Chapter 26

Best Practices in Free Energy Calculations for Drug Design

Michael R. Shirts

Abstract

Free energy calculations are increasingly of interest for computing biophysical properties of novel small molecules of interest in drug design, such as protein–ligand binding affinities and small molecule partition coefficients. However, these calculations are also notoriously difficult to implement correctly. In this article, we review standard methods for computing free energy differences via simulation, discuss current best practices, and examine potential pitfalls for computational researchers without extensive experience in such calculations. We include a variety of examples and tips for how to set up and conduct these calculations, including applications to relative binding affinities and absolute binding free energies.

Key words: Free energy calculation, Alchemical methods, Thermodynamic integration, Bennett acceptance ratio, MBAR, Drug design, Binding free energy

1. Introduction

The term *free energy calculation* describes a large family of simulation procedures to calculate the free energy difference between two thermodynamic states. Calculating the free energy difference between two states is extremely useful in simulations of biomolecular interactions in drug design. If we can calculate the free energy difference between two arbitrary molecular systems, we can determine small molecule transfer free energies and partition coefficients, and thus predict the concentration of the molecule in each phase. Perhaps the most relevant free difference for drug design is the free energy of binding of a small molecule to a receptor which can be directly related to the inhibition constant of the receptor.

Over the last decade, there has been a increasing enthusiasm in the potential for free energy calculations as a useful tool in drug design. First, methodological innovations make the calculations easier and more robust. Second, implementation of these methods

Riccardo Baron (ed.), *Computational Drug Discovery and Design*, Methods in Molecular Biology, vol. 819,
DOI 10.1007/978-1-61779-465-0_26,

makes them available to many users. Finally, steady increases in computer power enable application to a broader range of systems. These improvements bring these techniques close to providing reliable free energy estimates for biophysical systems. But free energy calculations are still difficult, and deciphering the proper techniques can be confusing. This methodological review is designed to help researchers who have some familiarity with molecular simulations and knowledge of statistical mechanics, but are looking for more guidance on performing free energy calculations.

Free energy calculations are among the most difficult types of biomolecular simulations to carry out for several reasons. Most simulation packages require extensive manual adjustments to input files to carry out free energy calculations. Calculation between two thermodynamic states can be extremely sensitive to choices of parameters unimportant for simulations of a single thermodynamic state. Additionally, a vast number of methodologies available can lead to a bewildering number of choices.

Standard computational methods for calculating free energies use molecular simulations to generate independent samples from the equilibrium distribution of the molecular system. Then, the information from these samples is analyzed using statistical tools to obtain an estimate of the free energy difference. Because of the statistical nature of this analysis, free energy calculations give *estimated* free energies, and repeating the calculation from different starting configurations or different random seeds will give different free energy estimates. To emphasize; free energy results are *not* exact results; they are statistical estimates obtained from sampling molecular probability distributions. Consequently, error analysis must always be performed to identify the statistical noise in the calculation, and no free energy calculation should ever be used or published without a statically robust uncertainty estimate.

Free energy calculations provide an estimate of the correct free energy difference between a thermodynamic process *given a particular set of parameters and physical assumptions*, not necessarily an estimate of the value that would be observed experimentally. The goal of good free energy methods is to converge to the unique free energy for that model. This is the "correct" free energy for the calculation. Only after this free energy is determined accurately can parameters of a model be improved, though substantial care must be taken to avoid overfitting. This review does not address finding or developing the best molecular model for a particular problem.

In this methods survey, we focus primarily on calculating binding free energies of ligands to proteins. A chapter entitled "An Introduction to Best Practices in Free Energy Calculations" in the book *Biomolecular Simulations: Methods and Protocols* in this same "Methods in Molecular Biology" series covers more

general aspects of free energy calculations for biophysical methods, and readers are encouraged to also review that chapter for an expanded discussion of many of the topics discussed here.

We first discuss what should and what should not be expected when performing free energy calculations of drug binding. We then cover basic theoretical principles behind free energy calculations of binding affinity. Finally, we outline the steps that must be performed for typical free energy calculations, including setup, running the simulation, and data analysis. We conclude with specific examples of these calculations for absolute binding free energies and relative ligand binding affinities.

1.1. What One Can Expect in Calculating Binding Free Energies of Drugs

At this time, it is not yet possible to accurately and reliably calculate free energies of protein-ligand binding using molecular simulation. Any such notion should be disabused up front. Tests of free energy calculations have thus far been insufficient to demonstrate if such calculations are truly predictive. This lack of testing has partly been because of the lack of good data sets, and partly because the computational expense to run such a comparisons has generally been too large. Nevertheless, there is almost universal agreement that the two most important factors are the inadequacy of current classical force field models to capture the biophysical properties of small molecules and proteins, and the difficulty of sampling all relevant configurations for macromolecular binding.

Full ligand binding free energy calculations should therefore not yet be seen as a useful screening technique because of the computational cost involved and the lack of validation of the underlying force fields. It can take hundreds of ns of simulation to compute a binding free energy with even moderate statistical convergence. The often difficult manual setup for most free energy calculations also creates a substantial barrier to performing high-throughput calculations.

However, there are a number of ways in which free energy calculations of binding affinities can still be useful. These calculations can have some predictive value if sufficient care is taken; for example, a recent study was able to perform accurate blind predictions to the apolar site of T4 lysozyme (1). More fundamentally, performing free energy calculations can allow physical identification of the molecular interactions contributing to binding in a dynamic macromolecular system with fluctuations, information that is very difficult to calculate in any other way. Full protein flexibility and explicit water molecules are naturally included in full statistical mechanical free energy calculations, so that one can discover locations of bound waters and correlations of ligand orientations and protein conformations.

Calculating free energies of binding through full statistical mechanical calculations provides a clear path for continual

improvement because of the fully physical approach. If the models are sufficiently physically accurate, and we perform sufficient sampling, the answers must be correct because of the underlying physics. Thus, the long-term prospects for free energy calculations as a predictive tool are significantly more encouraging than the current status may indicate.

1.2. Theoretical Principles

In this discussion, we assume standard classical molecular mechanics models, including harmonic bond and angle terms, periodic dihedral terms, and nonbonded terms consisting of point charges and Lennard-Jones repulsion/dispersion terms. We do not address free energies with polarizable models or with mixed QM/MM simulations (2), since these are not well developed enough to be of interest in most drug binding calculations yet.

All of the standard approaches for calculating free energies are variations of the same statistical sampling procedure. Samples are collected from simulations of thermodynamic ensembles and then analyzed to obtain a free energy difference. The main difference between approaches are in the types of energy data collected from simulation, and in the analysis performed with this data.

Free energy differences between states are directly related to the probabilities of those states. Specifically, the free energy difference is the log of the ratio of the partition coefficients of the thermodynamic states of interest. Rigorously, the free energy difference between two thermodynamic states in a constant volume ensemble is as follows:

$$\Delta A_{ij} = -k_{\mathrm{B}} T \ln \frac{Q_j}{Q_i} = -k_{\mathrm{B}} T \ln \frac{\int_{V_i} \mathrm{e}^{-\frac{U_i(\vec{q})}{k_{\mathrm{B}} T}} \mathrm{d}\vec{q}}{\int_{V_j} \mathrm{e}^{-\frac{U_i(\vec{q})}{k_{\mathrm{B}} T}} \mathrm{d}\vec{q}}, \qquad (1)$$

where ΔA_{ij} is the Helmholtz free energy difference between state j and state i, k_{B} the Boltzmann constant, Q the canonical partition function, T is the temperature in Kelvin, U_i and U_j are the potential energies as a function of the coordinates and momenta $\vec{q}$ for two states, and V_i and V_j are the *phase space volume* of $\vec{q}$ over which we sample. The phase space volume is the total set of coordinates and momenta in which the system has nonzero probability of being found. In this survey, we assume that this phase space volume is the same for both molecules, which is reasonable for most systems (see Note 1). For ease of notation, we also use $k_{\mathrm{B}} T = \beta$ in this article (see Note 2). We also assume that the masses of the particles do not change, and we use the potential energy U instead of the more general Hamiltonian H for clarity.

From this basic definition, we note that we are always calculating free energy *differences*, not absolute free energies. All of the quantities that are of interest in biophysical measurements are free energy differences between two thermodynamic states, so we

must always specify two states. Even "absolute" free energies of drug binding are still free energy differences between two states, specifically (1) a ligand in the binding site and (2) a ligand and a host separated from each other.

We can easily modify the above discussion to deal with the Gibbs free energy (G) instead. If we replace U_i and U_j with $U_i + P_iV$ and $U_j + P_jV$ respectively, and integrate over all system volumes V (not to be confused with the phase space volumes) in addition to integrating over the coordinates $\vec{q}$, then we will get the Gibbs free energy G instead of A and the isobaric-isothermal partition function Ξ instead of Q. All the derivations presented in this review can be extended directly with this substitution. At constant pressure, the change in free energy related to changes in average volume will be small at physiological pressures. This is only an approximation, as it ignores fluctuations, but illustrates that we can generally neglect the PV component to the free energy and perform calculations at NVT if we are careful to make sure that the simulation is actually at the average volume for the state (see Note 3). To make clear our discussions of the NVT case, we use the Helmholtz free energy difference ΔA. Again, any simulation method that includes proper isobaric–isothermal sampling of volumes can simply insert $U + PV$ in place of U, where P is the applied (not instantaneous) pressure and all the subsequent derivations will hold.

1.3. Simulation Methods Useful for Calculating Free Energy Differences

In this section, we discuss the need for having a pathway of intermediates connecting two states, and review the most common and/or useful method for computing free energy, the Zwanzig relationship, thermodynamic integration, the Bennett Acceptance ratio (BAR), the weighted histogram method (WHAM), and the multistate ( ). We use the term "alchemical transformation" to using these methods to compute the difference of a process that changes the chemical identity of our molecule (see Note 4).

1.4. Multiple Intermediates

In most instances where the states of interest have very little phase space overlap, the transformation can be broken into a series of intermediate states that do have good phase space overlap. By good *phase space overlap* between two states, we mean that the number of configurations that both states visit is some moderate percentage of each state's total phase space. Without good phase space overlap, it is impossible to compute the free energy differences between two states.

Consider $K - 1$ free energy calculations spanning a series of K states that *do* have phase space overlap, where $k = 1$ and $k = K$ are our states of interest. Mathematically it is as follows:

$$\Delta A_{1,K} = \sum_{i=1}^{K-1} \Delta A_{i,i+1}.$$

A separate free energy calculation is then performed for each of the individual ΔA's, simulating the two neighboring states. Since we care specifically about the free energies of only the end states, we do not care about precise form of the intermediates. This leaves us free to choose intermediate states that have high phase space overlap with one another, which means we can choose completely unphysical states if they lead to less overall error. Statistical uncertainty is a very steep function of the amount of phase space overlap, so the total uncertainty decreases quickly as a function of the number of intermediates. Common examples of nonphysical intermediates include atoms without charges, an atom with van der Waals parameters that are part-way between a carbon and a nitrogen, or a "softened" atomic site that solvent molecules can penetrate.

It is both useful conceptually and mathematically convenient to think of these intermediate states as lying along a pathway connecting the initial and final states. The parameterized distance along this path connecting the initial and final states is traditionally called λ, with $\lambda = 0$ corresponding to the initial state and $\lambda = 1$ corresponding to the final state. Since these states are often unphysical, we call them alchemical states. We can then think of the potential describing the system as a function of both λ and $\vec{q}$, writing this as $U(\lambda, \vec{q})$. We must then perform simulations of $U(\lambda, \vec{q})$ at a series of λ values, generating samples that will allow us to estimate each of the $\Delta A_{i,i+1}$ free energy differences.

1.5. Zwanzig Relationship

The most historically well-known method for calculating free energy differences from simulations is the Zwanzig relationship (3). This method is sometimes called *free energy perturbation* or *exponential averaging*. We refer to this method as EXP, for exponential averaging. The free energy between two potentials $U_0(\vec{q})$ and $U_1(\vec{q})$ over a coordinate and momentum space $\vec{q}$ can be calculated as:

$$\Delta A = \beta^{-1} \ln \left\langle e^{-\beta(U_1(\vec{q}) - U_0(\vec{q}))} \right\rangle_0 = \beta^{-1} \ln \left\langle e^{-\beta \Delta U(\vec{q})} \right\rangle_0. \quad (2)$$

Although the equation is exact for standard molecular models, except in the case of rather small changes, EXP converges very poorly as a function of the number of samples collected. Free energy differences that appear to have converged may only indicate very poor phase space overlap between the two states (4, 5). Except for very specific cases, where the difference between potential energy distributions is known to always be very small for all $\vec{q}$, on the order of $(1 - 2kT)$, EXP should generally *not* be used. There are some cases where all potential energy differences are known to be small; some of these cases are discussed in previously mentioned chapter of *Biomolecular simulations: methods and protocols* in the "Methods in Molecular Biology" series.

If a sufficiently large number of intermediate states are used, then EXP can give correct results, but it is usually significantly less efficient than other methods.

1.6. Thermodynamic Integration

By taking the derivative of the free energy with respect to the variable λ describing the distance along the series of intermediate alchemical states, we find the following:

$$\begin{aligned} \mathrm{d}A/\mathrm{d}\lambda &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\int \mathrm{e}^{-\beta U(\lambda,\vec{q})}\mathrm{d}\vec{q} = \left\langle \frac{\mathrm{d}U(\lambda,\vec{q})}{\mathrm{d}\lambda}\right\rangle_{\lambda} \\ \Delta A &= \int_0^1 \left\langle \frac{\mathrm{d}U(\lambda,\vec{q})}{\mathrm{d}\lambda}\right\rangle_{\lambda} \mathrm{d}\lambda. \end{aligned} \tag{3}$$

Computing free energies using this formula is called *thermodynamic integration*, abbreviated as TI in this chapter, and is often done using numerical integration. Since we can only simulate a limited number of intermediates, we must use some type of numerical integration of the integral. By definition, numerical integration introduces bias, which must be minimized sufficiently that it is well beneath the level of statistical noise.

Various numerical integration schemes are possible, but the trapezoid rule provides a simple, flexible, and robust scheme. All types of numerical integration can be written as follows:

$$\Delta A \approx \sum_{k=1}^{K} w_k \left\langle \frac{\mathrm{d}U(\lambda,\vec{q})}{\mathrm{d}\lambda}\right\rangle_k,$$

where the weights w_k correspond to a particular choice of numerical integration. Researchers have tried a large number of different integration schemes (6–8). Many other integration routines require specific choices of λ to minimize bias, which makes them unsuitable when the intermediates have widely varying levels of uncertainty. For starting researchers, we, therefore, recommend a simple trapezoidal rule scheme, as it allows for maximal flexibility in which values of λ are simulated. (see Notes 5 and 6).

TI can be extremely simple to apply for some paths, but most paths require derivatives with respect to λ to be calculated in the code itself. If the pathway is chosen such that $U(\lambda,\vec{q}) = (1-\lambda)U_0(\vec{q}) + \lambda U_1(\vec{q})$, then $\frac{\mathrm{d}U}{\mathrm{d}\lambda} = U_1(\vec{q}) - U_0(\vec{q})$, which can be easily calculated in post-processing by evaluating the same configuration at the initial and final states. If the pathway is not linear in the potential, then the derivative must be calculated analytically in the code. Unfortunately, most problems of interest require using pathways that are not linear, as we discuss later. However, if the code does compute $\frac{\mathrm{d}U}{\mathrm{d}\lambda}$, then TI is perhaps the simplest method to use, as it involves a very little postprocessing, and the analysis requires only simple averages and sums. As long as care is taken to make sure that enough intermediates are used to reduce bias in

the integration well below the statistical noise, then TI gives very robust energy results. The more curvature $\left\langle \frac{dU}{d\lambda} \right\rangle$ has, the more intermediates will be required.

1.7. Bennett Acceptance Ratio

Measurements of potential energy differences can be used in a statistically optimal way to compute free energy differences. The difference between the potential energy of the same configuration $\vec{q}$ for two different states along the pathway is $\Delta U_{ij}(\vec{q})$. There is a very robust, statistically optimal way to use this potential energy differences collected from both states i and j together to obtain an improved estimate of the free energy difference between two states. Bennett's original derivation started with a simple relationship for the free energies:

$$\Delta A_{ij} = -\ln kT \frac{Q_j}{Q_i} = kT \ln \frac{\left\langle \alpha(\vec{q}) \exp[-\beta \Delta U_{ij}(\vec{q})] \right\rangle_1}{\left\langle \alpha(\vec{q}) \exp[-\beta \Delta U_{ji}(\vec{q})] \right\rangle_0}, \quad (4)$$

which is true for any function $\alpha(\vec{q})>0$ for all $\vec{q}$. Bennett then used variational calculus to find the choice of $\alpha(\vec{q})$ that minimizes the variance of the free energy (9), resulting in an implicit function of ΔA easily solvable numerically:

$$\sum_{i=1}^{n_i} \frac{1}{1+\exp(\ln(n_i/n_j) + \beta \Delta U_{ij} - \beta \Delta A))} - \sum_{i=1}^{n_j} \frac{1}{1+\exp(\ln(n_j/n_i) - \beta \Delta U_{ji} + \beta \Delta A))} = 0, \quad (5)$$

where n_i and n_j are the number of samples from each state. A separate derivation shows that the same formula provides a maximum likelihood estimate of the free energy given the samples from the two states (10). Both derivations give the same robust estimate for the variance and uncertainty of the free energy. Studies have demonstrated both the theoretical and practical superiority of BAR over EXP in molecular simulations (4, 5), and EXP can be shown to converge to EXP in the limit that all samples are from a single state (9, 10). Significantly less overlap between the configurational space of each state is required to converge results than in the case of EXP, though some overlap must still exist. Many simulation packages have tools to compute the BAR estimator automatically, so it usually does not need to be implemented.

It is difficult to compare TI and BAR on a theoretical basis because the two approaches use different information. However, practical experience indicates BAR generally performs more efficiently. More precisely, given a amount of simulation, fewer intermediate states are required for BAR than for TI to give equivalent level of statistical precision. TI can be as efficient as BAR under conditions where the integrand is very smooth (4, 11), such as charging or small changes in bonded or nonbonded parameters.

In other cases, such as the pathways required to remove large numbers of atomic sites, BAR is much more efficient than TI or EXP for free energies of larger molecular changes (4, 5, 12). Additionally, no analytical computation of $du/d\lambda$ is required to use BAR.

1.8. Weighted Histogram Analysis Method

The *WHAM* provides a way to use information from all of the intermediate λ values in computing free energy differences between states. Most free energy calculations require simulation at a number of different intermediates, and we would prefer to use as much thermodynamic information as possible from all of these simulations simultaneously to save computational cycles. Histogram weighting techniques were first introduced by Ferrenberg and Swendsen (13) to capture all of the thermodynamic information from all sampled states in computations of free energies and other observables. WHAM is a histogram reweighting technique introduced in 1992 by Kumar and collaborators for alchemical simulations (14). WHAM is the lowest uncertainty method for calculating free energies using samples collected from discrete states. However, it introduces biases for continuous distributions, such the energies of atomistic simulations, because all variables must be discretized into bins. Other variations of WHAM based on maximum likelihood (15) and Bayesian methods (16) have also been developed. Beginners should generally not write their own WHAM implementation, because solving the nonlinear equations correctly can be very challenging. The CHARMM molecular mechanics package includes WHAM-based free energy calculations (17, 18), and several other stand-alone WHAM implementations are available, so new development of tools is not necessary, other than for pedagogical reasons.

One can reduce the WHAM equations to a simpler form by shrinking the width of the histograms to zero (14, 17), yielding a set of iterative equations that estimate the free energies from K states simultaneously.

$$A_i = -\beta^{-1} \ln \sum_{k=1}^{K} \sum_{n=1}^{N_k} \frac{\exp[-\beta U_i(\vec{q}_{kn})]}{\sum_{k\prime=1}^{K} N_{k\prime} \exp[\beta A_{k\prime} - \beta U_{k\prime}(\vec{q}_{kn})]}, \qquad (6)$$

where i runs from 1 to K, the A_i are the free energies of each state, $\vec{q}_{kn}$ is the nth sample from the kth state, and the U_i are the potentials of these K states. Although this looks like a formula for absolute free energies, not a formula for free energy differences, the equations are only unique up to an additive constant, so we must fix one of the free energies as a reference states. We are then effectively calculating free energy differences from that reference state. The derivation of this approximation is somewhat suspicious for finite numbers of samples, as the derivation involves

finding the weighting factors that minimize the variance in the occupancy of the bins, which becomes undefined as the bind width and therefore the average number of samples per bin goes to zero.

1.9. Multistate Bennett Acceptance Ratio

A multistate extension of BAR called the multistate Bennett's Acceptance Ratio, or MBAR (19) was recently introduced which overcomes the binning issues with WHAM. In this approach, a series of $K \times K$ weighting functions $\alpha_{ij}(\vec{q})$ are derived to minimize the uncertainties in free energy differences between all K states considered simultaneously. The lowest variance estimator is exactly the WHAM equation in the limit of zero-width histograms (6). WHAM can be therefore be interpreted as a histogram-based approximation to MBAR. This MBAR derivation additionally gives the statistical uncertainty of the calculated free energies, which is not available in WHAM. MBAR has no histogram bias and is guaranteed to have lower bias and variance than WHAM. However, in many cases, the bins are small enough so that the difference in free energies between the two methods is negligible compared to the statistical precision required. If WHAM is implemented directly in the code, it may not be worth the additional gain to switch to MBAR, as the statistical uncertainty can be obtained by alternate methods that we describe below. MBAR is still not standard in molecular simulation, but a MBAR implementation can be downloaded at https://simtk.org/home/pymbar.

1.10. Nonequilibrium Methods

Nonequilibrium simulations can also be used to compute free energies. In a physical or alchemical process where thermodynamic variables change over some interval of time, some amount of work W required to make this change. If this is done infinitely slowly, the process is reversible, and W will be the free energy difference between the end states. However, if the change is performed in a finite amount of time, this process will not be reversible and hence the work will not equal to the free energy. Jarzynski noticed that the free energy of the transformation can be written as the average of the nonequilibrium trajectories that started from an equilibrium ensemble.

$$\Delta G = \beta^{-1} \ln \left\langle e^{-\beta W} \right\rangle_0. \tag{7}$$

If the switching is instantaneous, then (7) is identical to EXP because the instantaneous work is simply the change in potential energy Δ_{ij}. A version of BAR (though not MBAR) can be constructed with the nonequilibrium work (10, 20).

Several studies have compared nonequilibrium pathways to the equilibrium pathways (21, 22). It appears that under most circumstances, equilibrium simulations are about the same or slightly more efficient than free energies calculated from ensembles

of nonequilibrium simulations. It is thus not yet clear the extent to which free energy calculations using Jarzynski's relationship will be useful in ligand binding calculations in the future, because of the extra complications of running many separate trials. This is an area of intense research, partly because this formalism has proven useful in treating nonequilibrium experiments as well as simulations, and partly because there are still some tantalizing possibilities for substantially increasing the efficiency in free energy calculations with such simulations. However, we do not recommend that beginners use these methods, as they add an extra degree of complication to both the simulation and the analysis. Further information on how to implement such calculations can be found in other reviews (23).

2. Methods

2.1. Outline of Free Energy Calculations

Fundamentally, calculating a free energy requires a molecular simulation package that generates samples from the equilibrium distribution of the states of interest, as well as from any intermediate states that might be required, and extracts basic energetic information from those states.

Several key ingredients in a simulation package can help make free energy calculations much more convenient. The key features of a code that makes it easy to calculate free energies efficiently are (1) the ability to simulate nonphysical intermediate states along a low variance pathway (2) automatic and computationally efficient calculation of the required energetic information (either ΔU_{ij} or $\frac{\mathrm{d}U}{\mathrm{d}\lambda}$) and (3) some degree of automation of the analysis of this information. Many types of free energy calculations can be performed with any molecular mechanics or Monte Carlo code, though calculations sufficiently efficient to study large and complicated systems require code specifically set up to support good free energy practices.

In what follows, we discuss how to conduct free energy calculations, striving to avoid specifics about particular codes and tools. It is impossible to give full descriptions of proper steps for all simulation packages, as the free energy capabilities of virtually all simulation packages are evolving rapidly. Most of the common packages used for biological simulation (AMBER, CHARMM, GROMACS, GROMOS, DL-POLY, LAMMPS, Desmond) have at least one of the standard free energy functionalities built in, but certainly not all programs have all free energy functionalities. With collaborators, have developed a Web site, http://www.alchemistry.org, which is intended to provides more in-depth information on with code specific instructions, points too detailed to include in this present format, and example files and results.

All methods for computing free energies differences presented in this review consist of the following steps:

1. Construct a thermodynamic cycle that allows easy calculation of the free energy of interest, and determine the end states for each calculation required by the thermodynamic cycle.
2. Choose a sequence of intermediate states connecting the two end states for each free energy calculation.
3. Perform equilibrium simulations of the states of interest and any required intermediate states to collect uncorrelated, independent samples.
4. Extract the information of interest required for the desired free energy method from the sampled configurations.
5. Analyze the information from the simulations to obtain a statistical estimate for the free energy, including an estimate of statistical error.

2.1.1. Construct Thermodynamic Cycles and Choose End States

The free energy is a state function, and any series of transformations connecting the two end points gives the correct free energy. In many cases, it will be significantly more effective to use less direct paths that are more efficient. Whenever performing a free energy calculation, it is important to construct the appropriate thermodynamics cycle to clearly visualize the transformation being performed.

For example, perform relative binding free energies can be simply understood by drawing the appropriate thermodynamics cycle. Relative free energies can be computed by performing two separate calculations of free energies of binding for two different molecules, and subtracting them (see Figure 1). However, free energies of binding can require extremely long simulation times, because they require removing the entire ligand from the environment of the protein and the solvent, a process that can require either prohibitively large amounts of computer power, or fairly involved constraining methods to improve convergence. However, as can be seen in Figure 1, this same free energy difference can be written as the difference of two different nonphysical processes, the changing of molecule A to molecule B while bound, and the changing of molecule A to B while unbound:

$$\Delta A_{bind} = \Delta A^{B}_{bind} - \Delta A^{A}_{bind} = \Delta A^{bound}_{A\to B} - \Delta A^{unbound}_{A\to B}.$$

Since the unbound protein is the same in both cases, no simulation needs to be performed of the unbound protein.

The next step is to determine which simulations correspond to the end states of the free energies differences of interest. This must be done carefully. For example, for computing a solvation free energy of a small molecule solute, the initial state the

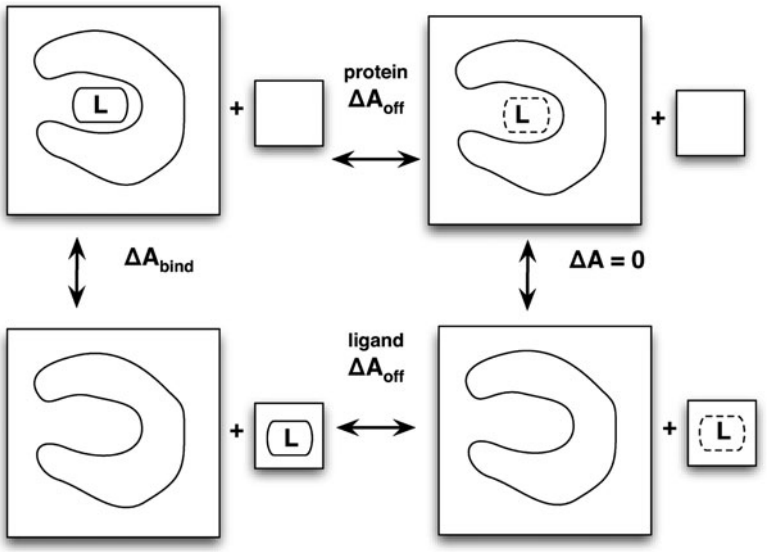

Fig. 1. The thermodynamic cycle for the relative binding affinities of ligand *A* and *B* to a host molecule.

calculation of the solvation consists of the solute and some quantity of solvent in a specified volume. The final state consists of the same small molecule solute in vacuum in the same volume as in the initial state, and plus the same number of solvent molecules as the initial state, also in the same volume as the initial state. A typical beginner's error is to use a final end state with all energetic terms of the solute turned off, which is not correct; in the vapor phase, the intramolecular interactions of the solute should remain turned on. Only the *intermolecular* interactions should be turned off.

2.1.2. Choose a Series of Intermediate States

If the end states of the transformation of interest do not have significant overlap in phase space, a series of intermediate states is required. The judicious choice of these intermediates is one of the most complicated aspects of free energy calculations.

It is important to clarify some of the terms used in free energy calculations. When performing *equilibrium* simulations of intermediate states along a pathway, any distinction between "forward" and "backward" is arbitrary. If one state contains an atom in state A that is not present in state B, then interpreting A as the initial state and B as the final state means that this atom is disappearing or being destroyed or annhiliated, whereas treating state B as the initial state means that the same atom is being created or introduced into the system. The choice of words to describe this change is entirely semantic. We generally refer to either of these changes as *decoupling*, where only intermolecular interactions are turned off, or *annihilation*, which refers to turning off all interactions with the system, both intermolecular and intramolecular, rather than creation or coupling.

The simplest choice for most transformations between two potential functions U_0 and U_1 is the linear pathway. For example:

$$U(\lambda, \vec{q}) = (1 - \lambda) U_0(\vec{q}) + \lambda U_1(\vec{q}) + U_{unaffected}(\vec{q}), \qquad (8)$$

where $U_{unaffected}(\vec{q})$ is the potential due to interactions which do not change as a function of intermediate state. For annihilation, it will be the solvent–solvent interactions; for decoupling, it will be the solvent–solvent *and* solute–solute interactions.

A significant problem with this approach is that equal spacing in λ does not actually lead to equal spacing in phase space overlap. If a Lennard-Jones function is used to for atomic exclusion and dispersion interactions, as is typical for biomolecular interactions, then when $\lambda = 0.1$, nearly at one end state, the excluded volume for a OPLS-AA united methane sphere (i.e., the volume with energy above 2–3 $k_B T$) will still be 60–70% of the original volume.

More severely, this choice of parameterization with a r^{-12} leads to a singularity in $\langle dU/d\lambda \rangle$ at $r = 0$, which then cannot be integrated numerically. Some studies try to approximate this difference by extrapolation, but this is extremely unreliable and error prone. Therefore, a linear pathway in energy should not be used to annihilate or decouple atoms (see Note 7).

Fortunately, there are now standard ways to handle the decoupling of atomic sites in an efficient way, the "soft core potential" approach (24, 25). In this approach, the infinity at $r = 0$ of the r^{-12} interaction is "smoothed out" in a λ dependent way. The most common form of the pairwise potential is:

$$H(\lambda, r) = 4\epsilon\lambda^n \left[\left(\alpha(1-\lambda)^m + \left(\frac{r}{\sigma}\right)^6 \right)^{-2} + \left(\alpha(1-\lambda)^m + \left(\frac{r}{\sigma}\right)^6 \right)^{-1} \right], \qquad (9)$$

where ε and σ are the standard Lennard-Jones parameters, α is a constant (usually 0.5), with the original choice of $n = 4$ and $m = 2$ (24). Further research has shown that using $n = 1$ and $m = 1$, with α fixed at 0.5, noticeably improves the variance (26–28).

To turn off intermolecular interactions between a molecule and its surroundings requires decoupling both the charge and the Lennard-Jones interactions. One highly reliable, relatively high efficiency pathway for annihilation or decoupling of atoms is to turn off the charges of these atoms linearly and then afterwards turn off the Lennard-Jones terms of the uncharged particles using the soft core approach. The same pathway can be followed in reverse for atomic sites that are introduced (18, 27). This ensures that when the repulsive cores with infinite positive energy at $r = 0$ are eliminated, there are no negative infinities energies at $r = 0$ due to Coulombic attraction between unlike charges.

Another similar approach is to turn off both the Coulombic and the dispersion term first, and then in a separate step turn off the repulsive term. There appears to be little difference in efficiency between these two approaches; both work well. It is possible to turn off both the Coulombic term and the van der Waals term at the same time using soft core potentials with both, (24, 29), but it can be difficult to choose parameters for these approaches that are transferable between systems. We highly recommend using the soft core for only the van der Waals interactions, after charges have been turned off separately.

Constructing alchemical pathways between two molecular end states involves one of two main approaches. These are the *single topology* approach and the *dual topology* approach (see Note 2). In the single topology approach (a, upper) a single topology has sites that correspond to atoms in both molecules. At one end state, two hydrogens are turned into "dummies" that have no nonbonded interactions with the rest of the system, and the upper heavy atom is an oxygen, while at the other end state, all atoms present, and the upper heavy atom is now a carbon. The alternative dual topology approach differs in that no atoms change their type; they merely change back and forth from being dummies to being fully interacting particles (b, lower). In this case, at the ethanol end state, the methyl group is noninteracting, while in the ethane end state, the hydroxyl group is noninteracting.

One advantage to dual topology approach is that the moieties which change are free to sample the configurational space while decoupled. This can help increase the sampling if the simulations at different intermediates are coupled in a way that allow exchanges, such as expanded ensemble or Hamiltonian exchange simulations. However, for a dual topology approach, more atoms or molecules must be annihilated or decoupled from the environment, which will require more intermediates. In many cases, the convergence time may be the limiting factor, so a dual topology approach can be more efficient.

Dummy atoms can in principle affect free energies, but handled correctly, their effects can often be neglected. Although the end states shown in Fig. 2 have the correct nonbonded interactions for both ethane and ethanol, they are clearly different molecular objects, as they have nonphysical dummies bonded to the carbon or oxygen. Because these dummy atoms affect the system, we need to account for their free energy contributions. The easiest solution is to perform the transformation in both vacuum and in the molecular surroundings. In the rigid-rotor approximation, where all bonds are fixed in length, the effect on the free energy of these nonphysical dummies cancel out (30). If the bonds are not constrained, then there will be slight differences, but they appear to be small enough (less than 0.01 kcal mol^{-1}) to be neglected in any problem of real interest.

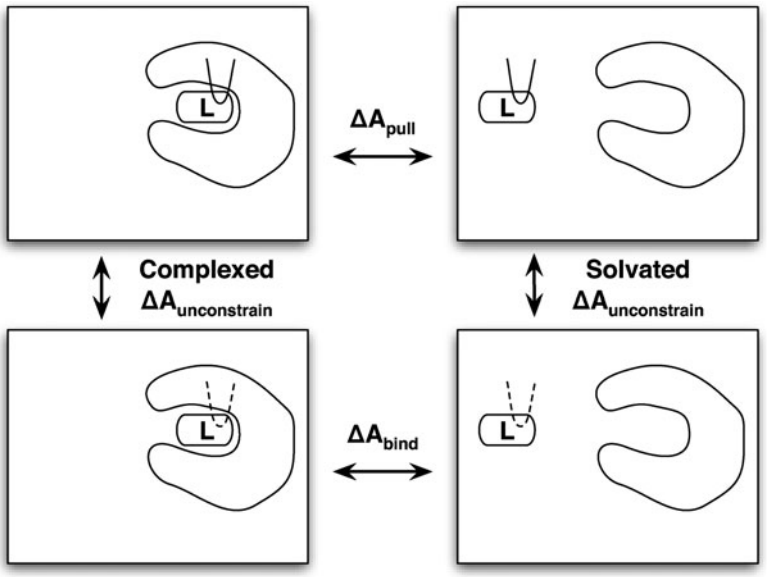

Fig. 2. Single topology (**a**, *upper*) and dual topology (**b**,*lower*) approaches to constructing an alchemical path between ethane and ethanol. D represents noninteracting dummies, while M represents nonphysical intermediate atoms. In a dual topology approach, no atoms change type, only have their interactions turned off from the rest of the system; however, more atoms need to be altered to go from the initial to the final state.

In many cases, we need to modify the bonded interactions of the molecule. This can be handled in a straightforward way. For example, in the single topology transformation of ethane to ethanol, the angle and dihedral terms involving the changing heavy atom are clearly different in the two end states. We must change these bonded interactions in addition to the nonbonded interactions. The variance due to changes in the bonding terms is not generally a problem; although the energy changes for these terms can be quite large, the time scale of the motions means that they converge quite quickly. Pathways that are linear in the bonded parameters (such as harmonic spring constants and equilibrium bond lengths or angles) are perfectly adequate. However, care must also be taken for constrained bonds. There is no phase space overlap between bonds constrained to two different lengths, and so an approach that only constrains hydrogen bonds is be much preferred to avoid correction terms that can be difficult to compute (31).

The choice of single versus dual topology will depend on the simulation code used – individual simulation packages may only support one or the other. Both cases will lead to correct final results. Notice that in neither case did we give an example with opening or closing rings; Both require removing bonds, which is problematic; it is much better to appear or disappear rings entirely, even if they are large. We, therefore, recommend never breaking rings in calculations.

2.1.3. Pulling Methods

A completely different choice of pathway for the free energy of protein ligand association is to physically pull the molecule away from the protein. If the final state is sufficiently far from the original protein, the free energy of this process is the free energy of binding. This can be done either by nonequilibrium simulations, using the Jarzynski equation as discussed earlier (32), or by computing a PMF using umbrella sampling with different overlapping harmonic oscillators at specified distances from the binding site (33–35)

There are a number of complications with pulling methods. Pulling a ligand out of a completely buried site can have high statistical error because of the lack of an direct pathway, and it can be difficult to pull the ligand sufficiently far away from the protein with a simulation box of tractable size. Additionally, a pulling pathway must be chosen. In the case of reasonable box sizes, some analytical or mean-field approximation must be applied for the free energy of pulling the ligand to infinity, and there has not been extensive research on the reliability of such corrections. Some researchers have argued that pulling may be more efficient for highly charged ligands (34). However, because of the difficulty of choice of pathway, pulling pathways are not recommended for beginners.

2.1.4. Rules of Thumb for Constructing Intermediate States

There are a number of other small points that are worth taking into account when deciding on a series of intermediates states, not all of which can be fully described in limited space, but we list as many as possible here, as well as summaries of the discussions above that require emphasis.

- Bonded terms, such as angle or bond force constants can be changed or turned off linearly. Changes in bond distances, if they are not constrained, can also be performed linearly.
- Constrained bonds should not generally change length, as there are free energy terms associated with these changes that cannot be neglected (36).
- Choose a pathway that maximizes the similarity between two states. Remove or decouple fewer atoms when possible.
- Do not open or close rings. There are some fundamental theoretical problems with changing the number of degrees of freedom in changing thermodynamic states. It is much better to make entire rings disappear and appear, even if it involves more atoms changing.
- Given a fixed number of intermediate states, the states should be chosen such that the statistical uncertainty of the free energy difference between any neighboring pair of states is equal. This is not simply an empirical rule of thumb; mathematically, it will lower the overall variance (37).

- *Changes* in parameters can be calculated using a simple linear function. Introduction or deletion of atoms should always be done with a "soft-core" type potential.
- Charges on any atoms to be created or annihilated should be completely off before the atomic repulsive terms are turned off. Otherwise, the simulation will rapidly crash as charges of opposite charge will approach to zero distance, crashing the simulation.
- The variance shrinks very quickly as a function of state spacing. Until the free energy differences between intermediates are lowered to approximately kT, and if sufficient CPU's are available, it is better to use more states than fewer states. If limited by the number of CPU's available, fewer states can be used, but it may end up being less statistically efficient in the end, more uncorrelated states will be required from each simulation.
- For a given scheme, the shape of the variance curve as a function of λ does not change significantly with the number of atoms (38). This means that if the same alchemical pathway is used for two different molecules, then both molecules will require tighter spacing of lambda in the same places, though of course more total intermediates will be required for a larger molecule.
- Quickly prototyping possible intermediate states with short simulations is highly recommended. The rough magnitude of variance of free energy differences can be estimated with very short simulations, frequently as quickly as 100 ps. Occasionally, simulations may get stuck in metastable states, and the true variance when the simulation is allowed to escape from such states may be larger than that observed in a short simulation.
- The total charge of the simulation should be maintained across all values of λ. Free energy calculations with charged molecules are fine, as long as the total charge of the system remains the same. Most methods for computing long-range electrostatics make approximations, such as a uniform neutralizing charge, which are reasonable if the total charge of the system remains the same. However, when the overall charge of the system changes as a function of λ, these approximations can lead to significant differences in the overall free energy. Simulations with changing charges will still give useful qualitative information, but the extent of the errors are not known, and they cannot be considered quantitatively reliable in most cases (39).

2.1.5. Perform Simulations of the States of Interest

The heart of the free energy calculation is conducting *equilibrium* simulations of *the states of interest* and any required intermediate states to collect *uncorrelated, independent* samples. There are several important topics to cover to ensure reliable, repeatable results.

- The simulations must be at equilibrium. Even for nonequilibrium work simulations, the initial states must be in equilibrium. Sufficient time must be given for the system to reach equilibrium before samples are collected. Because many free energy methods effectively give large weight to rare events, a small amount of unequilibrated data can have an outsized contribution to the overall free energies.
- The system must reach equilibrium *at each value* of λ. One efficient way to start each system is to run a series of short (10–100 ps) simulations at each λ state, restarting the next state from the final state of the previous simulation. This gives the system time to partly relax to the new intermediate state's potential and avoid instabilities in simulations. Changes in the volume occupied by the changing molecule or molecules can affect the total energy. As λ changes, the pressure should be allowed to adjust as well so that the solvent density of the system does not change as the effective volume of the molecule changes. Small changes in V can cause problems, not because the PV term becomes significant in calculating the free energy (see Note 3) but because liquids are nearly incompressible, and a small change in average volume leads to a large change in thermodynamic properties. To obtain the most consistent results, if the final simulations at each λ are run at NVT, they should use the average volume of the system, as different fluctuations in the box volume can lead differences of 0.1–0.3 kcal mol^{-1} in the final free energy. However, it can take 100's ps or several ns, or even longer in some cases (40) to relax to an equilibrium distribution in the new intermediate state. Significant simulation time should be allowed for this relaxation to occur. The required time varies drastically from system to system, and no hard and fast rule can be given. For solvation of smaller molecules, it may take only 100–500 ps, but for systems that are started out of equilibrium and have long correlation times, it could be hundreds of ns. The average energy of the simulation, $\left\langle \frac{dU}{d\lambda} \right\rangle$, as well as structural observables, must be carefully monitored for convergence. The number of hydrogen bonds to a small molecule is one useful observable to watch for convergence of a simulation because it can exhibit relatively slow equilibration rates (41).
- The samples must be collected at the state of interest. In all simulation codes, different choices of simulation parameters can result in changes in the potential energy surface. If such a change move the entire potential energy surface up by a constant amount, or affect the relative depths of wells by less than a few tenths of kT, then simulations at a given intermediate may appear to be unaffected. However, if these choices result in changes to the potential surface as a function of λ, it

can lead to significant modifications of the free energy of the end states of interest.

- To give just one example, for simulations done with the standard particle mesh Ewald (PME) treatment of long-range electrostatics, PME parameters that are sufficient for "standard" MD can give significant errors in the free energy for modifying partial charges on a molecule, up to 4 kcal mol^{-1} for some small molecules. So, when doing free energy calculations, it is in general not a good idea to assume that particular settings are not important. If the potential could possibly be affected, the dependence on this parameter should be checked.
- The samples must be *independent*, meaning they are uncorrelated in time. All of the analysis methods presented here assume independent samples. But for all but the simplest of systems, completely independent samples can be very difficult to generate. For protein-ligand binding affinities, the time scale for some motions may be hundreds of ns, meaning truly uncorrelated samples may be impossible to generate in a reasonable amount of time with today's simulation technology. In this case, free energy calculations *might* provide some useful information, but will only be approximations to the correct free energy for that model, and cannot be considered reliable.
- Monitor the simulations for changes in important degrees of freedom. For large ligand binding simulation, movement of most solvent degrees of freedom will happen quickly. However, there are a number of degrees of freedom that might not move quickly. This include tightly bound waters, ions, dihedrals of both side chains and the ligand, and large scale protein domain motions.

2.1.6. Extract Information from the Samples

Once samples and energies are obtained, then we can apply the analysis methods discussed above. The data required from the sample will depend on the method used.

- TI requires the value of $\frac{\mathrm{d}U(\vec{q})}{\mathrm{d}\lambda}$.
- EXP requires either the energy difference $\Delta U_{k,k+1}(\vec{q})$ or $\Delta U_{k,k-1}(\vec{q})$, where k is the state of that sample, depending on which direction the free energy is calculated.
- BAR requires both the energy difference $\Delta U_{k,k+1}(\vec{q})$ and $\Delta U_{k,k-1}(\vec{q})$ at each sample.
- WHAM and MBAR both require the set of energy differences $\Delta U_{k,j}(\vec{q})$, where $j = 1, \ldots, K$ runs over all states along the pathway, though this information must be binned for WHAM.

For BAR, MBAR, and WHAM, this information can either be computed directly during the simulation, or in post-processing. It is obviously preferable to have this information automatically computed during the simulation, as it removes additional work from simulation setup, avoids errors that might result from these additional steps, and reduces the amount of data that must be kept. It is recommended to use information computed during the simulation if at all possible, as it is faster and involves fewer potential human errors that could be introduced during sampling.

However, if configurations from each simulated state k are stored sufficiently frequently, and with sufficient precision, then single point energy calculations can be run using each of these configurations as input to produce the quantities $\Delta U_{k,j}(\vec{q})$. For BAR, only three single point calculations (at $k+1$, k, and $k-1$) need to be performed for each saved configuration. while for MBAR and BAR, K single point calculations need to be performed. Although technically $\Delta U_{k,k}(\vec{q})$ does not need to be computed, as it should be zero, it is highly recommended to compute this quantity. First, it allows a check of whether the energy obtained for that configuration during the original simulation at state k is the same as the energy obtained in the reevaluation. If the difference between the two energies is greater than could be explained by numerical precision issues, then the simulation setup should be rigorously checked for self-consistency; such errors can easily lead to large free energy differences. The precision in the coordinates of the output files must be greater than the precision in standard pdb files. Coordinates stored as binary format are of course greatly preferred, but precision to within 10^{-5} Å may be a sufficient compromise depending on the software used. In any case, specific choices must be carefully validated.

In some special cases where $U(\lambda, \vec{q})$ is a separable function of λ and $\vec{q}$ like the linear case, $U(\lambda, \vec{q}) = (1-\lambda)U_0(\vec{q}) + \lambda U_1(\vec{q})$, TI can be computed in postprocessing using the single point energies of the end – points. In other cases, such as for soft core potentials, $\frac{\mathrm{d}U(\vec{q})}{\mathrm{d}\lambda}$ cannot be computed in postprocessing and must be computed directly in code.

Once the data have been assembled, independent subsets of the data must be identified. This process involves an analysis of autocorrelation times. The autocorrelation time measures the time between effectively uncorrelated samples, and there are a number of approaches for computing it. Assume that we have an observable A gathered over a simulation of time T. If we write $\delta A(t) = A(t) - T^{-1}\int_{t=0}^{T} A(t)\mathrm{d}t$, or the instantaneous value minus the average over the interval then:

$$C_A(\Delta t) = \frac{\int_{\tau=0}^{T} \delta A(\tau)\delta A(\tau+\Delta t)\mathrm{d}\tau}{\int_{\tau=0}^{T} \delta A(\tau)^2 \mathrm{d}\tau}.$$

If the $C_A(\Delta\, t) = 0$ at and after $\Delta\, t$, then two samples separated by $\Delta\, t$ are uncorrelated, and can be treated as independent samples.

For a series of N samples, occurring time $\delta\, t$ apart, $C_A(\delta\, T)$ will be defined at i distinct points. Since $\delta A(i) = A(i) - \frac{1}{N}\sum_{i=0}^{N} A(i)$, then:

$$C_A(i) = \frac{\sum_{j=0}^{N} \delta A(j)\delta A(j+i)}{\sum_{j=0}^{N} \delta A(j)^2}$$

Under standard assumptions, samples can be considered effectively uncorrelated if they are spaced by 2τ.

In many circumstances, the autocorrelation function can be fit to an exponential, in which case τ is simply the relaxation time of the exponential function. Alternatively, τ can be computed as the integral under the $C_A(t)$ curve, though care must be taken as it becomes noisy at long times, especially at more than half the total simulation time. As a rule of thumb, a total time of 50τ should be simulated to feel confident about an estimate of τ, as very long correlation times may not be detected by shorter simulations. Many mature simulation packages have tools to compute these correlation times, sometimes at a more sophisticated level than that presented here. In any case, some tools for computing correlation times should emphatically be used, or the calculated statistical uncertainty will be lower than it should be.

It appears that for solvation free energies of small molecules, the time scales involved are often not particularly long. The longest time scales are those for water rearrangement and torsions. Some unpublished test give the correlation times of $\frac{\mathrm{d}U}{\mathrm{d}\lambda}$ for small rigid molecules are around 5–30 ps. However, if there are explicit torsional barriers in the molecule, which are particularly high, such as boat–chair transitions or slow rotations of internal torsions (such as the hydroxyl orientation in carboxylic acids, for example), this correlation time can be many nanoseconds (42).

Once the correlation time is calculated, there are two possible ways to use the information to obtain answers from independent data. For methods that compute averages from single states, like TI, the average over all samples can be used as the mean, and the variance then multiplied by $\sqrt{2\tau}$ to obtain an effective variance. Alternatively, the data set can be *subsampled*, with a set of samples mutually separated by 2τ being selected to analyze (see Note 8). If the correlation time is estimated accurately, we are not actually throwing away information by discarding data, since this discarded data duplicates information contained in the retained data.

Technically, we are only sampling independent configurations if all coordinates are uncorrelated between each sample, not just the energies. In most cases, independent sampling of the energies also implies uncorrelated sampling of the configurations.

However, there are a number of situations in which energies appear to be sampled approximately independently within the limit of the noise, but the configurational space is only partly sampled. For example, if there is a second binding pose that has similar binding affinity, but which the ligand only travels to occasionally, this might not show up when inspecting the correlation time of the energetic components alone.

This problem can be partially solved by also monitoring structural correlation times. For example, for a small molecule solvation energy, the correlation times of slow dihedrals can be computed. For a binding affinity problem, the autocorrelation time of the distance between a given point on the protein and the ligand, or the ligand dihedral angle between a bond in the protein and a bond in the ligand can be computed to verify that sufficient sampling is indeed happening on the time scale of the simulation.

2.1.7. Analyze the Information from the Samples to Obtain a Statistical Estimate for the Free Energy

Once we have a set of independent samples of energy data from a series of equilibrium simulations, we can analyze this data to obtain an estimate of the free energy and the error associated with its estimate. The exact form of the analysis will depend on the method being used, so we look at different methods individually.

Data Analysis for TI Given a set of N_k samples of $\frac{\mathrm{d}U}{\mathrm{d}\lambda}$ from equilibrium at each of k states, $\left\langle \frac{\mathrm{d}U}{\mathrm{d}\lambda} \right\rangle_k$ can be computed from the simple averages $\left\langle \frac{\mathrm{d}U}{\mathrm{d}\lambda} \right\rangle_k = N_k^{-1} \sum_{i=1}^{N_k} \frac{\mathrm{d}U}{\mathrm{d}\lambda}$ at each state k. To compute the free energy ΔA, we then perform numerical integration:

$$\Delta A \approx \sum_{k=1}^{K} w_k \left\langle \frac{\mathrm{d}U}{\mathrm{d}\lambda} \right\rangle_k,$$

where the w_k are weighting factors corresponding to different types of numerical integration (see Note 5). As discussed previously, the trapezoidal rule is the most robust and most recommended for beginners, since it easily allows for unequal spacing in λ, which is required to minimize the variance. Although alternative methods can yield lower integration error, these methods require significant problem specific information, and are not recommended for beginners. In almost all cases, it is simpler to identify regions of high curvature, and run more simulations in these areas.

Computing the overall variance of TI is straightforward, though it involves one pitfall. It is important to calculate the overall variance of the integration, rather than calculating the variance of each individual $\Delta A_{i,i+1}$ and assuming the variances add independently. They do not. Instead, since each of the $\left\langle \frac{\mathrm{d}U}{\mathrm{d}\lambda} \right\rangle$ results is independent of the others, since they are generated from different simulations, and therefore $var(\delta A) = \sum_{i=1}^{K} w_k^2 var\left(\frac{\mathrm{d}U}{\mathrm{d}\lambda}\right)_k$. In the case of simple trapezoidal rule, we can see that

$$var(\Delta A_{1,K}) = \sum_{k=1}^{K} w_k^2 \; var\left(\frac{dU}{d\lambda}\right)_i$$
$$= \frac{1}{4} var\left(\frac{dU}{d\lambda}\right)_1 + var\left(\frac{dU}{d\lambda}\right)_2 + \cdots + var\left(\frac{dU}{d\lambda}\right)_{K-1} + \frac{1}{4} var\left(\frac{dU}{d\lambda}\right)_K.$$

This is very different than if we calculated the variance for each $\Delta A_{i,i+1}$, and then added these variances directly, which would result in the following:

$$var(\Delta A_{i,i+1}) = \frac{1}{4} var\left(\frac{dU}{d\lambda}\right)_i + \frac{1}{4} var\left(\frac{dU}{d\lambda}\right)_i$$
$$var(\Delta A_{1,N}) = \sum_{i=1}^{N-1} var(\Delta A_{i,i+1})$$
$$= \frac{1}{4} var\left(\frac{dU}{d\lambda}\right)_1 + \frac{1}{2} var\left(\frac{dU}{d\lambda}\right)_2 + \cdots + \frac{1}{2} var\left(\frac{dU}{d\lambda}\right)_{K-1} + \frac{1}{4} var\left(\frac{dU}{d\lambda}\right)_K.$$

As discussed above, the standard error can then be computed as $\sqrt{var(\Delta A_{1,N})}$ from all samples, then multiplied by $\sqrt{2\tau}$ to obtain a corrected variance that corresponds to the correlation time.

Alternatively, averaging and integrating can be performed on the subsampled data set. For alchemical changes that result in smooth, low curvature sets of $\langle dU/d\lambda \rangle$, TI can be accurate using a relatively small number of points. However, if the curvature becomes large, as is frequently the case for alchemical simulations where Lennard-Jones potentials are turned on or off, then the bias introduced by discretization of the integral can become large (4, 24, 38). Even in the case of small curvature (i.e., charging of small polar molecule in water) reasonably large errors can be introduced (i.e., 5–10% of the total free energy with 5λ values). The basic conclusion is that TI is an adequate method for most purposes, but a researcher *must* verify that enough states are included such that the free energy is essentially independent of the number of states. If a molecule is being annhiliated, TI might require a large number of states to give accurate results, as the curvature of such decoupling paths is large. Large variance at a given states indicate large curvatures, so λ should be chosen to minimize variance.

Data Analysis for EXP Free energy propagation from EXP can be analyzed in the same way as TI, using the correlation time to either subsample or to calculate an effective sample number.

Since EXP produces free energies differences between intermediates that depend only on samples from one state, variance estimates for individual $\Delta A_{i,i+1}$ values are independent, and the total variances will add.

Data Analysis for BAR and MBAR For BAR, the mathematical details are somewhat more complicated, since they involve solving a set of iterative equations. The variance estimate from BAR computes the variance between two states. As with TI, the variances of consecutive intervals $k-1$ to k and k to $k+1$ are correlated, since they both involve samples from the state k. However, the relationship between these values is more complicated than with TI. Alternative methods, such as bootstrap sampling described below, must be used to obtain an accurate error estimate. MBAR involves solving complex systems of linear equations to compute the variances. However, for MBAR, all correlations between data are taken into account. Implementations of both BAR and MBAR, with examples for free energy calculations, can be found at http://www.simtk.org/home/pymbar if other tools are not available.

One straightforward statistical method that can be used for all methods is bootstrap sampling. In bootstrap sampling, we take N random samples generated with replacement from the original data, and calculate the variance over these N different values. Further details on bootstrap sampling can be found in a number of sources (43), and a simple tutorial is contained in the *Biomolecular simulations: methods and protocols* in the "Methods in Molecular Biology" series. The great power of bootstrap sampling is that it can be used with any statistical estimator. However, it does requires the additional overhead of calculating the function of the samples F repeated M times. In most cases, this time will be negligible compared to the time used to generate the data, perhaps 5–10 min for MBAR, seconds for TI.

2.2. Accelerating Sampling

We have presented a series of robust methods for calculating the free energy of a given system. However, in many cases of interest, this may require significant investment of computational resources, beyond that which can be obtained by most researchers. In this chapter, we therefore also examine additional tools for accelerating the sampling. Because of space limitations, we do not go deeply into all these methods. They are not needed to carry out free energy calculations, but may be required to converge calculations in complex systems with slow dynamics. Many of these techniques are relatively new, and may not be available in all simulation packages.

2.2.1. Using Umbrella Sampling for Convergence

One standard method for improving sampling in atomistic simulations is umbrella sampling (44), where bias terms are added to constrain the simulation in some way, and the effect of these

restraints is then removed in the analysis. This procedure can be used to either lower potential energy barriers or to restrain simulations to slow-interconverting configurations that are relevant to the binding affinity (for example, different torsional states), allowing each of the component free energies to be properly computed and then combined (1, 18, 45). Sometimes this can even be necessary for hydration free energy calculations (42) Another application of umbrella sampling is computing the free energy of constraining the free ligand into the bound conformation directly before computing the free energy of binding, and then computing the free energy of releasing these restraints. This usually decreases the correlation times for sampling of the intermediate states and thus increasing the efficiency of the simulation (18, 34).

2.2.2. Expanded Ensemble, Hamiltonian Exchange, and λ-Dynamics

It is possible to bring all the intermediates together in a single simulation system, either as series of coupled simulations of the intermediate states, usually called Hamilton exchange simulation, or as a single simulation that simultaneously samples both intermediate states and separate coordinates, called expanded ensemble simulation. A number of studies have shown that Hamiltonian exchange simulations can speed up simulations by allowing the system avoid kinetic barriers by going through alchemical states where those barriers are not as pronounced, significantly speeding up free energy simulations (46–51). Alternatively, the alchemical variable λ can be treated as a dynamical variable, which adds complications by introducing a fictitious mass corresponding to the λ degree of freedom, but is essentially equivalent to Monte Carlo techniques (49, 52–54). There are a number of variations of sampling in λ that may show promise in the future, but such methods are still in the preliminary stages of development (55–59).

At the current time, although they are extremely promising, we cannot recommend expanded ensemble and λ-dynamics methods to most practitioners. The methodology and implementations are not always robust and require tweaking additional parameters to obtain proper convergence. However, we do recommend Hamilton exchange methods. Most codes implementing Hamiltonian exchange methods do so on top of well tested temperature replica exchange routines, and no additional analysis is needed; the outputs of Hamiltonian exchange simulations can be analyzed in the same way as the outputs of K uncoupled simulations. These simulations are guaranteed to decorrelate as fast or faster than standard simulations, though the exact amount of improvement depends on the system. The analysis of correlation times can be somewhat complicated by such simulations; computation of correlation times should be computed along trajectories that sample different states, not along single states that might be switching back and forth along very different trajectories.

2.2.3. Verification, Verification, Verification

There are a number of ways that free energy simulations can go wrong, and the lists presented here cannot cover all possible problems. The best defense is to consistently evaluate the validity of each step of the process. For example, it is generally a very good idea to start out by calculating free energies which are well-known. The free energy of solvation of OPLS methane in TIP3P water is known to be -2.2 ± 0.1 kcal mol^{-1}, and has been replicated a number of times in different simulation programs. The hydration free energy of toluene in TIP3P water with AMBER/AMBER-GAFF parameters and HF 6-31G* RESP charges has also been the object of multiple studies and has been reported as -0.41 ± 0.2 and -0.7 ± 0.1 kcal mol^{-1} (27, 60). The Web site http://www.alchemistry.org maintains a number of these examples to test. If using any software suite to calculate free energies for the first time, it is highly recommended to first reproduce a simple solvation energy to verify that the approach is being performed correctly before moving to more complicated calculations.

One of the most common problems that can occur is that the input files and/or options used to perform the free energy calculations are different than the input files used to perform standard calculations. In virtually every free energy enabled code, this leads to the possibility that the state set up for free energy calculations no longer corresponds to the same state when free energy options are turned off. To avoid this, you should always verify that the potential energy of the system with free energy options turned off in the initial state is *exactly* the same as the potential energy at $\lambda = 0$ with the free energy options turned on.

Likewise, you should always verify that the potential energy of the system with free energy options turned off in the final state is *exactly* the same as the potential energy at $\lambda = 1$ with the free energy options turned on. "Exactly," in this case means that any differences should be no more than those caused by numerical rounding from differences in order of operations. Anything larger than this indicates some breakdown in the computation that could potentially result in propagated error significantly affects the results.

Another common problem is human error in setting up simulations. If humans are involved in editing topology and other input files for the initial and final states, it is easy to accidentally set up one atom to have an incorrect final state, or mistype a key parameter. This typically means human input is a bad idea and calculation setup should be done by script or program, since bugs are then reproducible. New tools for calculation setup should be carefully tested on cases with known results to ensure that the setup process is functioning correctly.

Poor convergence, undetected by uncertainty analysis, can also wreak havoc on results. There are several methods for validating convergence, such as checking that thermodynamic cycles sum

to zero, when conducting relative free energy calculations, or ensuring computed free energies are consistent when beginning from markedly different starting structures, which is applicable for both relative and absolute binding calculations, as well as for hydration free energy calculations (1, 42,61).

As a simple but very useful check, simulation trajectories should always be visually inspected. Visual inspection of trajectories can often catch errors that are hard to otherwise notice. For example, if the calculation is of the relative binding affinity of ligands that are tight binders, but the composite ligand is somehow ejected into the solvent, or adopts unnaturally high energy configurations, then there is likely an error in the simulation setup. If the molecules move visibly between very different configurations on a long time scale, it indicates either the system is not yet equilibrated, or that the correlation times for the system may be slow in a way that does not yet show up in the energetic analysis. Visual inspection during (and not just after simulations have completed) allows error to be recognized before too much computational time is wasted. Performing simulations for multiple physically reasonable starting configurations is also a useful technique on this subject we refer the reader to the Independent-Trajectory Thermodynamic Integration (IT-TI) approach described in Chapter 27 (62).

We now give some specific examples for implementing this description of free energy calculations. The examples are the relative free energy of phenol and toluene binding to T4 lysozyme (1, 45, 63), and the absolute free energy of binding toluene to T4 lysozyme. A discussion of small molecule solvation in water can be found in the free energy chapter of *Biomolecular simulations: methods and protocols* in the this same "Methods in Molecular Biology" series.

2.3. Example 1: Relative Free Energies of Binding

As a first example of free energy calculations for small molecule binding affinities, we first look at the difference in the relative free energies of binding of toluene and phenol in the apolar cavity of T4 lyzosyme.

- *What is the thermodynamic cycle?* We compute the free energy to turn phenol into toluene in the protein cavity and compute the free energy to turn phenol into toluene in solution, as described in Fig. 1.
- *What are the end states?* The end states for the first calculation calculation is T4 lysozyme, in water, with a intermediate molecule with nonbonded parameters that look like toluene at one end state and that look like phenol in the another end state. There are a number of choices for even this. It is likely simplest to choose a dual state topology, such as in Fig. 3; an *ortho* or *meta* arrangement be used just as easily.

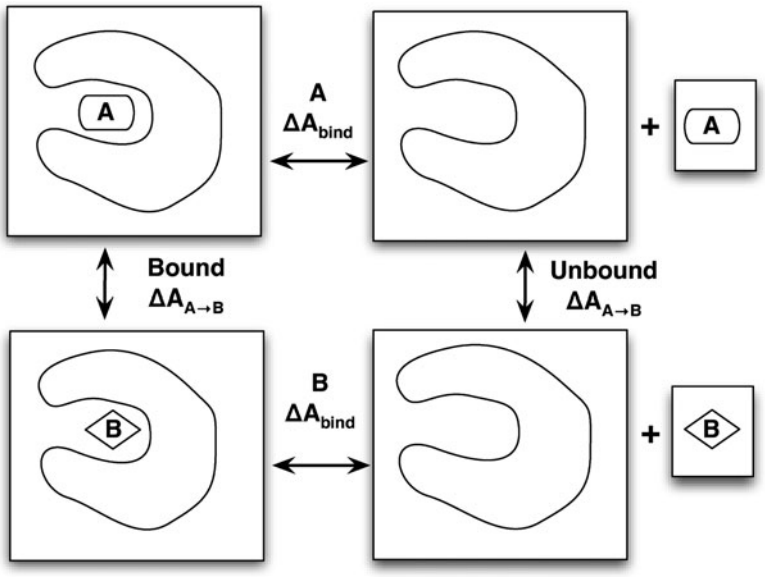

Fig. 3. A sample dual topology design for the transformation of phenol (**a**) to toluene (**b**) as described in the text. The choice of *para* arrangement is arbitrary; *ortho* or *para* arrangements would work as well.

- *Which series of intermediates?* Arbitrarily selecting phenol as the initial state, the OH moiety must disappear, and the methyl must appear. A good approach would be to first turn the charge on the OH and the *para* H to zero. It should be done keeping the overall system at the same total charge at each intermediate.

  Once these charges are turned, then the Lennard-Jones ε of the hydroxyl and that *para* can be turned to zero. At the same time, and angle, bonded, and torsional terms can be turned off linearly. Then the LJ ε of the methyl group and its *para* hydrogen can be turned on, while its bonded terms are turned on, and finally, the charges of the methyl and its *para* hydrogen can be turned on. How many intermediates will this require? In practice, using BAR or MBAR, perhaps 2–4 intermediates for turning off the charges, and 4–6 intermediates for turning off the Lennard-Jones terms and bonded terms. However, this assumes that intermediates are spaced to minimize the statistical error, and BAR or MBAR is used to calculate free energies. If equal spacing in λ is used, the number of states might be significantly higher, perhaps 10 for the Coulombic terms and 20 or more for Lennard-Jones terms. Therefore, finding an appropriate spacing equalizing variance between states is important for efficiency. There is a trade-off in adjusting the λ spacing; it obviously requires more processors to sample more intermediate states, but the decrease in variance compensates for this until spacing is relatively close. TI would likely require even more intermediates, the exact number depending on the level of uncertainty required.

In general, the choice of spacing will depend on the availability of processors, the correlation times of the system under study, and the level of precision required. Linear changes of bonded terms generally are well behaved.

If all time scales were the same as obtained with small molecule solvation, then this sort of spacing would result in statistical uncertainty in the 0.1–0.05 kcal mol^{-1} range for 5 ns simulation at each λ state is performed, as used in previous large scale studies (26). would require perhaps 2–3 times as many intermediates. However, this is the lower bounds on the statistical uncertainty; in most cases, it would be much larger, as there will be long time scale correlation times due to the motion of the protein. It is possible to perform both Lennard--Jones transformations simultaneously, but in this case, it would be necessary to remove the improper torsions on the rings, as they would force the substituents to collide with each other.

- *What simulations to run?* Equilibrium simulations must be run at each of the intermediates. Typically, one could start with the fully interacting state at all intermediates, and run for several nanoseconds, to allow the system to equilibrate at that intermediate. Even for relatively rigid molecules, such as FKBP, experience has demonstrated that equilibration typically requires 2–4 ns, though this will vary from system to system.

  The simulation box should be large enough for the solvated molecules not to interact with themselves, so the width of the box should be at least twice the cutoff plus the longest width of the protein plus ligand. The simulation time required will depend on the accuracy of the simulation. For a protein, all torsional exchanges tend to be slowed down, and one would expect something more like 20 ns. But more generally, the simulation time required will depend on the accuracy of the simulation; for a molecule this size, simulation times of perhaps 50 ns may be necessary to get consistent results. At the present time, even with relatively inflexible proteins, getting results that have statistical uncertainty of less than 0.5 kcal mol^{-1} is difficult.

  Deciding on what simulations to run also means deciding which starting configuration to use. The choice of which starting configurations is difficult, since the environment is a protein binding site, not a homogeneous liquid. The ideal starting structure is a crystal structure of the ligand bound to the site, or at least a homologous ligand to which the ligand of interest can be modeled without distorting the structure of either the ligand or the protein. If the binding site is not known, then obtaining an accurate free energy is not likely; docking is not necessarily reliable for picking the single true experimental binding site. If the binding site is known but a crystal structure is not available, then docking can be used to generate a range of

potential starting locations. Initial simulations several nanoseconds in length, tens of nanoseconds if possible, should be used to test if these configurations interconvert. If they do not, it may be necessary to run multiple simulations for each of the binding sites (1).

Once starting configurations are selected, one would again generally start with the fully interacting state at all intermediates, and run short simulations at each λ to allow the system to partially equilibrate at each new λ value, followed by long equilibrations for each λ state with constant pressure simulations to find the equilibrium density. In this case, at least 2–4 ns should be used to equilibrate. Even for a crystal structure, several starting configurations (perhaps obtained from multiple crystal structures, if available) should be used, and examined to see if the ligands sample the same conformational states in all simulations (1, 61). It is possible that even closely related ligands may not bind in the same orientation, presenting some sampling problems (61).

- *How do we analyze the data?* First, assume that we are using BAR, and that the code does not automatically print out the energy differences. In that case, the potential energy differences must be generated by single point simulations. This can be done by saving configurations every N steps, where N will depend on the correlation times of the potential energy. Typically, for a small rigid molecule, it would be around 1 ps, though if there are slow degrees of internal freedom, it could take longer. We would then take those configurations, a run a series of single point energy calculations. These calculations should be identical to the ones performed to generate the runs, but each configuration will be evaluated at the λ value of the neighboring intermediate. For each interval, we will have two energy differences, from state i to $i+1$ sampled from state i, and from $i+1$ to i, sampled from state $i+1$. The BAR calculation is performed for each interval, giving an estimate for the free energy difference. We could then apply bootstrap sampling to the data set of evaluated energies to obtain an error estimate. If the energy differences are printed out, then we can skip all but the bootstrap sampling and BAR calculation, greatly simplifying the analysis.

  Now assume that we are now performing thermodynamic integration. In this case, we expect that the values of $\frac{\mathrm{d}U}{\mathrm{d}\lambda}$ are printed at each step. It would be impossible to generate this TI data in postprocessing, assuming we are using the recommended soft core potentials. We simply average the values from each simulation, and perform numerical integration and error estimation from the formulas above. The free energies of two transformations, in aqueous solution and in the presence of the protein are then subtracted to obtain the final result.

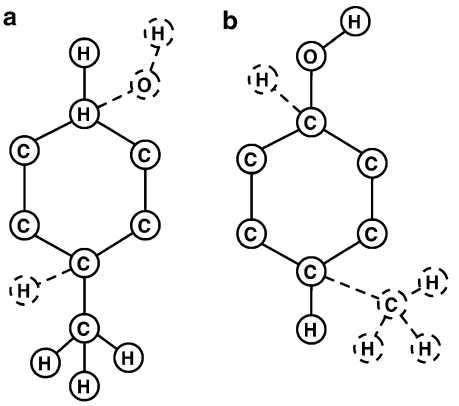

Fig. 4. A thermodynamic cycle for the absolute binding affinity of a ligand *L* by an "alchemical decoupling" pathway. The free energy of transfer of the decoupled molecule from protein to solvent is zero, resulting in a full thermodynamic cycle for computing the binding free energy ΔA_{bind}.

- *Anything else to watch out for?* Visualizing the simulation is always a good idea just to make sure nothing strange is happening. Note that the details of the protein was hardly mentioned in the discussion; the protein, in most respects, is just different external environment than the water. One difference that occasionally has some relevance is the location of the binding state. Assuming the ligand is a tight binder, then the ligand will always remain tightly localized around the binding site, and the definition of the binding site becomes pretty much irrelevant. See Note 9 for further discussion of weak binding. In most standard cases, determining precise binding affinities of weak binders is not required. Rather, the scientific questions will be to distinguish between tight binding ligands, or to tell whether a ligand is a tight or weak binder.

2.4. Example 2: Absolute Binding Affinity of Toluene to T4 Lysozyme

What is the thermodynamic cycle? There are several different potential ways to construct a thermodynamic cycle. Two of the most useful potential cycles are shown in Fig. 4 and Fig. 5. The alchemical decoupling pathway is shown in Fig. 4 and a pulling path is shown in Fig. 5.

In the alchemical decoupling path (Fig. 4), we start with a bound complex, and then turn off the interactions of the ligand with its environment. Since there are now no interactions between ligand and the rest of the system, we can transfer this "ghost" ligand from the solvated protein box to a solvent box with $\Delta A = 0$. We can then turn the intermolecular interactions back on while the ligand is in the pure solvent box; the binding free energy completes the cycle, since we now have a pathway that we can simulate from the complexed ligand to the solvated, unbound ligand.

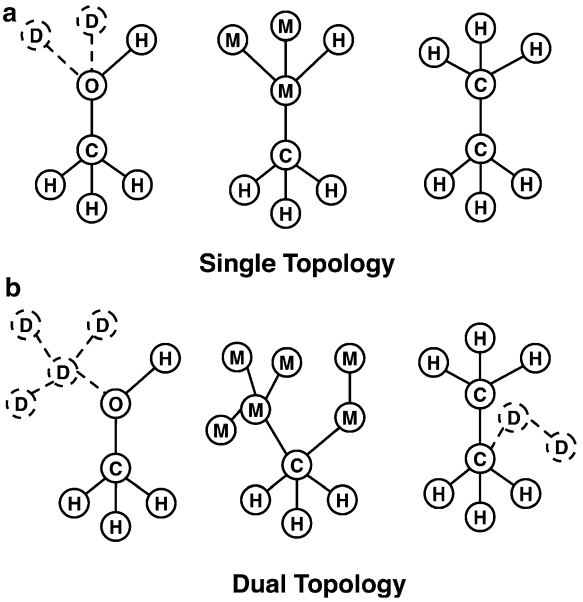

Fig. 5. A thermodynamic cycle for the absolute binding affinity of a ligand *L* by a pulling pathway. The ligand is pulled away by a harmonic potential attached to the center of mass of the ligand. The harmonic restraints are then turned off; via a series of equilibrium simulations in the complexed ligand, and analytically in the case of the solvated ligand.

In the pulling pathway (Fig. 4), we start again with the bound complex. We then turn on a harmonic restraint to the center of mass of the ligand, and then move the center of the harmonic restraint from inside the binding site to a location sufficiently far away from the the site. The harmonic restraint in solution can then be removed analytically (see Note 10), and which can be calculated as

$$\Delta A = kT \ln \left(\frac{3}{V} \left(\frac{\pi kT}{K} \right)^{3/2} \right), \tag{10}$$

where K is the force constant of the spring, and V is the volume of reference state. The average volume per molecule for a reference concentration of 1 M is 1.661×10^3 Å^3/molecule (see Note 11).

One problem with the alchemical decoupling pathway as described above is that the ligand can come free to wander throughout the entire simulation box, or get stuck in different parts if the protein in near-ghost states. These motions can lead to extremely long time scale motions, making it very difficult to collect uncorrelated samples, and thus making computation of accurate free energies. Additionally, this pathway neglects any knowledge of the standard state of the system; equilibrium constants are only defined up to the standard state. One standard

solution to both of these problems is to harmonically restrain the ghost state ligand to be to a certain location, defined by the geometry of the protein. Once in the ghost state, the harmonic restraint can be analytically removed before transferring the ligand to the solvent box, identical to the harmonic potential in the pulling pathway.

Unlike the pulling pathway, the ligand is harmonically constrained not to a fixed point in space, but to a geometric point in the cavity defined by the protein, as close as possible to the average bound location. At minimum, the ligand should be restrained in translational space; however, there is substantial evidence that it is also computationally efficient to restrain the orientational configuration of the ligand during the coupling (31, 63). Note that in the free state, the ligand is always free to move; the constraining potentials are added along the chain of intermediates. Other constraints schemes are possible (see Note 12). The rotational and translation degrees of freedom, instead of needing to be sampled along all intermediate states, only need to be sampled as the harmonic restraints are turned on. This means the correlation times along these degrees of freedom in the coupled state becomes very short, allowing data to be collected efficiently.

For translational harmonic restraints, the free energies can be computed using the formula $U(x_{\text{cm-ligand}}) = K/2\ (x_{\text{cm-ligand}} - x_0)^2$, where x_0 is the anchor point, $x_{\text{cm-ligand}}$ is the center of mass of the ligand, and K is the spring constant. For restraints in orientational space, then harmonic potentials are placed on six degrees of freedom; one distance, two angles, and three torsions, determined by the locations of three ligand and three protein atoms (31, 63) (See also Note 13).

What are the end states? For the alchemical pathway (Fig. 4), we have a four stage thermodynamic cycle. We use three free energies to compute the fourth, so we have potentially six end states to simulate. However, one of the free energy changes is zero, so we only need to worry about two calculations, with two end states for each. For the first calculation, the simulation is simply of the bound state. This is exactly the same as the initial state for relative free energies of binding. The information we need to simulate the end state for an absolute binding free energy is the same as that required for the protein-ligand coupled states of the relative free energy.

The final state of this first computation is the protein in "complex" with the decoupled ligand. There are two choices for our decoupled ligand end state. Once choice for the end state has all of its intramolecular interactions intact, and only has the ligand-environment calculations turned off. Alternatively, we could choose a ligand end state that has some of the intramolecular interactions turned off. Either method will work, as the "ghost" state is the same in both the protein and pure solvent

decoupling stages. Turning off the intramolecular interactions will result in larger free energy differences, but might aid convergence, since the decoupled state can sample configuration space more easily. However, only the nonbonded and proper torsional terms should be turned off. Turning off bonds, angles and improper terms could lead to geometric distortions in the molecule that could cause convergence problems.

Since the solvent box remains the same through the first transformation, we don't actually need to do perform any simulation of this box. It remains essentially a bookkeeping tool that allows us to see that the entire thermodynamic cycle is indeed complete. Similarly, for the second calculation in the alchemical pathway, the protein simulation is not affected, we do not need to perform any simulations with the protein in this step; including it is again bookkeeping tool to see the full cycle.

For the second transformation, the end states are the decoupled "ghost" ligand in solvent, and the fully coupled ligand. The "ghost" state must be the same as in the first simulation for the free energy to be zero. If the decoupled state only has intermolecular interactions turned off, then this free energy is the free energy of solvation. If the free energy of solvation is known, then this additional check might motivate using an end state with full intermolecular interactions.

For the pulling pathway, the initial state is again the fully interacting ligand and protein. The end state of the first calculation is a harmonically coupled ligand; turning on harmonic terms linearly appears to be an appropriate pathway. For the second calculation, the center of the harmonic spring is moved gradually, so that there is overlap between the volumes sampled by the "pinned" ligand in neighboring states, until the ligand is sufficiently far from the protein. Finally, the harmonic term can be removed analytically if when are sufficiently far from the protein. How far one must be depends on the system; some preliminary examples indicate that as little as 10 Å away from the nearest approach to the proteins may be sufficient for systems without large charge–charge interactions between the ligand and the protein.

Which series of intermediates? More atoms are changing in the absolute binding free energies than with the relative free energy calculations, which make it more important to choose a high efficiency pathway. For the alchemical pathway, the standard high efficiency pathway is again turning off all charges linearly and then turning off all the Lennard-Jones interactions using the soft core pathway. Turning off all ligand–ligand interactions as well as the ligand-environment interactions will result in a larger total free energy change for this part of the cycle, which will then be canceled out in the ligand–solvent calculation. However, the correlation times of the motion of the intramolecular interactions

will generally be lower than the correlation times of the intermolecular interactions, so the efficiency may not be very different. The specific choice of pathway is mostly independent of the protein–ligand and pure–solvent ligand simulations; an efficient pathway for one of the simulations will also be an efficient pathway for the other type of simulation (see Note 14).

What simulations to run? The prescription of exactly which experimental simulations to run are again similar for the simulations in relative binding affinities; again, the initial state is exactly the same, and the intermediate states are similar. If Hamiltonian exchange or expanded ensemble techniques are used, then the absolute binding case may actually converge more quickly than the relative binding affinities, as the ligands can more easily diffuse and escape from configurational traps when it is in the decoupled state. For relative free energy binding calculations, all intermediate states have a sizable number of atoms completely coupled to the system, and thus the "hybrid" ligand cannot move freely around the binding site.

How do we analyze the data? In the case of absolute binding affinities, the analysis is very similar to what it is for relative free energies. The data coming out of the simulations, whether we use BAR, TI, or MBAR, are in the same format as with relative free energies. This is one of the advantages of free energy simulations; the analysis methods do not care what simulations are actually being done, so the analysis will be the same. As before, we take the fact that the thermodynamic cycle must have a free energy that sums to zero, and compute the final free energy as sum or difference of the computed transformations.

Anything else to watch out for? Again, the problems to watch out for are similar for absolute binding affinities and for relative binding affinities. Visualizing the simulations are important, and choice of initial state is extremely important. Because absolute binding affinities are significantly larger than relative binding free energies, there can frequently be much less cancellation of error when comparing two relative binding affinities. If the underlying model is incorrect, this may mean that it is harder to extract understandable information from absolute binding affinities than for relative binding affinities.

2.5. Summary

Free energy calculations are a sophisticated, powerful set of tools for finding properties such as solvation free energies in arbitrary solvents and binding free energies. However, they give only a statistical estimate of free energy differences between two thermodynamic states whose accuracy and precision depend on careful choices of parameters, pathways of intermediates, and methods. Additionally, they can only give the free energies of the model, not the true experimental system; the molecular parameters of the

system under study must be sufficiently accurate for the correct free energy for the model to match the free energy of the system.

For complicated systems with long correlation times, free energy methods are not always reliable, because of the difficulty of collecting uncorrelated samples. Relative or absolute free energies of binding to proteins must, therefore, be taken with some degree of caution. As can be seen from the extensive set of notes and qualifications in the methods presented here, free energy "black-box" methods will not "automagically" determine free energies without significant investment in the physics, chemistry, and biology of the system under study.

However, such calculations are certainly closer to providing utility to biophysical researchers than they were in the past. As we present in this chapter, the methods used for free energy calculations are changing rapidly. Major biomolecular simulation packages, such as AMBER, CHARMM, NAMD, GROMACS, and GROMOS all are undergoing major improvements and changes in the features used to compute binding free energies. Although these changes will likely greatly improve the ability to perform free energy calculations in the near future, ongoing changes make it difficult to put together stable work flows for preparing ligands and simulation structures and determining ideal free energy protocols without significant human effort. It is difficult to recommend particular codes for the easiest use at the present time; we instead recommend using the code with which the user is most comfortable, as long as it supports one of the methods discussed here.

Because of the scope of free energy calculations, a single review article cannot be hoped to capture all possible problems or issues; for further information, readers are encouraged to read a number other of reviews on the subject of free energy calculations (64–71), particularly several very recent reviews (23, 72–74), as well as several useful books (64, 75–78).

3. Notes

1. The total phase space volume is the same for most molecular models because the only configurations with nonzero probability are where two atoms are directly on top of each other, with infinite positive energy. The effective phase space volume is lower, since there are a large number of configurations that might take a *very* long time to reach, because the atoms are partially overlapping. In that case, although (1) is true, it may take far more sampling to converge than could be done in any practical simulation.

2. The temperature T in (1) does not have a subscript, because we are considering free energy differences at the same temperature. It is possible to compute a free energy difference between systems at two different temperatures, $\Delta A_{ij} = -k_B T_j \ln Q_j + k_B T_i \ln Q_i$, but is no longer a ratio of partition functions, and it should never be necessary in systems of biological interest. When a researcher thinks of temperature dependence of the free energy, he or she is usually thinking about the effect of temperature on the free energy difference between two states, both of which are at the same temperature when the difference is calculated or measured.
3. For example, a change in average volume corresponding to the elimination of a 1 nm sphere would result in a PV work contribution to free energy of 0.032 kJ mol^{-1} or 0.008 kcal mol^{-1}, which is generally smaller than the error in all but the most precise experiments.
4. "Free energy perturbation" is a common term for these methods that directly compute the free energy difference as a function of changing molecular structure. "Perturbation" usually refers to an approximate theory that can be written as a series expansion. Free energy perturbation, however, is exact. The term "perturbation" here refers to the changes in the *chemical identity*, since simulations frequently involve changes in chemical identity, such as an amine to an alcohol, or a methyl group to a chlorine, or the disappearance of atoms.
5. If we are using simple trapezoidal rule, with equal lambda spacing, this becomes $w_1 = w_K = \frac{1}{2(K-1)}$, while $w_k = \frac{1}{K-1}$ for $i \neq 1,K$. For the trapezoidal rule with uneven spacing, $w_1 = \frac{\lambda_2 - \lambda_1}{2}$, $w_K = \frac{\lambda_K - \lambda_{K-1}}{2}$, and $w_k = \frac{\lambda_{k+1} - \lambda_{k-1}}{2}$ for $k \neq 1,K$.
6. Although the trapezoid rule is very robust, some improvements can be made by using a fit of the data to a polynomial fit (8) or to some other functional forms (7). Since fits to higher order polynomials can have numerical stabilities for some underlying functions, and alternate functional forms might only be appropriate with some transformations, some expertise and experience is required.
7. By using a power of $\lambda \geq 4$ instead of a strictly linear parameterization (such as $U(\lambda) = (1-\lambda)^4 U_0 + \lambda^4 U_1$) then the integral of $\langle dU/d\lambda \rangle$ will converge. However, it will converge rather slowly in number of samples, and can cause numerical instabilities (27, 28). For any nonzero λ, whatever the power, there will be small "fence posts," particles with a small impenetrable core (27). One possible way to avoid issues with these "fenceposts" has been to shrink the entire molecular structure. However, this can create problems with nonbonded interactions as

the molecular framework shrinks, causing instabilities in integration in molecular dynamics (27, 79, 80) and is generally not practical for large numbers of bonds.

8. For example, assume we are using BAR to compute the free energy between the 1st and 2nd states, and we have collected 5 ns of simulation, with snapshots every 10 ps, for a total of 500 samples. Then, we need to take the time series $\Delta U_{1,2}$ and $\Delta U_{2,1}$ and compute the autocorrelation function and correlation time. If we assume that the correlation time for $\Delta U_{1,2}$ is 20 ps, and the correlation time for $\Delta U_{2,1}$ is 40 ps, then we should take every fourth sample (or 2τ) from the $\Delta U_{1,2}$ data series and every eighth sample from the $\Delta U_{2,1}$ data series, and do subsequent analysis only with this reduced data set.
9. A ligand that is a weak binder ($K_d > 100$ μm) spends more time outside the binding site, and the terms "binding site" becomes more difficult to define. However, this difficulty of definition occurs in both experiment and simulation. For weak binder, one must carefully examine the physical phenomena leading to signaling of binding, as different signals may or may not be triggered by weak binding, and getting quantitative results is complicated.
10. We can remove this harmonic restraint analytically when we are in pure solvent because away from the binding site, the only part of potential energy of the ligand that depends on location is the harmonic restraint. The partition function can then be separated to the harmonic potential, acting only on the ligand center of mass, and the rest of the potential energy, which does not depend on the ligand center of mass. These free energies are thus independent.
11. We have examined a number of different ways to define the attachment point for the harmonic oscillators to the protein system, and a number of different spring strengths. The free energies appeared to be consistent, relatively independent of the spring strength (tested for 10–5,000 kcal mol^{-1}Å^{-2}) and the location (tested a range of harmonic anchor points 0,2,3,5 and 10 Å along the vector projected outward from the binding cavity from the average center of mass of the bound ligand).
12. It is also possible to add ligand conformation restraints (34, 35) to the ligand, which can reduce the correlation times even further. This is somewhat of a more advanced topic, since it is not clear if ligand constraints can be sampled as easily as harmonic restraints while being imposed.
13. One possibility for adding the restraints to the ligand is to add the restraints *before* decoupling the ligand from the environment, rather than during the decoupling. This will require more intermediates, but may be more efficient, as the

correlation times for sampling will be lower once the restraints are implemented, since only configurations near the restraints are allowed. Alternate configurations are only collected while the restraints are being turned on. However, it is not yet clear what the best choice of these two options are in general. If the restraints are turned on too quickly, or insufficient sampling of the states where the torsions are turned on is done, then the results will not converge to the correct answer.

14. A good pathway of intermediates when complexed to the protein is similar to a good path in solution, since in both cases, the ligand experiences a large number of charge interactions and large number of Lennard-Jones interactions, with approximately the same density of particles in both cases. So, although the quantitative result will depend on the parameters, the qualitative behavior of the free energy as a function of distance along the alchemical pathway will be the same in both environments.

Acknowledgments

The author wishes to thank John Chodera (UC-Berkeley) and David Mobley (University of New Orleans) for ongoing discussions of reliability and usability for free energy calculations.

References

1. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K., and Dill, K. A. (2007) Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol. 371*, 1118–1134.
2. Woods, C. J., Manby, F. R., and Mulholland, A. J. (2008) An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *J. Chem. Phys. 128*, 014109.
3. Zwanzig, R. W. (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys. 22*, 1420–1426.
4. Shirts, M. R., and Pande, V. S. (2005) Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys. 122*, 144107.
5. Lu, N. D., Singh, J. K., and Kofke, D. A. (2003) Appropriate methods to combine forward and reverse free-energy perturbation averages. *J. Chem. Phys. 118*, 2977–2984.
6. Resat, H., and Mezei, M. (1993) Studies on free energy calculations. I. Thermodynamic integration using a polynomial path. *J. Chem. Phys. 99*, 6052–6061.
7. Jorge, M., Garrido, N., Queimada, A., Economou, I., and Macedo, E. (2010) Effect of the Integration Method on the Accuracy and Computational Efficiency of Free Energy Calculations Using Thermodynamic Integration. *J. Chem. Theo. Comp. 6*, 1018–1027.
8. Shyu, C., and Ytreberg, F. M. (2009) Reducing the bias and uncertainty of free energy estimates by using regression to fit thermodynamic integration data. *Journal of Computational Chemistry 30*, 2297–2304.
9. Bennett, C. H. (1976) Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys. 22*, 245–268.
10. Shirts, M. R., Bair, E., Hooker, G., and Pande, V. S. (2003) Equilibrium free energies from

nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett 91*, 140601.

11. Ytreberg, F. M., Swendsen, R. H., and Zuckerman, D. M. (2006) Comparison of free energy methods for molecular systems. *J. Chem. Phys. 125*, 184114.
12. Rick, S. W. (2006) Increasing the efficiency of free energy calculations using parallel tempering and histogram reweighting. *J. Chem. Theory Comput. 2*, 939–946.
13. Ferrenberg, A. M., and Swendsen, R. H. (1989) Optimized Monte Carlo Data Analysis. *Phys. Rev. Lett 63*, 1195–1198.
14. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A., and Rosenberg, J. M. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem. 13*, 1011–1021.
15. Bartels, C., and Karplus, M. (1997) Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations. *J. Comput. Chem. 18*, 1450–1462.
16. Gallicchio, E., Andrec, M., Felts, A. K., and Levy, R. M. (2005) Temperature weighted histogram analysis method, replica exchange, and transition paths. *J. Phys. Chem. B 109*, 6722–6731.
17. Souaille, M., and Roux, B. (2001) Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput. Phys. Commun. 135*, 40–57.
18. Wang, J., Deng, Y., and Roux, B. (2006) Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophys. J. 91*, 2798–2814.
19. Shirts, M. R., and Chodera, J. D. (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys. 129*, 129105.
20. Crooks, G. E. (2000) Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E 61*, 2361–2366.
21. Oostenbrink, C., and van Gunsteren, W. F. (2006) Calculating zeros: Non-equilibrium free energy calculations. *Chem. Phys. 323*, 102–108.
22. Oberhofer, H., Dellago, C., and Geissler, P. L. (2005) Biased Sampling of Nonequilibrium Trajectories: Can Fast Switching Simulations Outperform Conventional Free Energy Calculation Methods? *J. Phys. Chem. B 109*, 6902–6915.
23. Pohorille, A., Jarzynski, C., and Chipot, C. (2010) Good Practices in Free-Energy Calculations. *J. Phys. Chem. B 114*, 10235–10253.
24. Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R., and van Gunsteren, W. F. (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett. 222*, 529–539.
25. Zacharias, M., Straatsma, T. P., and McCammon, J. A. (1994) Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Phys. Chem. 100*, 9025–9031.
26. Shirts, M. R., and Pande, V. S. (2005) Solvation free energies of amino acid side chains for common molecular mechanics water models. *J. Chem. Phys. 122*, 134508.
27. Steinbrecher, T., Mobley, D. L., and Case, D. A. (2007) Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *J. Chem. Phys. 127*, 214108.
28. Pitera, J. W., and van Gunsteren, W. F. (2002) A comparison of non-bonded scaling approaches for free energy calculations. *Mol. Simulat. 28*, 45–65.
29. Blondel, A. (2004) Ensemble variance in free energy calculations by thermodynamic integration: theory, optimal Alchemical path, and practical solutions. *J. Comput. Chem. 25*, 985–993.
30. Boresch, S., and Karplus, M. (1999) The Role of Bonded Terms in Free Energy Simulations. 2. Calculation of Their Influence on Free Energy Differences of Solvation. *J. Phys. Chem. A 103*, 119–136.
31. Boresch, S., Tettinger, F., Leitgeb, M., and Karplus, M. (2003) Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. A 107*, 9535–9551.
32. Ytreberg, F. (2009) Absolute FKBP binding affinities obtained via nonequilibrium unbinding simulations. *J. Chem. Phys. 130*, 164906.
33. Lee, M. S., and Olson, M. A. (2006) Calculation of Absolute Protein-Ligand Binding Affinity Using Path and Endpoint Approaches. *Biophys. J. 90*, 864–877.
34. Woo, H.-J., and Roux, B. (2005) Calculation of absolute protein-ligand binding free energy from computer simulation. *Proc. Natl. Acad. Sci. 102*, 6825–6830.
35. Gan, W., and Roux, B. (2008) Binding specificity of SH2 domains: Insight from free energy simulations. *Proteins 74*, 996–1007.
36. Boresch, S., and Karplus, M. (1996) The Jacobian factor in free energy simulations. *J. Chem. Phys. 105*, 5145–5154.

37. Shenfeld, D. K., Xu, H., Eastwood, M. P., Dror, R. O., and Shaw, D. E. (2009) Minimizing thermodynamic length to select intermediate states for free-energy calculations and replica-exchange simulations. *Phys. Rev. E 80*, 046705.
38. Shirts, M. R., Pitera, J. W., Swope, W. C., and Pande, V. S. (2003) Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys. 119*, 5740–5761.
39. Kastenholz, M. A., and Hünenberger, P. H. (2006) Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. *J. Chem. Phys. 124*, 224501.
40. Fujitani, H., Tanida, Y., Ito, M., Shirts, M. R., Jayachandran, G., Snow, C. D., Sorin, E. J., and Pande, V. S. (2005) Direct calculation of the binding free energies of FKBP ligands. *J. Chem. Phys. 123*, 084108.
41. Smith, L. J., Daura, X., and van Gunsteren, W. F. (2002) Assessing equilibration and convergence in biomolecular simulations. *Proteins: Struct., Funct., Bioinf. 48*, 487–496.
42. Klimovich, P. V., and Mobley, D. L. (2010) Predicting hydration free energies using all-atom molecular dynamics simulations and multiple starting conformations. *J. Comp. Aided Mol. Design 24*, 307–316.
43. Efron, B., and Tibshirani, R. J. *An Introduction to the Bootstrap*; Chapman and Hall/CRC: Boca Raton, FL, 1993.
44. Torrie, G. M., and Valleau, J. P. (1977) Nonphysical Sampling Distributions in Monte-Carlo Free-Energy Estimation : Umbrella Sampling. *J. Comput. Phys. 23*, 187–199.
45. Mobley, D. L., Chodera, J. D., and Dill, K. A. (2007) Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput. 3*, 1231–1235.
46. Okamoto, Y. (2004) Generalized-ensemble algorithms: Enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J. Mol. Graph. Model. 22*, 425–439.
47. Roux, B., and Faraldo-Gómez, J. D. (2007) Characterization of conformational equilibria through Hamiltonian and temperature replica-exchange simulations: Assessing entropic and environmental effects. *J. Comput. Chem. 28*, 1634–1647.
48. Woods, C. J., Essex, J. W., and King, M. A. (2003) Enhanced Configurational Sampling in Binding Free Energy Calculations. *J. Phys. Chem. B 107*, 13711–13718.
49. Banba, S., Guo, Z., and Brooks III, C.L. (2000) Efficient sampling of ligand orientations and conformations in free energy calculations using the lambda-dynamics method. *J. Phys. Chem. B 104*, 6903–6910.
50. Bitetti-Putzer, R., Yang, W., and Karplus, M. (2003) Generalized ensembles serve to improve the convergence of free energy simulations. *Chem. Phys. Lett. 377*, 633–641.
51. Hritz, J., and Oostenbrink, C. (2008) Hamiltonian replica exchange molecular dynamics using soft-core interactions. *J. Chem. Phys. 128*, 144121.
52. Guo, Z., Brooks III, C.L., and Kong, X. (1998) Efficient and flexible algorithm for free energy calculations using the λ-dynamics approach. *J. Phys. Chem. B 102*, 2032–2036.
53. Kong, X., and Brooks III, C. L. (1996) λ-dynamics: A new approach to free energy calculations. *J. Chem. Phys. 105*, 2414–2423.
54. Li, H., Fajer, M., and Yang, W. (2007) Simulated scaling method for localized enhanced sampling and simultaneous "alchemical" free energy simulations: A general method for molecular mechanical, quantum mechanical, and quantum mechanical/molecular mechanical simulations. *J. Chem. Phys. 126*, 024106.
55. Zheng, L., Carbone, I. O., Lugovskoy, A., Berg, B. A., and Yang, W. (2008) A hybrid recursion method to robustly ensure convergence efficiencies in the simulated scaling based free energy simulations. *J. Chem. Phys. 129*, 034105.
56. Zheng, L., and Yang, W. (2008) Essential energy space random walks to accelerate molecular dynamics simulations: Convergence improvements via an adaptive-length self-healing strategy. *J. Chem. Phys. 129*, 014105.
57. Min, D., and Yang, W. (2008) Energy difference space random walk to achieve fast free energy calculations. *J. Chem. Phys. 128*, 191102.
58. Li, H., and Yang, W. (2007) Forging the missing link in free energy estimations: lambda-WHAM in thermodynamic integration, overlap histogramming, and free energy perturbation. *Chem. Phys. Lett. 440*, 155–159.
59. Min, D., Li, H., Li, G., Bitetti-Putzer, R., and Yang, W. (2007) Synergistic approach to

improve "alchemical" free energy calculation in rugged energy surface. *J. Chem. Phys. 126*, 144109.

60. Mobley, D. L., Dumont, È., Chodera, J. D., and Dill, K. A. (2007) Comparison of charge models for fixed-charge force fields: Small-moleculehydration free energies in explicit solvent. *J. Phys. Chem. B 111*, 2242–2254.
61. Boyce, S. E., Mobley, D. L., Rocklin, G. J., Graves, A. P., Dill, K. A., and Shoichet, B. K. (2009) Predicting Ligand Binding Affinity with Alchemical Free Energy Methods in a Polar Model Binding Site. *J. Mol. Biol. 394*, 747–763.
62. Lawrenz, M., Baron, R., and McCammon, J. A. (2009) Independent-Trajectories Thermodynamic-Integration Free-Energy Changes for Biomolecular Systems: Determinants of H5N1 Avian Influenza Virus Neuraminidase Inhibition by Peramivir. *J. Chem. Theo. Comput. 5*, 1106–1116.
63. Mobley, D. L., Chodera, J. D., and Dill, K. A. (2006) On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys. 125*, 084902.
64. Shirts, M. R., Mobley, D. L., and Chodera, J. D. (2007) Alchemical Free Energy Calculations: Ready for Prime Time? *Annu. Rep. Comput. Chem. 3*, 41–59.
65. Huang, N., and Jacobson, M. P. (2007) Physics-based methods for studying protein-ligand interactions. *Curr. Opin. Drug Di. De. 10*, 325–31.
66. Gilson, M. K., and Zhou, H.-X. (2007) Calculation of Protein-Ligand Binding Affinities. *Annu. Rev. Bioph. Biom. 36*, 21–42.
67. Meirovitch, H. (2007) Recent developments in methodologies for calculating the entropy and free energy of biological systems by computer simulation. *Curr. Opin. Struc. Bio. 17*, 181–186.
68. Rodinger, T., and Pomès, R. (2005) Enhancing the accuracy, the efficiency and the scope of free energy simulations. *Curr. Opin. Struc. Bio. 15*, 164–170.
69. Jorgensen, W. L. (2004) The many roles of computation in drug discovery. *Science 303*, 1813–1818.
70. Chipot, C., and Pearlman, D. A. (2002) Free energy calculations. The long and winding gilded road. *Mol. Simulat. 28*, 1–12.
71. Brandsdal, B. O., Österberg, F., Almlöf, M., Feierberg, I., Luzhkov, V. B., and Åqvist, J. (2003) Free Energy Calculations and Ligand Binding. *Adv. Prot. Chem. 66*, 123–158.
72. Steinbrecher, T., and Labahn, A. (2010) Towards Accurate Free Energy Calculations in Ligand Protein-Binding Studies. *Curr. Med. Chem. 17*, 767–785.
73. Michel, J., and Essex, J. W. (2010) Prediction of protein–ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *J. Comput. Aided. Mol. Des. 24*, 639–658.
74. Christ, C. D., Mark, A. E., and van Gunsteren, W. F. (2010) Basic Ingredients of Free Energy Calculations: A Review. *J. Comp. Chem. 31*, 1569–1582.
75. Chipot, C., and Pohorille, A., Eds. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Springer, 2007; Vol. 86.
76. Frenkel, D., and Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: San Digeo, CA, 2002.
77. Reddy, M. R., and Erion, M. D., Eds. *Free Energy Calculations in Rational Drug Design*; Kluwer Academic, 2001.
78. Leach, A. R. *Molecular Modelling: Principles and Applications*; Addison Wesley Longman Limited: Harlow, Essex, England, 1996.
79. Pearlman, D. A., and Connelly, P. R. (1995) Determination of the differential effects of hydrogen bonding and water release on the binding of FK506 to native and TYR82 → PHE82 FKBP-12 proteins using free energy simulations. *J. Mol. Biol. 248*, 696–717.
80. Wang, L., and Hermans, J. (1994) Change of bond length in free-energy simulations: Algorithmic improvements, but when is it necessary? *J. Chem. Phys. 100*, 9129–9139.