

## 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 Arie 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 relative solvation 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 subtilisin [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 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 Nubuhiko 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,  $\xi$ , 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  $\xi$ . Wilfred van Gunsteren [70] hypothesized that the thermodynamic force was equal to the constraint force acting along  $\xi$ . It, however, soon became apparent that this conjecture was incorrect whenever  $\xi$  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  $\xi$ . 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, Alessandro Laio and Michele Parrinello [77] introduced a metadynamics 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 reason, 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.  $\text{H}^+$ ,  $\text{Li}^+$ ,  $\text{Na}^+$  and  $\text{HO}^-$  [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,  $\Omega$ , 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  $\Omega$  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.,  $\Omega(\mathcal{E})$  gives the number of microstates with energy  $\mathcal{E}$ . 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:

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

Here,  $\delta$  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  $\mathcal{E}$ . 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  $\mathcal{E}$ , 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:

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

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  $\Omega(\mathcal{E})d\mathcal{E}$  measures the volume of microstates of energy  $\mathcal{E} \pm d\mathcal{E}/2$ . 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  $\Omega(\mathcal{E}_1)$  is twice  $\Omega(\mathcal{E}_2)$ , 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  $\mathcal{E}_1$  or  $\mathcal{E}_2$ . 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  $\Omega$  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,  $S = k_B \ln \Omega$ , 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:

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

$\beta = (k_B T)^{-1}$ . 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  $\delta$ -function in (1.2) as a convolution integral involving two  $\delta$ -functions of each energy term:

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

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

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

and

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

Here  $m$  is the mass of a particle and  $\Gamma$  is the  $\Gamma$  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  $\delta$ -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:

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

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  $\Omega$  or introduced into the reweighting of results. The use of  $\Omega_{\text{ex}}$  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,  $\Omega_{\text{con}}(N, V, \mathcal{E})$  remains proportional to the number of microstates with given  $N, V, \mathcal{E}$ .

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  $\mathcal{E}'' = \mathcal{E} - \mathcal{E}'$ , it follows that

$$\begin{aligned} Q(N, V, T) &= \left[ \int e^{-\beta \mathcal{E}'} \Omega_{\text{ig}}(N, V, \mathcal{E}') d\mathcal{E}' \right] \left[ \int e^{-\beta \mathcal{E}''} \Omega_{\text{ex}}(N, V, \mathcal{E}'') d\mathcal{E}'' \right] \\ &= \frac{V^N}{\Lambda(\beta)^{3N} N!} \int e^{-\beta \mathcal{E}} \Omega_{\text{ex}}(N, V, \mathcal{E}) d\mathcal{E} \\ &= \frac{1}{\Lambda(\beta)^{3N}} \int e^{-\beta \mathcal{E}} \Omega_{\text{con}}(N, V, \mathcal{E}) d\mathcal{E} \end{aligned} \quad (1.8)$$

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

$$\begin{aligned} Q(N, V, T) &= \frac{1}{N! h^{3N}} \int \exp[-\beta \mathcal{H}(\mathbf{q}, \mathbf{p}_x)] d\mathbf{q} d\mathbf{p}_x \\ &= \frac{1}{\Lambda^{3N} N!} Z(N, V, T) \end{aligned} \quad (1.9)$$

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:

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

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  $\Omega$ . 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):

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

where  $\theta$  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  $\mathcal{E}'$  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  $\Omega$ , and the weight each configuration should carry is:

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

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]:

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

where it is assumed that the initial energy,  $U_0$ , is less than  $\mathcal{E}$ . 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:

$$A = -\beta^{-1} \ln Q(N, V, T). \quad (1.14)$$

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*,  $\Delta A$ , 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:

$$\Delta A = -\beta^{-1} \ln Q_1/Q_0 \quad (1.15)$$

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  $Z_1$

$$\Delta A = -\beta^{-1} \ln Z_1/Z_0. \quad (1.16)$$

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,  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , 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:

$$Q_0 = \frac{1}{N!h^{3N}} \int_{\Gamma_0} \exp[-\beta \mathcal{H}(\mathbf{x}, \mathbf{p}_x)] \, d\mathbf{x} \, d\mathbf{p}_x \quad (1.17)$$

where  $\Gamma_0$  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  $\Gamma_0$  and  $\Gamma_1$  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  $\Delta A$  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:

$$\begin{aligned}
\Delta A &= -\beta^{-1} \ln \frac{\int \exp[-\beta U_0(\mathbf{x})] d\mathbf{x}}{\int \exp[-\beta U_1(\mathbf{x})] d\mathbf{x}} \\
&= -\beta^{-1} \ln \exp\{-\beta [U_1(\mathbf{x}) - U_0(\mathbf{x})]\} P_0(\mathbf{x}) \\
&= -\beta^{-1} \ln \langle \exp\{-\beta [U_1(\mathbf{x}) - U_0(\mathbf{x})]\} \rangle_0
\end{aligned} \tag{1.18}$$

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.  $\beta_0$  and  $\beta_1$  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:

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

An interesting feature of (1.18) is that  $\Delta A$  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  $\Delta A$  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:

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

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  $\lambda$  defined such that  $\lambda_0$  and  $\lambda_1$  correspond to systems 0 and 1, respectively. This parameter could be a macroscopic variable – viz. the temperature, a parameter that transforms  $\mathcal{H}_0$  to  $\mathcal{H}_1$ , 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:

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

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

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

This equation provides a prescription for calculating  $\Delta A$ . The probability distribution function,  $P(\lambda)$ , for the range of  $\lambda$  comprised between  $\lambda_0$  and  $\lambda_1$  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  $d\Delta A/d\lambda$  rather than  $\Delta A$  directly. Differentiating (1.14) yields:

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

If  $\lambda$  is a parameter in the Hamiltonian, we obtain:

$$\frac{dA}{d\lambda} = \frac{\int \frac{\partial \mathcal{H}}{\partial \lambda} \exp(-\beta \mathcal{H}) \, d\mathbf{x} \, d\mathbf{p}_x}{\int \exp(-\beta \mathcal{H}) \, d\mathbf{x} \, d\mathbf{p}_x} = \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_\lambda \quad (1.24)$$

and the free energy difference between system 0 and system 1 is evaluated by integrating the average derivative of the Hamiltonian with respect to  $\lambda$ , which is in units of the force, in the range extending from  $\lambda_0$  to  $\lambda_1$ . For this reason, the method is called thermodynamic integration. If  $\lambda$  is a function of the positions of the particles, derivation of the formula for  $dA/d\lambda$  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  $\lambda$  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,  $\Delta A$  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

$$\overline{f(\mathbf{q}_0, \mathbf{p}_0)} = \frac{1}{T} \int_0^T f[\mathbf{q}(\mathbf{t}), \mathbf{p}(\mathbf{t})] dt \quad (1.25)$$

Similarly, we can define the ensemble average:

$$\langle f \rangle = \int f P(\mathbf{x}, \mathbf{p}_x) d\mathbf{x} d\mathbf{p}_x, \quad (1.26)$$

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  $T \rightarrow \infty$ , such that the time average of  $f$  is equal to its ensemble average

$$\lim_{T \rightarrow \infty} \overline{f(\mathbf{q}_0, \mathbf{p}_0)} = \langle f \rangle. \quad (1.27)$$

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  $\Omega$ . In a stratified method, this domain is first partitioned into a number of disjoint regions  $\Omega_i$ , called strata, such that their union covers the whole domain. In the region  $\Omega_i$ ,  $x_i$  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  $\Omega$  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  $\lambda$ . We further assume that these two states are separated by a high-energy barrier that corresponds to a value of  $\lambda$  between  $\lambda_0$  and  $\lambda_1$ . 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  $\lambda$  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  $\Delta A$  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  $\lambda$ , as described above. In lieu of sampling from the true probability distribution,  $P(\lambda)$ , we design simulations in which  $\lambda$  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:

$$P'(\lambda) = P(\lambda) \exp[\beta\eta(\lambda)] \quad (1.28)$$

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

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

The fundamental issue in implementing importance sampling in simulations is the proper choice of the biased distribution, or, equivalently, the weighting factor,  $\eta$ . 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  $\rightarrow$  ethane, oxonium  $\rightarrow$  ammonium, glycine  $\rightarrow$  alanine, and alanine  $\rightarrow$  phenylalanine in aqueous solution and to  $H_3O^+(H_2O)_3 \rightarrow NH_4^+(H_2O)_3$  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 alcohols 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 numerical 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.,  $\lambda$ -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 of ions, 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.; Wrigger, W.; Schulten, K., Steered molecular dynamics, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Deuffhard, 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.; Samuelsson, 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 interactions 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

Free Energy Calculations

Theory and Applications in Chemistry and Biology

Chipot, C.; Pohorille, A. (Eds.)

2007, XVIII, 518 p., Hardcover

ISBN: 978-3-540-38447-2