makes the method computer-intensive. This is especially serious when one is interested in a large set of ligands binding to a particular receptor, as is the case in a typical drug design project. Therefore, a great deal of effort has been put into the development of more-efficient techniques that can be more approximate, as long as they give enough accuracy to allow the selection of a subset of ligands for further study, perhaps with a high-level MDFE protocol. An extreme example of this strategy is the use of high-throughput virtual screening, where thousands or millions of potential ligands are considered, using a very fast and approximate 'scoring' function to evaluate receptor binding [10]. Here, we focus on an intermediate area, closer to MDFE than to high-throughput screening. The free energy difference between two systems is computed from one or two simulations, such as simulations of a protein–ligand complex and of a ligand or protein alone in solution (the two ends of the vertical legs in Fig. 12.1).

We present and analyze the most important simplified free energy methods, emphasizing their connection to more-rigorous methods and the underlying theoretical framework. The simplified methods can all be superficially defined by their use of just one or two simulations to compare two systems, as opposed to many simulations along a complete connecting pathway. More importantly, the use of just one or two simulations implies a common approximation of a near-linear response of the system to a perturbation. Another important theme for simplified methods is the use, in many cases, of an implicit description of solvent: usually a continuum dielectric model, often supplemented by a simple description of hydrophobic effects [11].

To illustrate these methods, we consider the main biological problems that have motivated their development. The problems that have received the most attention are the receptor–ligand binding problem [12–16] and the calculation of proton binding affinities (pK_a shifts) [17–20]. The methods described can also be applied to many related problems, such as redox protein behavior, protein–protein association, protein folding, or membrane insertion.

We begin by recalling briefly the basic equations of free energy perturbation theory, including approximate perturbation formulae and equations for ligand binding. As an application, we describe several 'single-step' perturbation methods for ligand binding. Next, we give a short review of linear response theory and its application to proton binding. We discuss the physical basis of widely used implicit solvent models, including models to describe hydrophobic effects and continuum electrostatic models. Then, as a first class of applications, we describe 'linear interaction methods,' which have become popular for the receptor–ligand binding problem, and which can be performed with either an explicit or an implicit solvent treatment. As a second class of applications, we consider receptor–ligand binding studies with the 'MM/PBSA' method, which usually uses a dielectric continuum solvent. We describe some recent protein–ligand binding studies. We analyze the calculation of pK_a shifts in some detail, as an important and illuminating example of the possibilities and limitations of continuum electrostatic models.

12.2 Thermodynamic Perturbation Theory and Ligand Binding

12.2.1 Obtaining Thermodynamic Perturbation Formulas

Free energy calculations rely on a well-known thermodynamic perturbation theory [6, 21, 22], which is recalled in Chap. 2. We consider a molecular system, described by the potential energy function $U(r^N)$, which depends on the coordinates of the N atoms: $r^N = (r_1, r_2, \ldots, r_N)$. The system could be a biomolecule in solution, for example. We limit ourselves to a classical mechanical description, for simplicity. Practical calculations always consider differences between two or more similar systems, such as a protein complexed with two different ligands. Therefore, we consider a change in the system, such that the potential energy function becomes:

$$U' = U + V, \tag{12.2}$$

where V is a new, 'perturbing,' potential energy term. The corresponding Helmholtz free energy change can be arranged to read [23, 24]

$$A' - A = -k_{\mathrm{B}}T \ln \left\langle \exp\left(-\frac{V}{k_{\mathrm{B}}T}\right)\right\rangle \tag{12.3}$$

$$= \langle V \rangle - k_{\mathrm{B}}T \ln \left\langle \exp\left(-\frac{\Delta V}{k_{\mathrm{B}}T}\right)\right\rangle . \tag{12.4}$$

The brackets on the right indicate an average over the ensemble of the starting system, i.e., with Boltzmann weights $\exp(-\frac{U}{k_{\mathrm{B}}T})$, and $\Delta V = V - \langle V \rangle$. Expanding in powers of $\delta V/k_{\mathrm{B}}T$ gives free energy perturbation formulas. While higher-order terms are difficult to calculate because of sampling problems, expansions to low orders (one to four) are often more robust numerically than the original formula (12.3), and are especially useful for treating many small perturbations of a single reference system [25]. Since $\langle e^{-\delta V/k_{\mathrm{B}}T}\rangle$ has the form of a moment-generating function [26], the coefficients of the expansion involve the cumulants C_n of δV:

$$A' - A = \langle V \rangle - k_{\mathrm{B}}T \sum_{i=2}^{\infty} \frac{C_n}{n!} \left(\frac{-1}{k_{\mathrm{B}}T}\right)^n . \tag{12.5}$$

The cumulants [26] are simple functions of the moments of the probability distribution of δV: $C_2 = \langle (V - \langle V \rangle)^2 \rangle$, $C_3 = \langle (V - \langle V \rangle)^3 \rangle$, $C_4 = \langle (V - \langle V \rangle)^4 \rangle - 3C_2^2$, etc. Truncation of the expansion at order two corresponds to a linear-response approximation (see later), and is equivalent to assuming V is Gaussian (with zero moments and cumulants beyond order two). To this order, the mean and width of the distribution determine the free energy; to higher orders, the detailed shape of the distribution contributes.

12.2.2 Ligand Binding: General Framework

The study of receptor–ligand binding is one of the most important applications of free energy simulations [1]. To approach this problem theoretically, one must first partition the conformational space into *bound* and *unbound* states. There is no unique way to do this, but in practical situations there is often a natural choice. The equilibrium binding constant is

$$K_b = \frac{c_{\mathrm{RL}}}{c_{\mathrm{R}} c_{\mathrm{L}}}, \tag{12.6}$$

where c_{R}, c_{L}, and c_{RL} are the concentrations (or number densities) of receptor, ligand, and complex, and K_b has units of volume. The chemical potential of each species in solution is [1, 23]

$$\mu_{\mathrm{S}} = k_{\mathrm{B}}T \ln \frac{c_{\mathrm{S}}}{c^o} - k_{\mathrm{B}}T \ln \frac{Z_{\mathrm{S}}}{Z_0 V c^o}, \tag{12.7}$$

where $S = RL$, R, or L; c^o is the standard state concentration, V the volume of the system, Z_S the partition function of S in solution, and Z_0 the partition function of the solution without S. The condition for equilibrium is

$$-k_B T \ln K_b \rho^o = \Delta S_b^o, \tag{12.8}$$

where $\Delta A_b^o = \mu_{RL}^o - \mu_R^o - \mu_L^o$ is the standard binding free energy – the free energy to bring two single molecules R and L together to form a complex RL when the concentrations of all species are fixed at c^o.

To relate the standard binding free energy to free energies that can be obtained from simulations, we use

$$\Delta A_b^o = -k_B T \ln \frac{Z_{RL} Z_0 V \rho^o}{Z_R Z_L} = -k_B T \ln \frac{Q_{RL} Q_0 V \rho^o}{Q_R Q_L}$$
$$= -k_B T \ln \frac{Q_{RL} \rho^o}{Q_R Q_{L0}/V} + k_B T \ln \frac{Q_L}{Q_{L0} Q_0}, \tag{12.9}$$

where the second equality takes into account a cancelation of the velocity partition functions and Q_{L0} is the configuration integral of the ligand alone (i.e., in the gas phase). The second term in (12.9) is the free energy to 'annihilate' L in solution, i.e., the free energy to reversibly turn off its interactions with the surrounding solution, effectively transferring it to the gas phase. The first term is the free energy to 'annihilate' the ligand in the binding site, with its center of mass fixed [27, 28]. The standard concentration ρ^o appears explicitly here. This free energy takes the form of an average over all positions in the active site (see [28]).

Many applications are only concerned with binding free energy differences. Comparing the binding of two ligands, L and L', to the receptor R, we have

$$\Delta \Delta A_b^o(L, L') = \Delta A_b^o(RL') - \Delta A_b^o(RL) = -k_B T \ln \frac{Z_{RL'}}{Z_{RL}} + k_B T \ln \frac{Z_{L'}}{Z_L}. \tag{12.10}$$

Thus, the standard state concentration cancels from the double free energy difference. The calculation can be done by mutating L to L' both in the complex and in solution (the horizontal legs of Fig. 12.1).

12.2.3 Applications of Thermodynamic Perturbation Formulas

Ligand Binding

In the early days of protein free energy calculations, computational efficiency was extremely important. Thermodynamic perturbation formulas were viewed as a promising route toward time-saving schemes, because free energy differences could be obtained (in principle) from a single simulation of a reference system (12.3). For example, one of the earliest ligand binding free energy studies considered two small molecule inhibitors binding to trypsin: benzamidine and parafluorobenzamidine [29]. The difference in binding free energies was obtained by transforming one ligand

into the other, in solution and in complex with the protein. Only simulations of the reference systems, benzamidine and the benzamidine:trypsin complex, were performed. The perturbation corresponds to changing the parahydrogen in the benzamidine ring into a fluorine. This is accomplished by specific changes in a few force-field parameters, reflected in a perturbing term V in the energy function. The Zwanzig perturbation formula (12.3) was applied, giving a result in reasonable agreement with experiment for this small chemical change.

Systematic Sensitivity Analysis

This idea can be expanded on, using a single simulation of the reference molecule to explore many perturbations at once. For example, one could explore many small modifications of a lead molecule in a drug design project. The information obtained can then be used to guide further simulations. One implementation of this idea is a standard engineering technique known as 'systematic sensitivity analysis' [15, 30]. A protein–ligand complex is simulated, and the derivatives of the free energy are computed with respect to parameters of interest, such as individual atomic charges or van der Waals radii. The key parameters that determine the binding free energy are thus identified, and used to guide the design of improved ligands. The coupling between system parameters is also of interest. For example, for protein–ligand interactions, one would like to identify groups on the protein and ligand that interact favorably. These can be characterized by appropriate second derivatives of the free energy, $\partial^2 A/\partial \lambda_i \lambda_j$, where λ_i and λ_j are force field parameters (charges, radii) corresponding to two atoms i and j. The derivatives are easily calculated

$$\frac{\partial A}{\partial \lambda} = \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_\lambda \tag{12.11}$$

$$\frac{\partial^2 A}{\partial \lambda_i \partial \lambda_j}(\lambda_1, \lambda_2, ...) = \left\langle \frac{\partial^2 U(r^N; \lambda_1, \lambda_2, \dots)}{\partial \lambda_i \partial \lambda_j} \right\rangle_{\lambda_1, \lambda_2, ...} \tag{12.12}$$

$$- \frac{1}{k_{\mathrm{B}}T} \left[\left\langle \frac{\partial U}{\partial \lambda_i} \frac{\partial U}{\partial \lambda_j} \right\rangle_{\lambda_1, \lambda_2, ...} - \left\langle \frac{\partial U}{\partial \lambda_i} \right\rangle_{\lambda_1, \lambda_2, ...} \left\langle \frac{\partial U}{\partial \lambda_j} \right\rangle_{\lambda_1, \lambda_2, ...} \right].$$

In these equations, the λ_i play the role of coupling parameters.

Single-Step Perturbations to Multiple Ligands

The same technique can be used in some cases to obtain accurate estimates of binding free energy differences for a set of ligands of interest [25, 31–34]. The molecule taken as the reference need not be a real molecule. Indeed, the reference molecule could be 'intermediate' between a large set of molecules of interest, so that conformations that are sufficiently representative of them all are sampled in the reference simulation. The justification for this approach is discussed in detail in Chap. 6. To achieve this for a variety of substituted phenols, Liu et al. [25] added dummy atoms to the ring at the sites they wished to substitute. Such dummy atoms can be 'softer'

than the real substituents one wants to consider, as if the substituents were 'half present.' Indeed, the free energy to introduce a new particle or delete an existing one is a sharply varying function, and it is difficult to do an 'all-to-nothing' extrapolation. Suppose the substituent is removed by applying a coupling parameter λ to a substituent–solvent van der Waals interaction energy: $u(r) = \lambda\epsilon(\frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6})$, and letting λ go to zero. Using diagram techniques from liquid theory, one can show that the free energy derivative scales as $\lambda^{-3/4}$ when $\lambda \to 0$, while the free energy scales as $\lambda^{1/4}$ [35, 36]. To avoid the resulting singularity and facilitate the introduction/removal of a large variety of substituents, Liu et al. used soft-core van der Waals potentials for selected atoms

$$U(r_{ij}) = \epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{(r_{ij}^6 + \alpha\sigma_{ij}^6)^2} - \frac{\sigma_{ij}^6}{r_{ij}^6 + \alpha\sigma_{ij}^6} \right). \tag{12.13}$$

Here, r_{ij} is the distance between a *softened* atom i on the reference molecule and another atom j; ϵ_{ij} and σ_{ij} are the van der Waals interaction parameters for this pair, and α is a parameter that softens the potential and prevents it from going to infinity at very short distances. Recently, the method was applied to polychlorinated biphenyls (PCBs) binding to the estrogen receptor [33, 34], and to protein-kinase inhibitors [37]. The method can yield reasonable qualitative results when functional groups up to about three atoms are deleted [38]. For larger transformations, the sampling problems associated with creating or annihilating atoms become too large.

λ-Dynamics

A further extension is to allow the reference molecule to wander freely within a predefined chemical space, and change spontaneously in whichever direction it prefers. This idea has been implemented by simulating many potential ligands simultaneously. Each ligand i is associated with its own coupling parameter or weight, λ_i, and with a term $\lambda_i U_i$ in the energy function. The coupling parameters are included in the simulation as coordinates participating in the molecular dynamics, with artificial masses, akin to 'pseudoparticles' [39, 40]. Because of this, the method has been referred to as λ-dynamics. The different weights obey $\sum_i \lambda_i = 1$. As the system evolves, the weights tend to adjust spontaneously in such a way that the most favorable ligand has the largest weight. Alternatively, the ligands can be made equiprobable by incorporating their free energies A_i into the energy function: each term $\lambda_i U_i$ is replaced by $\lambda_i(U_i - A_i)$. A_i is not known ahead of time, but can be determined iteratively [39]. This provides a new route to determine the relative solvation or binding free energies of two or more ligands, which was found to be more efficient than traditional thermodynamic perturbation or integration approaches in applications to simple systems. The variation of the λ_i with time implies that the system is never truly at equilibrium; to limit this effect, sufficient pseudomasses are needed for the λ_i.

As an example, consider a solution mixture of two molecules, 1 and 2. The system is described by the hybrid potential energy function:

$$U = U_0 + \lambda U_1 + (1 - \lambda)U_2, \qquad (12.14)$$

where λ is a coupling parameter (treated as a coordinate with an associated pseudo mass). U_1 (respectively, U_2) describes the interactions of molecule 1 (respectively, 2) with the solvent; U_0 describes solvent–solvent interactions. At equilibrium, the mean weights of 1 and 2 can be interpreted as relative concentrations, which obey the law of mass action [39, 40]

$$\frac{1 - \langle \lambda \rangle}{\langle \lambda \rangle} = \exp(-A_{12}/k_{\mathrm{B}}T). \qquad (12.15)$$

A_{12} is the free energy to transform 1 into 2 in solution. If we replace the U_i by $(U_i - A_i)$, we effectively change the nature of the solutes so that the transformation free energy A_{12} is now zero. The equilibrium value of $\langle \lambda \rangle$ is then seen to be 1/2: the two solutes have the same average population.

Electrostatic Perturbations

Single-step perturbation methods have also been applied to electrostatic processes. One study probed the dielectric properties of several proteins at a microscopic level [41, 42]. Test charges were inserted at many different positions within or around each protein, and a dielectric relaxation free energy was computed, which is related to a microscopic dielectric susceptibility (see Sect. 12.3).

Other electrostatic processes studied include proton binding [43] and changing the molecular charge distribution [44]. The free energy expansion formula (12.5) was used, including terms up to second order

$$\Delta A = \langle V \rangle - \frac{1}{2k_{\mathrm{B}}T} \langle (V - \langle V \rangle)^2 \rangle. \qquad (12.16)$$

This second-order, linear response approximation gave good accuracy for comparing the ground state and two excited states of tryptophan in solution [45]. The problem of proton binding to proteins [43] is more difficult. Recent studies of several proteins have shown that, for proton binding or electron binding, a second-order expansion of the free energy can be fairly accurate, but the sampling in a typical simulation is not sufficient to determine accurately the mean and especially the variance of V, which determine the expansion coefficients [46–48]. This makes it impossible to extrapolate the free energy accurately. For quantitative estimates of proton binding free energies, it is probably necessary in most cases to use simulations of two states, at least; preferably the initial and final states for the transformation of interest. This two-state, linear response approach is an important practical tool that is presented in the next section.

12.3 Linear Response Theory and Free Energy Calculations

12.3.1 Linear Response Theory: The General Framework

The dielectric response of a solvated protein to a perturbing charge, such as a redox electron or a titrating proton, is related to the equilibrium fluctuations of the unperturbed system through linear response theory [49, 50]. In the spirit of free energy

simulations, let us gradually introduce a perturbing charge density ρ_p, with the help of a coupling parameter λ that we vary gradually from zero to one. For a given value of λ, the perturbing charge density is $\lambda\rho_p$. It contributes a term

$$\Delta U = \lambda \int \frac{\rho(\mathbf{r})\rho_p(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{12.17}$$

to the Hamiltonian, where ρ represents the charge density of the protein and solvent (everything except the perturbing charge). In practice, the charge density will usually be located on atoms, so that $\mathbf{r}$ and $\mathbf{r}'$ will be atomic positions. This term can actually be rearranged into a more compact, vector form, which is better suited to our discussion

$$\Delta U = -\lambda \, \mathbf{f_p} \cdot \mathbf{P}. \tag{12.18}$$

Here, $\mathbf{f_p}$ is the field due to ρ_p; $\mathbf{P}$ is a polarization density, which is related to the electric field $\mathbf{E}$ produced by the remaining charge density ρ: $\mathbf{E} = -4\pi\mathbf{P}$. Finally, a dot represents the dot product between functions, $\int \mathbf{f_p}(\mathbf{r}) \cdot \mathbf{P}(\mathbf{r}) \, \mathrm{d}\mathbf{r}$, where the integration is over all space. We also introduce:

$$\delta\mathbf{P} = \mathbf{P} - \langle\mathbf{P}\rangle_{\mathbf{0}}, \tag{12.19}$$

where the brackets $\langle\rangle_0$ indicate a Boltzmann average over the unperturbed system (without ρ_p). This quantity is important, and justifies the effort to derive (12.18), because its Boltzmann average $\langle\delta\mathbf{P}\rangle_\lambda$ over the perturbed system represents the mean, microscopic, density of polarization charge induced by the perturbing field $\lambda\mathbf{f_p}$. This is precisely the 'response' to the perturbation. From (12.18), (12.19) and the definition of the Boltzmann average, we can derive the relation between the response and the perturbation. For example, consider the mean x component of $\delta\mathbf{P}$ at a given position $\mathbf{r}$ in space. We have

$$\langle\delta P_x\rangle_\lambda = \frac{1}{Q_\lambda} \int \delta P_x \exp(-\beta U_0) \exp(\beta\lambda\mathbf{f_p} \cdot \mathbf{P}) \, \mathrm{d}r^N. \tag{12.20}$$

Here, Q_λ is the configuration integral when the coupling parameter equals λ; r^N represents all the conformational degrees of freedom. To simplify the notations, we have not made explicit the dependency of $\mathbf{P}$ and δP_x, either on the position $\mathbf{r}$ or on the instantaneous value of r^N. (Remember that the dot product in the right-hand exponential represents an integral over all space; see earlier.) Following the usual linear response method [49, 50], we compute the right-hand expression to the first-order with respect to $\delta\mathbf{P}$. To this order, Q_λ can be replaced by Q_0 and $\mathbf{P}$ can be replaced by $\delta\mathbf{P}$ (because $\langle\delta P_x\rangle_0$ is zero)

$$\langle\delta P_x\rangle_\lambda \approx \frac{1}{Q_0} \int \delta P_x \exp(-\beta U_0)(1 + \beta\lambda\mathbf{f_p} \cdot \delta\mathbf{P}) \, \mathrm{d}r^N. \tag{12.21}$$

The constant of unity in the parentheses on the right can be dropped. Expanding the dot product on the right, we have

$$\langle \delta P_x \rangle_\lambda \approx \lambda \beta \frac{1}{Q_0} \int \mathrm{d}r^N \exp(-\beta U_0) \delta P_x \int_{r'} (\delta P_x f_x + \delta P_y f_y + \delta P_z f_z) \mathrm{d}\mathbf{r}'$$

$$\approx \lambda \beta \int_r [\langle \delta P_x(\mathbf{r}) \delta P_x(\mathbf{r}') \rangle_0 f_x(\mathbf{r}') + \langle \delta P_x(\mathbf{r}) \delta P_y(\mathbf{r}') \rangle_0 f_y(\mathbf{r}')$$

$$+ \langle \delta P_x(\mathbf{r}) \delta P_z(\mathbf{r}') \rangle_0 f_z(\mathbf{r}')] \, \mathrm{d}\mathbf{r}'. \tag{12.22}$$

The spatial dependencies have been made explicit in the last equation. Finally, we see that:

$$\langle \delta \mathbf{P} \rangle_\lambda = \lambda \alpha \, \mathbf{f} \, (1 + O(\beta \Delta U)) \tag{12.23}$$

$$\alpha(\mathbf{r}i, \mathbf{r}'j) = \beta \langle \delta P(\mathbf{r})_i \delta P(\mathbf{r}')_j \rangle_0 \tag{12.24}$$

where $O(\cdot\cdot)$ represents quantities of first order or more in $\beta \Delta U$ and i, j represent cartesian components x, y, or z. The matrix quantity α is the *dielectric susceptibility* [51], which completely characterizes the response of the system to the perturbation. Importantly, it does not depend on λ. We see that to lowest order in $\beta \Delta U$, the response scales linearly with the perturbing field. These last equations have the classic form of linear response theory:

–   The response appears as a linear function of the perturbation (12.23)
–   The relation involves a susceptibility operator α
– α is determined by the fluctuations of the unperturbed system (12.24). Equation (12.24) is an example of the fluctuation–dissipation theorem [49, 51].

To the same order in $\beta \Delta U$, the free energy $A(\lambda)$ is parabolic

$$\frac{\partial A}{\partial \lambda} = \left\langle \frac{\partial \Delta U}{\partial \lambda} \right\rangle_\lambda = -\mathbf{f} \cdot \langle \mathbf{P} \rangle_\lambda = -\mathbf{f} \cdot \langle \mathbf{P} \rangle_0 - \lambda \mathbf{f} \cdot \alpha \, \mathbf{f} \, (1 + O(\beta \Delta U)) \tag{12.25}$$

$$A(\lambda) - A(0) = -\lambda \mathbf{f} \cdot \langle \mathbf{P} \rangle_0 - \frac{\lambda^2}{2} \mathbf{f} \cdot \alpha \, \mathbf{f} \, (1 + O'(\beta \Delta H)) \tag{12.26}$$

$$= \langle \Delta U \rangle_0 - \frac{\beta}{2} (\langle \Delta U^2 \rangle_0 - \langle \Delta U \rangle_0^2)(1 + O'(\beta \Delta U)). \tag{12.27}$$

The last equation is the beginning of the well-known expansion of the free energy into its cumulants (12.5). Equations (12.17)–(12.27) are exact within classical statistical mechanics. A parabolic free energy (i.e., a negligible $O'(\beta \Delta U)$ in (12.26)–(12.27) implies that cumulants of ΔU of order > 2 sum to zero, i.e., ΔU has Gaussian fluctuations, and also that $\langle \delta \mathbf{P} \rangle_\lambda$ in (12.23) is linear. The reverse is obviously true, so that a parabolic free energy, a Gaussian ΔU, and the linearity of $\langle \delta \mathbf{P} \rangle_\lambda$ are seen to be equivalent properties, accurate to the same order in $\beta \Delta U$.

If $\Delta U(1)$ has Gaussian fluctuations, this means that the free energy is also a parabolic function of $\Delta U(1)$ itself. $\Delta U(1)$ is known as the energy gap [52–54]; we will denote it by η

$$\eta \overset{\mathrm{def}}{=} \Delta U(1). \tag{12.28}$$

The energy gap is the energy difference between the reactant and product states for a given instantaneous conformation of the system; i.e., it is the energy change (*not* the free energy change) associated with introducing a 'virtual' charge density ρ_p. The free energy curves $A_0(\eta)$ and $A_1(\eta)$, corresponding to the reactant and product states, respectively, are known in electron transfer theory as the 'diabatic' free energy curves. It is easy to show [52] that they have the same curvature, and that their intersection occurs exactly at the point $\eta = 0$. Fig. 12.2 illustrates the linear response of a protein–solvent medium in the case of a purely electrostatic perturbation. In this figure are shown recent simulation data for electron transfer to cytochrome c in solution [47]. The free energy derivative $\partial A/\partial \lambda$ is shown in the top panel. Evidently, this system responds to the redox electron as a linear medium. The diabatic free energy curves are shown in the bottom panel as a function of η.

The concept of the energy gap suggests a natural decomposition of the reaction free energy, introduced by Marcus in the development of electron transfer theory [54, 55]. We will see later that it leads to a practical method for pK_a calculations. It is illustrated in Fig. 12.2 (bottom panel). The idea is to introduce the perturbation in two steps, corresponding to two distinct free energy components. The first, static component ΔA_{stat} corresponds to the vertical arrow in the figure: introducing the perturbing charge with the system constrained to stay in its unperturbed structure. For electron transfer, this step corresponds to the electron hopping event, which is so fast that the environment does not have time to adjust. In a second step, the constraint is gradually released, yielding a relaxation, or reorganization free energy ΔA_{rlx}. We have the relations

$$\Delta A_{\text{stat}} = \langle\eta\rangle_0 = -\mathbf{f} \cdot \langle\mathbf{P}\rangle_0, \tag{12.29}$$

$$\Delta A_{\text{rlx}} = -\frac{1}{2}\mathbf{f} \cdot \langle\delta\mathbf{P}\rangle_\lambda = -\frac{1}{2}\mathbf{f} \cdot \alpha\,\mathbf{f}\,(1 + O'(\beta\eta)) \tag{12.30}$$

$$= -\frac{\beta}{2}\,\text{Variance}(\eta)\,(1 + O'(\beta\eta)). \tag{12.31}$$

If the medium is linear, the reactant and product state parabolas have the same curvature. In that case, one can show that the free energies to impose the constraints at the beginning and to remove them later exactly cancel, and we obtain the useful relations [56–58]

$$\begin{cases} \Delta A &= \Delta A_{\text{stat}} + \Delta A_{\text{rlx}} \\[2mm] \Delta A &= \dfrac{\langle\eta\rangle_0 + \langle\eta\rangle_1}{2} \\[2mm] \Delta A_{\text{rlx}} &= \dfrac{\langle\eta\rangle_1 - \langle\eta\rangle_0}{2}. \end{cases} \tag{12.32}$$

The subscripts '0' and '1' refer to the endpoints of the reaction, where $\lambda = 0$ or 1, respectively. The relaxation $\langle\eta\rangle_1 - \langle\eta\rangle_0$ of the energy gap due to the perturbation is determined (12.31) by its fluctuations in the absence of the perturbation: another example of a fluctuation–dissipation relation.
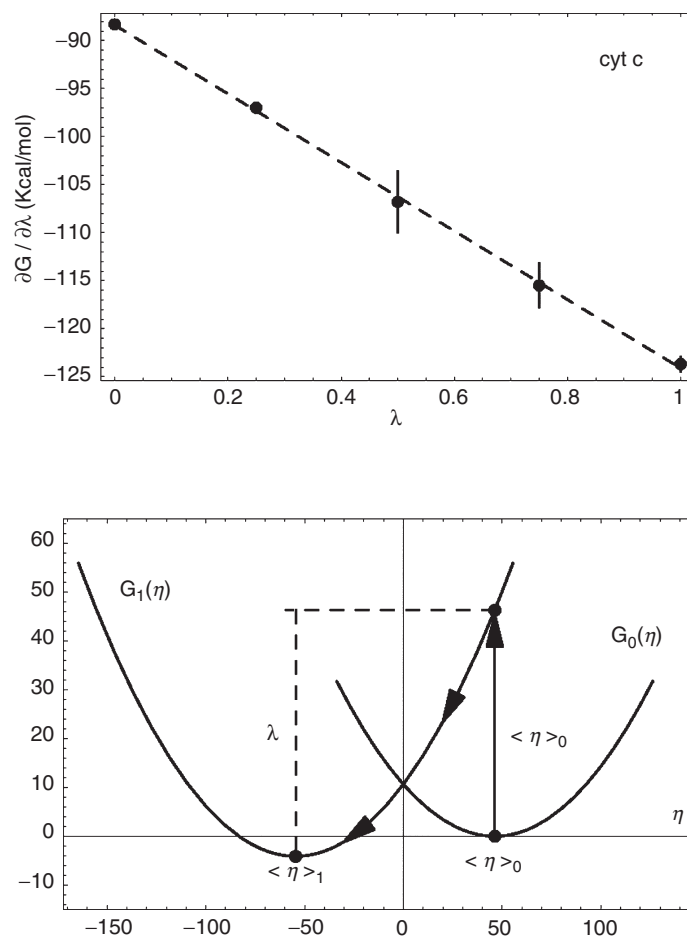
Fig. 12.2. Free energy data for electron transfer between the protein cytochrome c and the small acceptor microperoxidase-8 (MP8), from recent simulations [47]. *Top*: Gibbs free energy derivative $\frac{\partial G}{\partial \lambda}$ versus the coupling parameter λ. The data correspond to solvated cytochrome c; the MP8 contribution is not shown (adapted from [47]) *Bottom*: the Marcus diabatic free energy curves. The simulation data correspond to cyt c and MP8, infinitely separated in aqueous solution. The curves intersect at $\eta = 0$, as they should. The reaction free energy is decomposed into a static and relaxation component, using the two steps shown by arrows: a static, vertical step, then relaxation into the product state. All free energies in kcal mol^{-1}. Adapted with permission from reference [88]

12.3.2 Linear Response Theory: Application to Proton Binding and pK_a Shifts

We now turn to the problem of proton binding to proteins, an important area for simplified free energy methods. The linear response formalism earlier underlies most of the methods used today. It leads directly to one of the more useful practical methods, the so-called 'LRA,' or linear response approximation method [59], presented here.

Proton binding is usually modeled as a purely electrostatic process, where the atomic charges of a protein side chain are modified to represent the binding proton [60]. We assume for now that a single perturbing point charge, $q = +e$, is added to a single side-chain atom, and we focus on the corresponding free energy change ΔA. Further on, we will go into the practical details of a more realistic implementation.

The idea is to do simulations of the system before and after the proton binding; i.e., to simulate the reactant and product states. With the assumption of linear response, these provide all the information needed to compute ΔA. Indeed, the free energy to introduce a fractional charge λq into the reactant state ($0 \leq \lambda \leq 1$) is a parabolic function of λq, which can be written:

$$\Delta A_\lambda = \lambda q \Delta A_{\text{stat}}^{\text{reac}} + \lambda^2 q^2 \Delta A_{\text{rlx}}^{\text{reac}}. \tag{12.33}$$

The subscripts have the same meaning as in (12.29), (12.30). Indeed, (12.29), (12.30) show that in the parabolic free energy function, the static free energy is the linear term with respect to λq, while the relaxation free energy is the quadratic term. Thus, $\Delta A_{\text{stat}}^{\text{reac}}$ and $\Delta A_{\text{rlx}}^{\text{reac}}$ are, respectively, the static and the relaxation free energies to insert a unit charge into the reactant state.

The charging process can be completed by inserting the charge $-(1 - \lambda)q$ into the 'product state' at the same site. The corresponding free energy change can be written, with analogous notations

$$\Delta A_\lambda' = -(1 - \lambda)q \Delta A_{\text{stat}}^{\text{prod}} + (1 - \lambda)^2 q^2 \Delta A_{\text{rlx}}^{\text{prod}}. \tag{12.34}$$

We have seen that the free energy curves for the reactant and product states have the same curvature, so that the relaxation free energy is the same in the reactant and product states: $\Delta A_{\text{rlx}}^{\text{prod}} = \Delta A_{\text{rlx}}^{\text{reac}}$. This equality reflects the fact that the dielectric susceptibility α (12.24) does not depend on the perturbing field or charge, and is the same in the reactant and product states. We then obtain

$$\Delta A = \Delta A_\lambda + \Delta A_\lambda' = q\lambda \Delta A_{\text{stat}}^{\text{reac}} + q(1 - \lambda)\Delta A_{\text{stat}}^{\text{prod}} + (2\lambda - 1)\Delta A_{\text{rlx}}^{\text{reac}}. \tag{12.35}$$

Taking $\lambda = 1/2$, for example, gives back (12.32)

$$\boxed{\Delta A = \frac{q}{2}(\Delta A_{\text{stat}}^{\text{reac}} + \Delta A_{\text{stat}}^{\text{prod}})} \tag{12.36}$$

Requiring that ΔA be independent of λ (as it should be) gives back (12.32). From (12.29) and (12.17), the static free energies can be written

$$\Delta A_{\text{stat}}^{\text{reac}} = q\langle V \rangle_0 \tag{12.37}$$

$$\Delta A_{\text{stat}}^{\text{prod}} = q\langle V \rangle_1, \tag{12.38}$$

where V is the electrostatic potential at the charge insertion site and the brackets $\langle \rangle_0$ (respectively, $\langle \rangle_1$) represent a Boltzmann average in the absence (presence) of the perturbation. In practice, these average potentials are obtained from an MD or MC

simulation of the reactant and product states. An even simpler relation holds for the midpoint state, where one half of the perturbing charge has been inserted

$$\Delta A = \Delta A_{\text{stat}}^{\text{midpoint}}. \tag{12.39}$$

So far, we have considered a single perturbing point charge q. In a more realistic treatment, the proton that binds is modeled, not as a single perturbing charge but as a set of incremental charge shifts, $\{\Delta q_i\}$, that are inserted onto selected side-chain atoms. Equation (12.36) is replaced by the more general form

$$\Delta A = \frac{1}{2} \sum_{i \in A} \Delta q_i (\langle V_i^{\text{reac}} \rangle_0 + \langle V_i^{\text{prod}} \rangle_1). \tag{12.40}$$

Here, the subscript i refers to a side-chain atom and the sum is over the side chain, A, of interest. These charge shifts will normally correspond to the differences between the atomic partial charges for the neutral and ionized forms of the side chain of interest in a particular force field.

Equations (12.36) and (12.40) are the basis of the LRA method for calculating pK_a shifts [59]. Indeed, to obtain the pK_a shift due to the protein environment, we perform the same calculation for the protein and for a small molecule in solution, analogous to the side chain of interest. For a histidine side chain, for example, one would choose imidazole or methylimidazole in solution as a model compound. The pK_a shift due to the protein environment will then have the form:

$$pK_{a,\text{prod}} - pK_{a,\text{model}} = \frac{1}{2.303 k_B T} (\Delta A_{\text{prod}} - \Delta A_{\text{model}}), \tag{12.41}$$

where the subscripts, prod, model, refer to properties computed for the protein environment and the model compound in solution, respectively. If the pK_a of the model compound is known experimentally, we can obtain the pK_a in the protein.

12.4 Potential of Mean Force and Simplified Solvent Treatments

12.4.1 The Concept of Potential of Mean Force (PMF)

A key element of many simplified free energy methods is the use of an implicit description of the solvent. Implicit solvent models are based on the concept of a PMF, presented briefly in this section; see [11] for a detailed review; see Chap. 4 for applications of the PMF concept that are not related to implicit solvation.

We consider a biological macromolecule in solution. Let $\mathbf{X}$ and $\mathbf{Y}$ represent the degrees of freedom of the solute (biomolecule) and solvent, respectively, and let $U(\mathbf{X}, \mathbf{Y})$ be the potential energy function. The thermal properties of the system are averages over a Boltzmann distribution $P(\mathbf{X}, \mathbf{Y})$ that depends on both $\mathbf{X}$ and $\mathbf{Y}$. To obtain a reduced description in terms of the solute only, the solvent degrees of freedom must be integrated out. The reduced probability distribution P is

$$
\begin{aligned}
P(\mathbf{X}) &= \int d\mathbf{Y} \, P(\mathbf{X}, \mathbf{Y}) \\
&= \frac{e^{-U(\mathbf{X},\mathbf{Y})/k_{\mathrm{B}}T}}{\displaystyle\int d\mathbf{X} d\mathbf{Y} \, e^{-U(\mathbf{X},\mathbf{Y})/k_{\mathrm{B}}T}} \\
&\overset{\mathrm{def}}{=\!=} \frac{e^{-W(\mathbf{X})/k_{\mathrm{B}}T}}{\displaystyle\int d\mathbf{X} \, e^{-W(\mathbf{X})/k_{\mathrm{B}}T}}.
\end{aligned}
\tag{12.42}
$$

The function $W(\mathbf{X})$ is called the PMF; it was first introduced by Kirkwood to describe the structure of liquids [61]. It plays the role of a free energy surface for the solute. Notice that the dynamics of the solute on the free energy surface $W(\mathbf{X})$ do not correspond to the true dynamics. Rather, an MD simulation on $W(\mathbf{X})$ should be viewed as a method to sample conformational space and to obtain equilibrium, thermally averaged properties.

To construct an approximate PMF, we make the reasonable assumption that the potential energy has the form

$$
U(\mathbf{X}, \mathbf{Y}) = U_{UU}(\mathbf{X}) + U_{UV}(\mathbf{X}, \mathbf{Y}) + U_{VV}(\mathbf{Y}),
\tag{12.43}
$$

where the first term represents solute–solute interactions, the second, solute–solvent interactions, and the third, solvent–solvent interactions. This form is used in many molecular mechanics force fields for biomolecular simulations. From (12.42), the PMF then splits into two terms

$$
W(\mathbf{X}) = U_{UU}(\mathbf{X}) + \Delta W(\mathbf{X})
\tag{12.44}
$$

with

$$
e^{-\Delta W(\mathbf{X})/k_{\mathrm{B}}T} = \int e^{-U_{UV}(\mathbf{X},\mathbf{Y})/k_{\mathrm{B}}T} e^{-U_{VV}(\mathbf{Y})/k_{\mathrm{B}}T} \, d\mathbf{Y}.
\tag{12.45}
$$

Multiplying and dividing on the right by $e^{-U_{UU}(\mathbf{X})/k_{\mathrm{B}}T}$, we see that

$$
e^{-\Delta W(\mathbf{X})/k_{\mathrm{B}}T} = \langle e^{-U_{UV}(\mathbf{X},\mathbf{Y})/k_{\mathrm{B}}T} \rangle_{\mathbf{X}}.
\tag{12.46}
$$

This is precisely the Zwanzig perturbation free energy formula (12.4), with U_{UV} as the perturbation. Thus, $\Delta W(\mathbf{X})$ turns out to be the free energy to 'turn on' the interactions between the solute (biomolecule) and the solvent (water). Equivalently, it is the free energy to transfer the solute from the gas phase into solution, i.e., $\Delta W(\mathbf{X})$ is the solvation free energy of the solute, artificially maintained in the conformation $\mathbf{X}$.

We assume, furthermore, that the solute–solvent coupling has a 'nonpolar' component, U_{UV}^{np}, and an electrostatic component, U_{UV}^{elec}. Indeed, solute–solvent forces are dominated by short-range, repulsive interactions arising from Pauli's exclusion principle and long-range electrostatic interactions, arising from the nonuniform charge distribution. Attractive, dispersion interactions arising from electron correlation are weaker (except for purely nonpolar solutes or solvents, such as saturated

alkanes). In a molecular mechanics context, the Pauli and dispersion contributions would be represented by Lennard-Jones interactions and the longer-range electrostatic contributions would be represented by Coulomb interactions between partial charges on the solute and solvent particles.

The free energy $W(\mathbf{X})$ can then be computed in two steps. In the first, we reversibly introduce the nonpolar coupling, U_{UV}^{np}. In the second, we introduce the electrostatic coupling. This leads to

$$W(\mathbf{X}) = U_{UU}(\mathbf{X}) + \Delta W(\mathbf{X}) = U_{UU}(\mathbf{X}) + \Delta W^{\mathrm{np}}(\mathbf{X}) + \Delta W^{\mathrm{elec}}(\mathbf{X}) \quad (12.47)$$

with

$$\mathrm{e}^{-\Delta W^{\mathrm{np}}(\mathbf{X})/k_{\mathrm{B}}T} = \frac{\displaystyle\int \mathrm{d}\mathbf{Y}\,\mathrm{e}^{-[U_{VV}(\mathbf{Y}) + U_{UV}^{\mathrm{np}}(\mathbf{X},\mathbf{Y})]/k_{\mathrm{B}}T}}{\displaystyle\int \mathrm{d}\mathbf{Y}\,\mathrm{e}^{-U_{VV}(\mathbf{Y})/k_{\mathrm{B}}T}} \quad (12.48)$$

$$\mathrm{e}^{-\Delta W^{\mathrm{elec}}(\mathbf{X})/k_{\mathrm{B}}T} = \frac{\displaystyle\int \mathrm{d}\mathbf{Y}\,\mathrm{e}^{-[U_{VV}(\mathbf{Y}) + U_{UV}^{\mathrm{np}}(\mathbf{X},\mathbf{Y}) + U_{UV}^{\mathrm{elec}}(\mathbf{X},\mathbf{Y})]/k_{\mathrm{B}}T}}{\displaystyle\int \mathrm{d}\mathbf{Y}\,\mathrm{e}^{-[U_{VV}(\mathbf{Y}) + + U_{UV}^{\mathrm{np}}(\mathbf{X},\mathbf{Y})]/k_{\mathrm{B}}T}}. \quad (12.49)$$

The first step corresponds to the formation of a Lennard-Jones cavity with the shape of the solute; the charges are included in the second step. This free energy decomposition is, of course, path dependent: different (divergent) results would be obtained if the electrostatic coupling were included first.

To exploit the concept of PMF to represent solvent in free energy calculations, practical approximations must be constructed. A common approach is to treat the two components $\Delta W^{\mathrm{np}}(\mathbf{X})$ and $\Delta W^{\mathrm{elec}}(\mathbf{X})$ separately. Approximations for the nonpolar term are usually derived from geometric considerations, as in scaled particle theory, for example [62]. The electrostatic contribution is usually derived from continuum electrostatics. We consider these two contributions in turn.

12.4.2 The Nonpolar Contribution to the Potential of Mean Force

We begin by considering a spherical particle in water. We introduce the solute–solvent interactions gradually, through a coupling parameter λ that varies from zero to one

$$U(\lambda) = \lambda U_{UV}^{\mathrm{np}}(\mathbf{X}, \mathbf{Y}) + U_{VV}(\mathbf{Y}). \quad (12.50)$$

We assume for simplicity that the solvent is pure water, and that only the water–oxygen atoms have explicit Lennard-Jones interactions with the solute (this is typical of several common water models). We have seen that ΔW^{np} can be viewed as the free energy to change λ from zero to one. Therefore, a well-known thermodynamic integration formula gives

$$\frac{\partial W^{\mathrm{np}}}{\partial \lambda}(\lambda) = \left\langle \frac{U_{UV}^{\mathrm{np}}}{\partial \lambda} \right\rangle_{\lambda} = \int U_{UV}^{\mathrm{np}}\,\mathrm{e}^{-U_{VV}(\mathbf{Y})/k_{\mathrm{B}}T}\,\mathrm{d}\mathbf{Y}. \quad (12.51)$$

Due to the spherical symmetry of the system around the solute, this can be rewritten [63]

$$\frac{\partial W^{\mathrm{np}}}{\partial \lambda}(\lambda) = \int_0^\infty 4\pi r^2 \langle \rho(r) \rangle_\lambda u_{\mathrm{np}}(r)\, \mathrm{d}r, \qquad (12.52)$$

where $\langle \rho(r) \rangle_\lambda$ is the mean number density of water oxygen atoms at a distance r from the solute for a given value of λ and u_{np} represents the interaction energy between the solute and a single water oxygen. The functions in the integral are plotted in Fig. 12.3 for the case of an argon particle in liquid water. For intermediate values of λ, $0 < \lambda < 1$, the water density falls to zero inside the solute, while u_{np} is dominated by strong solute–solvent repulsive interactions at short range. The result is a product that has a peak close to the solute surface (thick lines in Fig. 12.3). This property is used to obtain approximate forms of the nonpolar contribution to the free energy solvation in 'scaled-particle theory' (SPT) [62, 64, 65] and 'solvent-exposed surface area' models [66], described next.

Scaled Particle Theory

A simple approach was proposed by Reiss et al. [62], Stillinger [64], and others [65] to describe the free energy of inserting a nonpolar repulsive sphere into a solvent. The
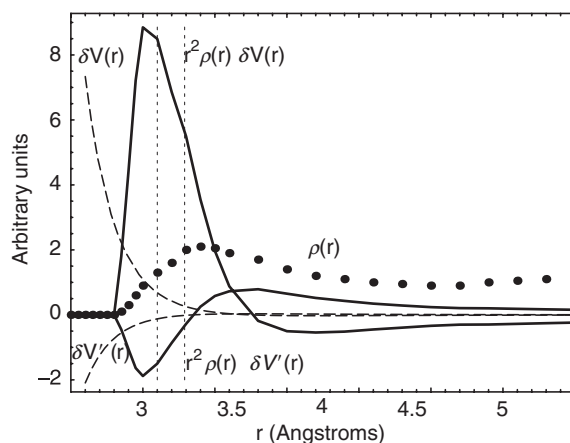
Fig. 12.3. Mutation of argon into the larger xenon or the smaller neon in aqueous solution. The mutation consists in changing the van der Waals parameters of the solute from those of argon to those of xenon or neon. The perturbing energy term is $\delta V = U_{\mathrm{vdW}}(\mathrm{xenon}) - U_{\mathrm{vdW}}(\mathrm{argon})$, the difference between the solute–solvent interaction calculated with the argon and xenon van der Waals parameters; similarly for neon ($\delta V'$). The vertical lines indicate the van der Waals radius of argon, xenon, and neon (3.29, 3.57, and 3.10 Å). The mean density ρ of water at a distance r from the solute is shown (*black dots*; arbitrary units), as well as δV and $\delta V'$. Integrating the product $r^2 \rho \delta V$ (respectively, $r^2 \rho \delta V'$) gives the free energy to expand the solute into xenon (respectively, to shrink it to neon). These functions are shown as thick black curves. They represent the radial density of free energy for each transformation (expansion, shrinkage), seen to be concentrated close to the solute surface

approach is called 'scaled particle theory' (SPT) because it is based on arguments involving the scaling of the repulsive sphere radius. The reversible work $W(R)$ to produce a spherical cavity of radius R can be calculated exactly for a hard-sphere liquid of bulk density $\bar{\rho}$ as long as $2R \leq a$, the hard sphere diameter

$$W(R) = -k_{\mathrm{B}}T \ln \left(1 - \frac{4}{3}\pi R^3 \bar{\rho}\right). \tag{12.53}$$

For a nonpolar solute in liquid water, a is assigned a value of 2.75 Å, corresponding to the distance of closest contact in the oxygen–oxygen radial distribution function of liquid water [64]. For a soft-sphere solute interacting with the solvent through $u(r) = \lambda A r^{-n}$, R is an equivalent hard-sphere radius related to the second virial coefficient. Generalizations to van der Waals and associated liquids have been made by introducing experimental densities and virial coefficients.

In the limit of a large cavity or solute particle, thermodynamic considerations [67] lead to

$$W(R) = \frac{4}{3}\pi R^3 p + 4\pi R^2 \gamma_{\mathrm{v}} \left(1 - \frac{4\delta}{R}\right) + \ldots, \tag{12.54}$$

where p is the pressure, γ_{v} the surface tension of the solvent, and δ is a molecular length scale. This expression implies that the microscopic surface tension coefficient depends on the radius of curvature, i.e., $\gamma(R) = \gamma_{\mathrm{v}}(1 - 4\delta/R)$. For water, Stillinger estimated that δ is approximately equal to 0.5 Å [64]. In the intermediate R range, $2R \geq a$, $W(R)$ can be expanded in powers of R; the first three expansion coefficients are obtained by matching the function and its first two derivatives, given at $R = a$ by (12.53). The third derivative is discontinuous at $R = a$ and cannot be used in this way. However, from (12.54), the R^3 term is likely to be negligible, and in any case, volume and surface area will often be correlated in practice, so that the R^3 term can be included in the surface term. Thus, the expansion effectively has the same form as (12.54) to third order, and can be considered an extension of the surface tension concept to molecular dimensions. In practice, the pV-like term is expected to be negligible. SPT has been compared with results from molecular dynamics simulations and free energy perturbation calculations for nonpolar rare gases [68, 69]. More recently, cavity formation in aqueous and nonaqueous solvents has been studied extensively by molecular dynamics simulations [70], and simulations combined with information theory [71]; see Pratt and Pohorille [72] for an extensive review.

Solvent-Exposed Area

SPT provides a conceptual basis relating the nonpolar free energy contribution to the solvent-exposed surface area. An attractive approximation is to ignore curvature effects and write

$$\Delta W^{(\mathrm{np})}(\mathbf{X}) = \gamma_{\mathrm{v}} \mathscr{A}_{\mathrm{tot}}(\mathbf{X}). \tag{12.55}$$

This description of the nonpolar contribution to the free energy has been extensively used in biophysical applications [72–75]. In practice, the surface tension γ_{v} is usually obtained from experimental transfer free energies of small organic molecules

between different solvents. The limitations of the surface area model are illustrated by vapor-to-water transfer free energies of saturated alkanes in Fig. 12.4 (from [74]). Proportionality of the solvation free energy to solute area is good for linear alkanes, but poor for saturated cyclic alkanes. The proportionality coefficient, or 'surface tension,' for linear alkanes is about $6 \, \text{cal} \, \text{mol} \, \text{Å}^{-2}$ for vapor-to-water transfer (Fig. 12.4) and $25 \, \text{cal} \, \text{mol} \, \text{Å}^{-2}$ for cyclohexane-to-water transfer.

Recent work revealed that the poor correlation between alkane surface areas and solvation free energies (Fig. 12.4) is due to the attractive solute–solvent dispersion interactions [76]. Indeed, for solutes of the same size that interact with solvent through a purely repulsive potential, $u(r) = A/r^{12}$, the solvation free energy is a nearly linear function of solute area. The solute–solvent dispersion energy, on the other hand, is much more dependent on the specific compound. Interestingly, introducing the dispersion interactions after the solute cavity is formed has a small effect on the solvent structure, and is an almost purely enthalpic process, contributing little to the solvation entropy. These data suggest [76] that the cavity term can be accurately described by a simple surface area term, whereas the dispersion contribution requires a term with a more complicated functional form.

12.4.3 Classical Continuum Electrostatics

Continuum electrostatics approximations in which the solvent is represented as a featureless dielectric medium are an increasingly popular approach for the electrostatic
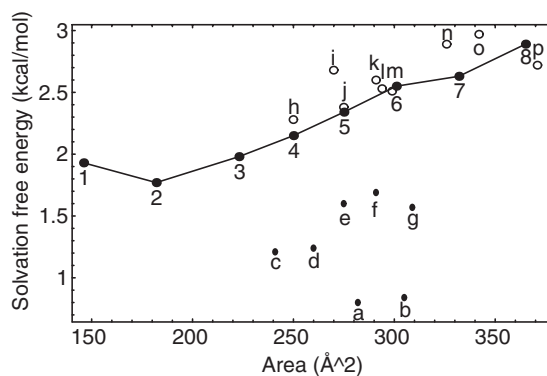
Fig. 12.4. Vapor-to-water transfer data for saturated hydrocarbons as a function of accessible surface area, from [131]. Standard states are 1 M ideal gas and solution phases. Linear alkanes (*small dots*) are labeled by the number of carbons. Cyclic compounds (*large dots*) are: a = cyclooctane, b = cycloheptane, c = cyclopentane, d = cyclohexane, e = methylcyclopentane, f = methylcyclohexane, g = *cis*-1,2-dimethylcyclohexane. Branched compounds (*circles*) are: h = isobutane, i = neopentane, j = isopentane, k = neohexane, l = isohexane, m = 3-methylpentane, n = 2,4-dimethylpentane, o = isooctane, p = 2,2,5-trimethylhexane. Adapted with permission from [74]. Copyright 1994, American Chemical Society

term in the PMF [11, 20, 77]. Lattice models of the solvent provide an alternative approach, extensively developed by Warshel and coworkers [77]; in practice, they contain similar, though not identical physics. The history of continuum electrostatic treatments in chemistry goes back to Born [78], Kirkwood [79], and Onsager [80]. These approaches are surprisingly successful in reproducing the electrostatic contribution to the solvation free energy of small solutes [73, 74, 77, 81, 82]. Continuum electrostatic approximations are based upon the Poisson equation for macroscopic media [83, 84]

$$\nabla \cdot [\epsilon(\mathbf{r}) \nabla \mathbf{V}(\mathbf{r})] = -4\pi \rho_{\mathrm{u}}(\mathbf{r}), \tag{12.56}$$

where $V(\mathbf{r})$ is the electrostatic potential at a point $\mathbf{r}$, $\rho_{\mathrm{u}}(\mathbf{r})$ represents the permanent charge density of the solute, and $\epsilon(\mathbf{r})$ is the position-dependent dielectric constant. The Poisson equation (12.56) can be solved numerically by mapping the system onto a discrete grid and using a finite-difference algorithm [85–88]. In applications to proteins, it is generally assumed that the dielectric constant is uniform within the protein and within the solvent, with two distinct values.

The main effect of the permanent charges of the solute is to polarize the solvent. The induced charge density ρ_{ind} in the solvent is related to the solvent polarization density $\mathbf{P}(\mathbf{r})$ [83, 84]

$$\langle \rho_{\mathrm{ind}}(\mathbf{r}) \rangle = -\nabla \cdot \mathbf{P}(\mathbf{r}). \tag{12.57}$$

At any point $\mathbf{r}$, the polarization $\mathbf{P}(\mathbf{r})$ and the total electrostatic field $\mathbf{E}^{\mathrm{tot}}(\mathbf{r})$ are assumed to be linearly related,

$$\mathbf{P}(\mathbf{r}) = \frac{\epsilon(\mathbf{r}) - 1}{4\pi} \mathbf{E}^{\mathrm{tot}}(\mathbf{r}); \tag{12.58}$$

this relation is usually taken as the definition of the dielectric constant ϵ [83, 89]. We see that the dielectric constant describes in an approximate, implicit way the polarization of the medium in response to fields or charges (such as the partial atomic charges carried by a ligand molecule, or the negative charge carried by a redox electron; see Sect. 12.3). We also see that a dielectric of one corresponds to a medium with no implicit polarizability, since $\epsilon = 1$ implies that $\mathbf{P}$ is always zero, even if electric fields are present. For a protein responding to a redox electron or a new ligand, choosing a dielectric greater than one is thus one way to model the induced polarization. Another way is to explicitly simulate the dynamics of the system at a microscopic level, through MD simulation; see Sect. 12.3. If the polarization is modeled explicitly by simulating the molecular dynamics of the solute, there is no need to add continuum dielectric polarization [89]. A dielectric constant of one will be appropriate for the solute in this case. In other cases, a larger value will be needed.

In practical cases, it is the solute charges that are modeled explicitly, and treated as permanent source charges. In contrast, the whole solvent medium is usually treated as a continuum, without any explicit, permanent, source charges. (This is reasonable for a solvent made of small, neutral molecules; ionic liquids would obviously need a different treatment.) Since there are no permanent charges in the solvent,

the divergence of the polarization density in the solvent is zero except near the dielectric boundary [83], and the solvent charge density is a sharply peaked function localized at the solute–solvent interface. Integrating the solvent charge density along an axis perpendicular to the surface over an infinitesimal range and making the width of the boundary go to zero, one recovers an expression for the surface charge density $\sigma(\mathbf{r})$, which is the basis of boundary-element formulations of the problem [90, 91].

The free energy of the system has the form [84]

$$F = \frac{1}{2} \int \rho_{\mathrm{u}}(\mathbf{r}) \mathbf{V}(\mathbf{r}) \, \mathrm{d}\mathbf{r}, \tag{12.59}$$

where $V(\mathbf{r})$ is the total electrostatic potential at $\mathbf{r}$ and the integration is over all space. If the permanent charge density ρ_{u} is made up of atomic point charges q_i, F takes the discrete form

$$F = \frac{1}{2} \sum_i q_i V_i, \tag{12.60}$$

where V_i is the total potential at the site of the atomic charge q_i.

12.5 Linear Interaction Energy Approaches

Linear response approximations have also been applied to the protein–ligand binding problem. Several applications were discussed earlier (Sect. 12.2.3). We turn now to a more systematic approach that has become popular in recent years, known as the linear interaction energy (LIE) method [13, 92]. Protein–ligand binding can be discussed with the thermodynamic cycle in Fig. 12.5. The LIE method is based on the following approximation for the binding free energy of a ligand L:

$$
\begin{aligned}
\Delta A_{\mathrm{bind,L}} = {} & \alpha \left(\langle V_{\mathrm{L}}^{\mathrm{vdw}} \rangle_{\mathrm{prot}} - \langle V_{\mathrm{L}}^{\mathrm{vdw}} \rangle_{\mathrm{solv}} \right) \\
& + \beta \left(\langle V_{\mathrm{L}}^{\mathrm{elec}} \rangle_{\mathrm{prot}} - \langle V_{\mathrm{L}}^{\mathrm{elec}} \rangle_{\mathrm{solv}} \right) \\
& + \gamma
\end{aligned}
\tag{12.61}
$$

where $\langle \cdots \rangle_{\mathrm{prot}}$, $\langle \cdots \rangle_{\mathrm{solv}}$ denote ensemble averages when the ligand is bound to the protein or free in solution, respectively. The quantity averaged is an interaction energy between the ligand and its surroundings; "vdw" indicates a van der Waals energy term; "elec" indicates a Coulombic electrostatic energy term; α, β, and γ are constants. To justify this approximation, the usual argument focuses on the leftmost, vertical leg of the cycle in Fig. 12.5, which corresponds to the binding reaction, P + L $\rightarrow$ PL. The binding reaction transforms the ligand environment from a pure solvent medium into a mixed solvent/protein medium. This can be accomplished by gradually changing the energy function, so as to switch off the ligand–environment interactions in the first medium and switch them on in the second [see (12.9)]. A corresponding energy gap η can be defined [see (12.29)–(12.32)], which takes the

form of a difference between the ligand interactions in the initial and final media, including both Coulombic and van der Waals contributions. If η is assumed to have a Gaussian probability distribution, the free energy change $\Delta A_{\text{bind},\text{L}}$ will indeed have the form given in (12.36), (12.32), which is the LIE form with $\alpha = \beta = 1/2$ as the numerical coefficients. The term γ has been dropped, because a constant γ will cancel out when two ligands are compared. Thus, LIE follows if the system obeys a linear response throughout the binding process, both for the van der Waals and the Coulombic energy terms. By taking γ as constant it is assumed that all effects other than van der Waals and Coulomb interactions are the same for all ligands.

This argument is not very compelling in the above form. The linear response approximation has been tested for modest charge rearrangements in several proteins and for small molecules in solution, and shown to give approximate, but reasonable results; see [88] and references therein. But the ligand binding reaction involves a much larger perturbation, where the entire environment of the ligand is changed from pure water to a heterogeneous protein–water mixture. Not only will the polarization density $\mathbf{P}$ reorganize; a protein–solvent interface is created, and the whole structure of the ligand environment changes. The relevant energy gap, η in (12.29)–(12.32), now contains both electrostatic and van der Waals contributions. There is no direct evidence that the van der Waals contribution has Gaussian fluctuations for any known ligand or protein–ligand complex (with the same variance in the bound and unbound states). On the contrary, practitioners of free energy simulations know that the energy gap for insertion of a new van der Waals particle into a condensed medium is usually distinctly non-Gaussian.

In addition, the constant γ must be inferred from experimental data for one or more ligands. This term mainly reflects the loss of rotational and translational entropy upon complex formation, as well as changes in vibrational entropy of the protein and protein–ligand complex. The rotational and translational entropies vary with the logarithm of the moments of inertia and the mass of the ligand [63], according to well-known formulae. But the vibrational entropy is more complex, and probably cannot be predicted precisely from first principles. Although normal-mode methods have been used abundantly for this purpose in the last few years, the approximations involved (harmonic fluctuations, crude solvent models in most cases, structures obtained by energy minimization in vacuum in many cases) are so severe that is not yet possible to decide whether they have any real predictive capability for the present problem. (Note that normal-mode calculations with an accurate implicit solvent model will provide a way forward [93].) As long as γ is obtained from experimental data, any claim to predict absolute binding free energies is unfounded.

Very similar difficulties arise if one considers an alchemical pathway, where a ligand L is transformed into another ligand L', both in complex with the protein and in solution.

It is interesting to consider another approximate derivation, which uses the implicit solvent models discussed earlier (Sect. 12.4). Indeed, we can decompose the binding reaction into the steps shown in Fig. 12.5 [94]: first, the ligand charges are switched off in pure solvent, leaving a nonpolar solute; second, the attractive

Restore ligand charges   Restore dispersion interactions

Remove ligand charges   Remove dispersion interactions
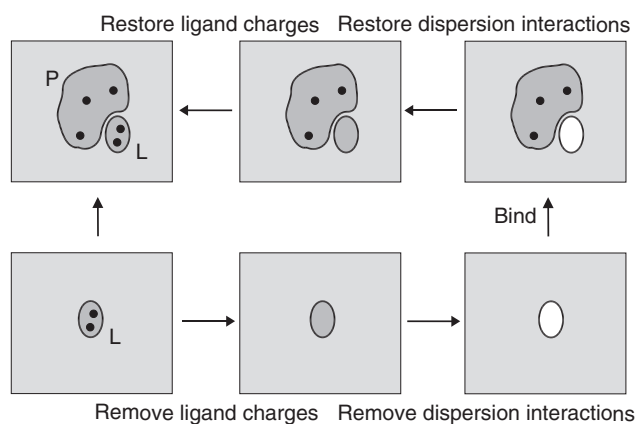
Fig. 12.5. Thermodynamic cycle for ligand binding. In the left-hand vertical leg, the ligand L simply binds to the protein P. The complicated path to the right has the following steps: (1) the electrostatic interactions of the ligand and its environment are switched off (this is schematized by removing its charges, shown as *black dots*), (2) the attractive dispersion interactions of the ligand are switched off, giving a nonpolar, repulsive cavity (*white ellipse*), (3) the cavity is transferred into the protein binding site (vertical leg on the right), (4) the ligand dispersion interactions are restored, and (5) the electrostatic interactions of the ligand with its environment are restored

dispersion interactions of the ligand are switched off; then the nonpolar solute is transferred into the protein binding site, and the ligand dispersion interactions and charges are switched back on. The contribution of the charging/uncharging steps to the double free energy difference $\Delta\Delta A$ can be expected to obey linear response, giving the electrostatic LIE term, with $\beta = 1/2$. For the vertical, transfer step on the right of Fig. 12.5, we can view the starting and final environments as two different solvents. We saw earlier that, for saturated alkanes of various sizes, the van der Waals energy alone is not sufficient to reconstruct an accurate transfer free energy between different solvents. However, the transfer free energy can be approximated empirically by a sum of terms, proportional to the solute surface area and the solute–solvent van der Waals interactions energies, respectively, [76]. Obviously, the simple correlations of free energy with surface area and van der Waals interactions that were discussed earlier (Sect. 12.4.2) corresponded to small solutes in a pure solvent; these correlations may change or break down when pure solvent is replaced by a heterogeneous protein–water mixture. If we adopt, nevertheless, this treatment for the heterogeneous binding site in the protein, we obtain a modified LIE free energy [95–97], which includes a new term, proportional to the amount of ligand surface area that is buried upon binding to the protein [98]. The coefficient for this term is akin to a molecular 'surface tension' (Sect. 12.4.2). This coefficient and the van der Waals coefficient α are both expected to differ from 1/2.

With this pathway, the separation between the electrostatic and van der Waals contributions in the LIE equation (12.61) is only approximate. Indeed, the vertical,

solute transfer step (right of Fig. 12.5) strips water away from the ligand binding site and substantially alters the protein–solvent electrostatic interactions. The corresponding free energy change may well display correlations with the ligand–solvent electrostatic interactions. Thus, there can be a 'mixing' of free energy contributions, and the LIE coefficient β can deviate from the 'theoretical,' linear response value of 1/2, *even if the medium is perfectly linear* for purely electrostatic perturbations.

Despite these limitations, the LIE method has given reasonable results in several applications [13–16]. The proteins studied include dihydrofolate reductase [99], thrombin [100], and others [95–97, 101–103]. In practice, the van der Waals coefficient α is usually adjusted empirically to fit the experimental data for a subset of ligands; β is sometimes adjusted, too. The resulting model is then used to make predictions for other, similar ligands. By using simulations of both the bound and the unbound state, induced fit conformational changes are included in the model. Additional free energy terms have been included in some cases, such as terms counting the number of solute–solvent hydrogen bonds [102, 103]. As more terms are included and more coefficients are fit, the method becomes more empirical, effectively like a Quantitative Structure–Activity Relationship (QSAR) treatment [104], but with descriptors deduced partly from computer simulations. This may be the most useful way to view the method. As an example, for 60 inhibitors of human factor Xa, Jorgensen et al. evaluated about 40 possible descriptors and obtained finally an empirical, or 'extended' LIE model that reproduced experimental inhibition constants with an RMS error of less than $1\,\mathrm{kcal\,mol^{-1}}$. The final model used two descriptors: the solute–environment van der Waals energy (12.61) and the number of hydrogen bonds lost by the ligand upon binding [102].

The LIE studies above rely on MD or MC simulations of ligands and protein:ligand complexes in solution, with a costly, explicit solvent representation. A logical further step is to replace the explicit solvent by a modern implicit solvent model [105, 106]. This raises two issues. First, the microscopic information on the solute–solvent van der Waals interactions is lost, and must be included in the implicit treatment. We saw earlier that similar assumptions underlie the most convincing LIE derivation (which uses the complicated pathway in Fig. 12.5), so that this may not affect the accuracy very much. Second, the Poisson–Boltzmann or Generalized Born (GB) implicit solvent models commonly used for the electrostatic contribution are not pairwise additive [107]. Nevertheless, under certain conditions, the electrostatic binding free energy can be subdivided into components corresponding to subsystems such as the protein or the ligand [108, 109]; see later. In the method of Zhou et al. [105] it is assumed the ligands occupy a similar space in the binding pocket, so that they 'desolvate' the protein to the same extent. The necessary free energy contributions can then be obtained easily; details are give in Sect. 12.6.

12.6 Free Energy Methods Using an Implicit Solvent: PBFE, MM/PBSA, and Other Acronyms

Many competing effects can contribute to ligand–receptor binding free energies: changes in rotational, translational, conformational, and vibrational entropy of the

partners, entropy changes associated with solvent ordering around hydrophobic or charged groups, solute conformational strain, changes in electrostatic and van der Waals interactions within and between the partners and the solvent, counterion reorganization. All of these effects can be accounted for automatically in explicit solvent simulations, as long as sufficient conformational sampling is performed. All of them have been included in simplified continuum or semimicroscopic models in the past [110], with different flavors and varying degrees of success.

The simplest case is the calculation of relative binding free energies in systems dominated by electrostatics. The continuum electrostatic free energies of the bound and free states are simply subtracted. The key model ingredients are the dielectric constant(s) used for the solutes and the structures used to model the various states (e.g., bound, unbound ligand). We begin by discussing this case. We refer to these calculations as PBFE. Then, we go on to models that include nonelectrostatic effects, primarily through a nonpolar free energy component. These latter models are usually referred to as MM/PBSA models.

12.6.1 Thermodynamic Pathways and Electrostatic Free Energy Components: The PBFE Method

The best starting point to appreciate the approximations of the PBFE method for ligand binding problems is to decompose the free energy into components. We focus therefore on electrostatic contributions to the protein–ligand binding free energy. We start from the continuum electrostatic expression for A_{pl}, the free energy of the protein:ligand complex in solution. It can be written

$$A_{\mathrm{pl}} = \frac{1}{2} \sum_i q_i V_i^{\mathrm{pl}} = \frac{1}{2} \sum_{i \in \mathrm{lig}} q_i V_i^{\mathrm{pl}} + \frac{1}{2} \sum_{i \in \mathrm{prot}} q_i V_i^{\mathrm{pl}}, \qquad (12.62)$$

where the first sum is over all protein and ligand atoms; q_i is the partial charge of atom i; V_i^{pl} is the total electrostatic potential on atom i in the complex, and the sums on the right are over ligand and protein atoms, respectively. From the linearity of continuum electrostatics, the potential on atom i can be expressed as a sum over all ligand and protein atoms

$$V_i^{\mathrm{pl}} = \sum_j V_{j \to i}^{\mathrm{pl}} = \sum_{j \in \mathrm{lig}} V_{j \to i}^{\mathrm{pl}} + \sum_{j \in \mathrm{prot}} V_{j \to i}^{\mathrm{pl}}, \qquad (12.63)$$

where $V_{j \to i}^{\mathrm{pl}}$ is the potential at atom i when only the partial charge q_j is present in the protein. $V_{j \to i}^{\mathrm{pl}}$ is known as a Green's function [83]. Using the very general reciprocity relation $q_i V_{j \to i}^{\mathrm{pl}} = q_j V_{i \to j}^{\mathrm{pl}}$ [84], we have

$$A_{\mathrm{pl}} = \frac{1}{2} \sum_{i \in \mathrm{lig}, j \in \mathrm{lig}} q_i V_{j \to i}^{\mathrm{pl}} + \sum_{i \in \mathrm{prot}, j \in \mathrm{lig}} q_i V_{j \to i}^{\mathrm{pl}} + \frac{1}{2} \sum_{i \in \mathrm{prot}, j \in \mathrm{prot}} q_i V_{j \to i}^{\mathrm{pl}}. \quad (12.64)$$

The first and third sums on the right each include terms of the form $\frac{1}{2} q_i V_{i \to i}$, representing the Born 'self-energy' of each charge q_i [107]. To obtain the binding free

energy ΔA_{bind}, we subtract the analogous expressions for the separated protein and ligand

$$
\begin{aligned}
\Delta A_{\mathrm{bind}} &= A_{\mathrm{pl}} - A_{\mathrm{p}} - A_{\mathrm{l}} \\
&= \sum_{i \in \mathrm{prot}, j \in \mathrm{lig}} q_i V_{j \to i}^{\mathrm{pl}} + \frac{1}{2} \sum_{i \in \mathrm{lig}, j \in \mathrm{lig}} q_i [V_{j \to i}^{\mathrm{pl}} - V_{j \to i}^{\mathrm{l}}] \\
&\quad + \frac{1}{2} \sum_{i \in \mathrm{prot}, j \in \mathrm{prot}} q_i [V_{j \to i}^{\mathrm{pl}} - V_{j \to i}^{\mathrm{p}}].
\end{aligned}
\tag{12.65}
$$

The first term on the right of (12.65) represents direct interactions between the ligand and the protein residues in the complex, screened by solvent. It will be referred to as the 'direct interaction term.' The second term ('the ligand desolvation term') includes the change in intraligand interactions upon binding, due to changes in the ligand geometry or charge distribution, as well as changes in the interaction of the ligand with polarization charge in the surrounding dielectric media. The polarization charge is spread over the solute–solvent interface [83]. If the ligand is assumed to have the same geometry and partial charges in the bound and free state, then the intraligand interactions do not contribute, and this term arises entirely from ligand interactions with the polarization charge. It corresponds therefore to decreased ligand–solvent interactions. The third term on the right of (12.65) has an identical interpretation. If the protein is assumed to have the same structure in the bound and free state, then this term represents the effect of displacing solvent from the binding site by inserting the ligand, decreasing the protein–solvent interactions. It will be referred to as the 'protein desolvation term.' Each term corresponds to one step in the binding pathway shown in Fig. 12.5. Similar, multistep pathways to analyze binding were proposed as early as 20 years ago [94].

Solvent contributions are present implicitly in all three terms. Furthermore, each protein or ligand group contributes to ΔA_{bind} through all three terms in (12.65). Thus, the protein contributes to the ligand desolvation term: even though the protein charges do not appear explicitly in this term, each protein residue occupies space around the ligand and contributes to its desolvation [107]. Similarly, the ligand contributes to the protein desolvation term, even though its charges do not appear explicitly. If one compares two ligands that occupy exactly the same space in the active site, the protein desolvation terms will be identical for the two ligands.

The contribution $\Delta A_{\mathrm{R}}^{\mathrm{DI}}$ of residue R to the direct interaction term has the form

$$
\Delta A_{\mathrm{R}}^{\mathrm{DI}} = \sum_{i \in \mathrm{R}, j \in \mathrm{lig}} q_i V_{j \to i}^{\mathrm{pl}}.
\tag{12.66}
$$

The quantity on the right can be obtained from a calculation of the electrostatic potential arising from the partial charges of the ligand at the positions of the partial charges q_i of the protein:ligand complex. Subtracting the results for the PL and PL' complexes, we obtain the contribution of residue R to the direct interaction term in the binding free energy difference.

An atom-by-atom decomposition is not possible for the desolvation terms. Indeed, these are the terms that make the continuum electrostatic free energy a many-body function [107]. This can be understood by considering an interacting pair of protein side chains. The strength of their interaction depends mainly on the extent of dielectric shielding by high-dielectric solvent. This, in turn, depends on the presence or absence of other nearby protein groups, since these can occupy nearby space, exclude solvent, and limit dielectric shielding. Fortunately, the decomposition of most interest is not an atomic decomposition, but the separation of protein and ligand desolvation, already evident in (12.65). The protein charges do not appear explicitly in the ligand desolvation term; rather, the protein atoms contribute by occupying space around the ligand in the protein:ligand complex, replacing high-dielectric solvent by the lower-dielectric protein medium. For two ligands that bind in similar positions, but have different charge distributions, this term can contribute significantly to $\Delta\Delta A$.

In the context of LIE with a GB solvent, Zhou et al. assumed that the ligands occupied similar positions in the protein binding pocket, desolvating the protein to the same extent [105]. The two other contributions to ΔA_{bind} can then be calculated easily. The direct interaction term is given by the GB screened interaction energy term. The ligand desolvation term corresponds to the GB self-energy of the ligand. Each term is a sum over the pairs i, j, $i \in \mathrm{lig}$, $j \in \mathrm{prot}$. Thus, the decomposition in (12.65) is sufficient to construct LIE methods with a continuum solvent.

12.6.2 Other Free Energy Components: MM/PBSA Methods

Ligand binding also involves nonelectrostatic contributions. Combined approaches have been developed for at least 30 years [111–113]. In Sect. 12.4, we discussed at length the path dependency of free energy components, and the need to formulate approximate free energy methods using a specific pathway for the binding reaction. Like LIE, MM/PBSA can be formulated using the pathway in Fig. 12.5. The main difference between LIE and MM/PBSA is that with the latter, electrostatic components will be obtained from a dielectric continuum model. Recent flavors of MM/PBSA combine continuum electrostatic solvation with a ligand–receptor van der Waals energy, intramolecular, stereochemical energy terms, a nonpolar solvation term proportional to buried surface area, and sometimes an estimate of vibrational entropies [14, 15, 75, 114, 115]. The free energy of each state (ligand, receptor, complex) has the form

$$A = \langle E_{\mathrm{intra}} \rangle + \langle E_{\mathrm{vdw}} \rangle + \langle E_{\mathrm{elec}} \rangle + \langle E_{\mathrm{SA}} \rangle - T \langle S_{\mathrm{vib}} \rangle, \qquad (12.67)$$

where the successive terms represent energies of intramolecular interactions (intra), van der Waals interactions between ligand and receptor (vdw), electrostatic interactions (elec; including the effect of continuum solvent), the nonpolar surface area term (SA), and a vibrational entropy term (vib). The continuum solvent and nonpolar terms have been abundantly tested and discussed. The van der Waals term is expected to be correlated with the surface that is buried upon binding and E_{SA}, as

discussed [76]; therefore it is partly redundant. The vibrational entropy calculations have often been done with normal-mode calculations with a crude solvent model, which are of questionable accuracy and precision.

There is still no clear statistical demonstration of the predictive value of all the terms in MM/PBSA. In some cases, the overall free energy function may be more accurate than the individual terms. One indication of this is the good accuracy of PBFE, in several cases, for the binding free energy difference between different ligands, despite a very poor accuracy for binding enthalpy and entropy, taken separately. Binding enthalpy and entropy are quite difficult to compute accurately in aqueous solvent, largely because of the particular properties of liquid water at room temperature. One often observes experimentally [116] that a small chemical change in a solute can have a large effect on its aqueous solvation enthalpy and entropy, but a small one on its solvation free energy. The robustness of the free energy can be seen as the result of entropy/enthalpy cancelation in aqueous solvent. This robustness makes the free energy an easier target for simple models, compared to the individual enthalpy and entropy. Indeed, for a number of systems, PBFE and MM/PBSA have yielded free energy differences in good agreement with experiment and/or higher-level computations (MDFE with explicit solvent). Some examples are reviewed in Sect. 12.6.3.

12.6.3 Some Applications of PBFE and MM/PBSA

Continuum models are being increasingly used to study protein–ligand recognition [115]. Most studies have considered series of similar ligands or protein mutants and focussed on binding free energy differences. This leads to partial cancelation of some troublesome contributions, especially rotational/translation/vibrational entropy of the solutes.

In one study, detailed molecular dynamics free energy simulations (MDFE) were done to determine the microscopic basis for the substrate specificity of an aminoacyl-tRNA synthetase enzyme. A Poisson–Boltzmann free energy approach (PBFE) was then used to study several point mutants of the enzyme. Specifically, aspartate and asparagine binding to aspartyl–tRNA synthetase (AspRS) were compared through MDFE. Several point mutations in the amino acid binding pocket were then considered. For each mutant, an MD simulation with explicit solvent was performed to generate structural models. For each ligand (Asp, Asn), free energies of the bound and separated states were calculated by the finite-difference Poisson–Boltzmann method and subtracted. The difference between Asp and Asn gives the electrostatic contribution to the binding free energy difference $\Delta\Delta A$.

A crucial parameter in these calculations is the dielectric constant ϵ of the solutes, assumed to be the same for all solutes and all states. It is difficult [117], though not impossible [58, 115, 118] to evaluate the dielectric 'constant' of a protein; in fact it does not have to be spatially constant [118], and its effective value may depend strongly on the set of atomic charges used (the force field) and the process considered [58]. Therefore, an empirical approach is usually preferred. The strategy used by Archontis et al. [109] was to adjust ϵ to reproduce with PBFE the $\Delta\Delta A$ obtained from MDFE for Asp/Asn binding to native AspRS (15 kcal mol^{-1}). This led to $\epsilon \approx 4$,

a reasonable value which has been used in many, though not all [114, 119] other studies. With the same value for ϵ, good agreement was also found for the Lys198Leu mutant of AspRS.

Another key ingredient is the set of structural models used in the calculations. It is important to use, for each state, structures corresponding to that state. If the native structure was used to calculate properties of the Lys198Leu mutant AspRS, say, large errors were obtained; specifically, the Asp binding decreased by $10\,\text{kcal}\,\text{mol}^{-1}$. It was also important to average results for each state over several structures, taken from a simulation of that state. Indeed, the PBFE results are very sensitive to the details of the structure, so that free energies from instantaneous structures (or the X-ray structure) can deviate by 4–$5\,\text{kcal}\,\text{mol}^{-1}$ from the ensemble average [109].

Hünenberger et al. obtained good agreement with experimental dissociation constants using a single solute dielectric of two for inhibitors binding to cAMP-dependent protein kinase [120]; crystal structures of the various protein:inhibitor complexes were used. Chong et al. used a solute dielectric of one and MD structures of the bound states to study hapten binding to a mature and germ-line antibody [119]; good agreement with the experimental binding free energy difference was obtained. A related, 'alanine-scanning' approach was applied to a fragment of p53 binding to deletion mutants of the oncoprotein Mdm2 [121]. The idea is to study a large number of mutations using the structure of only one variant, e.g., the native protein. A correlation between calculated binding free energies and an experimental mutational tolerance was observed.

Component analyses of Poisson–Boltzmann binding free energies (PBFE) have been proposed by many authors, but a fully systematic treatment [summarized earlier; (12.65)–(12.66)] was developed only recently [108], and applied to GCN4 leucine zipper formation and to amino acid binding by aspartyl–tRNA synthetase [109]. It distinguishes two main effects: partial desolvation of each molecule due to its association with the other, and direct interactions between the charges on the two ligands, screened by the surrounding dielectric media. The direct interaction term is readily decomposed into residue or group contributions. For the desolvation contributions, several approximate decompositions are possible. In the GCN4 dimer [108], electrostatic effects were found to disfavor dimerization by as much as $15\,\text{kcal}\,\text{mol}^{-1}$, due to desolvation of charged and polar groups, insufficiently compensated by direct interactions between monomers in the bound state. Desolvation always reinforces interactions within each ligand, a net favorable effect in this system, but still insufficient to compensate loss of interactions with solvent. In aspartyl–tRNA synthetase [109], the component analysis was used to identify groups that discriminate between binding of the substrate Asp and the analog Asn to the native and mutant proteins. This was especially useful because the most important interactions were not all evident from visual inspection of the structures. Sims et al. [122] analyzed protein kinase–inhibitor binding, and underlined the importance of solvent-mediated intramolecular interactions. Roux and McKinnon [123] analyzed the selectivity of the KcsA potassium channel in terms of particular free energy components, including desolvation components and direct interactions of the ion with the channel helices. The partial ion desolvation that occurs in the channel

favors monovalent over multivalent ions; the field of the pore helices, enhanced by the low-dielectric membrane environment, favors cations over anions. Thus, the channel appears to confer selectivity for monovalent cations by simple electrostatic principles.

12.6.4 The Choice of Dielectric Constant: Proton Binding as a Paradigm

The most important model parameter in PBFE and MM/PBSA is the dielectric constant used for the solutes. Most studies have taken an empirical approach, viewing the dielectric constant as an adjustable parameter. While this seems plausible, it is prudent to analyze the physical problem in more detail, because, in some cases, the experimental data can be fit by models that are distinctly unphysical, despite some plausible features. We therefore come back to the simplest possible PBFE calculation: the important problem of proton binding, or pK_a shifts. We discuss a 'nonempirical' model that attempts to avoid parameter fitting and that gives insights into the limitations of simplified continuum electrostatic free energy methods.

The PB/LRA Method for pK_a Shifts

We saw in Sect. 12.3 that linear response theory leads to a practical method for pK_a calculation. The basic equation, (12.40), relates the proton binding free energy to the electrostatic potentials in the endpoint states (unprotonated and protonated). These can be obtained from MD simulations with explicit solvent, giving the so-called LRA method [59]. An alternative is to do MD simulations with explicit solvent, then discard the solvent molecules and calculate the electrostatic potentials from a continuum model. This is the so-called PB/LRA method [124, 125]. The advantage is that, while a continuum model will usually lead to somewhat poorer protein structures, it can actually give a superior estimate of the electrostatic potentials, partly due to the difficulty in adequately sampling the solvent polarization with explicit solvent methods. Since the protein structures before and after proton binding are sampled by MD, most of the dielectric relaxation of the protein is accounted for explicitly. With fixed-charge force fields, the protein's electronic polarizability is not represented explicitly, so the electronic relaxation is still accounted for implicitly. Therefore, the PB calculations should be performed with a low dielectric constant, presumably between one (if electronic relaxation is not important) and two (if electronic relaxation plays a role). Larger values amount to double counting of the protein's dielectric relaxation.

Using atomic charges borrowed from the force field used in the MD, in combination with a protein dielectric of one or two, the PB/LRA approach gave good agreement with experiment for two highly shifted pK_as, and a larger error of 3 pK_a units for an unshifted carboxylate in thioredoxin [125]. Although no parameter adjustment was done, the results were of the same quality as explicit solvent MDFE results [48]. The level of agreement is thus qualitatively good, and this may be the best that can be expected for a macroscopic continuum model at the molecular scale.

A larger protein dielectric constant of four was used by Eberini et al. [124] to fit the experimental pK_a, in a case where the protein structural relaxation upon protonation was especially large. The need for a larger protein dielectric suggests a breakdown of the linear response assumption for this system. It may be preferable in such a case to simulate an additional point along the reaction pathway, such as the midpoint, rather than shifting to what is effectively a parameter-fitting approach.

Relaxation Free Energy and Internal Consistency of the Model

In the current literature, a different route is usually used to calculate pK_a shifts with continuum models [18–20]. Structures from only one endpoint of the proton binding reaction are used. For an aspartate side chain, for example, a crystal structure with the ionized Asp is typically available. The free energy is not decomposed into a static and a relaxation term; the total binding free energy ΔA is calculated directly. Nevertheless, a component analysis of the method is very instructive, and provides a powerful test of the internal consistency or inconsistency of the model. Indeed, the static and relaxation free energies can always be computed separately, then added to give ΔA (12.32). The static free energy has already been considered in detail (12.29), (12.37), (12.38). Despite its rather abstract form (12.30), (12.31), the relaxation free energy is just as easy to compute. In fact, the relaxation free energy is equal to the Born self-energy of the charge increments that are used to model the proton [58]. Thus, ΔA_{stat} and ΔA_{rlx} are readily obtained from the structure of the reactant state (with ionized Asp in our example). The relaxation free energy computed in this way has an essential dependency on the protein dielectric constant. It has no dependency whatsoever on the set of atomic charges used in the model, other than the charge increments that represent the inserted proton. Exactly the same results will be obtained with two different force fields. In contrast, the static free energy depends strongly on the choice of atomic charges for the whole protein.

Importantly, another expression also exists for ΔA_{rlx}, (12.32). This expression makes explicit use of both the reactant and product structures, and so contains an explicit representation of the structural relaxation. The implicit (12.31) and explicit (12.32) representations of the relaxation are equivalent

$$\Delta A_{\text{rlx}} = -\frac{\beta}{2} \, \text{Variance}(\eta) = \frac{\langle \eta \rangle_1 - \langle \eta \rangle_0}{2}. \tag{12.68}$$

As mentioned, this equivalence is a consequence of the fluctuation–dissipation theorem (the general basis of linear response theory [51]). In (12.68), we have dropped nonlinear terms and we have not indicated for which state $\text{Variance}(\eta)$ is computed (because the reactant and product state results only differ by nonlinear terms). We see that ΔA, ΔA_{stat}, and ΔA_{rlx} are all linked and are all sensitive to the model parameters, with different computational routes giving a different sensitivity for ΔA_{rlx}.

In most pK_a calculations with the 'standard' method, a moderate to high ($\epsilon_p =$ 4–20) protein dielectric is used. Often, a molecular mechanics charge set is used. In cases where protonation induces a large protein conformational relaxation, this combination is likely to give a poor consistency between the two underlying free energy

components. That is, if ΔA_{stat} and ΔA_{rlx} are computed separately, using the same model parameters and the same structure, then the consistency relation (12.68) will be violated [58, 125]. For ionization of Aspartate26 in thioredoxin, the relaxation free energy from MDFE was about $-56\,\text{kcal}\,\text{mol}^{-1}$, corresponding to a large protein reorganization upon ionization. To reproduce this value with a continuum model, a protein dielectric of three is needed. This dielectric value is typical for a protein interior [88, 117, 118]. In contrast, for the same ionization reaction, a dielectric of 1–2 was optimal to reproduce the equilibrium potentials and the static free energy obtained by MDFE. A dielectric of three gives an error of over $10\,\text{kcal}\,\text{mol}^{-1}$ for ΔA_{stat}. This apparent discrepancy is not surprising, since the molecular mechanics charge set used for MDFE was optimized with a dielectric of one. For cases such as this, where protein reorganization is large, the 'standard' method combines a charge set and a dielectric constant that are mutually inconsistent.

In contrast, for cases where the protein is more rigid, the standard continuum approach can give excellent results. A striking example is the case of photosystems and redox proteins, where a low reorganization is needed to maintain fast charge-transfer kinetics. For these systems, carefully parameterized continumm models can give an accurate picture of redox potentials and their coupling to acid/base reactions [126–128].

12.7 Conclusions

When MDFE methods were first developed for proteins, they raised great hopes as a potential new tool for improving lead molecules in drug design [129]. Unfortunately, MDFE methods are costly and complex, and these particular hopes have never been fulfilled. Rather, computational studies for drug design are always based on simplified methods [10]. Nevertheless, MDFE has begun to play an indirect role in drug design, fulfilling its early promise in two ways. On the one hand, it has largely inspired the simplified free energy methods discussed in this chapter. On the other hand, it has provided an essential benchmark for the simplified methods, complementary to experimental data. For example, MDFE can be used both to parameterize a PBFE model and to test some of its specific predictions. The rigorous and simplified methods can employ the same charge set and the same molecular structures, obtained from the same MD or MC simulations.

Simple free energy methods usually decompose the free energy into a few distinct contributions, such as the van der Waals and electrostatic terms in the LIE equation (12.61). They usually make use of only one or two simulations, and rely on an assumption of linear response, at least for some of the free energy contributions. In fact, the individual free energy contributions depend on the choice of a specific reaction pathway. For a ligand binding reaction, an instructive pathway is the one shown in Fig. 12.5, where the purely nonpolar interactions, the attractive dispersion interactions, and the Coulomb interactions are all treated in separate steps. With this pathway, it is apparent that linear response is unlikely to hold for all the steps, and so no rigorous, analytical free energy formula exists. Rather, we can take an empirical

view, inspired by the scaling behavior of solvation free energies in simple liquids, where nonpolar and dispersion terms have a simple behavior. This leads to the modified, or extended LIE methods that have become popular [13–16]. With careful parametrization, and in combination with other experimental and computational methods, they provide a useful and predictive tool.

PBFE and MM/PBSA provide a related, but different set of tools that can give good qualitative accuracy if they are carefully parameterized. They are especially useful for processes dominated by electrostatics, such as comparisons between a charged and a neutral ligand [12]. However, they must be used with caution, particularly when protein reorganization is significant. If a very high protein dielectric constant ($\epsilon \sim 20$–80) is needed to reproduce an experimental result, it is possible that the physical basis of the continuum model has broken down for the problem at hand. For example, the linear response approximation may not be verified. In the case of acid/base reactions with unusual pK_a values, it is highly recommended to obtain structures for both endpoints of the ionization reaction (possibly from MD simulations), and to check that calculations using either of the two endpoints give similar results. If not, one may end up reproducing an experimental result simply through parameter fitting with an unphysical model.

As computers become even faster and force fields expand to cover more types of molecules, both MDFE and simplified methods will be increasingly useful. The most fruitful approach will be to use a spectrum of methods, both for ligand design and for fundamental studies of biomolecular thermodynamics.

References

1. Gilson, M.; Given, J.; Bush, B.; McCammon, J.A., The statistical-thermodynamic basis for computation of binding affinities: a critical review, *Biophys. J.* **1997**, *72*, 1047–1069
2. Fersht, A., *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, Freeman: New York, 1999
3. Tembe, B.; McCammon, J.A., Ligand–receptor interactions, *Comput. Chem.* **1984**, *8*, 281–283
4. Reinhardt, W.P.; Miller, M.A.; Amon, L.A., Why is it so difficult to simulate entropies, free energies, and their differences? *Acc. Chem. Res.* **2001**, *34*, 607–614
5. Hummer, G.; Szabo, A., Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 3658–3661
6. Simonson, T. Free energy calculations. in *Computational Biochemistry & Biophysics*, Becker, O.; Mackerell Jr., A.; Roux, B.; Watanabe, M., Eds. Marcel Dekker: New York, 2001, ch. 9
7. Miyamoto, S.; Kollman, P., Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches, *Proteins* **1993**, *16*, 226–245
8. Lamb, M.L.; Jorgensen, W.J., Computational approaches to molecular recognition, *Curr. Opin. Chem. Biol.* **1997**, *1*, 449–457
9. Woo, H.J.; Roux, B., Calculation of absolute protein ligand binding free energy from computer simulations, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6825–6830

10. Alvarez, J.; Shoichet, B., *Virtual Screening in Drug Discovery*, CRC: West Palm Beach, FL, USA, 2005

11. Roux, B.; Simonson, T., Implicit solvent models, *Biophys. Chem.* **1999**, *78*, 1–20

12. Simonson, T.; Archontis, G.; Karplus, M., Free energy simulations come of age: the protein–ligand recognition problem, *Acc. Chem. Res.* **2002**, *35*, 430–437

13. Aqvist, J.; Luzhkov, V.B.; Brandsal, B.O., Ligand binding affinities from MD simulations, *Acc. Chem. Res.* **2002**, *35*, 358–365

14. Aqvist, J.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V.B.; Brandsal, B.O., Free energy calculations and ligand binding, *Adv. Prot. Chem.* **2003**, *66*, 123–158

15. Wong, C.F.; McCammon, J.A., Protein simulation and drug design, *Adv. Prot. Chem.* **2003**, *66*, 87–121

16. Jorgensen, W.L., The many roles of computation in drug discovery, *Science* **2003**, *303*, 1813–1818

17. Warshel, A.; Levitt, M., Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, *J. Mol. Biol.* **1976**, *103*, 227

18. Bashford, D.; Karplus, M., The pK_a's of ionizable groups in proteins: atomic detail from a continuum electrostatic model, *Biochemistry* **1990**, *29*, 10219–10225

19. Antosiewicz, J.; McCammon, J.A.; Gilson, M., The determinants of pK_a's in proteins, *Biochemistry* **1996**, *35*, 7819–7833

20. Schaefer, M.; Vlijmen, H.W.T. van; Karplus, M., Electrostatic contributions to molecular free energies in solution, *Adv. Prot. Chem.* **1998**, *51*, 1–57

21. Beveridge, D.; DiCapua, F., Free energy via molecular simulation: applications to chemical and biomolecular systems, *Ann. Rev. Biophys. Chem.* **1989**, *18*, 431–492

22. van Gunsteren, W.; Beutler, T.C.; Fraternali, F.; King, P.M.; Mark, A.E.; Smith, P.E., Computation of free energy in practice: choice of approximations and accuracy limiting factors. In *Computer Simulation of Biomolecular Systems*, van Gunsteren, W.; Weiner, P.; Wilkinson, A., Eds. Escom Science: Leiden, 1993, pp. 315–348

23. Fowler, R.H.; Guggenheim, E.A., *Statistical Thermodynamics*, Cambridge University Press: Cambridge, 1939

24. Zwanzig, F., High-temperature equation of state by a perturbation method. I. Non-polar gases , *J. Chem. Phys.* **1954**, *22*, 1420

25. Liu, H.; Mark, A.; van Gunsteren, W.F., Estimating the relative free energy of different molecular states with respect to a single reference state, *J. Phys. Chem.* **1996**, *100*, 9485–9494

26. von Mises, R., *Mathematical Theory of Probability and Statistics*, Academic: New York, 1964

27. Jorgensen, W.; Buckner, K.; Boudon, S.; Tirado-Rives, J., Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water, *J. Chem. Phys.* **1988**, *89*, 3742–3746

28. Roux, B.; Nina, M.; Pomes, R.; Smith, J., Thermodynamic stability of water molecules in the Bacteriorhodopsin proton channel: a molecular dynamics and free energy perturbation study, *Biophys. J.* **1996**, *71*, 670–681

29. Wong, C.; McCammon, J.A., Dynamics and design of enzymes and inhibitors, *J. Am. Chem. Soc.* **1986**, *108*, 3830–3832

30. Wong, C.F.; Thacher, T.; Rabitz, H., Systematic sensitivity analysis. In *Reviews in Computational Chemistry*, Lipkowitz, K.; Boyd, D., Eds. Wiley: New York, 1998, pp. 281–326

31. Radmer, R.J.; Kollman, P.A., The application of three approximate free energy calculations methods to structure based ligand design: trypsin and its complex with inhibitors., *J. Comput. Aided Mol. Des.* **1998**, *12*, 215–227

32. Schäfer, H.; Mark, A.; van Gunsteren, W.F., Estimating relative free energies from a single ensemble: hydration free energies, *J. Comput. Chem.* **1999**, *20*, 1604–1617

33. Oostenbrink, C.; Pitera, J.W.; van Lipzig, M.M.H.; Meerman, J.H.N.; van Gunsteren, W.F., Simulations of the estrogen receptor ligand-binding domain: affinity of natural ligands and xenoestrogens, *J. Med. Chem.* **2000**, *43*, 4594–4605

34. Oostenbrink, C.; van Gunsteren, W.F., Free energies of binding of polychlorinated biphenyls to the estrogen receptor in a single simulation, *Proteins* **2004**, *54*, 237–246

35. Simonson, T., Free energy of particle insertion. An exact analysis of the origin singularity for simple liquids, *Molec. Phys.* **1993**, *80*, 441–447

36. Resat, H.; Mezei, M., Studies on free energy calculations. II. A theoretical approach to molecular solvation, *J. Chem. Phys.* **1994**, *222*, 6126–6140

37. Wong, C.F.; Hunenberger, P.H.; Akamine, P.; Narayana, N.; Diller, T.; McCammon, J.A.; Taylor, S.; Xuong, N.H., Computational analysis of PKA–balanol interactions, *J. Med. Chem.* **2001**, *44*, 1530–1539

38. Mordasini, T.Z.; McCammon, J.A., Calculations of relative hydration free energies: a comparative study using thermodynamic integration and an extrapolation method based on a single reference state, *J. Phys. Chem. B* **2000**, *104*, 360–367

39. Kong, X.; Brooks, C.L., λ-dynamics: a new approach to free energy calculations, *J. Chem. Phys.* **1996**, *105*, 2414–2423

40. Pomes, R.; Eisenmesser, E.; Post, C.B.; Roux, B., Calculating excess chemical potentials using dynamic simulations in the fourth dimension, *J. Chem. Phys.* **1999**, *111*, 3387–3395

41. Simonson, T.; Perahia, D.; Brünger, A.T., Microscopic theory of the dielectric properties of proteins., *Biophys. J.* **1991**, *59*, 670–690

42. Simonson, T.; Perahia, D., Microscopic dielectric properties of cytochrome c from molecular dynamics simulations in aqueous solution, *J. Am. Chem. Soc.* **1995**, *117*, 7987–8000

43. Del Buono, G.S.; Figueirido, F.E.; Levy, R., Intrinsic pK_a's of ionizable residues in proteins: an explicit solvent calculation for lysozyme, *Proteins* **1994**, *20*, 85–97

44. Simonson, T.; Wong, C.; Brünger, A.T., Classical and quantum simulations of tryptophan in solution, *J. Phys. Chem. A* **1997**, *101*, 1935–1945

45. Simonson, T.; Archontis, G.; Karplus, M., Continuum treatment of long-range interactions in free energy calculations. Application to protein–ligand binding, *J. Phys. Chem. B* **1997**, *101*, 8349–8362

46. Ceccarelli, M.; Marchi, M., Simulation and modeling of the *Rhodobacter spaeroides* bacterial reaction center, *J. Phys. Chem. B* **2003**, *107*, 1423–1431

47. Simonson, T., Gaussian fluctuations and linear response in an electron transfer protein, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 6544–6549

48. Simonson, T.; Carlsson, J.; Case, D.A., Proton binding to proteins: pK_a calculations with explicit and implicit solvent models, *J. Am. Chem. Soc.* **2004**, *126*, 4167–4180

49. Chandler, D., *Introduction to Modern Statistical Mechanics*, Oxford University Press: Oxford, 1987

50. Hansen, J.P.; McDonald, I., *Theory of Simple Liquids*, Academic: New York, 1986

51. Landau, L.; Lifschitz, E., *Statistical Mechanics*, Pergamon: New York, 1980

52. Warshel, A., Dynamics of reactions in polar solvents. Semiclassical trajectory studies of electron transfer and proton transfer studies, *J. Phys. Chem.* **1982**, *86*, 2218–2224

53. Gehlen, J.N.; Marchi, M.; Chandler, D., Dynamics affecting the primary charge transfer in photosynthesis, *Science* **1994**, *263*, 499

54. Marcus, R., Electron transfer reactions in chemistry: theory and experiment. In *Protein Electron Transfer* (1996), Bendall, D., Ed., BIOS Scientific: Oxford, pp. 249–272

55. Marcus, R., Chemical and electro-chemical electron transfer theory, *Ann. Rev. Phys. Chem.* **1964**, *15*, 155–196

56. Marcus, R., On the theory of shifts and broadening of electronic spectra of polar solutes in polar media, *J. Chem. Phys.* **1965**, *43*, 1261–1274

57. Muegge, I.; Qi, P.X.; Wand, A.J.; Chu, Z.T.; Warshel, A., Reorganization energy of cytochrome c revisited, *J. Phys. Chem. B* **1997**, *101*, 825–836

58. Simonson, T.; Archontis, G.; Karplus, M., A Poisson-Boltzmann study of charge insertion in an enzyme active site: the effect of dielectric relaxation, *J. Phys. Chem. B* **1999**, *103*, 6142–6156

59. Sham, Y.Y.; Chu, Z.T.; Warshel, A., Consistent calculations of pK_a's of ionizable residues in proteins: semi-microscopic and microscopic approaches, *J. Phys. Chem. B* **1997**, *101*, 4458–4472

60. Warshel, A.; Sussman, F.; King, G., Free energy changes in solvated proteins: microscopic calculations using a reversible charging process, *Biochemistry* **1986**, *25*, 8368–8372

61. Kirkwood, J., Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **1935**, *3*, 300–313

62. Reiss, H., Scaled particle methods in the statistical thermodynamics of fluids, *Adv. Chem. Phys.* **1965**, *9*, 1–84

63. McQuarrie, D., *Statistical Mechanics*, Harper and Row: New York, 1975

64. Stillinger, F., Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory, *J. Sol. Chem.* **1973**, *2*, 141–158

65. Pierotti, R.A., A scaled particle theory of aqueous and nonaqueous solutions, *Chem. Rev.* **1976**, *76*, 717–726

66. Tanford, C., *The Hydrophobic Effect*, Wiley: New York, 1980

67. Tolman, R.C., Consideration of the Gibbs theory of surface tension, *J. Chem. Phys.* **1948**, *16*, 758–774

68. Postma, J.; Berendsen, H.; Haak, J., Thermodynamics of cavity formation in water. A molecular dynamics study, *Far. Symp. Chem. Soc.* **1982**, *17*, 55–67

69. Straatsma, T.; Berendsen, H.; Postma, J., Free energy of hydrophobic hydration: a molecular dynamics study of noble gases in water, *J. Chem. Phys.* **1986**, *85*, 6720

70. Pratt, L.; Pohorille, A., Theory of hydrophobicity: transient cavities in molecular liquids, *Proc. Natl Acad. Sci. USA* **1992**, *89*, 2995–2999

71. Hummer, G.; Garde, S.; Garcia, A.E.; Pohorille, A.; Pratt, L.R., An information theory model of hydrophobic interactions, *Proc. Natl Acad. Sci. USA* **1996**, *93*, 8951–8955

72. Pratt, L.R.; Pohorille, A., Hydrophobic effects and modeling of biophysical aqueous solution interfaces, *Chem. Rev.* **2002**, *102*, 2671–2691

73. Sitkoff, D.; Sharp, K.; Honig, B., Accurate calculation of hydration free energies using macroscopic solvent models, *J. Phys. Chem.* **1994**, *98*, 1978–1988

74. Simonson, T.; Brünger, A.T., Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment, *J. Phys. Chem.* **1994**, *98*, 4683–4694

75. Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D.A.; Cheatham, T.E., Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc. Chem. Res.* **2000**, *33*, 889–897

76. Gallicchio, E.; Kubo, M.M.; Levy, R.M., Enthalpy–entropy and cavity decomposition of alkane hydration free energies: numerical results and implications for theories of hydrophobic hydration, *J. Phys. Chem. B* **2000**, *104*, 6271–6285

77. Shurki, A.; Warshel, A., Structure/function correlations of proteins using MM, QM/MM and related approaches: methods, concepts, pitfalls, and current progress, *Adv. Prot. Chem.* **2003**, *66*, 249–313

78. Born, M., Volumen und hydrationswärme der ionen, *Zeit. Phys.* **1920**, *1*, 45–48

79. Kirkwood, J., Theory of solutions of molecules containing widely separated charges with special application to zwitterions, *J. Chem. Phys.* **1934**, *2*, 351–361

80. Onsager, L., Electric moments of molecules in liquids, *J. Am. Chem. Soc.* **1936**, *58*, 1486

81. Jean-Charles, A.; Nicholls, A.; Sharp, K.; Honig, B.; Tempzyck, A.; Hendrickson, T.; Still, W.C., Electrostatic contributions to solvation energies: comparison of free energy perturbation and continuum calculations., *J. Am. Chem. Soc.* **1991**, *113*, 1454–1455

82. Nina, M.; Beglov, D.; Roux, B., Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations, *J. Phys. Chem. B* **1997**, *101*, 5239–5248

83. Jackson, J.D., *Classical Electrodynamics*, Wiley: New York, 1975

84. Landau, L.; Lifschitz, E., *Electrodynamics of Continuous Media*, Pergamon: New York, 1980

85. Warwicker, J.; Watson, H., Calculation of the electrostatic potential in the active site cleft due to α helix dipoles, *J. Mol. Biol.* **1982**, *157*, 671–679

86. Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B., Focussing of electric fields in the active site of Cu–Zn superoxide dismutase, *Proteins* **1986**, *1*, 47

87. Holst, M.J.; Baker, N.A.; Wang, F., Adaptive multilevel finite element solution of the Poisson–Boltzmann equation I: algorithms and examples, *J. Comp. Chem.* **2000**, *21*, 1319–1342

88. Simonson, T., Electrostatics and dynamics of proteins, *Rep. Prog. Phys.* **2003**, *66*, 737–787

89. Fröhlich, H., *Theory of Dielectrics*, Clarendon: Oxford, 1949

90. Juffer, A.; Botta, E.; van Keulen, B.; van der Ploeg, A.; Berendsen, H., The electric potential of a macromolecule in a solvent: a fundamental approach, *J. Comp. Phys.* **1991**, *97*, 144

91. Ghosh, A.; Rapp, C.S.; Friesner, R.A., Generalized Born model based on a surface area formulation, *J. Phys. Chem. B* **1998**, *102*, 10983–10990

92. Åqvist, J.; Median, C.; Samuelsson, J.E., A new method for predicting binding affinity in computer-aided drug design, *Prot. Eng.* **1994**, *7*, 385–391

93. Basu, G.; Kitao, A.; Kuki, A.; Go, N., Protein electron transfer reorganization energy from normal mode analysis. 1. Theory, *J. Phys. Chem. B* **1998**, *102*, 2076–2084

94. Gilson, M.; Honig, B., Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis, *Proteins* **1988**, *4*, 7–18

95. Carlson, H.A.; Jorgensen, W.L., An extended linear response method for determining free energies of hydration, *J. Phys. Chem.* **1995**, *99*, 10667–10673

96. Jones-Hertzog, D.K.; Jorgensen, W.L., Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method, *J. Med. Chem.* **1997**, *40*, 1539–1549

97. Lamb, M.L.; Tirado-Rives, J.; Jorgensen, W.L., Estimation of the binding affinities of FKBP12 inhibitors using a linear response method, *Bioorg. Med. Chem.* **1999**, *7*, 851–860

98. McDonald, N.A.; Carlson, H.A.; Jorgensen, W.L., Free energies of solvation in chloroform and water from a linear response approach, *J. Phys. Org. Chem.* **1997**, *10*, 563–576

99. Graffner-Nordberg, M.; Kolmodin, K.; Aqvist, J.; Queener, S.F.; Hallberg, A., Design, synthesis, computational prediction, and biological evaluation of ester soft drugs as inhibitors of dihydrofolate reductase from *Pneumocystis carinii*, *J. Med. Chem.* **2001**, *44*, 2391–2402

100. Ljungberg, K.B.; Marelius, J.; Musil, D.; Svensson, P.; Norden, B.; Aqvist, J., Computational modelling of inhibitor binding to human thrombin, *Eur. J. Pharm. Sci.* **2001**, *12*, 441–446

101. Wall, I.D.; Leach, A.R.; Salt, D.W.; Ford, M.G.; Essex, J.W., Binding constants of neuraminidase inhibitors: an investigation of the linear interaction energy method, *J. Med. Chem.* **1999**, *42*, 5142–5152

102. Ostrovsky, D.; Udier-Blagovic, M.; Jorgensen, W.L., Analyses of activity for factor $X\alpha$ inhibitors based on Monte Carlo simulations, *J. Med. Chem.* **2003**, *46*, 5691–5699

103. Kroeger-Smith, M.B.; Hose, B.M.; Hawkins, A.; Lipchock, J.; Farnsworth, D.W.; Rizzo, R.C.; Tirado-Rives, J.; Arnold, E.; Zhang, W.; Hughes, S.H.; Jorgensen, W.L.; Michedja, C.J.; Smith, R.H., Molecular modeling calculations of HIV-1 reverse transcriptase nonnucleoside inhibitors: correlation of binding energy with biological activity for novel 2-aryl-substituted benzimidazole analogues, *J. Med. Chem.* **2003**, *46*, 1940–1947

104. Kubinyi, H.; Folkers, G.; Martin, Y.C., *3D QSAR in Drug Design: Volume 2: Ligand–Protein Interactions and Molecular Similarity*, Springer: Berlin, Heidelberg, New York, 2000

105. Zhou, R.; Friesner, R.A.; Ghosh, A.; Rizzo, R.C.; Jorgensen, W.L.; Levy, R.M., New linear interaction method for binding affinity calculations using a continuum solvent model, *J. Phys. Chem. B* **2001**, *105*, 10388–10397

106. Sham, Y.Y.; Shao, Z.T.; Tao, H.; Warshel, A., Examining methods for calculations of binding free energies: LRA, LIE and PDLD/S-LRA calculations of ligand binding to an HIV protease, *Proteins* **2000**, *39*, 393–407

107. Schaefer, M.; Froemmel, C., A precise analytical method for calculating the electrostatic energy of macromolecules in aqueous solution, *J. Mol. Biol.* **1990**, *216*, 1045–1066

108. Hendsch, Z.; Tidor, B., Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding, *Prot. Sci.* **1999**, *8*, 1381–1392

109. Archontis, G.; Simonson, T.; Karplus, M., Binding free energies and free energy components from molecular dynamics and Poisson–Boltzmann calculations. Application to amino acid recognition by aspartyl–tRNA synthetase, *J. Mol. Biol.* **2001**, *306*, 307–327

110. Schapira, M.; Totrov, M.; Abagyan, R., Prediction of the binding energy for small molecules, peptides and proteins, *J. Molec. Recog.* **1999**, *12*, 177–190

111. Huron, M.J.; Claverie, P., Calculation of the interaction energy of one molecule with its whole surroundings. I. Method and application to pure nonpolar compounds., *J. Phys. Chem.* **1972**, *76*, 2123–2133

112. Claverie, P.; Daudey, J.P.; Langlet, J.; Pullman, B.; Piazzola, D.; Huron, M.J., Studies of solvent effects. 1. Discrete, continuum, and discrete-continuum models and their comparison for some simple cases: NH_4^+, CH_3OH, and substituted NH_4^+, *J. Phys. Chem.* **1978**, *82*, 405–418

113. Floris, F.M.; Tomasi, J.; Pascal-Ahuir, J.L., Dispersion and repulsion contributions to the solvation energy: refinements to a simple computational model in the continuum approximation, *J. Comp. Chem.* **1991**, *12*, 784–791

114. Srinivasan, J.; Cheatham, T.E.; Cieplak, P.; Kollman, P.A.; Case, D.A., Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices, *J. Am. Chem. Soc.* **1998**, *120*, 9401–9409

115. Simonson, T., Macromolecular electrostatics: continuum models and their growing pains, *Curr. Opin. Struct. Biol.* **2001**, *11*, 243–252

116. Jencks, W.P., *Catalysis in Chemistry and Enzymology*, Dover: New York, 1986

117. Schutz, C.N.; Warshel, A., What are the dielectric 'constants' of proteins and how to validate electrostatic models?, *Proteins* **2001**, *8*, 211–217

118. Simonson, T.; Perahia, D., Internal and interfacial dielectric properties of cytochrome c from molecular dynamics simulations in aqueous solution, *Proc. Natl Acad. Sci. USA* **1995**, *92*, 1082–1086

119. Chong, L.T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P., Molecular dynamics and free energy calculations applied to affinity maturation in antibody 48G7, *Proc. Natl Acad. Sci. USA* **1999**, *96*, 14330–14335

120. Hunenberger, P.; Helms, V.; Narayana, N.; Taylor, S.S.; McCammon, J.A., Determinants of ligand binding to cAMP-dependent protein kinase, *Biochemistry* **1999**, *38*, 2358–2366

121. Massova, I.; Kollman, P., Computational alanine scanning to probe protein–protein interactions: a novel approach to evaluate binding free energies, *J. Am. Chem. Soc.* **1999**, *121*, 8133–8143

122. Sims, P.A.; Wong, C.F.; Vuga, D.; McCammon, J.A.; Sefton, B.F., Relative contributions of desolvation, inter- and intramolecular contributions to binding affinity in protein kinase systems, *J. Comput. Chem.* **2005**, *26*, 668–681

123. Roux, B.; MacKinnon, R., The cavity and pore helices in the KcsA K^+ channel: electrostatic stabilization of monovalent cations, *Science* **1999**, *285*, 100–102

124. Eberini, I.; Baptista, A.M.; Gianazza, E.; Fraternali, F.; Beringhelli, T., Reorganization in apo- and holo-β-lactoglobulin upon protonation of Glu89: molecular dynamics and pK_a calculations, *Proteins* **2004**, *54*, 744–758

125. Archontis, G.; Simonson, T., Proton binding to proteins: a free energy component analysis using a dielectric continuum model, *Biophys. J.* **2005**, *88*, 3888–3904

126. Baptista, A.M.; Martel, P.J.; Soares, C.M., Simulation of electron–proton coupling with a Monte-Carlo method: application to cytochrome c_3 using continuum electrostatics, *Biophys. J.* **1999**, *76*, 2978–2998

127. Ullmann, M., The coupling of protonation and reduction in proteins with multiple redox centers: theory, computational method and application to cytochrome c_3, *J. Phys. Chem. B* **2000**, *104*, 6293–6301

128. Ishikita, H.; Morra, G.; Knapp, E.W., Redox potential of quinones in photosynthetic reaction centers from *Rhodobacter sphaeroides*: dependence on protonation of Glu-L212 and Asp-L213, *Biochemistry* **2003**, *42*, 3882–3892

129. Kollman, P.A., Free energy calculations: applications to chemical and biochemical phenomena, *Chem. Rev.* **1993**, *93*, 2395

130. Mackerell, A.D.; Bashford, D.; Bellott, M.; Dunbrack, R.L.; Evanseck, J.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D.T.; Prodhom, B.; Reiher, W.E.; Roux, B.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., An all-atom empirical potential for molecular modelling and dynamics study of proteins, *J. Phys. Chem. B* **1998**, *102*, 3586–3616

131. Ben Naim, A.; Marcus, Y., Solvation thermodynamics of nonionic solutes, *J. Chem. Phys.* **1984**, *81*, 2016–2027

13

Applications of Free Energy Calculations to Chemistry and Biology

Christophe Chipot, Alan E. Mark, Vijay S. Pande, and Thomas Simonson

13.1 Introduction

A complete understanding of most chemical and biochemical processes requires a careful examination of the underlying free energy behavior. Solvation and transport properties, protein–ligand binding, proton and electron transfer reactions are all of major interest, and evidently cannot be understood or predicted without a knowledge of the associated free energy changes. The ability to determine a priori the associated physical constants with a reasonable level of reliability, using statistical simulations, is within reach today. Large, realistic, physical and biological assemblies remain a challenge for modern theoretical chemistry. However, enormous progress has been made since the first attempts, published over two decades ago. Developments on several fronts, including theory, software, and hardware have helped to bring free energy calculations to the level of robust and well-characterized modeling tools, while broadening their field of applications. Taking advantage of massively parallel architectures, for example, cost-effective, precise and accurate free energy calculations can help rationalize experimental observations, and, in some cases, be predictive. In this chapter, we review applications of free energy calculations in chemistry and biology, ranging from the estimation of solvation free energies, protein–ligand binding constants, partition coefficients, conformational equilibrium constants, and acid–base and redox constants. Our goal is not to provide an exhaustive account of all possible simulations published hitherto, but rather highlight a number of applications that illustrate why one might want to perform free energy calculations, what one may learn from these that is not easily learned from experiment, and what new insight into interesting chemical or biological processes one may gain. Conclusions on the role played by free energy calculations in the molecular modeling community are drawn, with a prospective look into their promising future.

13.2 Protein–Ligand Association

13.2.1 Relative Protein–Ligand Binding Constants

The cost of a molecular dynamics (MD) free energy study depends very much on both the system and the goal of the study. If the goal is to reproduce qualitatively an experimental number and interpret it in terms of microscopic interactions, and if the systems of interest (e.g., native and mutant protein) are very similar, then only limited conformational sampling will be needed in most cases, and a few short runs with a small model may suffice.

If the goal is to predict accurately and precisely an unknown free energy difference, or if the transformation involves large conformational changes, MD free energy calculations can be more costly. Not only are many long simulations needed, but comparisons between different force fields may be necessary to assess the accuracy. One such study examined the binding of the native Raf protein and its Arg89Lys mutant to the signaling protein Ras [1]. Experimentally, only a lower bound (3 kcal/mol) was known for the reduction in binding due to the mutation; this number was measured [1] after the MD free energy study was performed. MD simulations showed that two very different conformations of the region around Arg89 are populated in both the native and mutant Raf. By using multiple MD free energy runs with three different force fields, and by identifying and exploring the important conformers using biased sampling techniques, it was shown that the reduction in binding is 3 ± 2 kcal/mol, close to the experimental lower bound. Interestingly, the calculations showed that the weaker binding of the Arg89Lys mutant protein comes from a stronger solvation of the Lys89 side chain in the mutant protein in the unbound state. This is presumably a general effect, which contributes to the known, lower propensity of Lys (compared to Arg) to participate in protein–protein interfaces. This effect would have been difficult to observe by experiment alone.

A second example of a difficult case is the enzyme aspartyl–tRNA synthetase (AspRS). Aminoacyl–tRNA synthetases attach a specific amino acid to a tRNA that bears the appropriate anticodon, establishing the amino acid–trinucleotide correspondence that forms the genetic code. Engineering amino acid specificity is an important goal in biotechnology, which has already led to bacteria with an extended or reduced genetic code [2, 3]. In particular, this provides a route for introducing artificial amino acids into proteins in vivo [3, 4]. AspRS was the object of a series of MD free energy studies, which aimed to understand and modify the specific binding of its substrate, L-aspartate (Asp), and its discrimination against chemical analogues such as asparagine (Asn) and D-aspartate [5–7]. Neither the binding constant for Asn nor the X-ray structures of the AspRS:Asn and AspRS:D-Asp complexes were known experimentally. This system represents a difficult electrostatic problem. The substrate Asp is charged, while Asn is neutral. The co-substrate ATP can bind two or three Mg^{2+} cations, for a total charge of either zero or two. Two nearby histidines can be charged or neutral. A flexible, 'flipping' loop can close over the amino acid, bringing a negative glutamate to coordinate the amino acid's ammonium group. To enhance Asn binding, mutations of several nearby residues were considered, modifying the net charge of the binding pocket further. All these charged groups can
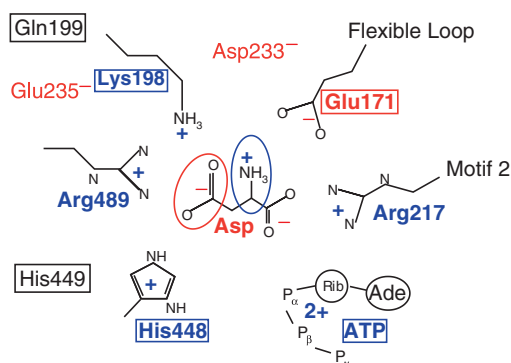
Fig. 13.1. Cartoon of the aspartyl–tRNA synthetase amino acid binding site. The aspartate ligand is shown, along with the most important recognition residues. Groups that have been mutated in free energy simulations are boxed or circled. 'Flexible loop' and 'Motif 2' refer to conserved motifs in the enzyme structure

couple to each other, so that a complex network of interactions and many alternate electrostatic states must be considered. Figure 13.1 shows a cartoon of the active site, including the most important specificity determinants.

In the MD free energy studies, over 15 different states were considered, corresponding to six different values of the net charge in the binding pocket (from -2 to $+3$). The inhomogeneous continuum reaction field method was used (see below) [8, 9], so that all electrostatic interactions were explicitly included. Experimental measurements were available or performed specifically to compare Asp, Asn, D-Asp, and succinate binding. For Asp versus Asn, the discrimination is very strong, so that only a lower bound could be obtained experimentally for the binding free energy difference, $\Delta\Delta G \geq 7$ kcal/mol. This was sufficient to support important features of the computational model. Good agreement was obtained for D-Asp and succinate, which bind less strongly than Asp, but less weakly than Asn. Thus validated, the computations revealed several interesting features. Long-range interactions electrostatically couple the amino acid ligand, ATP and its associated Mg^{2+} cations, a histidine side chain (His448) next to the amino acid ligand, and the 'flipping' loop, which closes over the active site in response to amino acid binding. Closing this loop brings a negative glutamate into the active site; this causes His448 to recruit a labile proton, which interacts favorably with Asp and accounts for most of the Asp/Asn discrimination. Co-binding of the second substrate, ATP, increases specificity for Asp further and makes the system robust towards removal of His448, which is mutated to a neutral amino acid in many organisms. Thus, AspRS specificity is assisted by a labile proton and a co-substrate, and ATP acts as a mobile discriminator for specific Asp binding to AspRS [5–7].

Alchemical transformations have also been applied to the challenging case of G protein-coupled receptors (GPCRs), for which little structural information is available experimentally at the atomic level. Starting from a template of a seven-helix

transmembrane domain, a model of the human cholecystokinin-1 receptor (CCK1R) was refined through successive stages, using a host of experimental data, including site-directed mutagenesis experiments of both the ligand and the receptor [10]. The in vacuo construct was then immersed in a realistic membrane environment and examined over a period of 30 ns, during which the receptor and the solvent were relaxed concomitantly. Free energy perturbation (FEP) calculations were carried out to mutate the agonist nonapeptide CCK9 at its N- and C-termini, both in the free and in the bound states [11]—viz. S-Tyr^3 to Tyr and Asp^8 to Ala. The very good reproduction of the experimental binding constants opens new vistas for the design of potent agonists and antagonists in GPCRs in the absence of experimentally resolved three-dimensional structures.

In the field of membrane proteins, free energy calculations may also help interpret inferences based on experiments. For instance, to understand how ion selectivity is controlled in voltage-gated potassium channels, Benoît Roux and coworkers applied FEP calculations to KcsA, artificially disrupting the interaction of the cation K$^+$ with the different carbonyl moieties lining the narrow pore region of the channel [12]. Comparing simulations performed in a fully flexible channel and in a frozen one, they shed new light on the dynamic nature of the pore region, which, combined with the intrinsic electrostatic properties of the participating carbonyl groups, ensures ionic selectivity — in particular K$^+$ versus Na$^+$ ions. It is worth noting that removal of carbonyl–carbonyl interactions immediately abolishes the selectivity at the center of the narrow pore region, with which sodium ions may now interact favorably.

These examples show that for difficult cases, and especially when a prediction is being made, a large number of simulations may be necessary. Today, the continuing increase in computer power has made such multiple simulations possible in a reasonable time frame. Several other recent studies illustrate the scope of molecular dynamics free energy for molecular recognition problems; they include studies of nucleic acids [13], proteins [14–16], and methodological studies of convergence and precision [17, 18]. Several recent reviews provide additional examples [19, 20].

13.2.2 Absolute Protein–Ligand Binding Constants

As has been discussed in the chapter on perturbation theory, the term *absolute free energy* is often abused. It was emphasized that the epithet *absolute* referred to our ability to determine $A = -1/\beta \ln Q_{NVT}$, thus implying that current computational methods could allow the modeler to access the full canonical partition function, Q_{NVT}. *This is not the case*. For example, determination of the Helmholtz solvation free energy of a given chemical species implies an accurate evaluation of Q_{NVT}, which requires that all possible configurations of the solvent around the solute be taken into account. This would be true for an ergodic system, in the hypothetical limit of infinite sampling. But numerical estimates of the partition function based on finite-length simulations are governed by Boltzmann sampling, which favors the low-energy regions of configurational space, and are, therefore, necessarily incomplete.

It still remains that the pervasive use of the expression *absolute free energy* in the literature requires that this concept be clarified in the case of protein–ligand

association. As can be seen in the thermodynamic cycle of Figs. 2.8 and 2.9, the direct, horizontal transformation that brings the ligand from its free, unbound state to its associated state, is not amenable to current, state-of-the-art statistical simulations. Protein–ligand binding phenomena generally occurs over time scales that are incommensurate with those characteristic of molecular dynamics simulations, because they involve a global search by the ligand of the optimal anchoring site in the protein, and, therefore, fall into the category of multiple-minimum problems. Free energy differences can be computed along a predefined order parameter. However, the definition of a useful, non ambiguous reaction coordinate connecting the initial and final states of the transformation is usually difficult. Recent work by Woo and Roux addresses this problem in terms of potentials of mean force, and opens new perspectives for the direct determination of binding constants [21]. In this approach, a series of independent stages are performed. The free ligand is first restrained in the conformation of the native, bound state, then translated into the binding pocket of the protein. Remarkable agreement with experiment was found for the binding of pYEEI peptide to the SH2 domain of the human Lck protein. It should be emphasized, however, that association occurs at the surface of the Lck protein. It is not clear whether the proposed method would remain applicable to ligands buried deeply in protein cavities, e.g., as in GPCRs. Binding free energies can also be computed using nonequilibrium, steering, numerical experiments [22], in which the ligand is pulled out of the protein pocket by means of an external force; attempts in this direction admittedly have remained scarce [23, 24].

The terminology *absolute* can, in a sense, be understood as the opposite of *relative*. As seen in Figs. 2.8 and 2.9, the computation of *absolute* binding constants constitutes a special case of relative free energy calculations. Instead of mutating the ligand into an alternate one and measuring a relative affinity toward a common protein, the ligand is annihilated in both the free, unbound state and in the associated state. Such free energy calculations, first proposed by Jorgensen et al. [25], are often referred to as double-annihilation simulations. Annihilation transformations cancel the interaction of the ligand with its environment by scaling either the non-bonded parameters or the interaction potential energy function—see Chap. 2. In annihilation simulations, as the protein–ligand interactions are reduced, the ligand may drift away from its original, native position. This leads to practical, sampling difficulties, since in theory, the ligand should explore the entire simulation volume, which would require an excessively long simulation. A better scheme is to lock the non-interacting, *ghost* ligand into the binding pocket by means of an appropriate set of restraints. Enforcing these positional restraints leads to a loss of translational and rotational entropy, and, hence, to a free energy contribution that must be taken into account in the thermodynamic cycle of Figs. 2.8 and 2.9 [26, 27]. For this, analytical methods can be used [21, 27, 28].

Arguably, the foundations for the calculation of protein–ligand binding constants were laid by the pioneering article of Tembe and McCammon [29], which demonstrated the usefulness of the FEP methodology using a naive, van der Waals–sphere representation of a receptor–ligand complex. Hermans and Shankar pushed the level of description one step further by examining the *absolute* free energy of

association of gaseous xenon to myoglobin [30]. Benefiting from a noteworthy increase of computational power, Merz was able to investigate over a somewhat more realistic time scale binding of carbon dioxide to carbonic anhydrase [31], and Lee et al. all the association of phosphoryl choline to antibody MP603 [32]. The following year, Miyamoto and Kollman tackled the hitherto unexplored interaction of biotin with avidin, one of the strongest known covalent binding free energies between a small peptide ligand and a protein. Employing the FEP machinery, they estimated the *absolute* free energy of association of biotin with the related streptavidin protein [33, 34]. Despite a simplified description of the protein–ligand complex, and somewhat short simulations, in comparison with the current standards for free energy calculations, the experimental binding constant [35] was reproduced satisfactorily and a hypothesis for van der Waals pre-organization in the binding pocket was proposed to play as significant a role as electrostatic contributions in protein–ligand association [34]. Interestingly, the prototypical biotin–streptavidin problem was revisited eight years later [36], employing a markedly more rigorous modeling of the tetrameric structure of streptavidin and covering longer time scales. Although the improvement over the Miyamoto and Kollman results was marginal, the more recent calculations suggested that electrostatic and van der Waals contributions were of comparable weights and underlined that the free energy term arising from positional restraints should not be ignored when growing a ligand into the binding pocket of a protein.

While free energy calculations of protein–ligand association have taken advantage of the enormously increased computer power available in recent years, they have also benefited greatly from methodological developments. For instance, Roux et al. [37] and Gilson et al. [27] clarified the theoretical bases for the computation of protein–ligand binding constants and explained how it can be compared with experiment. Hermans and Wang emphasized the critical role played by the loss of translational and rotational entropy in protein–ligand binding free energy calculations [26]. Boresch et al. pursued these efforts by proposing a general scheme for the computation of protein–ligand *absolute* binding free energies, using appropriate external biases to restrict the translational and rotational motion of bound ligands [28]. These authors presented a theoretical framework to take into account these restraints and relate them to the sometimes-overlooked standard state dependence of protein–ligand association equilibria. Also of interest, the investigation of Swanson et al. based on long MD trajectories sheds new light on the conformational modifications of both the ligand and the protein upon molecular association [38].

A problem closely related to that of protein–ligand binding is the question of how water molecules migrate in and out of the binding pocket as the ligand is accommodated by the protein [37]. Quantification of this phenomenon relies to a large extent on our ability to estimate the associated entropic contributions. Hamelberg and McCammon devised a numerical approach to measure the standard free energy for removing water molecules occupying the binding sites of protein–ligand complexes [39]. Woo et al. have applied grand canonical Monte Carlo simulations to the determination of the chemical potential of water molecules confined in binding pockets and in thermodynamic equilibrium with an external water reservoir [40].

13.2.3 Molecular Dynamics Free Energy Yields Structures and Free Energy Components

In experimental studies of molecular recognition, an important goal is to interpret the overall binding free energy in terms of specific structural groups, such as individual hydrogen-bond partners, and specific physical effects, such as electrostatic forces. Group contributions to the binding affinity can be estimated from point mutations, as in alanine scanning experiments [41, 42]. Double mutant cycles can be used to measure the contributions of individual residues to ligand binding specificity [42, 43]. For example, the binding affinities of a substrate and an inhibitor to a native and a mutant protein can be measured. The result is a triple free energy difference, $\Delta\Delta\Delta G = \Delta\Delta G_{\mathrm{mut}} - \Delta\Delta G_{\mathrm{nat}}$. Here, $\Delta\Delta G_{\mathrm{nat}}$ and $\Delta\Delta G_{\mathrm{mut}}$ are the binding free energy differences between the substrate and the inhibitor for the native and mutant proteins, respectively. If the mutation replaces a particular residue by an alanine, deleting the original side chain, $\Delta\Delta\Delta G$ can be interpreted as a measure of that side chain's contribution to the binding specificity.

Free energy simulations can provide the same information. We continue to use ligand binding specificity as an illustration. It is straightforward to compute $\Delta\Delta\Delta G$ values with FEP. However, as with experiments, they are usually costly to obtain. In some situations, an approximate method can be used to obtain the free energy changes $\Delta\Delta G_{\mathrm{mut}}$ for several mutations at a time from just one or two simulations; see [44] and Chap. 12. Usually, though, a complete set of simulations must be done for each side chain or group whose contribution $\Delta\Delta\Delta G$ is sought.

A much simpler and more efficient approach is to calculate free energy 'components' [45]. Let ΔG_{prot} and ΔG_{solv} be the free energies to transform the substrate into the inhibitor in the protein complex and in solution, respectively, so that $\Delta\Delta G_{\mathrm{nat}} = \Delta G_{\mathrm{prot}} - \Delta G_{\mathrm{solv}}$. Because molecular mechanics energy functions commonly take the form of sums over small groups of atoms, these two free energies can be expressed as a sum over groups of atoms [45]. In particular, ΔG_{prot} takes the form of a sum over individual protein residues and solvent. The contribution of a particular side chain or residue to $\Delta\Delta G_{\mathrm{nat}}$ is referred to as a free energy component. It provides another measure of the contribution of the side chain to the ligand specificity.

The interpretation of free energy components requires care, since they are not measurable quantities (unlike $\Delta\Delta\Delta G$) and depend on the details of the calculation. As discussed in Sect. 2.10, free energy components are not state functions, so that they depend on the integration path used to connect the endpoint states [46–50]. This is easy to see by considering a diatom A–B in solution. Let us first 'turn off', or decouple the van der Waals and Coulomb interactions of B from its environment; second, we lengthen the A–B bond to a large value; third, we restore the van der Waals and Coulomb interactions of B; fourth, we shorten the bond back to its original value. This closed thermodynamic cycle nevertheless has nonzero free energy components for both A and B [46, 47]. Indeed, the solvent distribution around B is different in steps 1 and 3, so that the corresponding van der Waals and electrostatic contributions do not cancel; meanwhile, there is a nonzero free energy component

for atom A in step 4, arising from the bond shortening, but not in the other steps (including step 2, where B is a ghost particle).

Nevertheless, previous experience has shown that free energy components provide a useful, qualitative measure of the importance of particular amino acids for ligand specificity and other properties. Their significance has been tested in several ways. Archontis et al. computed components for the aspartyl–tRNA case, and compared them to residue components obtained with a continuum electrostatics, or Poisson–Boltzmann (PB) free energy approach [6], for which the free energy components are pathway independent. Both methods identified the same amino acids as specificity determinants, and showed that contributions from several other nearby ionized amino acids approximately cancel. The amino acids with large components were found to be completely conserved among all known AspRS sequences. The MD free energy components were systematically larger than the PB components by a factor of about 4–10. The smaller magnitude of the PB components was due to the contribution of continuum solvent, which is 'folded into' the amino acid components in the case of PB. Another test was performed more recently for the same system [7]: $\Delta\Delta\Delta G$ values were computed for the four groups that have the largest free energy components in the AspRS ligand binding pocket. The $\Delta\Delta\Delta G$ values and free energy components agreed qualitatively, with the components being systematically larger by a factor of 3–8. As with the PB components, the $\Delta\Delta\Delta G$ values have a contribution from protein and solvent dielectric relaxation folded in, and indeed, the reduction factor is in good agreement with the mean dielectric constant estimated for the AspRS active site [51].

Finally, Mark and van Gunsteren studied the path dependency of free energy components for several systems, including the protein azurin, for which the effect of the Asn47Leu mutation on the copper oxidation potential was computed [49, 52]. This case is of interest because both legs of the thermodynamic cycle (analogous to Figs. 2.8 and 2.9) can be calculated. The horizontal, 'alchemical' and vertical, 'chemical' runs agreed approximately, but gave very different free energy components, because they corresponded to different physical processes.

Other free energy decompositions have also been found useful. These include calculations of dielectric relaxation free energies [51], of the free energy to freeze the ligand's rotational degrees of freedom in a protein–ligand binding reaction [21], of the van der Waals contribution to a protein solvation free energy [53], of the solvent contribution to a protein–protein association free energy [54], and of ligand binding entropies [55].

13.2.4 Electrostatic Treatments

When the mutation of interest involves a significant rearrangement of charge, it is critical to treat electrostatic interactions accurately. Since the systems of interest are macroscopic, a finite computer model is not normally sufficient: bulk solvent must be explicitly included at some stage of the calculation. This is especially important if a charge is introduced into or removed from the system.

Four recent studies provide a good overview of the strategies available today. Van den Bosch et al. [56] used periodic boundary conditions, with a protein fully solvated in a large box of water; long-range electrostatic interactions were approximated by a homogeneous continuum reaction field approach. Another, increasingly popular approach is to use Ewald summation (also with periodic boundary conditions) [57]. The main drawback of periodic boundaries is the cost associated with the large explicit solvent layer. Aqvist and Luzhkov [58] and Warshel et al. [59] both used large, but finite spheres with radii of 20–30 Å; within the spheres, all electrostatic interactions were calculated efficiently by use of a multipole approximation for distant groups [32, 60]. Since no net charges were created or deleted and the simulation models were fairly large, bulk solvent was not considered explicitly. Other work
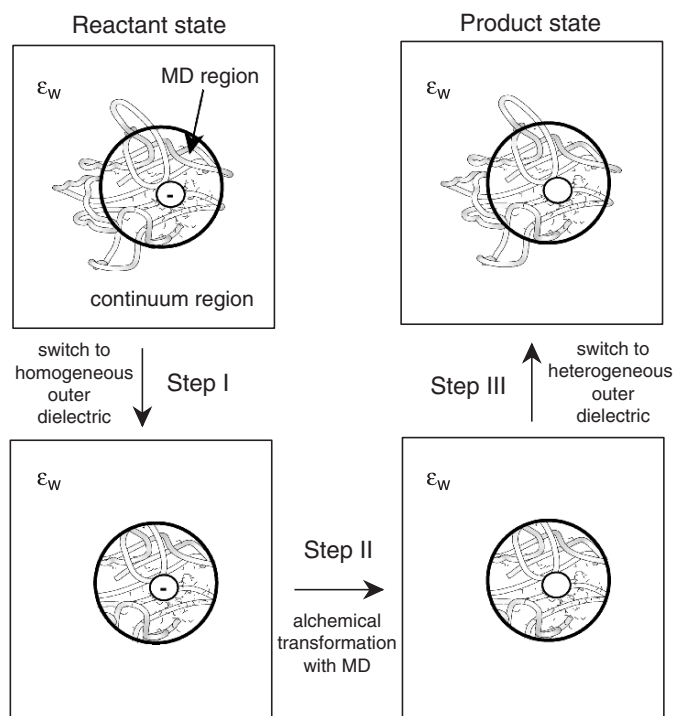
Fig. 13.2. Three-step ICRF scheme to compute the free energy change ΔG associated with a local transformation in a macromolecule in bulk solution. *Step I:* the dielectric constant of that portion of the macromolecule that lies outside a spherical inner region ('MD region') is switched to that of bulk solvent. *Step II:* the local transformation is carried out using a series of MD simulations. The transformation is schematized by the removal of a minus sign near the center of the inner, MD region. *Step III:* the dielectric constant of that portion of the macromolecule that lies outside the MD region is switched back to its original value. The free energy changes for steps I and III are calculated with a Poisson–Boltzmann continuum model; the free energy change for step II is calculated from thermodynamic perturbation theory. Reproduced from the *Journal of Physical Chemistry*

with spherical models has explicitly taken into account the free energy to transfer the finite simulation models into a bulk continuum solvent [61, 62]. In their early studies involving charge creation [63], Warshel and coworkers used 'shell' models, where regions close to the mutation are treated in detail, more-distant regions are treated as networks of polarizable dipoles, and the most distant regions are treated as a dielectric continuum. Those studies were the first to include a sophisticated model of bulk solvent. Simonson et al. used a new, inhomogeneous continuum reaction field method (ICRF) [8, 9] (Figure 13.2). This method employs a spherical system, including part of the protein and some explicit solvent, initially surrounded by either vacuum [8] or a homogeneous dielectric medium [9, 64]. The mutation is performed for this system using an MD free energy simulation. In a second stage, the finite model is transferred into the inhomogeneous environment formed by the complete protein and bulk solvent. The free energy for the transfer is obtained from continuum electrostatics. An efficient method was proposed recently by Roux et al. to include the inhomogeneous reaction field during the MD free energy step directly, eliminating the need for a transfer step [65, 66]. The ICRF approaches are similar in spirit to the shell models of Warshel et al., but employ a continuum dielectric environment rather than polarizable dipoles. In both cases, by combining MD simulations without truncation and a sophisticated treatment of bulk solvent, all the relevant electrostatic interactions are included. The efficiency of the ICRF methods makes them attractive alternatives to periodic boundary models and Ewald summation.

13.3 Recognition and Association: Following the Binding Reaction

Investigation of recognition and association processes involves the definition of an order parameter that delineates rigorously the approach of the two chemical species. Our ability to determine the underlying free energy changes along the order parameter is central to the study of these processes of fundamental chemical and biological interest. Definition of a nonambiguous order parameter, possibly a true reaction coordinate, constitutes the sine qua non condition to compute the free energy. There are noteworthy examples where a practical reaction coordinate is difficult to find. Such is the case of protein–ligand recognition and association, where the ligand must be brought from the free, unbound state, into the binding pocket. A number of authors have, however, considered the reverse, dissociative transformation, which consists of extracting the ligand from the protein towards the bulk aqueous medium by means of an external force applied within a reasonable pulling regime.

Historically, pioneering free energy calculations along an order parameter mainly relied on Monte Carlo simulations, using multistage sampling [67] and the so-called umbrella sampling (US) [68] numerical schemes, which are described in Sect. 3.3. In the latter, a bias is applied to the system to overcome the barriers and escape from the minima of the free energy landscape, thereby guaranteeing a uniform sampling along the order parameter. Biases are not known a priori and must be guessed intuitively by the modeler, which may be cumbersome for qualitatively new problems. Ideally,

to obtain a uniform sampling of conformational space, the bias should correspond to the negative of the free energy. US simulations are distinct from stratification strategies, which break the reaction path into a large number of narrow windows, in each of which sampling is necessarily expected to be fairly uniform. US simulations were pioneered by Patey and Valleau, who derived the free energy profile describing the interaction of an ion pair dissolved in a dipolar fluid [69]. The early successes prompted a growing community to follow this path for novel, exciting investigations of recognition and association phenomena.

Thus, Pangali et al. applied Monte Carlo simulations to examine the hydrophobic effect in a very simplistic model formed by two Lennard-Jones spheres in a bath of 214 water molecules [70]. Employing a multistage strategy, they were able to recover successfully the signature of the hydrophobic effect. Notably, in 1979, their free energy profile was in excellent agreement with predictions based on the Pratt and Chandler model [71]. Furthermore, it would be confirmed by subsequent, much longer simulations, using far more sophisticated models and potential energy functions [72, 73]. Following similar multistage strategies, Jorgensen and coworkers investigated a variety of association processes of chemical systems, including the mutual interaction of aromatic compounds [74], and that of aromatic compounds with chaotropes [75]. In 1987, Tobias and Brooks used thermodynamic perturbation theory to explore the free energy surface of two tagged argon atoms in liquid argon [76] and observed the expected features for such a system, viz. a contact and a solvent-separated minima. Since then, many authors studied free energies of association of small molecules in order to improve our general understanding of the nature of the primitive hydrophobic effect and wetting/dewetting phenomena [77, 78]. Employing thermodynamic integration, Peter Kollman and coworkers tackled the association of somewhat larger solutes in aqueous environments to decipher the underlying physical principles that drive π–π [73, 79] and cation–π [80] association.

Ion–ion recognition and association also represents an important field of applications for free energy calculations. The seminal articles in this area focused primarily on the potential of mean force delineating the interaction of oppositely charged, simple ions solvated in water, addressing the existence of contact and solvent-separated ion pairs [81, 82]. The demonstration that free energy calculations could provide a bridge between the structural detail of ion pairs and their associated energetics was rapidly fueled by several studies tackling a variety of solvated ionic systems, including like-charge ion pairs [83–85]. More recently, Rozanska and Chipot employed US simulations with different periodic boundary conditions to examine the reversible association in water of a guanidinium cation and an acetate anion along their $\mathscr{C}_{2v}$ axis [86]. Ewald boundary conditions yield the expected features of the free energy profile: A contact and a solvent-separated minimum, with the free energy tending towards $q_{Gdm+}q_{AcO-}/4\pi\epsilon_0\epsilon_1 r$ at large separations r. Here, q_{Gdm+} and q_{AcO-} are, respectively, the charges of the cation and the anion, ϵ_0 is the permittivity of vacuum and ϵ_1, that of water. In contrast, using a 12 Å cutoff leads to a physically unrealistic, repulsive profile. This result suggests that great care ought to be taken when simulating hydrated proteins that feature salt bridges between the charged amino acids, usch as arginine and aspartate.

Compared to US and its subsequent variants, the ABF method obviates the a priori knowledge of the free energy surface. As a result, exploration of ξ is only driven by the self-diffusion properties of the system. It should be clearly understood, however, that while the ABF helps progression along the order parameter, the method's efficiency depends on the (possibly slow) relaxation of the collective degrees of freedom orthogonal to ξ. This explains the considerable simulation time required to model the dimerization of the transmembrane domain of glycophorin A in a simplified membrane [54].

13.4 Free Energies of Solvation

Because of its simplicity, the calculation of the free energy of solvation was one of the first practical applications of the free energy perturbation and integration methodology [87, 88]. Today, solvation free energies remain of primary importance both for testing new methodology, for developing and testing force fields, and for specific applications such as predicting how molecular compounds will partition between different environments [89]. The free energy of solvation corresponds to the free energy of transferring a compound from one well-defined reference state (gas) to another (solution), allowing a direct comparison with experiment. In addition, as the interaction of a solute with its environment in the gas state (vacuum) is effectively zero, only the interactions of the solute with a particular solvent environment need be considered.

One of the most influential applications of free energy calculations over recent years has been the use of solvation free energies for the refinement and verification of empirical molecular force fields. Thermodynamic properties such as excess free energies and free energies of solvation have long been used in the parametrization of models for simple molecules and solvents such as water [90] and methanol. However, the parameter sets used to describe most compounds have historically been based primarily on structural properties and fitting to the results from quantum mechanical calculations. This is despite the fact that many of the properties of interest, especially in biomolecular systems, such as protein or peptide folding, depend on how compounds partition between different environments. The primary reason why thermodynamic properties have until recently not been more generally incorporated into force field parametrization was cost, and the difficulty in obtaining converged results.

Jorgensen and coworkers did nevertheless make extensive use of solvation free energies over many years for the verification of the OPLS force field [91–93]. In particular, solvation free energies have long been used to rationalize the choice of partial atomic charges. Although electron density can be determined from quantum mechanical calculations (and experiment), the assignment of partial atomic charges is an artificial construct and as such, subjective and highly model-dependent. Solvation free energies have been used both to justify the use of a particular charge model (such as the so-called restrained electrostatic potential (RESP) charges which

proved highly successful in reproducing free energies of solvation for a range of compounds [94]) or for the introduction of specific charge-scaling factors [95, 96].

A shift in force field parametrization came when Daura et al. [97] began to use solvation free energies as a primary input in the parametrization of aspects of the GROMOS force field. An immediate offshoot of this was a dramatic improvement in the accuracy with which folding of small peptides could be predicted in solution [98]. In 2003, Villa and Mark [99] published a systematic study of the free energy of solvation in water and hexane of neutral analogues of 18 of the 20 common amino acids based on the GROMOS96 force field using a thermodynamic integration approach. Notably, the hydration free energies of the analogues of the polar amino acids were shown to be too positive; in essence, the polar amino acids in the force field were too hydrophobic. Equivalent studies by Macallum and Tieleman [100] using the OPLS force field and Shirts et al. [101] using CHARMM [102], AMBER and OPLS all demonstrated a similar systematic underestimation of the interaction of polar solutes with water. Using a distributed computing approach and dedicating the equivalent of ca. 200 CPU years (viz. Celeron 1 Ghz) to the calculation, the solvation free energies for 15 amino acid side analogues were determined to unprecedented precision, i.e., three significant figures. This left no question that the observed deviation from experiment, greater than 2 kcal/mol in some cases, was significant. This work is leading to changes in force field parametrization. For example, the latest versions of the GROMOS force field, the 53a5 and 53a6 parameter sets, have specifically been developed to reproduce free enthalpy of hydration and solvation [103]. Notably, two parameter sets were published, as it proved impossible to reproduce simultaneously the properties of pure (low dielectric liquids) and hydration free energies using a single charge set. This provides compelling evidence for the need to include explicit polarization in empirical force fields. It also opened the question of the reliability of the solvent models used as a reference. Shirts and Pande have, for example, proposed a modified version of the TIP3P model of Jorgensen, specifically parameterized to better reproduce solvation free energies with existing force fields [104].

In addition to being used to verify and refine current empirical force fields, calculations of the free energy of solvation at times challenge our basic assumptions regarding chemical systems and the adequacy of current models used to describe them. For example, in contrast to simple hydrophobicity arguments, the successive methylation of amines does not lead to a monotonic increase in solvation free energy. At one level, such anomalous behavior highlights the inadequacy of predictions based on group contribution models and suggests that rigorous thermodynamic perturbation approaches should be applied even in apparently simple cases. However, despite their best efforts, a succession of workers [105–107] failed to reproduce the trend observed experimentally using perturbation approaches. Even the inclusion of polarization effects, which had been argued was the reason for the anomalous behavior of such amines, did not lead to a significant improvement in the agreement between the calculations and experiment [108]. We note that several workers have recently claimed to have resolved this issue by developing specific all-atom and united-atom models that reproduce the observed experimental trends [109, 110]. However, the need to introduce specific parameters for each compound in what is

in effect a homologous series simply underlines the limitations of our current ability to describe such systems. Nevertheless, despite questions in regard to how best to model specific systems there is little doubt that our growing ability to use free energy perturbation methods to reliably estimate thermodynamic properties such as solvation free energies is improving the reliability and driving the convergence of molecular force fields.

13.5 Transport Phenomena

Transport properties of molecular assemblies may also be investigated through the computation of free energy profiles along a representative order parameter. Here, the vocabulary *transport* embraces both the translocation of a solute across the interface separating two media of different dielectric permittivities and the permeation through integral membrane proteins. Simulation of transport processes are closely related to the previous topic of solvation free energies. In particular, studying its equilibrium aspects, such as partition coefficients may be viewed as a special case of solvation free energy calculations.

13.5.1 Partitioning Between Solvents

Differential solvation properties and partition coefficients may be investigated using free energy calculations [111–114], in particular perturbation theory, wherein the solute is either created or annihilated in two different solvents, i.e., by determining $\Delta\Delta A_{\mathrm{solvation}} = \Delta A^2_{\mathrm{creation}} - \Delta A^1_{\mathrm{creation}}$, where $\Delta A^i_{\mathrm{creation}}$ denotes the free energy change upon creation of the solute in solvent i, from which the partition coefficient, $\log_{10} P_{1,2}$, may be inferred. Whereas the partition coefficients, $\log_{10} P_{1,2}$, indicate the propensity of a given solute to translocate spontaneously between solvent 1 and solvent 2, they provide no information on the thermodynamics of this process and the underlying interfacial properties of the solute. This crucial information may, however, be accessed by computing the free energy profile along the direction normal to the surface that separates the two media [115].

One of the major shortcomings of the additive pairwise approximation used in classical potential energy functions lies in the absence of explicit induction effects, which may modulate significantly the computed free energy differences characterizing the translocation of a solute between two media of different dielectric permittivities. Electrostatic potential-derived point charges representing the isolated solute described by an unperturbed, ground-state wave function, ψ_0, and a Hamiltonian, $\hat{\mathscr{H}}_0$, correspond to an underestimated polarity when the latter is immersed in a polar environment, e.g., water. Symmetrically, point charges derived from the electrostatic potential obtained from a polarized wave function, ψ, and a perturbed Hamiltonian, $\hat{\mathscr{H}}$, often largely overshoot the molecular dipole moment. This excess of polarity may be characterized by a reversible work for 'inflating' the charges borne by the solute, known as the distortion energy [95, 113], $\Delta A_{\mathrm{distort}}$, and defined by:

$$\Delta A_{\text{distort}} = \left\langle \psi \left| \hat{\mathscr{H}}_0 \right| \psi \right\rangle - \left\langle \psi_0 \left| \hat{\mathscr{H}}_0 \right| \psi_0 \right\rangle. \tag{13.1}$$

Considering that, roughly speaking, the electrostatic component of the solvation free energy varies as the cube of the molecular dipole moment, it becomes obvious that the corrective term (13.1) should be taken into account in the determination of differential solvation properties of very polar solutes. In the computation of transfer free energies across an interface, it has been suggested that equation (13.1) be expressed as a function of the number density of one of the two media, so that the correction is zero in solvent 1 and $\Delta A_{\text{distort}}$ in solvent 2 [115].

Hydration may be viewed as a special case of differential solvation, where solvent 1 is described by a water lamella in equilibrium with the gas phase that corresponds to solvent 2. The hydration free energy is determined from the free energy profile that delineates the transfer of the solute across the water–air interface. Accurate reproduction of this quantity using effective intermolecular potential energy functions has proven to be a difficult task, because the solute is expected to undergo significant changes in its polarization upon translocation from the gas phase to the aqueous medium [95]. Equally challenging is the estimation of adsorption free energies, which may be inferred from the profile that describes the free energy changes along the direction normal to the aqueous interface, and compared readily to second-harmonic-generation experiments [116, 117]. Pohorille and Benjamin pioneered in the endeavor of accessing adsorption free energies from molecular statistical simulations [118]. From their US simulations, they predict the adsorption free energy of phenol at a water–air interface to be equal to -2.8 kcal/mol, to be compared to the experimental value of -3.8 kcal/mol. For the same system, Chipot reports a largely underestimated adsorption free energy of -1.7 kcal/mol, but a hydration free energy of -6.6 kcal/mol, which coincides almost perfectly with the experimental value of -6.5 kcal/mol [95]. That different theoretical calculations can yield good agreement with different experimental quantities reflects more the limitations of pairwise additive force fields than the difficulty to obtain converged free energy profiles for relatively simple molecular systems.

In principle, free energy calculations could serve as a predictive tool for estimating water–membrane partition coefficients of small drugs, in strong connection with the so-called blood–brain barrier (BBB). Along with the assessment of their toxicity, this would represent the ultimate step in a rational, de novo design of pharmacologically active molecules. Diffusion of small, organic solutes in lipid bilayers has been examined for a variety of molecular species ranging from benzene [119, 120] to more-complex anesthetics [121, 122]. Accessing partition coefficients through statistical simulations implies, however, the determination of the underlying free energy changes along an appropriately chosen order parameter, like the direction normal to the interface [123]. In the specific instance of inhaled anesthetics, analysis of the variations of the free energy for translocating the solute from the aqueous medium into the interior of a lipid bilayer suggests that potent anesthetics reside preferentially near the water–membrane interface. In sharp contrast with the dogmatic Meyer–Overton hypothesis [124], potency has been shown to correlate with the interfacial concentration of the anesthetic, rather than only its lipophilic properties [125].

Models of lipid bilayers have been employed widely to investigate diffusion properties across membranes through assisted and non-assisted mechanisms. Simple monovalent ions, e.g., Na^+, K^+, and Cl^-, have been shown to play a crucial role in intercellular communication. In order to enter the cell, the ion must preliminarily permeate the membrane that acts as an impervious wall towards the cytoplasm. Passive transport of Na^+ and Cl^- ions across membranes has been investigated using a model lipid bilayer that undergoes severe deformations upon translocation of the ions across the aqueous interface [126]. This process is accompanied by thinning defects in the membrane and the formation of water fingers that ensure appropriate hydration of the ion as it permeates the hydrophobic environment.

Free energies of solvation, in general, and the determination of partition coefficients, in particular, are also widely used to benchmark free energy perturbation methods against alternative approaches and the explicit treatment of solvent against continuum representation such as PB calculations or the increasingly popular generalized Born (GB) models. For instance, Best et al. [114] performed extensive calculations on a series of eight small organic compounds in water and water saturated octanol using the AMBER force field. They compared explicit free energy calculations and calculations based on a GB model to differences in solvation free energy and partition coefficients obtained experimentally. Using an explicit perturbation approach, the average unsigned error in the solvation free energies was 1.34 kcal/mol in water-saturated octanol, and 1.28 kcal/mol in water. The error in the octanol/water partition constants was significantly smaller at 0.74 kcal/mol which suggests some cancelation of error and small but systematic problems with the force field. Interestingly, using the more computationally efficient GB model the error in the partition constants was only 0.50 kcal/mol. The force field used in conjunction with the GB model was identical to that used in the explicit simulations. This begs the question of whether continuum methods provide a better representation of the electrostatic contributions to the solvation free energy or whether the GB agreement was partly fortuitous.

13.5.2 Assisted Transport in the Cell Machinery

The considerable free energy associated with the transfer of ions from the aqueous medium to the interior of the membrane [126] rationalizes the use of specific membrane channels and transporters by the cell machinery, facilitating and controlling selectively the passage of ionic species across the lipid bilayer [127–129]. For instance, gramicidin A, a prototypical channel for assisted ion transport across membranes, has been the object of thorough analyses on both experimental and theoretical fronts. MD has proven to be able to reproduce the key structural features observed experimentally [129], thereby suggesting that ion selectivity and binding in membrane carriers, as well as gating mechanisms are amenable to atomistic simulations, in general, and free energy calculations, in particular.

Schulten and coworkers coupled SMD [23] simulations with the Jarzynski identity [102] to derive the free energy profile for glycerol conduction in the facilitator GlpF, a channel that allows the selective passage of water and small, linear alcohols,

e.g., glycerol [106]. Beyond the computational challenge of performing free energy calculations on a large molecular assembly, this study rationalizes experimental observations by highlighting potential binding sites along the conduction pathway, together with energetic barriers.

The unique role played by the NPA motif of GlpF was the object of a related investigation aimed at deciphering the molecular mechanism that prevents proton translocation across aquaporins. Chakrabarti et al. determined the reversible work involved in: (i) the transfer of a proton across the single-file water wire formed in the pore, and (ii) the subsequent reorganization of that water wire [130]. In the absence of the proton, the second step is impeded by the bipolar orientation of water around the NPA motif. On the other hand, proton transfer is strongly disfavored by a large free energy barrier arising at the location of the NPA region. The electric field generated by the channel appears to act against proton translocation from the periplasm to the cytoplasm.

Aquaporins have been the object of active research targeted at understanding how nature controls the selective passage of small molecules across the biological membrane. Combining nonequilibrium MD simulations and the Jarzynski equality, Tajkhorshid and coworkers proposed to pinpoint the key features that distinguish *Escherichia coli* AqpZ, a pure water channel, from GlpF [131]. Coercing the passage of glycerol through AqpZ results in a free energy barrier approximately three times higher than that characterizing GlpF, which may be ascribed to steric hindrances in the narrow region of the selectivity filter. The computed free energy profiles also reveal differences in the periplasmic vestibule of the channels, the deeper minimum of GlpF being proposed to enhance the capture of glycerol.

Crystal structures of the potassium channel KcsA were determined recently, and a wide range of biophysical and electrophysiological data are available. A series of MD free energy studies have been performed, in which the role of the ligand is played by one or more of the transported ions [12, 58, 132, 133]. In particular, the ability of potassium and sodium to bind in the KcsA pore were compared. The selectivity of the pore was compared to a series of simpler systems — viz. liquid N-methylacetamide (essentially a liquid of peptide groups), the cyclic ionophore valinomycin, and several simpler model systems [12]. By considering systems of varying rigidity and polarity, it was shown that the pore selectivity arises from a balance between the attractive interactions between the ion and the backbone carbonyl groups, and repulsion between the carbonyl groups. This balance is sensitive to the size of the ion, and is robust with respect to the thermal fluctuations of the pore.

An earlier, structure-based hypothesis had proposed that the selectivity was determined by a precise and rigid size complementarity between the pore and potassium; this rigid complementarity is inconsistent with the plasticity of the KcsA structure and is not needed for selectivity. Additional simulations examined in detail the mechanism of ion conduction through the channel [58, 132, 134]. The translocation of K^+ ions in single file through the narrowest region of the pore is expected to be the rate-limiting step in the conduction mechanism. In both studies, ion conduction was found to involve transitions between two or three main states, with either two or three K^+ ions occupying the selectivity filter in the KcsA pore. Small differences

between the studies are compatible with the expected MD free energy uncertainty. The largest free energy barrier was about 2–3 kcal/mol, so that ion conduction is limited by diffusion [58, 132]. This example illustrates further the power of free energy simulations, where dynamical and thermodynamical data are obtained from the same set of computer experiments, and direct comparisons can be made to model systems, some of which cannot be studied experimentally.

13.6 Protein Folding and Stability

An often proposed and potentially very exciting application of the free energy perturbation methods is the prediction of the effect of mutations on protein stability. In the early 1990s, this attracted much attention, especially after calculations of Kollman and coworkers [135, 136] showed good correlation with experiment. This was despite limited simulation times and the very challenging nature of the calculations. To estimate the effect of amino acid substitution on protein stability, one must consider not just the effect within the folded protein, but also the effect of the mutation on the unfolded state. While the determination of the effect of a substitution in the folded state is fairly straightforward, the unfolded state is problematic, due to our inability to define an appropriate structural model. In many early studies, the unfolded state was simply approximated by a short peptide in an extended conformation. While this might be adequate if the residue in question was fully exposed to the solvent, it quickly became clear this approach was not generally applicable. Unfortunately, the agreement with experiment in many early studies appeared to have been simply fortuitous [137]. The rise and sudden fall in the use of free energy methods to predict changes in protein stability provides an important illustration of how to select successful applications. Free energy calculations perform well when used to determine the difference in free energy between two well-defined states. In the case of protein stability the problem is not simply that an unfolded protein cannot be adequately modeled as an extended chain, but that in principle, all possible unfolded states must be considered. Although attempts have been made to combine multiple states, for example a proposed unfolded state and a proposed transition state [138], the real utility of these calculations is questionable.

This said, well-chosen free energy calculations can be used to address important issues associated with protein stability. For example, Hodel et al. [139] investigated the effect of *cis–trans* proline isomerization on the stability of staphylococcal nuclease. In this case the unfolded reference state is the same and the difference in free energy dependent solely on the nature of the folded state. In other cases important insight can be gained despite the uncertainties in relation to the value of the calculated free energy an example being the effect of helix capping in lambda–cro protein [140]. Caution must, however, always be used when attempting to gain insight from simulations involving poorly defined end states or based on an assignment of components [45]. As noted above, the assignment of free energy components are force field and pathway dependent. As such, free energy components do not have a clearcut physical interpretation that can be directly related to experiment [49].

Rather, their value is for qualitative comparisons between series of similar systems, with free energies computed along similar pathways [141]. A practical illustration of this is the work of Prevost et al. [142] and of Sun et al. [143], which aimed to determine the contribution of hydrophobicity to the effect of the mutation Ile96Ala on the stability of barnase. This is a comparatively trivial mutation and both sets of workers obtain estimates of the change in stability in close agreement with experiment. However, whereas Prevost et al. [142] found that hydrophobicity makes a minor contribution to the change in stability, Sun et al. [143] found that it is the dominant contribution. This difference may in part be related to differences in the force field used, but is primarily attributable to differences in methodology with one group using a dual topology approach while the other scaled individual terms in the force field to effect the mutation. Another example was discussed above, where mutations in the protein azurin were performed along very different pathways, leading to different free energy components [49].

13.7 Redox and Acid–Base Reactions

13.7.1 The Importance of Electrostatics

In biochemical systems, acid–base and redox reactions are essential. Electron transfer plays an obvious, crucial role in photosynthesis, and redox reactions are central to the response to oxidative stress, and to the innate immune system and inflammatory response. Acid–base and proton transfer reactions are a part of most enzyme mechanisms, and are also closely linked to protein folding and stability. Proton and electron transfer are often coupled, as in almost all the steps of the mitochondrial respiratory chain.

Acid–base and redox reactions have also been used extensively to probe electrostatic interactions in proteins. Indeed, the proton binding constants of titratable groups in a protein are highly sensitive to electrostatic interactions with their surroundings, including other protein groups or cofactors, aqueous solvent, counterions, lipids, and other macromolecules. The same is true for the redox potential of a protein group or cofactor, such as a heme or an iron–sulfur cluster. Interactions with specific protein side chains can be probed by site-directed mutagenesis.

In both the cell and the test tube, proteins are surrounded by aqueous solvent and, in many cases, lipid bilayers. The dielectric properties of the protein interior contrast sharply with those of the highly polarizable aqueous solvent. The electrostatic interaction between two amino acids cannot be deduced from a simple Coulomb's law: dielectric shielding by protein and especially by solvent must be accounted for. In addition, when a charge such as a redox electron is transferred to a site within a protein, the protein and solvent relax, or reorganize; the corresponding reorganization free energy makes an important contribution to the redox potential. Finally, when a charge is transferred from a small molecule in solution to a protein interior, one must take into account not only the interactions within the protein, but also the attractive interactions with solvent that it experienced prior to the transfer. Within the protein,

the interactions with solvent are reduced, so that a partial desolvation of the charge has occurred. Polar groups within the protein may or may not be able to compensate effectively for the desolvation. The solvent contribution to the transfer free energy can be thought of as a 'desolvation penalty'. It represents an important contribution to the redox potential difference (or pK_a difference) between the protein and the small molecule. Similar considerations apply to the protein folding reaction.

These effects are difficult to analyze fully with experiments alone, and free energy simulations are a powerful complementary tool. One of the first applications of MD free energy to proteins was a study of an acid–base reaction [144]. For these problems, it is especially important to model long-range electrostatic interactions accurately and efficiently. Methods such as those described in the previous section (e.g., particle mesh Ewald) are now readily available in many simulation programs. With molecular dynamics free energy and an explicit treatment of solvent, the solvent shielding of protein charges is automatically included, as are protein and solvent dielectric relaxation. Recent studies have considered both redox and acid–base processes, including problems related to photosynthesis and respiration. An especially fruitful area is the study of enzyme reactions, where the reactive part of the system is treated quantum mechanically (QM/MM treatment). Another recent development is to perform MD simulations in an extended thermodynamic ensemble, where protons can freely bind to the protein or dissociate, effectively simulating a constant pH ensemble. This provides another route to obtain proton binding constants [145–147]. Recent studies have also illustrated two outstanding difficulties. One is the difficulty to obtain sufficient conformational sampling and convergence of the thermodynamic properties, because of the long dielectric relaxation times that are often involved. Another is the need for polarizable force fields in certain cases. In the next sections, we describe selected recent studies that illustrate these different aspects. Many earlier studies have been reviewed elsewhere; see for example [63, 148, 149].

13.7.2 Redox Reactions and Electron Transfer

In the MD free energy context, electron transfer has usually been modeled as a simple, classical mechanical, electrostatic process [56, 150–152]. The redox electron is treated as a set of point charges. The transfer reaction consists of shifting the charges from a small group of donor atoms to a group of acceptor atoms. This leads to a reasonable description of outer-sphere contributions to the redox potential, and to a particularly simple MD free energy setup. The model can then be improved in several ways. For example, quantum mechanical effects associated with inner-sphere groups can be treated separately, by performing ab initio calculations on a small fragment of the protein. A more sophisticated method is to include quantum effects during the MD free energy simulations, either through a path integral treatment of nuclear degrees of freedom [141, 153], or through a QM treatment of selected electronic degrees of freedom (the QM/MM method) [154], or both.

An important model system for electron transfer studies is the electron carrier cytochrome c (Cyt c). Its redox center is a heme, coordinated by a histidine and

a methionine. Oxido-reduction of yeast Cyt c and a small heme:peptide analogue, microperoxidase-8 (MP8) were compared recently by FEP, using both classical and quantum heme models, a large explicit solvent box, a particle mesh Ewald treatment of long-range electrostatics, and long simulations (a total of 23 ns) [148, 150]. Several computational points are worth noting. Oxidizing the heme group changes the net charge of the system. Since tin foil boundary conditions were used, a uniform, compensating, background charge density develops that exactly cancels the new charge. Since the background charge density is uniform, it does not contribute to the forces or perform any work, and can be ignored [155]. Since the system is periodic, a contribution to the free energy change arises from interactions between the redox electron and its images in other boxes. With the 59 Å box length used, this contribution was very small (<0.25 kcal/mol). Finally, despite the long simulation times, convergence of the free energy was only approximate. In particular, runs prepared independently, using either the oxidized or the reduced Cyt c crystal structures gave somewhat different free energy changes for Cyt c oxidation. The results from different runs spanned a 7 kcal/mol range. The average over all runs was about 6 kcal/mol (relative to the MP-8 value), fortuitously close to the experimental value of 7 kcal/mol. The scatter among runs could be attributed in part to the behavior of a buried water molecule; in some runs, it diffused to the surface of the protein and escaped, in disagreement with crystallographic data.

Despite the limited convergence and the simplicity of the overall model, this study illustrates the information that can be obtained by FEP. The dielectric response of the protein and solvent were found to be linear, and the reorganization free energy was found to be a more robust property than the redox potential (in agreement with experiment). The reorganization free energy for Cyt c oxidation was much smaller than that for MP-8 in solution. The former free energy could be reproduced with a dielectric continuum model, in which the heme was treated as a cavity, the rest of the protein as one dielectric medium, and surrounding solvent as another. The solvent dielectric constant was 80; that of the protein was just 1.1. This low value indicates that the protein reduces its dielectric reorganization almost to the theoretical limit (a dielectric of one). Note that the MD model did not include electronic polarizability, so that these dielectric values refer to atomic reorganization, not electronic polarization. Together with the Marcus estimate of the electron transfer rate, this result means that Cyt c is a 'dielectrically perfect' enzyme: by reducing its atomic reorganization to the theoretical limit, it speeds up electron transfer to the greatest possible extent.

Studies of ferredoxin [152] and a photosynthetic reaction center [151] have analyzed further the protein's dielectric response to electron transfer, and the protein's role in reducing the reorganization free energy so as to accelerate electron transfer [152]. Different force fields were compared, including a polarizable and a non-polarizable force field [151]. One very recent study considered the effect of point mutations on the redox potential of the protein azurin [56]. Structural relaxation along the simulated reaction pathway was analyzed in detail. Similar to the Cyt c study above, several slow relaxation channels were found, which limited the ability to obtain very precise free energy estimates. Only semiquantitative values were

obtained, as in [141]. Finally, a QM/MM MD free energy treatment was developed and applied to study FAD reduction in cholesterol oxidase [154].

13.7.3 Acid–Base Reactions and Proton Transfer

MD free energy studies of proton transfer present similar features and similar difficulties. Proton binding is usually modeled as a rearrangement of atomic charges (Fig. 13.3). Often, the goal is to compare proton binding to a protein side chain, such as an aspartate, and to an analogous, small molecule in solution, such as acetate. Comparison of the binding free energies yields the pK_a shift of the protein side chain, relative to the small molecule:

$$pK_{a,\text{prot}} = pK_{a,\text{model}} + \frac{1}{2.303\, k_B T}\, \Delta\Delta G. \tag{13.2}$$

Here, $pK_{a,\text{prot}}$ and $pK_{a,\text{model}}$ refer to the protein and the small molecule, respectively, and $\Delta\Delta G$ is the difference between the proton binding free energies to the protein and to the small molecule. Interactions between the positive charge of the proton and the surrounding protein and solvent are very strong, so that the protonation free energies are large, and $\Delta\Delta G$ is obtained as a difference between two large numbers.

In one recent study, the pK_a values of three aspartate side chains in two proteins were computed [156]. The aspartates had, respectively, a large pK_a upshift, a large downshift, and an unshifted pK_a. Long, repeated simulations were performed with two force fields. The upshifted aspartate was partly buried within the protein (thioredoxin). For this case, several slow relaxation channels were identified. They correspond to protein degrees of freedom that couple strongly to the reaction coordinate (the gradual protonation of the aspartate) and adjust on a nanosecond timescale. They include the rotation of the aspartate itself around its own χ_2 dihedral angle,
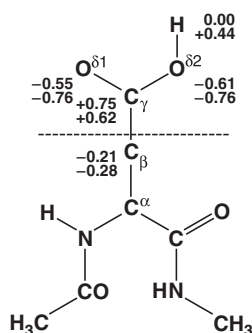
Fig. 13.3. Schematic view of an aspartate molecule showing the atomic charges that are changed going from the ionized form (lower values) to the neutral form (upper values). The specific values are taken from the CHARMM22 force field.

rearrangements of a chain of water molecules that runs alongside the aspartate, re-arrangements of side chains that line the water chain, and the rotation of a nearby lysine to form a salt bridge with the ionized aspartate. Thus, it was difficult to obtain adequate convergence of the corresponding pK_a shift. A similar, slow convergence was observed in another recent study [51]. Furthermore, the ionized state was un-derstabilized by both force fields (AMBER [157] and CHARMM [102]), leading to an error of 3 units for the computed pK_a. This is probably because the partly buried Lys–Asp salt bridge was not correctly described with the fixed atomic charges em-ployed. A polarizable force field may be necessary to capture the correct magnitude of the interaction. For the unshifted and the downshifted pK_a values, the experimen-tal behavior was captured more accurately.

Enzyme reactions frequently involve proton transfer and acid–base reactions. They can be studied with FEP, using either classical or quantum mechanical methods [63, 158, 159]. Many studies have been done with either semiclassical, empirical va-lence bond methods or with QM/MM methods using a semi-empirical Hamiltonian (e.g., AM1 or PM3). For instance, the multistate empirical valence bond model has recently been extended to describe the protonation of titrable amino acids and the determination of accurate of free energy profiles delineating acid ionization and the corresponding pK_a values [160], thereby opening the way to the simulation in pro-teins of proton transfer between protonable residues [161–163]. The QM/MM meth-ods automatically include electronic polarizability within the QM region. All these studies have played an essential role in elucidating the basic mechanisms of enzyme catalysis and the range and limitations of transition state theory [63, 158]. For exam-ple, Warshel and coworkers have accumulated a large body of evidence to document catalysis by electrostatic stabilization of the transition state. It is beyond the scope of this chapter to present the QM/MM methodology for FEP; we refer to earlier books [63, 164] and to several recent articles [154, 165, 166], which provide just a small, illustrative sample of the recent literature.

13.8 High-Performance Computing

Even with sophisticated, modern sampling methods, the calculation of free energy quantities with experimentally relevant precision and accuracy can still be an enor-mous challenge. Thus, there are many types of free energy calculations that can sig-nificantly benefit from high-performance computing. In this section, we will review some of these new methods and applications which efficiently utilize new, massively parallel computer architectures to improve the speed, precision, and accuracy of such calculations, concentrating on biomolecular free energies. While, biomolecular free energies are of course not the only free energies of interest to the scientific commu-nity, this field is sufficiently diverse to demonstrate a range of methods, approaches, and applications.

For example, calculating the affinity of ligand–protein binding (a natural first step in structure-based, computational drug design) is a canonical application of free energy methods. It is important to ask why are drugs not 'engineered' like other

pieces of technology in modern life, like bridges or automobiles. In all of these cases (drugs, bridges, automobiles), the fundamental physics is understood. However, the accuracy needed in assessing the ligand binding free energy of binding is surprisingly high — at least 1 kcal/mol and ideally 0.5 kcal/mol or better — and this level of accuracy would be extremely difficult to achieve, if possible at all. If possible, such a technique could revolutionize the early stages of drug design and has been a goal of free energy calculations for many years.

However, while one goal of free energy calculation must be high accuracy (and corresponding high precision), another important goal is computational tractability. Clearly if a calculation is impossible to perform (or effectively impossible due to exceedingly high computational requirements), such a calculation would not be useful for applications. Moreover, with the goals of screening over a large number (e.g., 100,000 to 1,000,000) by calculating the ligand binding free energy for each ligand bound to the desired protein, the computational tractability of a given calculation becomes of even greater importance.

Herein lie two fairly different approaches to this problem. One approach considers the computational tractability to be paramount, and in this case, one is willing to sacrifice accuracy for tractability. Protein–ligand docking is an example of this approach, and due to its tractability, has become a very prevalent means to quickly assay how well a ligand can bind to a protein.

The opposite approach has also been considered: to make accuracy paramount, independent of computational cost. For example, these cases typically employ the most accurate methods and complex sampling methods, both of which contribute to the computational complexity of the calculation. Clearly, just because a calculation is computationally demanding does not in itself demonstrate that it will be more accurate. However, elements which contribute to the accuracy of a calculation in terms of more accurate models — e.g., all-atom models, explicit solvation — or enhancing the degree of sampling would clearly be more computationally demanding.

While such highly accurate methods may not be universally practical today, on account of their large computational demands, they are useful for forecasting the future of free energy methods. Even after decades of exponential growth in computer speed, computers are still expected to grow exponentially faster for at least another 10–15 years, thereby yielding a 1000-fold increase in computer power. That, coupled with new methods in massively parallel and grid-based computation, suggests that the computational resources available to the average researcher in 15 years will be 1000- to 1,000,000-fold times greater than what is available today.

What could one expect from such computational power? The inherent challenge lies in the tradeoff that one must make: with a fixed computational budget, does one perform more rigorous sampling with a less detailed model, or less sampling with a more accurate model? However, without sufficient sampling (leading to a high-*precision* calculation), it is impossible to truly assess the *accuracy* of the models. Thus, enhancing sampling is a natural place to start in order to take the next steps.

To examine the accuracy of solvation free energies, first one needs to achieve a sufficient precision to compare with experiment. This was recently done extensively by Shirts et al. [101], where the `Folding@Home` distributed computing network

was used to perform extensive TI calculations of amino acid side-chain analogs. They achieved a high degree of statistical precision in their simulations, obtaining uncertainties for the free energy of hydration of 0.02–0.06 kcal/mol, equivalent to that obtained in experimental hydration free energy measurements of the same molecules. They found that the free energies were surprisingly accurate (with a typical RMS error of 0.6 kcal/mol). Moreover, they found that there was an overall trend to make amino acid side chains less soluble in solution, as pointed out above.

13.8.1 Enhancing Sampling: A Natural Role for High-Performance Computing

In many cases of interest, such as protein–ligand binding, the desired timescales for sampling (microseconds to milliseconds) are typically beyond what is possible with modern atomistic simulations (nanoseconds). How can one surpass this 1,000- to 1,000,000-fold impasse? One approach is to use existing supercomputing resources (typically thousands of CPUs) to work together to speed a single calculation. This more traditional approach has been used successfully in many areas, but has limitations in scaling — typically, even with the fastest networks, one can only scale to roughly 100 atoms per CPU. Thus, a 10,000 atom system could scale only to 100 CPUs. This limitation has led to the natural progression to study large systems (hundreds of thousands to millions of atoms) for a relatively short timescale (nanoseconds). While there are several problems which are well served by this method, such as water channels [167, 168] or ion channels [132, 169, 170], clearly complementary methods would serve to address systems which cannot be examined by these means.

Recent advances in computational methodology suggest an alternate approach. While a computer that is 100,000 times faster than the fastest CPUs does not exist, it is possible to obtain grid computing resources with 100,000 processors [171, 172]. However, 100,000 CPUs is not the same as a single CPU which is 100,000 times faster — one must make significant changes to the algorithms used. This paradigm shift is analogous to that which occurred in the early 1990s when supercomputers shifted to massively parallel machines, such as the Connection Machine CM-5 or Cray T3D. New algorithms were needed to run on this new hardware paradigm: a problem must be broken in multiple pieces, although fast communication between CPUs allow for a reasonably tight coupling of calculations.

However, as traditional supercomputers only have a few hundred to thousand CPUs, to obtain larger resources, distributed or grid computing methods have been utilized. These distributed computing resources lead to another paradigm shift: even more CPUs are present, but with much slower communication and typically with a large diversity of CPU speeds. Finally, with such a large CPU array, it is common for CPUs to join or leave the grid at any time, making dynamic load balancing and fault tolerance a critical challenge. Thus, with these new challenges, new algorithms are needed for this newer paradigm. While utilizing this unusual resource is a challenge, if such a resource could be efficiently utilized, one should be able to reach levels of sampling previously impossible [101, 104, 171–173].

13.8.2 Conformational Free Energy

A natural application of high-performance computing is to conformational sampling, which can be significantly enhanced using grid-based algorithms on large grids (e.g. 10,000 CPUs). Clearly, in most cases, running a single long simulation is not possible, due to the long timescales typical of conformational equilibration (milliseconds to seconds) compared with simulation (nanoseconds). Towards this end, there are a variety of approaches which have been employed, with differing goals, computational tractability, and outputs.

For example, Brooks and coworkers [174] pioneered the approach of calculating the free energy landscape (or potential of mean force) from the equilibrium population distribution. Because it is excessively time-consuming to reach equilibrium for high-dimensional protein molecules with conventional molecular dynamics, simulations are performed with US, already discussed in Sect. 1.3. A biasing potential is added to the original Hamiltonian; the biased simulations are subsequently recombined to remove the bias in a mathematically strict way [175]. The population distribution, $\mathscr{P}(q)$ can then be converted to the free energy with $\Delta A(q) = -1/\beta \ln \mathscr{P}(q)$. With this approach, Brooks and coworkers have obtained the free energy landscape and folding dynamics of an α-helical protein (viz. protein A [176]), an $\alpha\beta$ mixed protein (viz. GB1 [177, 178]), and a mostly β protein (viz. src-SH3 [179]) with numerous successful comparisons to experiment.

US studies can produce informative free energy landscapes but assume that degrees of freedom orthogonal to the surface equilibrate quickly. The MD time needed for significant chain or backbone movement could exceed the length of typical US simulations (which are each typically on the nanosecond timescale). However, in spite of this caveat, US approaches have been very successful. One explanation for this success lies in the choice of initial conditions: US simulations employ initial coordinates provided by high-temperature unfolding trajectories, which themselves have been found to yield predictive information about the nature of the relevant conformational space.

A number of other techniques have been developed to overcome the long timescales present in conformational sampling. For example, generalized ensemble methods have played a significant role in new methodology for free energy calculation [180]. We will focus on parallel tempering, or replica exchange MD, which has been widely used in protein folding simulations. In this approach, a number of simulations ('replicas') are performed in parallel at different temperatures. After a certain period (typically 0.1 to 1 ps), conformations are exchanged with a Metropolis probability. This criterion ensures that the sampling follows the canonical Boltzmann distribution at each temperature. Ideally, kinetic trapping at lower temperatures is avoided by exchanging conformations with higher temperature replicas. This method is easier to apply than other generalized ensemble methods because it does not require a priori knowledge of the population distribution. However, replica exchange is typically more computationally demanding than US, requiring more CPU days for a given calculation. Also, it requires processors more tightly coupled (in terms of their networking), thus making grid-based applications more difficult for replica

exchange MD. However, one potential benefit of replica exchange MD over US is that one need not specify order parameters a priori in replica exchange MD, allowing one to potentially perform more-complex analysis with the same set of simulation data.

Replica exchange MD has had a significant impact in the study of conformational free energy in protein folding. After Sugita and Okamoto demonstrated its effectiveness with a gas-phase simulation of a pentapeptide Met-enkephalin [181], Sanbonmatsu and García obtained the free energy surface of the same system using explicit water [182]. With 16 parallel replicas they observed enhanced sampling (at least five times) compared to conventional constant temperature molecular dynamics. Because the method is quite simple and because it is trivially parallelized in low-cost cluster environments, it gained wide application rapidly. Berne and coworkers applied this method to obtain a free energy landscape for β-hairpin folding in explicit water using 64 replicas with over 4,000 atoms [183]. With the equilibrium ensemble and the free energy landscape in hand, they reported that the β-hairpin population and the hydrogen-bond probability were in agreement with experiments, and proposed that the β-strand hydrogen bonds and hydrophobic core form together during the folding pathway. If care is taken to fully reach equilibrium [184], replica exchange MD becomes powerful for elucidating the folding landscape. For example, García and Onuchic applied the method to a relatively large system, protein A [185]. With 82 replicas for more than 16,000 atoms with temperatures ranging from 277 to 548 K, and with about 13 ns MD simulations for each replica, they reported convergence to the equilibrium distribution with quantitative determination of the free energy barrier of the folding.

García and coworkers studied two 21-residue helical peptides, for which we report equilibrium simulation results herein: the capped alanine homopolymer A_{21} (Ace–A_{21}–NMe), which is naturally insoluble in water, and the F_s peptide (Ace-$A_5[AAAR^+A]_3A$–NMe), a soluble α-helical arginine-substituted analog of A_{21}. Using replica exchange MD, with a total sampling time of $\sim$1.7 μs, they showed that AMBER-94 overstabilizes helical conformations in both peptides [186] by comparing the Lifson–Roig (LR) helix–coil parameters [187, 188] derived from simulation to the experimentally determined values. In response to the poor agreement resulting from that comparison, they introduced a modified potential, in which the ϕ and ψ torsion potentials in the original AMBER-94 are set to zero, and found much better agreement with experimental helix–coil parameters. In comparing the two sequences they reported a shielding of backbone carbonyl oxygen atoms from the surrounding aqueous media by the large arginine side chains four residues downstream acting to stabilize helical polyalanine-based peptides with such insertions, as suggested in previous studies [189, 190]. Additionally, Nymeyer and García compared generalized Born (GB) implicit solvation with an explicit (transferable intermolecular potential) representation of the solvent and showed that the implicit model significantly favors a non-native, compact helical bundle in simulations of F_s [185], suggesting that an explicit representation of the solvent may be needed to most accurately capture helix-coil dynamics in simulation.

Clearly, replica exchange MD is not a panacea, and there have been developments to generalize this approach as well as to examine other generalized ensemble methods. For example, a replica exchange MD variant was further developed to include exchanges in multidimensional Hamiltonian space in combination with US [191]. It was also adapted to a heterogeneous parallel cluster by multiplexing the replicas at each temperature [184]. Nevertheless, it suffers from one significant problem when it is applied to significantly large systems. As can be inferred from the examples described in the above (82 replicas for protein A versus 16 for Met-enkephalin), the major drawback of the original replica exchange MD is the dependence of the number of replicas on the degrees of freedom N in the system. To obtain a reliable result, each pair of adjacent replicas must have overlapping energy distributions [181]. Due to the variance of energies in statistical mechanics, a system N times larger requires $N^{1/2}$ times more replicas. As a remedy, alternate Hamiltonian replica exchange MD has been proposed where replicas are generated by varying parameters other than the temperature such as the degree of hydrophobicity of the polymer chain [192] and a scaling parameter for selected energy terms such as the dihedral energy and protein–protein non-bonded interactions [193].

Moreover, there have been concerns regarding the computational cost of obtaining converged data with replica exchange MD. To test convergence, one must show that the results obtained in a given calculation are independent of the initial conditions. However, this test is fairly expensive, since it requires multiple replica exchange MD calculations to test convergence. Moreover, typically an replica exchange MD calculation is seeded with conformations from throughout the free energy landscape, making tests of convergence more difficult. As this method matures, we expect that further tests for convergence and a push for quantitative results will help generalized ensemble methods to attain even further predictive capability.

Finally, applying replica exchange MD-like methods to grid computing is not always simple and direct. Unfortunately, replica exchange MD requires both a homogenous cluster as well as fault tolerance, two aspects which are not present in grid computing. Therefore, for cases where even this level of coupling is too strong, there have been other approaches suggested. For example, a natural first approach to take is a generalization of replica exchange MD: Multiplexed replica exchange MD [184] addresses the issues of fault tolerance and the inhomogeneity of the network by assigning multiple processors (and thus multiple simulations) to each temperature replica, and then performing swaps between replicas of neighboring temperatures as those simulations complete. By making copies of replicas at a given temperature, fault tolerance is achieved (since the entire multiplexed replica exchange MD simulation does not wait for any single CPU to complete) and the difference in CPU speeds is handled naturally (since processors simply swap when available in an asynchronous manner). Also, recent efforts towards Markovian state models [194] suggest new advances in sampling which should be useful for both kinetics and thermodynamics.

13.9 Conclusions and Future Perspectives for Free Energy Calculations

Free energy calculations constitute a direct link between theory and experiment, by quantifying at the thermodynamic level the physical phenomena modeled by statistical simulations. With over two decades of insight gained from methodological development and characterization, a wide variety of problems of both chemical and biological relevance can now be tackled with confidence. Among the advances achieved in recent years, a significant step forward has been made in the calculation of free energies along a well-delineated reaction coordinate. 'Alchemical transformations' employed to mimic site-directed mutagenesis experiments have also benefited from advances in the understanding of the methodology and its efficient application. Efforts to appraise in a rigorous fashion the associated errors have played an active role in turning free energy calculations into another tool in the arsenal of the modeler. Altogether, free energy calculations have come of age, and are now a predictive approach, and no longer a mere proof of concept. High-performance computing has played an important role in reaching this status. By improving sampling, one can determine free energies with a sufficiently high accuracy, laying the path for such calculations to become more commonplace in the future. With this increase in precision, quantitative comparison to experiment yields very detailed tests of the accuracy of our models.

Thus, in addition to having greater predictive power, we now can better understand the limits of applicability of our models and the natural next steps to improve them. What is the future for free energy computation in academia and industry? Will massively parallel computations play a role? While it is natural to use very large-scale methods to address free energy calculations, clearly one would not invest in such methods (considering their great expense financially and in human time) without knowledge that these methods would be useful. But, until such resources are available, one would not know how potentially useful they may be. This leads to a 'catch–22' situation: while it is tempting to consider using greater computational resources, such resources generally are not available to scientists unless there is clear proof that they would significantly improve their results.

With recent advances in both sampling and model accuracy going hand in hand, future studies will have much greater precision and accuracy than ever before. It is likely that with the increased access to massively parallel architectures, as the price/performance ratio of computer chips continues to fall, the bottleneck of free energy calculations will shift from a purely computational aspect to a human one, due to the need for qualified modelers to set these calculations up and interpret them. This explains why their use in industry, and particularly in the pharmaceutical world over the past years has remained scarce. Cutting-edge applications of free energy calculations emanate essentially from academic environments, where the focus is not so much on high throughput, but rather on well-delineated, specific problems that often require more human attention than computational power. Yet, it is envisioned that in the reasonably near future, free energy methods will become an unavoidable

element of screening pipelines, discriminating between candidates selected from cruder approaches, to retain only the best leads towards a given target.

References

1. Zeng, Jun; Fridman, M.; Maruta, H.; Treutlein, Herbert R.; Simonson, T., Protein–protein recognition: an experimental and computational study of the R89K mutation in Raf and its effect on Ras binding, *Prot. Sci.* **1999**, *8*, 50–64.

2. Pezo, V.; Metzgar, D.; Hendrickson, T.L.; Waas, W.F.; Döring, V.; Marlière, P.; Schimmel, P.; de Crécy-Lagard, V., Artificially ambiguous genetic code confers growth yield advantage, *Proc. Natl Acad. Sci. USA* **2004**, *101*, 8593–8597.

3. Xie, J.; Schultz, P.G., Adding amino acids to the genetic repertoire, *Curr. Opin. Chem. Biol.* **2005**, *9*, 548–554.

4. Strømgaard, A.; Jensen, A. A.; Strømgaard, K., Site-specific incorporation of unnatural amino acids into proteins, *ChemBioChem* **2004**, *5*, 909–916.

5. Archontis, G.; Simonson, T.; Moras, D.; Karplus, M., Specific amino acid recognition by aspartyl–tRNA synthetase studied by free energy simulations, *J. Mol. Biol.* **1998**, *275*, 823–846.

6. Archontis, G.; Simonson, T.; Karplus, M., Binding free energies and free energy components from molecular dynamics and Poisson–Boltzmann calculations. Application to amino acid recognition by aspartyl–tRNA synthetase, *J. Mol. Biol.* **2001**, *306*, 307–327.

7. Thompson, Damien; Plateau, Pierre; Simonson, Thomas, Free-energy simulations and experiments reveal long-range electrostatic interactions and substrate-assisted specificity in an aminoacyl-tRNA synthetase., *ChemBioChem* **Feb 2006**, *7*, 337–344.

8. Simonson, T.; Archontis, G.; Karplus, M., Continuum treatment of long-range interactions in free energy calculations. Application to protein-ligand binding, *J. Phys. Chem. B* **1997**, *101*, 8349–8362.

9. Simonson, T., Electrostatic free energy calculations for macromolecules: a hybrid molecular dynamics/continuum electrostatics approach, *J. Phys. Chem. B* **2000**, *104*, 6509–6513.

10. Archer-Lahlou, E.; Tikhonova, I.; Escrieut, C.; Dufresne, M.; Seva, C.; Clerc, P.; Pradayrol, L.; Moroder, L.; Maigret, B.; Fourmy, D., Modeled structure of a G-protein-coupled receptor: the cholecystokinin-1 receptor, *J. Med. Chem.* **2005**, *48*, 180–191.

11. Hénin, J.; Maigret, B.; Tarek, M.; Escrieut, C.; Fourmy, D.; Chipot, C., Probing a model of a GPCR/ligand complex in an explicit membrane environment. The human cholecystokinin-1 receptor, *Biophys. J.* **2006**, *90*, 1232–1240.

12. Noskov, S.Y.; Bernèche, S.; Roux, B., Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands, *Nature* **2001**, *431*, 830–834.

13. Dolenc, J.; Oostenbrink, C.; Koller, J.; van Gunsteren, W.F., Molecular dynamics simulations and free energy calculations of netropsin and distamycin binding to an AAAAA DNA binding site, *Nucl. Acids Res.* **2005**, *33*, 725–733.

14. Talhout, R.; Villa, A.; Mark, A.E.; Engeberts, J.B.F.N., Understanding binding affinity: a combined isothermal tritration calorimetry/molecular dynamics study of the binding of a series of hydrophobically modified benzamidinium chloride inhibitors to trypsin, *J. Am. Chem. Soc.* **2003**, *125*, 10570–10579.

15. Guo, Z.Y.; Durkin, J.; Fischmann, T.; Ingram, R.; Prongay, A.; Zhang, R.M.; Madison, V., Application of the lambda-dynamics method to evaluate the relative binding free energies of inhibitors to HCV protease, *J. Med. Chem.* **2003**, *46*, 5360–5364.

16. Archontis, G.; Watson, K.A.; Xie, Q.; Andreou, G.; Chrysina, E.D.; S.E. Zographos; Oikonomakos, N.G.; Karplus, M., Glycogen phosphorylase inhibitors: a free energy perturbation analysis of glucopyranose spirohydantoin analogues, *Proteins* **2005**, *61*, 984–998.

17. Villa, A.; Zangi, R.; Pieffet, G.; Mark, A.E., Sampling and convergence in free energy calculations of protein-ligand interactions: the binding of triphenoxypyridine derivatives to Factor Xa and Trypsin, *J. Comp. Aided. Mol. Des.* **2003**, *17*, 673–686.

18. Gouda, H.; Kuntz, I.D.; Case, D.A.; Kollman, P.A., Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and thermodynamic integration methods, *Biopolymers* **2003**, *68*, 16–34.

19. Jorgensen, W. L., The many roles of computation in drug discovery, *Science* **2004**, *303*, 1813–1818.

20. Wong, C.F.; McCammon, J.A. Drug design. in *Encyclopedia of Supramolecular Chemistry*, Atwood, J.; Steed, J., Eds. Marcel Dekker: New York, 2004.

21. Woo, H. J.; Roux, B., Calculation of absolute protein–ligand binding free energy from computer simulations, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6825–6830.

22. Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. Steered molecular dynamics. in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Deuflhard, P.; Hermans, J.; Leimkuhler, B.; Mark, A. E.; Skeel, R.; Reich, S., Eds., vol. 4, *Lecture Notes in Computational Science and Engineering*. Springer Verlag: Berlin, 1998, pp. 39–65.

23. Grubmüller, H.; Heymann, B.; Tavan, P., Ligand binding: molecular mechanics calculation of the streptavidin–biotin rupture force, *Science* **1996**, *271*, 997–999.

24. Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Biophysical, K. Schulten., Molecular dynamics study of unbinding of the avidin–biotin complex, *Biophys. J.* **1997**, *72*, 1568–1581.

25. Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J., Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water, *J. Chem. Phys.* **1988**, *89*, 3742–3746.

26. Hermans, J.; Wang, L., Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme, *J. Am. Chem. Soc.* **1997**, *119*, 2707–2714.

27. Gilson, M.K.; Given, J.A.; Bush, B.L.; McCammon, J.A., The statistical–thermodynamic basis for computation of binding affinities: a critical review, *Biophys. J.* **1997**, *72*, 1047–1069.

28. Boresch, S.; Tettinger, F.; Karplus, M., Absolute binding free energies: a quantitative approach to their calculation, *J. Phys. Chem. B* **2003**, *107*, 9535–9551.

29. Tembe, B.L.; McCammon, J.A., Ligand–receptor interactions, *Comp. Chem.* **1984**, *8*, 281–283.

30. Hermans, J.; Shankar, S., The free energy of xenon binding to myoglobin from molecular-dynamics simulation, *Isr. J. Chem.* **1986**, *27*, 225–227.

31. Merz Jr., K.M., Carbon dioxide binding to human carbonic anhydrase II, *J. Am. Chem. Soc.* **1991**, *113*, 406–411.

32. Lee, F. S.; Warshel, A., A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations, *J. Chem. Phys.* **1992**, *97*, 3100–3107.

33. Miyamoto, S.; Kollman, P. A., What determines the strength of noncovalent association of ligands to proteins in aqueous solution ?, *Proc. Natl Acad. Sci. USA* **1993**, *90*, 8402–8406.

34. Miyamoto, S.; Kollman, P. A., *Proteins: Structure, Function and Genetics* **1993**, *16*, 226–245.

35. Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R., Structural origins of high-affinity biotin binding to Streptavidin, *Science* **1989**, *243*, 85–88.

36. Dixit, S. B.; Chipot, C., Can absolute free energies of association be estimated from molecular mechanical simulations ? The biotin–streptavidin system revisited, *J. Phys. Chem. A* **2001**, *105*, 9795–9799.

37. Roux, B.; Nina, M.; Pomès, R.; Smith, J. C., Thermodynamic stability of water molecules in the Bacteriorhodopsin proton channel: a molecular dynamics and free energy perturbation study, *Biophys. J.* **1996**, *71*, 670–681.

38. Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A., Revisiting free energy calculations: A theoretical connection to MM/PBSA and direct calculation of the association free energy, *Biophys. J.* **2004**, *86*, 67–74.

39. Hamelberg, D.; McCammon, J. A., Standard free energy of releasing a localized water molecule from the binding pockets of proteins: Double-decoupling method, *J. Am. Chem. Soc.* **2004**, *126*, 7683–7689.

40. Woo, H. J.; Dinner, A. R.; Roux, B., Grand canonical Monte Carlo simulations of water in protein environments, *J. Chem. Phys.* **2004**, *121*, 6392–6400.

41. Fersht, A., *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, Freeman: New York, 1999.

42. Pons, J.; Rajpal, A.; Kirsch, J.F., Energetic analysis of an antibody/antigen interface: alanine scanning mutagenesis and double mutant cycles on the HyHEL-10/lysozyme interaction, *Prot. Sci.* **1999**, *8*, 958–968.

43. Horovitz, A., Double-mutant cycles: a powerful tool for analyzing protein structure and function, *Fold. Des.* **1996**, *1*, R121–126.

44. Liu, S. Y.; Mark, A. E.; van Gunsteren, W. F., Estimating the relative free energy of different molecular states with respect to a single reference state, *J. Phys. Chem.* **1996**, *9485–9494*, 1749.

45. Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M., Hidden thermodynamics of mutant proteins: A molecular dynamics analysis, *Science* **1989**, *244*, 1069–1072.

46. Simonson, T.; Brünger, A. T., Thermodynamics of protein–peptide binding in the ribonuclease S system studied by molecular dynamics and free energy calculations., *Biochemistry* **1992**, *31*, 8661–8674.

47. Hodel, A.; Simonson, T.; Fox, R. O.; Brünger, A. T., Conformational substates and uncertainty in macromolecular free energy calculations., *J. Phys. Chem.* **1993**, *97*, 3409–3417.

48. van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. Computation of free energy in practice: Choice of approximations and accuracy limiting factors. in *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J., Eds., vol. 2. Escom: The Netherlands, 1993, pp. 315–348.

49. Mark, A.; van Gunsteren, W.F., Decomposition of the free energy of a system in terms of specific interactions: implications for theoretical and experimental studies, *J. Mol. Biol.* **1994**, *240*, 167–176.

50. Boresch, S.; Archontis, G.; Karplus, M., Free energy simulations: the meaning of the individual contributions from a component analysis, *Proteins: Structure, Function and Genetics* **1994**, *20*, 25–33.

51. Archontis, G.; Simonson, T. Dielectric relaxation in an enzyme active site: molecular dynamics simulations intepreted with a macroscopic continuum model, *J. Am. Chem. Soc.* **2001**, *123*, 11047–11056.

52. Boresch, S.; Karplus, M., The meaning of component analysis: decomposition of the free energy in terms of specific interactions, *J. Mol. Biol.* **1995**, *254*, 801–807.

53. Pitera, J.W.; van Gunsteren, W.F., The importance of solute–solvent van der Waals interactions with interior atoms of biopolymers, *J. Am. Chem. Soc.* **2001**, *123*, 3163–3164.
54. Hénin, J.; Pohorille, A.; Chipot, C., Insights into the recognition and association of transmembrane α-helices. The free energy of α-helix dimerization in glycophorin A, *J. Am. Chem. Soc.* **2005**, *127*, 8478–8484.
55. Carlsson, J.; Aqvist, J., Absolute and relative entropies from computer simulation with applications to ligand binding, *J. Phys. Chem. B* **2005**, *109*, 6448–6456.
56. van den Bosch, M.; Swart, M.; Snijders, J.G.; Berensen, H.J.C.; Mark, A.E.; Oostenbrink, C.; van Gunsteren, W.F.; Canters, G.W., Calculation of the redox potential of the protein azurin and some mutants, *ChemBioChem* **2005**, *6*, 738–746.
57. Sagui, C.; Darden, T., Molecular dynamics simulations of biomolecules: long-range electrostatic effects, *Ann. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155–179.
58. Åqvist, J.; Luzhkov, V., Ion permeation mechanism of the potassium channel, *Nature* **2000**, *404*, 881–884.
59. Florian, J.; Goodman, M.F.; Warshel, A., Free energy perturbation calculations of DNA destabilization by base substitutions: the effect of neutral guanine–thymine, adenine–cytosine and adenine–difluorotoluene mismatches, *J. Phys. Chem. B* **2000**, *104*, 10092–10099.
60. Stote, R.; States, D.; Karplus, M., On the treatment of electrostatic interactions in biomolecular simulation, *J. Chim. Phys.* **1991**, *88*, 2419–2433.
61. Warshel, A., Energetics of enzyme catalysis, *Proc. Natl Acad. Sci. USA* **1978**, *75*, 5250.
62. Resat, H.; McCammon, J.A., Free energy simulations: correcting for electrostatic cutoffs by use of the Poisson equation, *J. Chem. Phys.* **1996**, *104*, 7645–7651.
63. Warshel, A., *Computer Modelling of Chemical Reactions in Enzymes and Solutions*, John Wiley: New York, 1991.
64. Beglov, D.; Roux, B., Finite representation of an infinite bulk system: solvent boundary potential for computer simulations, *J. Chem. Phys.* **1994**, *100*, 9050–9063.
65. Im, W.; Bernèche, S.; Roux, B., Generalized solvent boundary potential for computer simulations, *J. Chem. Phys.* **2001**, *114*, 2924–2937.
66. Banavali, N.K.; Im, W.; Roux, B., Electrostatic free energy calculations using the generalized solvent boundary potential method, *J. Chem. Phys.* **2002**, *117*, 7381–7388.
67. Valleau, J. P.; Card, D. N., Monte Carlo estimation of the free energy by multistage sampling, *J. Chem. Phys.* **1972**, *57*, 5457–5462.
68. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free energy estimation: umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199.
69. Patey, G. N.; Valleau, J. P., A Monte Carlo method for obtaining the interionic potential of mean force in ionic solution, *J. Chem. Phys.* **1975**, *63*, 2334–2339.
70. Pangali, C.; Rao, M.; Berne, B. J., A Monte Carlo simulation of the hydrophobic interaction, *J. Chem. Phys.* **1979**, *71*, 2975–2981.
71. Pratt, L. R.; Chandler, D., Theory of hydrophobic effect, *J. Chem. Phys.* **1977**, *67*, 3683–3704.
72. New, M. H.; Berne, B. J., Molecular Dynamics Calculation of the Effect of Solvent Polarizability on the Hydrophobic Interaction, *J. Am. Chem. Soc.* **1995**, *117*, 7172–7179.
73. Chipot, C.; Kollman, P. A.; Pearlman, D. A., Alternative approaches to potential of mean force calculations: Free energy perturbation versus thermodynamic integration. Case study of some representative nonpolar interactions, *J. Comput. Chem.* **1996**, *17*, 1112–1131.
74. Jorgensen, W. L.; Severance, D. L., Aromatic–aromatic interactions: free energy profiles for the benzene dimer in water, chloroform and liquid benzene, *J. Am. Chem. Soc.* **1990**, *112*, 4768–4774.

75. Duffy, E. M.; Kowalczyk, P. J.; Jorgensen, W. L., Do denaturants interact with aromatic hydrocarbons in water ?, *J. Am. Chem. Soc.* **1993**, *115*, 9271–9275.

76. Tobias, D. J.; Brooks III, C. L., Calculation of free energy surfaces using the methods of thermodynamic perturbation theory, *Chem. Phys. Lett.* **1987**, *142*, 472–476.

77. Pratt, L. R., Molecular theory of hydrophobic effects: "She is too mean to have her name repeated", *Annu. Rev. Phys. Chem.* **2002**, *53*, 409–436.

78. Pratt, L. R.; Pohorille, A., Hydrophobic effects and modeling of biophysical aqueous solution interfaces, *Chem. Rev.* **2002**, *102*, 2671–2692.

79. Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A., Benzene dimer: a good model for π–π interactions in proteins? A comparison between the benzene and the toluene dimers in the gas phase and in an aqueous solution, *J. Am. Chem. Soc.* **1996**, *118*, 11217–11224.

80. Chipot, C.; Maigret, B.; Pearlman, D. A.; Kollman, P. A., Molecular dynamics potential of mean force calculations: a study of the toluene–ammonium π–cation interactions, *J. Am. Chem. Soc.* **1996**, *118*, 2998–3005.

81. Berkowitz, M.; Karim, O. A.; McCammon, J. A.; Rossky, P. J., Sodium chloride ion pair interaction in water: computer simulation, *Chem. Phys. Lett.* **1984**, *105*, 577–580.

82. Belch, A. C.; Berkowitz, M.; McCammon, J. A., Solvation structure of a sodium chloride ion pair in water, *J. Am. Chem. Soc.* **1986**, *108*, 1755–1761.

83. Dang, L. X.; Pettitt, B. M., Solvated chloride ions at contact, *J. Chem. Phys.* **1987**, *86*, 6560–6561.

84. Jorgensen, W. L.; Buckner, J. K.; Huston, S. E.; Rossky, P. J., Hydration and energetics for *tert*-butyl chloride ion pairs in aqueous solution, *J. Am. Chem. Soc.* **1987**, *109*, 1891–1899.

85. Hummer, G.; Soumpasis, D. M.; Neumann, M., Computer simulations do not support Cl–Cl pairing in aqueous NaCl solution, *Mol. Phys.* **1993**, *81*, 1155–1163.

86. Rozanska, X.; Chipot, C., Modeling ion–ion interaction in proteins: a molecular dynamics free energy calculation of the guanidinium–acetate association, *J. Chem. Phys.* **2000**, *112*, 9691–9694.

87. Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M., *J. Chem. Phys.* **1986**, *85*, 6720.

88. Kollman, P. A., Free energy calculations: applications to chemical and biochemical phenomena, *Chem. Rev.* **1993**, *93*, 2395–2417.

89. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.

90. Hermans, J.; Pathiaseril, A.; Anderson, A., Excess free-energy of liquids from molecular-dynamics simulations — application to water models, *J. Am. Chem. Soc.* **1988**, *110*, 5982–5986.

91. Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L., Free-energies of hydration and pure liquid properties of hydrocarbons from the OPLS all-atom model, *J. Phys. Chem.* **1994**, *98*, 13077–13082.

92. Jorgensen, W. L.; Tirado-Rives, J., Free energies of hydration for organic molecules from Monte Carlo simulations, *Perspectives in Drug Discovery and Design* **1995**, *3*, 123–138.

93. Jorgensen, W. L.; Laird, E. R.; Nguyen, T. B.; Tirado-Rives, J., Monte Carlo simulations of pure liquid substituted Benzenes with OPLS potential functions, *J. Comput. Chem.* **1993**, *14*, 206–215.

94. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A., Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation, *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.

95. Chipot, C., Rational determination of charge distributions for free energy calculations, *J. Comput. Chem.* **2003**, *24*, 409–415.

96. Udier-Blagovic, M.; De Tirado, P. M.; Pearlman, S. A.; Jorgensen, W. L., Accuracy of free energies of hydration using CM1 and CM3 atomic charges, *J. Comput. Chem.* **2004**, *25*, 1322–1332.

97. Daura, X.; Mark, A. E.; van Gunsteren, W. F., Parametrization of aliphatic CHn united atoms of GROMOS96 force field, *J. Comput. Chem.* **1998**, *19*, 535–547.

98. Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E., Reversible peptide folding in solution by molecular dynamics simulation, *J. Mol. Biol.* **1998**, *280*, 925–932.

99. Villa, A.; Mark, A. E., Calculation of the free energy of solvation for neutral analogs of amino acid side chains, *J. Comput. Chem.* **2002**, *23*, 548–553.

100. Maccallum, J. L.; Tieleman, D. P., Calculation of the water–cyclohexane transfer free energies of neutral amino acid side-chain analogs using the OPLS all-atom force field, *J. Comput. Chem.* **2003**, *24*, 1930–1935.

101. Shirts, M.; Pitera, J.; Swope, W.; Pande, V., Extremely precise free energy calculations of amino acid chain analogs: comparison of common molecular mechanical force fields for proteins, *J. Chem. Phys.* **2003**, *119*, 5740–5761.

102. MacKerell Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prod-hom, B.; Reiher III, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

103. Oostenbrink, C.; Villa, A.; Mark, A. E; van Gunsteren, W. F., A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6., *J. Comput. Chem.* **Oct 2004**, *25*, 1656–1676.

104. Shirts, M. R.; Pande, V. S., Solvation free energies of amino acid side chain analogs for common molecular mechanics water models, *J. Chem. Phys.* **2005**, *122*, 134508.

105. Rao, B. G.; Singh, U. C., Hydrophobic hydration: a free energy perturbation study *J. Am. Chem. Soc.* **1989**, *111*, 3125–3133.

106. Morgantini, P. Y.; Kollman, P. A., Solvation free-energies of amides and amines — Dis-agreement between free-energy calculations and experiment, *J. Am. Chem. Soc.* **1995**, *117*, 6057–6063.

107. Ding, Y. B.; Bernardo, D. N.; Kroghjespersen, K.; Levy, R. M., Solvation free-energies of small amides and amines from molecular-dynamics free-energy perturbation simu-lations using pairwise additive and many-body polarizable potentials, *J. Phys. Chem.* **1995**, *99*, 11575–11583.

108. Meng, E. C.; Caldwell, J. W.; Kollman, P. A., Investigating the anomalous solvation free energies of amines with a polarizable potential, *J. Phys. Chem.* **1996**, *100*, 2367–2371.

109. Rizzo, R. C.; Jorgensen, W. L., OPLS all-atom model for amines: resolution of the amine hydration problem, *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.

110. Oostenbrink, C.; van Gunsteren, W. F., Free energies of ligand binding for structurally diverse compounds, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 6750–6754.

111. Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L., Relative partition coefficients for organic solutes from fluid simulations, *J. Phys. Chem.* **1990**, *94*, 1683–1686.

112. Essex, J. W.; Reynolds, C. A.; Richards, W. G., Theoretical determination of partition coefficients, *J. Am. Chem. Soc.* **1992**, *114*, 3634–3639.

113. Orozco, M.; Colominas, C.; Luque, F. J., Theoretical determination of the solvation free energy in water and chloroform of the nucleic acid bases, *Chem. Phys.* **1996**, *9*, 209–678.

114. Best, S. A.; Merz Jr., K. M.; Reynolds, C. H., Free energy perturbation study of octanol/water partition coefficients: Comparison with continuum GB/SA calculations, *J. Phys. Chem. B* **1999**, *103*, 714–726.

115. Bas, D.; Dorison-Duval, D.; Moreau, S.; Bruneau, P.; Chipot, C., Rational determination of transfer free energies of small drugs across the water–oil interface, *J. Med. Chem.* **2002**, *45*, 151–159.

116. Hicks, J. M.; Kemnitz, K.; Eisenthal, K. B.; Heinz, T. F., Studies of liquid surfaces by second harmonic generation, *J. Phys. Chem.* **1986**, *90*, 560–562.

117. Eisenthal, K. B., Liquid interfaces, *Acc. Chem. Res.* **1993**, *26*, 636–643.

118. Pohorille, A.; Benjamin, I., Molecular dynamics of phenol at the liquid–vapor interface of water, *J. Chem. Phys.* **1991**, *94*, 5599–5605.

119. Alper, H. E.; Stouch, T. R., Orientation and diffusion of a drug analogue in biomembranes: molecular dynamics simulations, *J. Phys. Chem.* **1995**, *99*, 5724–5731.

120. Bassolino-Klimas, D.; Alper, H. E.; Stouch, T. R., Drug–membrane interactions studied by molecular dynamics simulation: size dependence of diffusion, *Drug Des. Discov.* **1996**, *13*, 135–141.

121. Tu, K.; Klein, M.; Tobias, D. J., Constant-pressure molecular dynamics investigations of cholesterol effects in a dipalmitoylphosphatidylcholine bilayer, *Biophys. J.* **1998**, *75*, 2147–2156.

122. Koubi, L.; Tarek, M.; Klein, M. L.; Scharf, D., Distribution of halothane in a dipalmitoylphosphatidylcholine bilayer from molecular dynamics calculations, *Biophys. J.* **2000**, *78*, 800–811.

123. Pohorille, A.; Chipot, C.; New, M.; Wilson, M. A. Molecular modeling of protocellular functions, in *Pacific Symposium on Biocomputing '96*, Hunter, L.; Klein, T. E., Eds. World Scientific: Singapore, 1996, pp. 550–569.

124. Overton, E., *Studien über die Narkose zugleich ein Betrag zur allgemeinen Pharmakologie*, Verlag von Gustav Fischer: Jena, 1901.

125. Pohorille, A.; Wilson, M.A.; New, M.H.; Chipot, C., Concentrations of anesthetics across the water–membrane interface; the Meyer–Overton hypothesis revisited, *Toxicology Lett.* **1998**, *100*, 421–430.

126. Wilson, M. A.; Pohorille, A., Mechanism of unassisted ion transport across membrane bilayers, *J. Am. Chem. Soc.* **1996**, *118*, 6580–6587.

127. Pohorille, A. Private communication, 1999.

128. Tieleman, D. P.; Berendsen, H. J. C.; Samson, M. S. P., Voltage-dependent insertion of alamethicin at phospholipid/water and octane/water, *Biophys. J.* **2001**, *80*, 331–346.

129. Roux, B., Theoretical and computational models of ion channels, *Curr. Opin. Struct. Biol.* **2002**, *12*, 182–189.

130. Chakrabarti, N.; Tajkhorshid, E.; Roux, B.; Pomès, R., Molecular basis of proton blockage in aquaporins, *Structure* **2004**, *12*, 65–74.

131. Wang, Y.; Schulten, K.; Tajkhorshid, E., What makes an aquaporin a glycerol channel? A comparative study of AqpZ and GlpF, *Structure* **2005**, *13*, 1107–1118.

132. Bernèche, S.; Roux, B., Energetics of ion conduction through the K^+ channel, *Nature* **2001**, *414*, 73–77.

133. Allen, T.W.; Andersen, O.S.; Roux, B., Energetics of ion conduction through the gramicidin channel, *Proc. Natl Acad. Sci. USA* **2004**, *101*, 117–122.

134. Burykin, A.; Kato, M.; Warshel, A., Exploring the origin of the ion selectivity of the KcsA potassium channel, *Proteins* **2003**, *52*, 412–426.

135. Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A., Calculation of the relative change in binding free energy of a protein–inhibitor complex, *Science* **1987**, *235*, 574–576.

136. Dang, L. X.; Merz, K. M.; Kollman, P. A., Free-energy calculations on protein stability – Thr^{157}–Val^{157} Mutation of T4 lysozyme, *J. Am. Chem. Soc.* **1989**, *111*, 8505–8508.

137. Shi, Y. Y.; Mark, A. E.; Wang, C. X.; Huang, F. H.; Berendsen, H. J. C.; van Gunsteren, W. F., Can the stability of protein mutants be predicted by free-energy calculations, *Prot. Eng.* **1993**, *6*, 289–295.

138. Pan, Y. P.; Daggett, V., Direct comparison of experimental and calculated folding free energies for hydrophobic deletion mutants of chymotrypsin inhibitor. 2: Free energy perturbation calculations using transition and denatured states from molecular dynamics simulations of unfolding, *Biochemistry* **2001**, *40*, 2723–2731.

139. Hodel, A.; Rice, L. M.; Simonson, T.; Fox, R. O.; Brünger, A. T., Proline *cis–trans* isomerization in staphylococcal nuclease — multi–substate free-energy perturbation calculations, *Prot. Sci.* **1995**, *4*, 636–654.

140. Tidor, B., Helix–capping interaction in λ–Cro protein — a free-energy simulation analysis, *Proteins: Struct. Func. Genet.* **1994**, *19*, 310–323.

141. Simonson, T.; Archontis, G.; Karplus, M., Free energy simulations come of age: protein–ligand recognition, *Acc. Chem. Res.* **2002**, *35*, 430–437.

142. Prevost, M.; Wodak, S. J.; Tidor, B.; Karplus, M., Contribution of the hydrophobic effect to protein stability — analysis based on simulations of the Ile-96- Ala mutation in barnase, *Proc. Natl Acad. Sci. USA* **1991**, *88*, 10880–10884.

143. Sun, Y. C.; Veenstra, D. L.; Kollman, P. A., Free energy calculations of the mutation of Ile96 → Ala in barnase: contributions to the difference in stability, *Prot. Eng.* **1996**, *9*, 273–281.

144. Warshel, A.; Sussman, F.; King, G., Free energy changes in solvated proteins: microscopic calculations using a reversible charging process, *Biochemistry* **1986**, *25*, 8368–8372.

145. Börjesson, Ulf; Hünenberger, Philippe H., Explicit-solvent molecular dynamics simulation at constant pH: methodology and application to small amines, *J. Chem. Phys.* **2001**, *114*, 9706–9719.

146. Mongan, J.; Case, D.A.; McCammon, J.A., Constant pH molecular dynamics in generalized Born implicit solvent, *J. Comp. Chem.* **2004**, *25*, 2038–2048.

147. Lee, M.S.; Salsbury Jr., F.R.; Brooks III, C.L., Constant pH molecular dynamics using continuous titration coordinates, *Proteins* **2004**, *56*, 738–752.

148. Simonson, T., Electrostatics and dynamics of proteins, *Rep. Prog. Phys.* **2003**, *66*, 737–787.

149. Ichiye, T. Simulating redox proteins. in *Computational Biochemistry & Biophysics*, Becker, O.; Mackerell Jr., A.; Roux, B.; Watanabe, M., Eds. Marcel Dekker: New York, 2001.

150. Simonson, T., Gaussian fluctuations and linear response in an electron transfer protein, *Proc. Natl Acad. Sci. USA* **2002**, *99*, 6544–6549.

151. Sterpone, F.; Ceccarelli, M.; Marchi, M., Linear response and electron transfer in complex biomolecular systems and a reaction center protein, *J. Phys. Chem. B* **2003**, *107*, 11208–11215.

152. Tan, M.L.; Dolan, E.A.; Ichiye, T., Understanding intramolecular electron transfer in ferredoxin: a molecular dynamics study, *J. Phys. Chem. B* **2004**, *108*, 20435–20441.

153. Warshel, A.; Chu, Z.T.; Parson, W.W., Dispersed polaron simulations of electron transfer in photosynthetic reaction centers, *Science* **1989**, *246*, 112–116.

154. Li, G.H.; Zhang, X.D.; Cui, Q., Free energy perturbation calculations with combined QM/MM. Potential complications, simplifications, and applications to redox potential calculations, *J. Phys. Chem. B* **2003**, *107*, 8643–8653.

500   C. Chipot et al.

155. Hummer, G.; Pratt, L.; Garcia, A. E., Free energy of ionic hydration, *J. Phys. Chem.* **1996**, *100*, 1206–1215.
156. Simonson, T.; Carlsson, J.; Case, D. A., Proton binding to proteins: pK_a calculations with explicit and implicit solvent models, *J. Am. Chem. Soc.* **2004**, *126*, 4167–4180.
157. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. C.; Kollman, P. A., A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
158. Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D.G., How enzymes work: analysis by modern reaction rate theory and computer simulations, *Science* **2004**, *303*, 186–195.
159. Mulholland, A., Modelling enzyme reaction mechanisms, specificity and catalysis, *Drug Disc. Today* **2005**, *10*, 1393–13402.
160. Maupin, C. M.; Wong, K. F.; Soudackov, A. V.; Kim, S.; Voth, G. A., A multistate empirical valence bond description of protonatable amino acids, *J. Phys. Chem. A* **2006**, *110*, 631–639.
161. Xu, J.; Voth, G. A., Computer simulation of explicit proton translocation in cytochrome c oxidase: the D-pathway, *Proc. Natl Acad. Sci. U S A* **2005**, *102*, 6795–6800.
162. Xu, J.; Voth, G. A., Free energy profiles for H$^+$ conduction in the D-pathway of Cytochrome c oxidase: a study of the wild type and N98D mutant enzymes, *Biochim. Biophys. Acta* **2006**, *1757*, 852–859.
163. Voth, G. A., Computer simulation of proton solvation and transport in aqueous and biomolecular systems, *Acc. Chem. Res.* **2006**, *39*, 143–150.
164. Lyne, P. Mixed quantum/classical methods. in *Computational Biochemistry & Biophysics*, Becker, O.; Mackerell Jr., A.; Roux, B.; Watanabe, M., Eds. , N.Y., New York, 2001.
165. Ridder, L.; Rietjens, I.M.; Vervoort, J.A.; Mulholland, A., Quantum mechanical/molecular mechanical free energy simulations of the glutathione S-transferase (M1-1) reaction with phenanthrene 9,10-oxide, *J. Am. Chem. Soc.* **2002**, *123*, 9926–9936.
166. Cummins, P.L.; Gready, J.E., Computational methods for the study of enzymic reaction mechanisms III: a perturbation plus QM/MM approach for calculating relative free energies of protonation, *J. Comp. Chem.* **2005**, *26*, 561–568.
167. Gumbart, J.; Wang, Y.; Aksimentiev, A.; Tajkhorshid, E.; Schulten, K., Molecular dynamics simulations of proteins in lipid bilayers, *Curr. Opin. Struct. Biol.* **Aug 2005**, *15*, 423–431.
168. Tajkhorshid, E.; Nollert, P.; Jensen, M. Ø.; Miercke, L. J. W.; O'Connell, J.; Stroud, R. M.; Schulten, K., Control of the selectivity of the aquaporin water channel family by global orientational tuning, *Science* **2002**, *296*, 525–530.
169. Roux, B., The art of dissecting the function of a potassium channel, *Neuron* **Sep 2005**, *47*, 777–778.
170. Roux, B., Ion conduction and selectivity in K$^+$ channels, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 153–171.
171. Bohannon, John, Distributed computing. Grassroots supercomputing., *Science* **May 2005**, *308*, 810–813.
172. Shirts, M.; Pande, V., Screen savers of the world unite!, *Science* **2000**, *290*, 1903–1904.
173. Shirts, M. R.; Pande, V. S., Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration, *J. Chem. Phys.* **Apr 2005**, *122*, 144107.
174. Shea, J. E.; Brooks III, C. L., From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding, *Annu. Rev. Phys. Chem.* **2001**, *52*, 499–535.

175. Ferrenberg, A. M.; Swendsen, R. H., Optimized Monte Carlo data analysis, *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.
176. Boczko, E.M.; Brooks III, C.L., First-principles calculation of the folding free energy of a three-helix bundle protein, *Science* **1995**, *269*, 393–396.
177. Sheinerman, F. B.; Brooks, C. L., Calculations on folding of segment B1 of streptococcal protein G., *J. Mol. Biol.* **May 1998**, *278*, 439–456.
178. Sheinerman, F. B.; Brooks, C. L., Molecular picture of folding of a small alpha/beta protein., *Proc. Natl Acad. Sci. USA* **Feb 1998**, *95*, 1562–1567.
179. Shea, J. E.; Onuchic, J. N.; III, C. L. Brooks, Probing the folding free energy landscape of the Src-SH3 protein domain, *Proc. Natl Acad. Sci. USA* **Dec 2002**, *99*, 16064–16068.
180. Mitsutake, A.; Sugita, Y.; Okamoto, Y., Generalized-ensemble algorithms for molecular simulations of biopolymers, *Biopolymers* **2001**, *60*, 96–123.
181. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding, *Chem. Phys. Lett.* **1999**, *314*, 141–151.
182. Sanbonmatsu, K. Y.; García, A. E., Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics, *Prot. Struct. Function Genetics* **Feb 2002**, *46*, 225–234.
183. Zhou, F. X.; Berne, B. J.; Germain, R., The free energy landscape of β hairpin folding in explicit water, *Proc. Natl Acad. Sci. USA* **2001**, *98*, 14931–14936.
184. Rhee, Y. M.; Pande, V. S., Multiplexed-replica exchange molecular dynamics method for protein folding simulation, *Biophys. J.* **2003**, *84*, 775–786.
185. Nymeyer, H.; García, A. E., Simulation of the folding equilibrium of α-helical peptides: a comparison of the generalized Born approximation with explicit solvent, *Proc. Natl Acad. Sci. USA.* **Nov 2003**, *100*, 13934–13939.
186. García, A. E.; Sanbonmatsu, K. Y., Exploring the energy landscape of a beta hairpin in explicit solvent., *Proteins* **2001**, *42*, 345–354.
187. Lifson, S.; Roig, A., Theory of helix–coil transition in polypeptides, *J. Chem. Phys.* **1961**, *34*, 1963–1974.
188. Qian, H.; Schellman, J.A., Helix–coil theories: a comparative study for finite length polypeptides, *J. Phys. Chem.* **1992**, *96*, 3987–3994.
189. Vila, J. A.; Ripoll, D. R.; Scheraga, H. A., Physical reasons for the unusual α-helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides, *Proc. Natl Acad. Sci. USA* **2000**, *97*, 13075–13079.
190. Wu, X.; Wang, S., Helix folding of an alanine-based peptide in explicit water, *J. Phys. Chem. B* **2001**, *105*, 2227–2235.
191. Sugita, A. Kitao; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations, *J. Chem. Phys.* **2000**, *113*, 6042–6051.
192. Fukunishi, O. Watanabe; Takada, S., On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction, *J. Chem. Phys.* **2002**, *116*, 9058–9067.
193. Jang, S. Shin; Pak, Y., Replica-exchange method using the generalized effective potentia, *Phys. Rev. Lett.* **2003**, *91*, 058305–4.
194. Singhal, N.; Snow, C. D.; Pande, V. S., Using path sampling to build better Markovian state models: predicting the folding rate and mechanism of a tryptophan zipper beta hairpin, *J. Chem. Phys.* **Jul 2004**, *121*, 415–425.

14

Summary and Outlook

Andrew Pohorille and Christophe Chipot

In this final chapter, we address a few issues that might have lingered in the reader's mind from the beginning of the book. In the first section, we will summarize and illustrate in several examples the central idea that was initially put forth in the Introduction. This idea is that the many different methods for free energy calculations have common theoretical underpinnings, and may, therefore, be understood in a unified framework. On the basis of our current state of knowledge, we will further provide general suggestions about which methods might be used successfully for different problems of both chemical and biological interest, and how to make free energy calculations more reliable. In the second section, we will turn toward the future and address two closely related topics that are important for further development in this research area. One is the relevance of free energy calculations to solving problems of interest to basic science and technology. The second regards the main conceptual and technical difficulties in the field and the most promising directions for resolving them.

14.1 Summary: A Unified View

A recurrent, unifying theme of this book is that almost all free energy calculations in chemistry and biology are based on just four approaches for estimating the ratio of the partition functions in (1.15) of Chap. 1. They are: (1) free energy perturbation (FEP), (2) probability density, i.e., histograms, (3) thermodynamic integration (TI), and (4) nonequilibrium work methods, and are all discussed at length in Chaps. 2–6. Conceptually, all these approaches are closely related. For instance, FEP and TI can be considered as two extreme cases of the nonequilibrium work method, in which the change in a system is either instantaneous or infinitely slow. Similarly, TI can be derived from FEP by taking the limit of an infinitely small perturbation. Different methods for calculating free energy differences can be framed in terms of one of these four generic formalisms, or the combination thereof.

At the core of practically all methods for free energy calculations are techniques for improving sampling of the relevant regions of phase space. These techniques

have proven to help overcome quasi-nonergodicity scenarios, which, unfortunately, are only too common in chemical and biological systems. Many of these techniques appear to be quite different and are often specifically tied to estimating free energy as either a time-averaged or a space-averaged quantity. Yet, after careful analysis, all of them can be reduced to two basic strategies: stratification, or multistaging, and importance sampling, introduced in Sect. 1.4 of Chap. 1. Since both strategies have turned out to be extremely powerful, they are often used together.

Another apparent difference between the various free energy methods lies in the treatment of order parameters. In the original formulation of a number of methods, order parameters were dynamical variables – i.e., variables that can be expressed in terms of the Cartesian coordinates of the particles – whereas in others, they were parameters in the Hamiltonian. This implies a different treatment of the order parameter in the equations of motion. If one, however, applies the formalism of metadynamics, or extended dynamics, in which any parameter can be treated as a dynamical variable, most conceptual differences between these two cases vanish.

Understanding and appreciating the connections between different methods is important not only from a theoretical standpoint, but also for evaluating the efficiency and the accuracy of each of them. In some instances, closely related methods were developed independently of each other, and the connection between them was discovered only later. Bennett's acceptance ratio method [1] and the simple overlap sampling (SOS) scheme of Lee and Scott [2], discussed in Chap. 6, can serve as a case example. Both methods rely on sampling configurations in the reference and the target states, and calculating energy differences between these two states and a suitably defined intermediate. In SOS, this intermediate is located halfway between the reference and the target state, whereas in Bennett's method, its location is optimized to minimize the statistical error. This immediately implies that SOS is simpler, albeit less accurate than the seminal acceptance ratio method.

Conversely, analysis of two methods that seem to be very closely related – viz. the adaptive biasing potential and the adaptive biasing force (ABF), described in Chaps. 3 and 4, respectively – reveals that they are in fact less similar than might have appeared. Although they are both, in their generic form, aimed at augmenting the Hamiltonian of the system such that all values of the order parameter between the initial and the final states are sampled uniformly, they belong to different families of free energy methods, namely probability density and TI, respectively. More importantly, they differ in the adaptation strategy. To modify the potential, it is required that the probability distribution be approximated in the full range of the order parameter, or at least within one stratum. In contrast, the derivative of the potential with respect to the order parameter – i.e., the average force – can be adapted locally. The latter procedure provides good estimates faster, and for this reason ABF is expected to be more efficient that the adaptive biasing potential.

A closer look at different methods helps us to understand which features are responsible for their success. Let us compare, for instance, the nonequilibrium work method with the adaptive equilibrium approaches, described above. In the most common implementation of the former, the instantaneous force acting on the system along the order parameter is always equal to zero. In contrast, in the adaptive

approaches, only the average forces at each value of the order parameter are modified to become equal to zero, yet fluctuations of the force remain unchanged. This difference appears to form the physical basis for the argument that free energies calculated using configurations from an equilibrium ensemble will be more efficient than free energies estimated from nonequilibrium simulations [3].

Even seemingly unrelated methods can be represented in a similar framework. Let us consider now parallel tempering, discussed in Chap. 8, in which separate molecular dynamics (MD) trajectories – or Monte Carlo (MC) walkers – are run for N different values on an order parameter, e.g., the temperature. Let us further assume that N increases within a given range. In the limit $N \rightarrow \infty$, the order parameter becomes a continuous parameter. It can then be treated as a dynamical variable in extended dynamics simulations, and the free energy can be evaluated using probability distributions or TI methods augmented, for example, by an adaptive algorithm. These considerations illuminate a feature of parallel tempering that makes this method efficient: the extension of the sampled space by an additional parameter, combined with possible large jumps in the phase space help overcome quasi-nonergodicity in the system. In general, this is a well-known strategy for improving the efficiency of optimizing complex functions of many variables. In the context of free energy calculations, the usefulness of this strategy was demonstrated for TI [4]. Conversely, reducing the dimensionality of phase space that occurs in constrained TI, or effectively in stratified simulations with very small windows, can exacerbate problems with quasi-nonergodicity. This shows that the same technical trick works for different methods.

Since different methods are conceptually related, and rely on similar enhanced sampling techniques, one might propose that they are approximately equally efficient if implemented optimally. Although this assertion has never been demonstrated formally, intuitively, it may not be far from true. In practice, however, at least at the current state of the art, different methods have not reached the same level of optimality. This is one reason why there are established preferences for applying specific free energy methods to different problems. Yet, these preferences are not always related to efficiency. Often, other criteria, such as simplicity and robustness, also play a role. For instance, assigning parameters to additional variables in extended dynamics, so that they are effectively coupled to the rest of the system, still remains somewhat of an art. For this reason, applications of this approach to free energy calculations have been rather limited, even though methods based on extended dynamics may ultimately prove to be very efficient.

Among the methods discussed in this book, FEP is the most commonly used to carry out 'alchemical transformations' described in Sect. 2.8 of Chap. 2. Probability distribution and TI methods, in conjunction with MD, are favored if there is an order parameter in the system, defined as a dynamical variable. Among these methods, ABF, derived in Chap. 4, appears to be nearly optimal. Its accuracy, however, has not been tested critically for systems that relax slowly along the degrees of freedom perpendicular to the order parameter. Adaptive histogram approaches, primarily used in Monte Carlo simulations – e.g., multicanonical, WL and, in particular, the transition matrix method – yield superior results in applications to phase transitions,

as commented on in Chaps. 3 and 10. Traditional applications of the potential distribution theorem (PDT), relying upon the particle insertion method and described at the beginning of Chap. 9, may yield unparalleled accuracy when estimating free energies of dissolving small solutes. Parallel tempering, discussed in Chap. 8, has proven its usefulness for problems, in which: (1) a simple order parameter cannot be conveniently defined, and (2) severe quasi-nonergodicity problems are encountered during sampling. Problems of this kind are frequently encountered in the exploration of protein structures, protein folding, or in the formation of clusters. If a suitable order parameter cannot be identified, but the initial and the final states of the system are known, transition path sampling introduced in Chap. 7 is the method of choice. Nonequilibrium methods, presented in Chap. 5, have proven to be extremely valuable for interpreting the so-called 'pulling' experiments carried out on single molecules. So far, however, they have not been found to be particularly efficient for other problems. It should, nonetheless, be stressed that these methods are much newer than the methods based on equilibrium simulations, and, quite obviously, their potential has not yet been fully realized.

Association of different methods and specific classes of problems should not be viewed as a ready-made recipe for deciding how to carry out free energy calculations. In fact, for each case listed above, there are documented instances, in which methods other than the popular, most common ones have been applied successfully to a given problem. Instead, it provides an initial guidance for choosing a method suitable for the problem at hand – a valuable help, especially for users less experienced with free energy calculations. It is unlikely that one would go wrong by selecting a method on the basis of this guide, but it might be possible to do even better by considering appropriate alternatives.

No matter which method is selected, it is essential to follow good practices, which can be considered as the computational equivalent of proper experimental protocols. At a very modest, additional computational effort, and sometimes even saving computer time, a properly chosen method can provide greatly improved free energy estimates. For instance, it is recommended to use overlap sampling or Bennett's method in conjunction with FEP. Similarly, combining the probability density or TI approach with any good technique that yields nearly uniform sampling of the relevant regions in phase space almost always guarantees improved accuracy of the results. If simulations with several biasing potentials are carried out, employing the weighted histogram analysis method (WHAM), or any similar scheme (see Sect. 6.6.3), is strongly encouraged. Another aspect of free energy calculations that needs to be closely controlled is the completeness of sampling. Although there is no foolproof defense against persistent quasi-nonergodicity scenarios, various indicators can be used to assess whether adequate sampling has been achieved. For example, in FEP calculations, one should routinely plot the integrand in (2.12) of Chap. 2 as a function of ΔU to provide evidence that it is sufficiently smooth to yield reasonable estimates of ΔA. The final, and critical element is a careful analysis of the associated error, which, as has been seen in Chap. 6, can be tricky. Without error bars on free energy estimates, the goal of achieving good agreement between theory and experiment – the Holy Grail of the field – becomes, however, meaningless.

14.2 Outlook: What is the Future Role of Free Energy Calculations?

The significance of free energy calculations in basic theoretical research on chemical or biological systems is firmly established. This is because free energies, or chemical potentials, are the central quantities that determine the behavior of these systems at or near equilibrium. Their knowledge is essential for elucidating the principles that explain how different systems work, or processes of interest proceed. Once these principles are understood, they can guide the experimental design of novel systems. Theoretical and computational research on the nature of hydrophobic effects represent a good example of this approach. After an extensive effort that has spanned three decades and involved numerous calculations of hydration free energies, potentials of mean force and conformational equilibria for solutes of different sizes, and under different thermodynamic conditions, a coherent molecular picture of hydrophobic phenomena has started to emerge [5–7].

The usefulness of free energy calculations to problems of technological interest is not as clear, even if the solution to these problems depends on the knowledge of thermodynamic equilibria at a microscopic level. This is particularly evident in biotechnology, nanotechnology, and computer-aided drug design. Doubts about free energy calculations persist, but not because there has not been any progress in this area. On the contrary, as we have argued in the Introduction, this progress has been impressive. It has been matched, however, by equally impressive progress in competitive approaches. One of them relies on high-throughput experiments. In drug design, it entails rapid screening of large libraries of ligands for protein binding to uncover potential lead compounds. In protein engineering, it involves similar techniques for finding proteins with novel functionalities such as phage display [8], DNA shuffling [9, 10], and in vitro evolution [11]. Another competitive approach is based on heuristic computational methods. In the earliest days of free energy calculations, only a few such methods – e.g., QSAR – were available, and their range of applicability was rather limited. Since then, several powerful methods, such as molecular docking [12–14], homology modeling [15, 16], and computational protein design [17–19] have been developed and matured.

For these reasons, it has sometimes been argued that ab initio free energy calculations are and will remain less effective than the alternative approaches. We feel that this point of view is based on false premises. All the methods mentioned above have their place in technology-motivated research. Quite likely, the most effective approach is to use them in combination, probably in a hierarchical fashion. In this spirit, molecular-level free energy calculations should be used to improve and refine initial designs obtained from high-throughput experiments or computations.

The continuing success of free energy calculations in developing a better understanding of complex chemical and biological systems and the usefulness of these calculations in aiding molecular-level technologies rely on further progress in several areas of theory and computation. Despite remarkable theoretical advances, especially in enhanced sampling techniques, several methodological issues that may influence dramatically the efficiency and the reliability of free energy calculations remain

unresolved. They are mainly connected with understanding the nature of different motions in complex systems. One example is the choice of an appropriate order parameter, or, more generally, a suitable reduction of a many-body problem to motion of the system on a low-dimensional hyperplane in a potential of force averaged over the remaining degrees of freedom. This problem is particularly acute if slow, collective motions in a system are responsible for quasi-nonergodicity. Currently, no general method has been put forth for identifying degrees of freedom that equilibrate slowly during the transformation of a system from the initial to the final state. Consequently, it is often difficult to define order parameters that facilitate their adequate sampling. For instance, insertion of a protein into a membrane involves slow, collective reorganization of the many surrounding phospholipids and extensive changes in the structure of the water–membrane interface near the site of insertion. Similarly, protein folding is associated with concerted changes of many torsional angles in both the backbone and the side chains. In spite of countless efforts, a low-dimensional description of either one of these transformations has not been yet developed. On several recent occasions, it has been realized that seemingly reasonable choices of order parameters are poor approximations to the actual reaction coordinates, which represent progress along reaction pathways. Such choices may lead to unreliable estimates of the free energy barriers associated to the process of interest, and to incomplete sampling of states of the system that are relevant to the accurate estimate of free energy differences.

The knowledge of an order parameter that approximates well the reaction coordinate is insufficient to guarantee a reliable estimate of the free energy. Once such an order parameter has been identified, enhanced sampling techniques almost always have to be used to ensure efficient free energy calculations. This reduces the time scale of motion along this parameter, and, hence, leads to a more efficient statistical averaging. There is, however, a natural limit to this reduction: sampling has to be sufficiently extensive to allow motions along other degrees of freedom in the system to be properly averaged statistically. This means that, if there are many slow motions that influence the state of the system during the transformation, benefits from enhanced sampling techniques are limited, unless all these motions are identified and adequately sampled. The existence of such motions in complex systems, for instance in proteins, is quite common, as brought to light by the so-called 'essential dynamics' [20–23], and normal-mode analysis [24–29].

Whenever a suitable definition of an order parameter is not feasible, the range of tools that can be used for efficient free energy calculations becomes quite limited. In some instances, it is possible to employ a 'generic' order parameter, such as temperature, for example in conjunction with parallel tempering described in Chap. 8. In the special case when both the initial and the final states are known, transition path sampling, discussed in Chap. 7, can be used profitably. At the present time, however, these methods are not as reliable and efficient as approaches requiring an order parameter. Our ability to calculate free energy differences becomes even more limited if the final state of the system is not known a priori, as is often the case in protein folding. In general, successful calculations of free energies associated with such complex processes as protein–protein and protein–DNA interactions,

binding of large, flexible ligands to macromolecules, or assisted transport of ions across membranes will, most likely, require the extension of the current theoretical framework into the areas of statistical mechanics and molecular simulations that have been traditionally considered as separate – for instance chemical kinetics and chemical dynamics.

In addition to theoretical advancements, the second ingredient that has driven progress in free energy calculations was the continuously increasing power of digital computers. This can be clearly seen by comparing the work of Tembe and McCammon [30], Miyamoto and Kollman [31], and that of Hénin et al. [32], published roughly 10 years apart. They all have in common the use of perturbation theory and molecular dynamics simulations to tackle protein–ligand association problems. Yet, whereas Tembe and McCammon investigated a rudimentary model consisting of a 'receptor' and a 'ligand' atom solvated by a bath of 50 solvent Lennard-Jones particles, Miyamoto and Kollman considered one hydrated monomer of the homotetramer streptavidin, to which the ligand biotin was bound, and Hénin et al. studied the seven-helix human G-coupled protein receptor (GPCR) of cholecystokinin [33] in a hydrated palmitoyloleylphosphatidylcholine (POPC) lipid bilayer, forming a total of 72,255 atoms.

In an effort to bridge computer simulations to reality, especially in biology and material science, the size of the systems modeled at the atomic level constantly increases. For example, the recent MD simulation of the movement of tRNA into the ribosome during decoding [34] – a fundamental step for information transfer from RNA to protein – involved a 2 ns trajectory of some 2,640,000 atoms. Such computational effort would have been unthinkable a decade ago. This creates large demands on free energy calculations. As the size of the system grows continuously, so does the range of motions that need to be sampled properly. Reliable free energy calculations for large systems thus also require longer simulation times – or, equivalently, longer MC walks. It is unlikely that simultaneous growth in both spatial and temporal scales of simulations can be achieved by simply relying on the increasing speed of a single processor, even if the latter continues to obey Moore's law in the coming 10 years. To meet the growing computational demands, it is necessary to take full advantage of massively parallel or massively distributed computing. The first steps in this direction have already been made, in particular in the pioneering work of Vijay Pande, described in the previous chapter.

The increasing focus on interprocessor communication rather than on floating-point operations, which is associated with massively parallel computing, has to be accompanied by the development of algorithms for molecular simulations and free energy calculations that minimize the time needed for transferring data between processors. Fortunately, most methods for obtaining free energies parallelize in a natural way. For example, individual trajectories in nonequilibrium work calculations or parallel tempering can be run on separate processors, or clusters of processors. A similar strategy can be applied to different strata in stratified calculations, discussed in Sects. 1.4, 2.6, and 3.3.1. It is also possible to carry out parallel free energy calculations for many different ligands interacting with a common macromolecule, or for a variety of single-point mutations in a protein. This would be analogous to

high-throughput experiments, albeit performed in silico. These strategies work well on tens of processors, but it is not clear whether they will remain equally efficient if implemented without modifications on hundreds or thousands of processors. New theoretical formulations of the problem and/or new algorithms might then be needed.

Methodological and computational advances should not be dichotomized from the issue of precision in the description of the forces acting in the system. In the early days of free energy calculations, the reliability of the results was essentially determined in terms of sampling accuracy. In fact, several high-impact, early simulations were later proven to be too short to yield accurate free energy estimates. As the field matured, it became, however, increasingly commonly accepted that the shortcomings of the molecular force fields rather than incomplete sampling constituted the main factor limiting the reliability of free energy calculations.

Calculations of forces may be improved in several ways. One is to pursue efforts towards the development of accurate classical, atomic-level force fields. A promising extension along these lines is to add nonadditive polarization effects to the usual pairwise additive description of interatomic interactions. This has been attempted in the past [35–39], but has not brought the expected and long-awaited improvements. This is not so much because polarization effects are not important, or pairwise additive models can account for them accurately in an average sense in all, even highly anisotropic environments. Instead, it seems more likely that the previously developed nonadditive potentials were not sufficiently accurate to offer an enhanced description of those systems in which induction phenomena play a crucial role.

Another direction for improving the description of intermolecular interactions is to include quantum effects. This is usually necessary whenever there is a significant charge redistribution during the transformation between the initial and the final state, and, in particular, when chemical bonds are formed or broken, or when a charge transfer occurs between chemical species. The problem arises, for example, in the calculation of free energies characterizing spontaneous or enzymatic reactions, or proton transfer between ionizable groups. If a system of interest can be reduced to a small-size model, purely quantum mechanical methods can be employed. In general, however, combination of quantum and statistical mechanics is needed. A variety of techniques, such as quantum mechanics/molecular mechanics (QM/MM) [40, 41], Car–Parrinello [42–45], or valence bond [46–48] methods have been devised for this purpose. The problem in applying these techniques is not conceptual; as has been shown in Chap. 11, free energy calculations can be integrated smoothly with quantum mechanical simulations. In practice, these calculations become prohibitively slow as the size of the system increases, even if only a reduced part of it is treated at the quantum mechanical level.

At the antipodes of the latter description, there is a continuous need for better 'low-resolution' models that involve, for instance, coarse graining of molecules, or implicit solvation. This need is motivated by the expectation that the free energy of a large system can be calculated with sufficient accuracy without requiring that all its components be described at the atomic level. In many cases, this is equivalent to the assumption that a mean-field approximation works, or that many fast degrees of freedom can be removed from the system, yet without any appreciable loss of

accuracy. Several ideas along these lines have been discussed in Chap. 12, but there is still plenty of room for improvement.

The requirements for additional details in the electronic structure, which are captured in quantum mechanical calculations, and for a reduced representation of the system appear to be at cross purposes. This is, however, not so, because both of these requirements can be accommodated in the rapidly developing area of multiscale simulations. They hold great promise, allowing large, complex systems to be described accurately in simulations of manageable size. Yet, this field is still in the early stages of development, and relatively little work has been done on integrating it with free energy calculations. One new approach that appears to be particularly compatible with multiscale modeling is based on the quasichemical approximation, which was discussed in Chap. 9. Without much doubt, the ability to obtain reliable free energy estimates on the basis of multiscale simulations would constitute a significant theoretical and practical advance.

The final element consists of providing access to state-of-the-art techniques of free energy calculations to a broad research community interested in applications of molecular simulations to chemistry and biology. This requires integration of these techniques with modern simulation software to a much greater extent than has been done so far. The goal is to create tools that would allow researchers, who are primarily focused on applications rather than on theoretical developments, to carry out free energy calculations with the same level of confidence and control that they currently have in electronic structure calculations or bioinformatics analyses. This does not, of course, mean that such calculations should be treated as 'black-box' routine jobs. As is the case in other areas of theoretical chemistry and biology, computational tools can complement, but not substitute for an insight into the problem at hand.

References

1. Bennett, C. H., Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **1976**, *22*, 245–268
2. Lee, C. Y., Scott, H. L., The surface tension of water: a Monte Carlo calculation using an umbrella sampling algorithm, *J. Chem. Phys.* **1980**, *73*, 4591–4596
3. Oberhofer, H., Dellago, C., Geissler, P. L., Biased sampling of non-equilibrium trajectories: can fast switching simulations beat conventional free energy calculation methods? *J. Phys. Chem. B* **2005**, *109*, 6902–6915
4. Bitetti-Putzer, R., Yang, W., Karplus, M., Generalized ensembles serve to improve the convergence of free energy simulations, *Chem. Phys. Lett.* **2003**, *377*, 633–641
5. Paulaitis, M. E., Pratt, L. R., Hydration theory for molecular biophysics, *Adv. Prot. Chem.* **2002**, *62*, 283–310
6. Pratt, L. R., Molecular theory of hydrophobic effects: "She is too mean to have her name repeated", *Annu. Rev. Phys. Chem.* **2002**, *53*, 409–436
7. Chandler, D., Interfaces and the driving force of hydrophobic assembly, *Nature* **2005**, *437*, 640–647
8. Petrenko, V. A., Smith, G. P., Phage display, *Chem. Rev.* **1997**, *97*, 381–410
9. Kolkman, J.A., Stemmer, W.P., Directed evolution of proteins by exon shuffling, *Nat. Biotechnol.* **2002**, *19*, 423–428

512   A. Pohorille and C. Chipot

10. Farinas, E.T., Bulter, T., Arnold, F.H., Directed enzyme evolution, *Curr. Opin. Biotechnol.* **2001**, *12*, 545–551
11. Keefe, A. D., Szostak, J. W., Functional proteins from a random-sequence library, *Nature* **2001**, *410*, 715–718
12. Shoichet, B., Bodian, D., Kuntz, I., Molecular docking using shape descriptors, *J. Comput. Chem.* **1992**, *13*, 380–397
13. Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R., Development and validation of a genetic algorithm for flexible docking, *J. Mol. Biol.* **1997**, *267*, 727–748
14. Claussen, H., Buning, C., Rarey, M., Lengauer, T., FLEXE: efficient molecular docking considering protein structure variations, *J. Mol. Biol.* **2001**, *308*, 377–395
15. Bowie, J.U., Luthy, R., Eisenberg, D., A method to identify protein sequences that fold into a known three-dimensional structure, *Science* **1991**, *253*, 164–170
16. Dunbrack, R. L., Sequence comparison and protein structure prediction, *Curr. Opin. Struct. Biol.* **2006**, *16*, 374–384
17. Street, A. G., Mayo, S. L., Computational protein design, *Structure* **1999**, *7*, R105–R109
18. Bonneau, R., Baker, D., Ab initio protein structure prediction: progress and prospects, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 173–189
19. Saven, J. G., Combinatorial protein design, *Curr. Opin. Struct. Biol.* **2002**, *12*, 453–458
20. Amadei, A., Linssen, A. B. M., Berendsen, H. J. C., Essential dynamics of proteins, *Proteins: Struct. Funct. Genet.* **1993**, *17*, 412–425
21. Amadei, A., Apol, M. E. F., Berendsen, H. J. C., The quasi-Gaussian entropy theory: free energy calculations based on the potential energy distribution function, *J. Chem. Phys.* **1996**, *104*, 1560–1574
22. Daidone, I., Amadei, A., Roccatano, D., Nola, A. D., Molecular dynamics simulation of protein folding by essential dynamics sampling: folding landscape of horse heart cytochrome c, *Biophys. J.* **2003**, *85*, 2865–2871
23. Chen, C., Xiao, Y., Zhang, L., A directed essential dynamics simulation of peptide folding, *Biophys. J.* **2005**, *88*, 3276–3285
24. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M., CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.* **1983**, *4*, 187–217
25. Go, N., Noguti, T., Nisikawa, T. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl Acad. Sci. USA* **1983**, *80*, 3696–3700
26. Levitt, M., Sander, C., Stern, P. S., Normal-mode dynamics of a protein: bovine pancreatic trypsin inhibitor, *Int. J. Quant. Chem: Quant. Biol. Symp.* **1983**, *10*, 181–199
27. Bahar, I., Atilgan, A. R., Erman, B., Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, *Fold. Des.* **1997**, *2*, 173–181
28. Hinsen, K., Analysis of domain motions by approximate normal mode calculations, *Proteins: Struct. Funct. Genet.* **1998**, *33*, 417–429
29. Schuyler, A. D., Chirikjian, G. S., Efficient determination of low-frequency normal modes of large protein structures by cluster-NMA, *J. Mol. Graph. Model.* **2005**, *24*, 46–58
30. Tembe, B. L., McCammon, J. A., Ligand–receptor interactions, *Comput. Chem.* **1984**, *8*, 281–283
31. Miyamoto, S., Kollman, P. A., What determines the strength of noncovalent association of ligands to proteins in aqueous solution? *Proc. Natl Acad. Sci. USA* **1993**, *90*, 8402–8406
32. Hénin, J., Maigret, B., Tarek, M., Escrieut, C., Fourmy, D., Chipot, C., Probing a model of a GPCR/ligand complex in an explicit membrane environment. The human cholecystokinin-1 receptor, *Biophys. J.* **2006**, *90*, 1232–1240

33. Wank, S. A., G protein-coupled receptors in gastrointestinal physiology. I. CCK receptors: an exemplary family, *Am. J. Physiol.* **1998**, *274*, G607–G613

34. Sanbonmatsu, K. Y., Simpson, J., Tung, C. S., Simulating movement of tRNA into the ribosome during decoding, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 15854–15859

35. Cieplak, P., Caldwell, J. W., Kollman, P. A., Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and *N*-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases, *J. Comput. Chem.* **2001**, *22*, 1048–1057

36. Ren, P., Ponder, J. W., Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations, *J. Comput. Chem.* **2002**, *23*, 1497–1506

37. Kaminski, G. A., Stern, H. A., Berne, B. J., Friesner, R. A., Cao, Y. X., Murphy, R. B., Zhou, R., Halgren, T. A., Development of a polarizable force field for proteins via ab initio quantum chemistry: first generation model and gas phase tests, *J. Comput. Chem.* **2002**, *23*, 1515–1531

38. Patel, S., Mackerell, A. D., Brooks, C. L., CHARMM fluctuating charge force field for proteins. II. Protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model, *J. Comput. Chem.* **2004**, *25*, 1504–1514

39. Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B., Mackerell, A. D., A polarizable model of water for molecular dynamics simulations of biomolecules, *Chem. Phys. Lett.* **2006**, *418*, 245–249

40. Gao, J., Truhlar, D. G., Quantum mechanical methods for enzyme kinetics, *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505

41. Friesner, R. A., Guallar, V., Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis, *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427

42. Car, R., Parrinello, M., Unified approach for molecular dynamics and density functional theory, *Phys. Rev. Lett.* **1985**, *55*, 2471–2474

43. Galli, G., Parrinello, M., Ab initio molecular dynamics: principles and practical implementation. In: *Computer Simulation in Material Science*, Meyer, M., Pontikis, V., Eds. Kluwer: Dordecht, 1991, p. 283

44. Marx, D., Hutter, J., Ab initio molecular dynamics: theory and implementation. In: *Modern Methods and Algorithms of Quantum Chemistry*, Grotendorst, J., Ed. John von Neumann Institute for Computing: Jülich, NIC Series, 2000, pp. 301–449

45. Jensen, M. O., Rothlisberger, U., Rovira, C., Hydroxide and proton migration in aquaporins, *Biophys. J.* **2005**, *89*, 1744–1759

46. Hwang, J.-K., Warshel, A., How important are quantum mechanical nuclear motions in enzyme catalysis?, *J. Am. Chem. Soc* **1996**, *118*, 11745–11751

47. Burykin, A., Warshel, A., What really prevents proton transport through aquaporin? Charge self-energy versus proton wire proposals, *Biophys. J.* **2003**, *85*, 3696–3706

48. Wu, Y., Voth, G. A., A computer simulation study of the hydrated proton in a synthetic proton channel, *Biophys. J.* **2003**, *85*, 864–875

Index