

Free energy methods in molecular simulation

David A. Kofke*

*Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York,
303 Furnas Hall, Buffalo, NY 14260-4200, USA*

Abstract

An overview is given of methods for calculating free energies by molecular simulation. Techniques may be categorized as density-of-states methods and work-based methods. Emphasis is given to the latter, with particular focus on issues of performance and accuracy. Two points of view are provided for understanding these issues. One is based on the distribution of work values observed in the calculation, and the other considers the relationship of the important regions of phase space for the systems. We present a metric that brings together the two viewpoints, and provide a demonstration that connects to the accuracy of a calculation as a function of sampling length. We also discuss staging methods and their appropriate formulation in consideration of phase-space relations.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Molecular simulation; Free energy perturbation; Phase space

1. Introduction

Free energy is the central quantity in thermodynamics, insofar as thermodynamics applies to the characterization of physical equilibrium [1]. The principle of free energy minimization (applied in the presence of appropriate constraints) underlies many of the standard approaches that engineers use to understand, model, and predict phase equilibria. There is indeed a very broad range of phenomena to which this general approach applies. The question can always be reduced (at least implicitly) to one involving the comparison of the relative stabilities of two or more systems. The answer to the question is obtained by comparing values of the free energies of the relevant systems; the stable system is the one with the lowest free energy at the conditions of interest. Examples are listed in Table 1. It should be apparent from the table that the notion of a “system” can be construed very broadly.

In some cases there may be continuum of conditions joining two systems, and the interest lies in the free energy profile

along the reaction coordinate defining the intermediate systems. This may be because the stable system lies somewhere along the reaction coordinate, such as in the case of chemical equilibrium. Alternatively, there may be a transition state along this path, and its characterization is needed to quantify the rate at which the system moves into the more stable condition [2]; here again the focus is on the free energy.

In this article, we provide an overview of methods for calculating the free energy in molecular simulations. Calculations are typically performed in the context presented above, and in such calculations one is interested in calculating free energy *differences*, or perhaps a free energy profile. It is the nature of such calculations in molecular simulation that the result so obtained is often incorrect, in the sense that it is prone to systematic error. We consider free energy calculations with particular emphasis given to this issue.

We should begin by pointing out that in some cases it may seem that thermodynamic stability is evaluated without recourse to free energy analysis. For example, prediction of the most stable polymorph of a molecular crystal [3] is often performed by exhaustive search for the configuration having the lowest *potential energy* (given a molecular model). Of course, this is in principle a free energy analysis, only with the free

* Tel.: +1 716 645 2911x2209; fax: +1 716 645 3822.
E-mail address: kofke@buffalo.edu.

Table 1

Examples of competing systems for which free energy evaluations are applied to determine relative stability

Vapor vs. liquid vs. solid vs. etc.	Mixed vs. unmixed
Reactant vs. product	Bound vs. unbound
Folded vs. unfolded	Structured vs. unstructured
Assembled/nucleated vs. dispersed	Adsorbed vs. free

energy given approximately by the potential energy. The neglect of entropic effects means that the results strictly apply only at zero temperature, leaving no recourse to examine the effect of temperature on the polymorphic behavior; improvements to this approach can be made by applying harmonic analysis. Another way to examine stability without calculating the free energy is through a phenomenological approach. Here one simply performs a simulation sufficiently lengthy to permit the system to equilibrate to its stable condition. This strategy has been used to good effect recently by the Glotzer group to examine structures formed by various models for nanoparticulate systems [4].

More often, stability is evaluated through detailed calculation of the free energy. This practice is made difficult by the tendency of free energy calculations to yield results that are precise but inaccurate: repeated calculations may yield a result that is reproducible but incorrect. The problem lies not with conceptual errors in the formulation of the methods, but instead arises from a bias inherent in the (otherwise correct) working formulas, causing them to give erroneous results unless an exceedingly impractical amount of configuration sampling is applied. The problem is not appropriately characterized as non-ergodicity; true, ergodic problems can also contribute to inaccuracy, but such pathologies in the sampling of a single system are not particular to free energy calculations. Free energy calculations have special difficulties because they require good sampling of *two* (or more) systems in a single simulation. Usually one of the systems is more stable than the other, so it is not a trivial problem to

ensure that both are adequately represented in a single simulation; further, it is not possible to gather the needed information from separate, disjoint simulations of the two systems. Failure to sample one of the systems of interest, while nevertheless sampling the other well (ergodically), produces the inconspicuous inaccuracies that we are concerned with here.

The nature of the problem is illustrated in Fig. 1. Presented are calculations performed to measure the chemical potential of a Lennard–Jones (LJ) fluid and water, respectively [5]. The two systems of interest are therefore each fluid with ($N + 1$) and without (N) one molecule interacting with all the others. For LJ the Widom insertion/deletion method is used [6], while for water a non-equilibrium work approach is applied (see below). The figures show the error in each calculation as the simulation progresses. Results are presented for two approaches to the calculation, one in which sampling is based on the N system (insertion) and the other in which sampling is based on the $(N + 1)$ system (deletion). The systematic error is clear, and its nature depends greatly on the choice of whether to perform the calculation as insertion or deletion. Also shown is the average of the insertion and deletion approaches. Such averaging is sometimes performed as a countermeasure to the inaccuracies of the calculations, recognizing that the inaccuracies will lie on opposite sides of the correct value. However, it should be clear from the figures that the inaccuracies of the two directions are not in general symmetric. For the LJ calculation, the insertion method yields a good value, so averaging it with a deletion calculation leads to a poor result. In contrast, for water the averaging does improve the outcome in comparison to either insertion or deletion taken individually. However, the overall result is still not very good.

It is hard to know for a specific free energy calculation if the insertion/deletion methods (or their equivalents to the system of interest) are in error, and if so, to what degree for each. Presently the only recourse is to perform extensive calculations, perhaps using alternate methods, to ascertain which free energy calculation is providing the correct results.

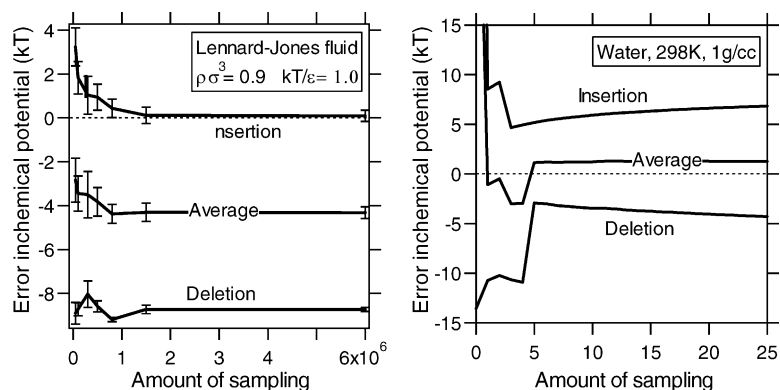


Fig. 1. Convergence plots for calculations of the chemical potential. Results are given for the error in the calculated chemical potential as the simulation progresses. Calculation for the LJ fluid used free energy perturbation (Eq. (4)), while those for water applied a non-equilibrium work approach (Eq. (3)); in both cases the free energy difference of interest is that between systems of 108 and 109 molecules [5]. “Amount of sampling” indicates number of FEP attempts (for LJ example) and number of NEW traversals (for water example). The latter has many fewer samples because each NEW traversal is more computationally expensive than a simple FEP attempt.

Clearly this is not the most efficient way to proceed. We do not solve the problem in this review, but we describe the thinking we believe is needed to make progress.

Methods for computing the free energy by molecular simulation can be broadly classified as either density-of-states methods or work-based methods [7]. The former has a more statistical–mechanical orientation, while the latter has a more classical thermodynamic viewpoint. Our interest of late has been in understanding issues related to work-based methods, so we will focus more on those approaches here. In the next section, we will provide a cursory overview of density-of-states methods, then in Section 3 turn to work-based methods. In Section 4, we consider distribution functions and phase space as complementary ways to describe the performance of work-based methods. In Section 5, we consider how these ideas impact the design and selection of staging methods, and we conclude in Section 6.

2. Density-of-states methods

Density-of-states methods calculate the free energy through its relation to the ensemble weights of different systems. If given complete freedom to explore two systems, A and B, a physically correct simulation algorithm will sample each in proportion to the Boltzmann factor of their free energies

$$\pi_i \propto \exp(-\beta F_i) \quad (1)$$

where π_i is the probability to sample system i , which has free energy F_i and $\beta = 1/kT$ the reciprocal temperature in energy units. The free energy difference between two systems is then easily obtained as the ratio. In practice, it is difficult to get sufficient sampling of all systems of interest to ensure that a meaningful result is obtained this way, so it becomes necessary to bias the process to better result in a uniform sampling of the systems of interest [6]. The bias may be easily removed afterwards. A significant amount of effort toward the development of density-of-states methods is directed toward efficient use of the accumulated simulation data to reformulate the bias to obtain such sampling as the simulation proceeds [8–12]. One strategy employs windows, in which sampling is restricted to a subset of all of the systems of ultimate interest. Other design issues related to this approach are concerned with how to join efficiently the data taken from separate but overlapping windows [13–16].

A significant advance with density-of-states methods has come with the advent of transition matrix Monte Carlo (TMCM) methods [17–21]. The focus of these approaches is on attempted transitions (be they successful or not) between different systems in a Monte Carlo process that samples across them. The transition matrix Π measured this way can be used to evaluate the desired density of states, through the balance condition

$$\pi \Pi = \pi \quad (2)$$

Attempts are also being made now to develop hybrids that take the best features of TMCM and visited states methods [22]. It is interesting to consider that TMCM methods introduce a connection to work-based methods (discussed next), in that the focus of TMCM is on the energy (work) involved in transitioning from one state to another. This point has been made in more detail by Fenwick and Escobedo [23].

3. Work-based methods

Work-based methods can be derived from a very interesting relation recently presented by Jarzynski [24–28]. The formula relates the free energy difference ΔF to an average of exponential work values

$$\exp(-\beta \Delta F) = \overline{\exp(-\beta W_{A \rightarrow B})} \quad (3)$$

Here, $W_{A \rightarrow B}$ is the work associated with a process that transforms the system A into system B. It is a remarkable feature of this result that the equilibrium free energy difference can be related to this work, regardless of the nature of the transformation. In particular, the transformation need not be performed reversibly, and the equation is valid regardless of the path taken. Importantly, the result applies not for any single realization of the transformation, but only as an average (indicated by the overbar) for an ensemble of transformations beginning from an equilibrated system A. Still, it is remarkable that the formula applies without detailed knowledge of the non-equilibrium processes accompanying the transformation.

Two well-known limiting cases follow from the non-equilibrium work (NEW) formula (Eq. (3)). First, if the transformation is performed reversibly, then classical thermodynamics tells us that all realizations of the work calculation will give the same result, and $W_{A \rightarrow B} = \Delta F$. This limit encompasses the methods such as thermodynamic integration and adiabatic switching [29]. Second, if the transformation is performed instantaneously, then the work is given directly by the energy difference between the A and B systems for the configuration in place when the transformation is performed. The formula then can be expressed as an ensemble average over the A system:

$$\exp(-\beta \Delta F) = \langle \exp(-\beta \Delta U) \rangle_A \quad (4)$$

This is free energy perturbation (FEP) formula due to Zwanzig [30], and it is widely used. Note that the role of the A and B systems can be interchanged (as in the deletion versus insertion methods for the chemical-potential calculation), and the quality of the result obtained might vary significantly with the choice.

There is an interesting paradox that arises in considering a trivial application of the NEW method. We can apply Eq. (3) to a process of simple translation of a molecule from one point in a fluid to another, as illustrated in Fig. 2. System A has the molecule fixed at one point, while system B has it fixed at

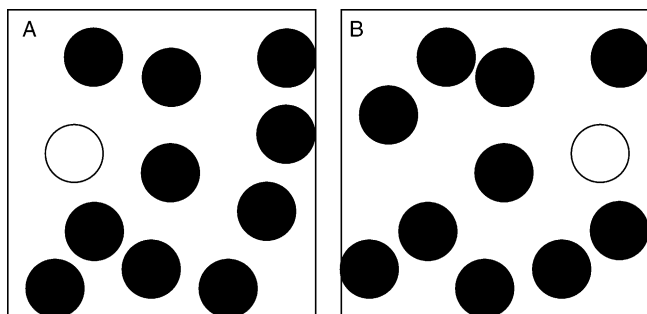


Fig. 2. Two systems differing only in the (fixed) position of one molecule (indicated by the white circle). Positions of the other molecules (colored black) are subject to ensemble averaging, and only representative configurations are indicated.

another. Obviously, for a homogeneous fluid in the absence of any external fields the free energy will be independent of the position of a single molecule, and $\Delta F = 0$.

The transformation from A to B can be accomplished in many ways. Let us consider specifically that the static particle is moved from its position in A to that in B by translating it at constant velocity. During this process, force is applied as needed to keep the particle's velocity constant as it interacts with the other molecules in the system (note that anything that happens after the molecule reaches its system B position does not contribute to the work and is irrelevant). The work then is the integral of this fluctuating force over the translation path. In the limit of a reversible process, the particle is moved infinitely slowly, and at each point along the way it encounters an ensemble-averaged net force of zero, and the $\Delta F = 0$ result is recovered.

One understands intuitively that if the particle is moved at a finite speed, it will encounter many more collisions on its "front" end, and it will experience a net retarding force as it moves along. Thus in the exponential average, by far the most common contribution is one in which $W_{A \rightarrow B}$ is positive (and perhaps large compared to kT), and consequently most contributions to Eq. (3) average will be near zero. If one does not perform a sufficient number of traversals, the average will be systematically in error, and a too large (nonzero and positive) ΔF will result. This error will be exacerbated as the molecule translation is performed at greater speeds, and it will also grow with the distance that the particle must translated as defined by the A and B systems. Yet Eq. (3) is correct, and it remains valid even in the face of these exacerbating factors that cause it to render an increasingly bad result in practice. This is the paradox. Part of the resolution is that the full realization of the average in Eq. (3) will include a few (perhaps a very few) terms in which the work is negative (and again perhaps large compared to kT). These terms make an enormous contribution to the average, and evidently just enough to offset the many near-zero contributions from the positive work realizations. A negative work can be explained as one in which the random motion of the fluid molecules produces enough "rear-end" collisions to result in a net "push" on the translated molecule. Failure to sample these rare but

very important contributions results in systematic error, or inaccuracy, in the calculated free energy.

This explanation has some merit, but we think it is incomplete. The pathology of molecule translation has much in common with that underlying the particle insertion and deletion inaccuracies presented in Section 1. In fact, in the limiting case of an infinite translation speed the source of the inaccuracy in the two cases must be the same. Consider first a coupled deletion/insertion calculation, in which the static molecule is deleted from its position in system A and immediately re-inserted at its position in system B. If the positions are sufficiently separated, and the system is large, the deletion and insertion averages will be independent, and the work average in Eq. (3) will reduce to a product of insertion and deletion averages each of the form given by Eq. (4). This product will be unity and in principle we recover $\Delta F = 0$. This calculation is equivalent to averaging the insertion and deletion results, and we learned from Fig. 1 that such a procedure is not reliably accurate. Yet there are no issues involved with respect to inadequate representation of contributions that involve many rear-end collisions—the molecule has no presence at points between its initial and final position, so there can be nothing to say about the types of events that might occur along such a path. All of the inaccuracy must lie with considerations arising exclusively at the point of insertion and deletion.

Indeed this is the case, and the matter is well understood [31,32]. Most of the problem lies on the deletion end (system A). Here very large contributions can be made by configurations in which one of the molecules overlaps the static molecule—the contribution is of order $\exp(+U_A/kT)$, where U_A is the (very large) configurational energy. The probability to observe such a configuration is $\exp(-U_A/kT)$, which precisely balances the contribution, and the product is a finite, non-negligible but never-seen contribution to the average.

Consider now the infinite-speed work process, in which the molecule actually does have a presence at intermediate points. However, we stipulate that it moves so fast that the fluid molecules in effect remain completely static as it completes its translation (this presumes the potential does not have a component involving true hard repulsion). In this case the work is—just as with the deletion/insertion method—given by the difference in energies at the start and end of the translation. All work contributions in between will sum to zero (the effect on the fluid molecules will be different, as they will experience a net acceleration in directions perpendicular to the path of the translating molecule; but these effects are not relevant to the work average). Given the equivalence of the calculation with and without the intermediate presence, we must conclude that the nature of the inaccuracy is the same. The error found at the infinite-speed limit must arise from the neglected overlap configurations at the outset of the translation process (and lesser effects at the other end). In fact, the manner that the error applies in this case can be seen in one of the ways the NEW formula is developed [28],

in which work values involving $\exp(+U_A/kT)$ contribute with probability $\exp(-U_A/kT)$.

Still, it is not easy to see how this error is attenuated as the limit of reversibility is approached. The same product of offsetting exponentials is present in the working formula, but it is clear that if the translation is performed slowly, they do not introduce inaccuracies. As the process becomes more reversible, the explanation for the inaccuracy must evolve to that involving the missed rear-end collisions.

Aside from the important conceptual advance represented by Jarzynski's formula, the NEW formalism to calculate the free energy is of special interest because it indicates an experimental approach to calculation of the free energy of molecular systems. So far, most attention has been paid to its application in molecule-pulling experiment, in which the (irreversible) work required to pull a long-chain molecule (such as RNA) is measured and used to evaluate the free energy of chain collapse and extension [33–36].

4. Distribution functions and phase space

It is useful to examine functions describing the distribution of work values observed in repeated realization of the work process described in Eq. (3) [37,38]. Distinct distributions apply for the $A \rightarrow B$ process and the $B \rightarrow A$ process, and we denote them $p_A(W)$ and $p_B(W)$; for both we define W to be in the sense of going from A to B, so $p_A(W) \equiv p_A(W_{A \rightarrow B})$ and $p_B(W) \equiv p_B(-W_{B \rightarrow A})$. In connection with Eq. (3), the free energy difference is given from these distributions by

$$\exp(-\beta\Delta F) = \int_{-\infty}^{+\infty} p_A(W) e^{-\beta W} dW \quad (5a)$$

$$\exp(+\beta\Delta F) = \int_{-\infty}^{+\infty} p_B(W) e^{+\beta W} dW \quad (5b)$$

Typical forms for the distributions are given in Fig. 3. The distributions are related [37,25,38]

$$p_A(W) e^{-\beta W} = p_B(W) e^{-\beta\Delta F} \quad (6)$$

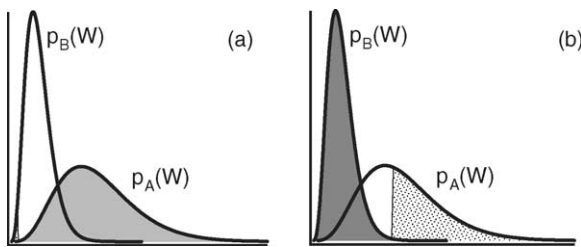


Fig. 3. Schematic depiction of non-equilibrium work distribution functions. Solid shaded regions indicate work values that will be well sampled in a finite-length simulation. Patterned shading describes the area of the conjugate distribution lying outside the region of the sampled distribution, and which characterizes the inaccuracy of the calculation. (a) Sampling is performed on the A system and the inaccuracy is small; (b) sampling is performed on the B system and the inaccuracy is large.

This equation is useful for understanding the asymmetry in the inaccuracy of the $A \rightarrow B$ and $B \rightarrow A$ NEW work calculations [39]. For example, when performing a $A \rightarrow B$ NEW calculation, the inaccuracy arises from neglect of the contributions in Eq. (5a) for large negative W , where the exponential is very large. We can apply a model then in which we assert that *all* of the inaccuracy is due to the neglect of contributions below a particular value of W , designated W^* , such that no sampling is contributed below this value, and perfect sampling is achieved for $W > W^*$. Then according to Eq. (6), the inaccuracy is given in terms of the area of the conjugate distribution, $p_B(W)$, lying below W^* . A simple probabilistic argument can be applied to evaluate the probability distribution for W^* given $p_A(W)$ in the case where M work measurements are performed toward Eq. (3). This is given by the probability density that the work value W^* is observed times the probability that all other $M - 1$ work measurements are greater than W^* :

$$p_A^*(W^*) = p_A(W^*) \left[\int_{W^*}^{+\infty} p_A(W) dW \right]^{M-1} \quad (7)$$

Then to compute the inaccuracy, one could take W^* as the mean or mode of p_A^* . Lu and Kofke [39] used the latter choice, for which W^* is given by the solution of

$$\left. \frac{d \ln p_A(W)}{dW} \right|_{W=W^*} = M p_A(W^*) \quad (8)$$

Then, given W^* , the model indicates that the inaccuracy δ in the simulation value of ΔF will be

$$\begin{aligned} \delta &\equiv (\Delta F^{\text{sim}} - \Delta F^{\text{exact}}) \\ &= -kT \ln \left[1 - \int_{-\infty}^{W^*} p_B(W) dW \right] \end{aligned} \quad (9)$$

A corresponding relation is given for the inaccuracy of the $B \rightarrow A$ work process. Thus from the work distributions, the expected performance of a simulation of given sampling length can be predicted. Tests have shown that this neglected-tail inaccuracy model (and improvements on it) works well in describing the performance of FEP calculations [39–41].

A qualitative conclusion can be taken from this model. Considering the distributions given by Fig. 3, we see that the inaccuracy is less when sampling is performed on the broader distribution. This is because it can encompass the narrower conjugate distribution more easily than the conjugate distribution encompasses it. Given that the inaccuracy is characterized by the amount of the conjugate distribution lying outside the sampling distribution, advantage is given to sampling the broader distribution [39,41].

Another perspective on the inaccuracy problem can be made by considering the phase space of the system of interest [31] (we use “phase space” to refer to the configurations of the molecules, ignoring effects related to momenta). We consider that the A and B systems are defined on the same phase space

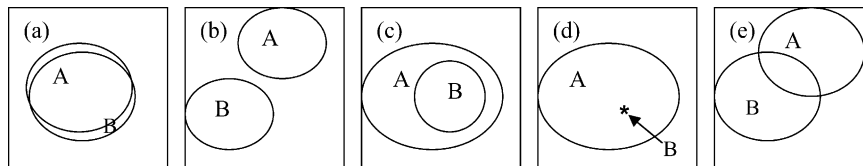


Fig. 4. Schematic representation of phase-space relations for two systems. Each circle represents the set of configurations important to the A and B systems, as indicated. (a) The important regions coincide; (b) no overlap between important regions; (c) configurations important to B are a subset of those important to A; (d) a special case of (c), in which the B-important region is a very small subset of the A-important region; (e) regions overlap without having a subset relation.

(if defined appropriately, this stipulation need not limit the generality of the discussion). We are interested in collecting all the phase-space points that are “important” to the A and B systems, respectively, and we consider how these sets of points relate to each other. The possibilities are illustrated schematically in Fig. 4 [42,31,7,39].

For the instantaneous-change FEP limit, only A and B are defined, and the full picture is as given by Fig. 4. However, one should consider that the work process takes the important phase space from A to B (or vice versa) though a sequence of systems having their own important phase-space regions, and these are not indicated in the figure. For the reversible limit, the sequence follows a smooth path, and one system differs only infinitesimally from the one preceding it; each pair of adjacent intermediates then have phase spaces relating as in Fig. 4(a). Otherwise, the intermediates may relate to each other as in any of the cases indicated by the figure. Inaccuracy results from situations in which adjacent work states have any phase-space relation other than a simple subset (Fig. 4(c) or (a)). In particular, in cases indicated by Fig. 4(b) and (e), there are important regions of one system that are practically inaccessible to the one preceding it. These unsampled conditions have very large, negative contributions to the work, and thus will make large (albeit unrealized) contributions to the free energy calculation. This problem was summarized above in the discussion of the molecule-translation paradox (for which Fig. 4(e) most likely applies). Here we simply present it in the context of phase-space overlap.

The phase-space viewpoint is useful because it can provide a basis for applying physical intuition when confronting a new free energy calculation. It also provides a conceptual framework for devising staging methods appropriate to a par-

ticular system. However, to make use of this picture it is necessary to know the phase-space relation for the system at hand. An appropriate metric can bring together the distribution function and the phase-space views. Information theory [43] tells us that a suitable set of measures for the “distance” between two distributions is

$$s_A \equiv \int d\Gamma p_A(\Gamma) \ln \left[\frac{p_A(\Gamma)}{p_B(\Gamma)} \right],$$

$$s_B \equiv \int d\Gamma p_B(\Gamma) \ln \left[\frac{p_B(\Gamma)}{p_A(\Gamma)} \right] \quad (10)$$

Here the distributions p_A and p_B are defined on ($3N$ -dimensional) phase space, so for example $p_A(\Gamma) \equiv e^{-\beta U_A(\Gamma)} / Q_A$ where Q_A is the partition function of the A system for which the configurational energy is $U_A(\Gamma)$. Using these measures, we can identify the cases indicated by Fig. 4: coincidence (a), both s_A and s_B near zero; incomplete overlap, (b) or (e), both s_A and s_B very large; and subset, (c) or (d), s_A finite and s_B very large. We expect that with some experience in using these quantities we can calibrate them to better distinguish the different cases.

It is easy to show that these measures are equal to the same measures defined on the (one-dimensional) energy-perturbation distribution space (i.e., $p_A(W)$ and $p_B(W)$ when applied to FEP, Eq. (4)):

$$s_A = \int d(\Delta U) p_A(\Delta U) \ln \left[\frac{p_A(\Delta U)}{p_B(\Delta U)} \right],$$

$$s_B = \int d(\Delta U) p_B(\Delta U) \ln \left[\frac{p_B(\Delta U)}{p_A(\Delta U)} \right] \quad (11)$$

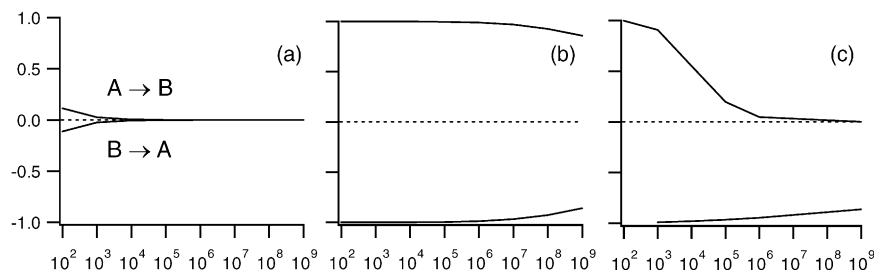


Fig. 5. Convergence of FEP averages. The abscissa is the number of work cycles or free energy perturbations performed. The ordinate is a measure of the inaccuracy, specifically $\pm[\exp(-\beta\delta) - 1]$ (see Eq. (9)), where the “-” is used for the A \rightarrow B path, and the “+” is used for B \rightarrow A. (a) $s_A = s_B = 0.25$; (b) $s_A = s_B = 25$; (c) $s_A = 42$, $s_B = 7.8$.

Table 2
Summary of staging methods for non-equilibrium work calculations

Method	Formula for $e^{-\beta\Delta F}$	Preferred staging potential, $e^{-\beta U_C}$
Umbrella sampling	$\frac{e^{-\beta W_{C \rightarrow B}}}{e^{-\beta W_{C \rightarrow A}}}$	$(e^{-2\beta(U_A - F_A)} + e^{-2\beta(U_B - F_B)})^{1/2}$
Overlap sampling	$\frac{e^{-\beta W_{A \rightarrow C}}}{e^{-\beta W_{B \rightarrow C}}}$	$(e^{+\beta(U_A - F_A)} + e^{+\beta(U_B - F_B)})^{-1}$
Funnel sampling	$e^{-\beta W_{A \rightarrow C}} \times e^{-\beta W_{C \rightarrow B}}$	No simple formulation

We can therefore connect the overlap measures to the expected inaccuracy of the calculations. Using a simple model involving perturbation between two systems of N independent harmonic oscillators, we can demonstrate expected convergence rates for different cases [5]. An example is provided in Fig. 5. Fig. 5(a) corresponds to a system of the type depicted in Fig. 4(a), Fig. 5(b) is a system of type Fig. 4(e), while Fig. 5(c) is of type Fig. 4(d). The model shows how both types of behavior displayed in Fig. 1 can arise. The LJ example is similar to Fig. 5(c) (and thus indicating the small-subset phase-space relation) and the water example looks like Fig. 5(b) (indicating a partial-overlap phase-space relation).

5. Staging methods

In practice, one is presented with a pair of systems A and B and desires to evaluate their free energy difference ΔF . If their phase-space relation is not as in Fig. 4(a) or (c), a straightforward application of an FEP calculation, and probably any irreversible NEW calculation, will yield inaccurate results. In such circumstances, it is necessary to break the calculation up into smaller parts, each involving a pair of systems that will yield an accurate result for the respective free energy difference. This is known as “staging”, or perhaps “stratification”. Often the term “umbrella sampling” is applied also, but this appellation is sometimes misused in this context.

If a single intermediate is employed to break the calculation into two stages, there are three ways in which the overall free energy might be calculated. Designating the intermediate stage as C, the options are summarized in Table 2. The preferred staging potential is formulated such that each intermediate free energy calculation is between systems having phase-space relations as in Fig. 4(a) or (c). The overlap sampling method [44–46], when applied using the given staging potential, is equivalent to Bennett’s method [37,6]. An

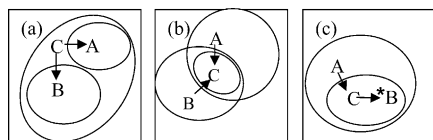


Fig. 6. Phase-space depiction of appropriate intermediate system (labeled C) for two-stage free energy calculation constructed between two systems A and B. (a) Umbrella sampling; (b) overlap sampling; (c) funnel sampling. Arrows from one letter to another indicate the direction of sampling for each stage.

optimization similar to that used in formulating this potential can be applied to the umbrella- and funnel-sampling schemes. The umbrella-sampling form is also simple, but that for funnel-sampling involves the solution to a cubic polynomial and does not seem to give a practical specification. The phase-space picture corresponding to the systems A, B, and C is depicted in Fig. 6.

The same line of thinking can guide the formulation of multiple-stage schemes. Here there may be several intermediates required to transition from A to B, and it is important that in each intermediate free energy calculation a subset relation is obeyed for the phase spaces of the relevant systems. Fig. 7 shows both inappropriate and appropriate choices. Part (a) is a scheme often described as “umbrella sampling”, but we prefer to reserve this designation for staging schemes of the type illustrated in Fig. 6(a). Fig. 7(b) shows how intermediate stages can be set up effectively in conjunction with overlap-sampling ideas. In this connection, it is interesting to remember that the Ferrenberg–Swendsen histogram-reweighting algorithm gives results equivalent to Bennett’s method when applied using only two histograms [14]. The NEW staging approach illustrated in Fig. 7(b) might be viewed as a generalization of overlap sampling that relates to the multi-histogram Ferrenberg–Swendsen density-of-states formulation.

The multistaging ideas presented here are traditionally developed in connection to FEP calculations, but they can also be applied in more general NEW calculations. In a NEW calculation one is free (at least in the context of molecular simulation) to choose the path defining the work process. The considerations above indicate that the path should be selected to adhere to the subset relation at each point along the work

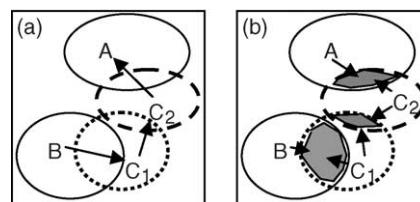


Fig. 7. Phase-space depiction of strategies for multistaging, using two explicit intermediate systems, labeled C_1 and C_2 . Circles represent regions of phase space important to each system (broken lines are used for intermediate to improve clarity of the figure). Arrows indicate the direction of the work or perturbation calculation. (a) An inappropriate staging scheme, in which intermediates do not satisfy a subset relation, and direct perturbations are performed from one to the next. (b) An appropriate staging scheme, in which overlap sampling is used to measure free energy differences between adjacent stages (thereby implicitly introducing three new stages as indicated by the shaded overlap regions).

trajectory. If the A and B systems have some phase-space overlap, then a NEW implementation of overlap sampling can be applied. The work path can be defined as (cf. Table 2)

$$\exp(-\beta U(\gamma)) = ((1 - \gamma)e^{+\beta U_A} + \gamma e^{+\beta U_B})^{-1} \quad (12)$$

where γ goes from 0 to 1 in traversing the path from A to B. The results from optimizing the overlap-sampling FEP method suggest that the best value of γ obeys

$$\frac{\gamma}{1 - \gamma} = \exp(-\beta \Delta F) \quad (13)$$

A typical NEW calculation involves repeated traversals of γ from 0 to 1 and back, so a complete $\Delta F(\gamma)$ profile (which in principle should be independent of γ) is available via the overlap-sampling formula in Table 2, and Eq. (13) can be solved self-consistently with the calculation of ΔF . This approach has been demonstrated to calculate free energies of ion charging in water, and for mutation of an adenosine molecule [46]. Issues regarding the NEW path have also been examined in relation to path integral methods [47].

6. Concluding remarks

We have reviewed methods for calculating free energies via molecular simulation. Our emphasis has been on understanding the performance of work-based methods, particularly their accuracy. Yet the ideas presented here might also find value in consideration of density-of-states methods, particularly in relation to the formulation of paths joining states of interest (to the extent that freedom exists to select the path in a given problem). For all approaches, further progress can be made by developing and applying methods to quantify phase-space overlap. Initial steps in that direction have been presented here. A good understanding of the phase-space relation for a system of interest can be helpful in guiding and interpreting the application of free energy methods used to study it. Such knowledge can contribute to the formulation of an effective staging scheme, one yielding accurate results with a minimum of computational effort.

Acknowledgments

This work is supported by the Division of Chemical Sciences, Office of Basic Energy Sciences of the U.S. Department of Energy (contract DE-FG02-96ER14677). I am grateful to Nandou Lu and Di Wu for contributions to this work. The molecule-translation paradox was brought to mind by Alan Mark. I also wish to acknowledge valuable conversations with Chris Jarzynski, particularly (but not only) regarding the measures introduced in Eq. (10) and their relation to Eq. (11).

References

- [1] H.B. Callen, *Thermodynamics and An Introduction to Thermostatistics*, Wiley, New York, 1985.
- [2] D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, Oxford, 1987.
- [3] R.J. Gdanitz, *Curr. Opin. Solid State Mater. Sci.* 3 (1998) 414.
- [4] Z.L. Zhang, M.A. Horsch, M.H. Lamm, S.C. Glotzer, *Nano Lett.* 3 (2003) 1341.
- [5] D. Wu, D.A. Kofke, *J. Chem. Phys.*, in preparation.
- [6] D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, San Diego, 2002.
- [7] D.A. Kofke, D. Frenkel, *Perspective: free energies and phase equilibria*, in: S. Yip (Ed.), *Handbook of Molecular Modeling*, Kluwer Academic Publishers, Dordrecht, 2005.
- [8] A. Laio, M. Parrinello, *Proc. Nat. Acad. Sci.* 99 (2002) 12562.
- [9] J. Lee, *Phys. Rev. Lett.* 71 (1993) 211.
- [10] F. Wang, D.P. Landau, *Phys. Rev. E* 64 (2001) 0561011.
- [11] F. Wang, D.P. Landau, *Phys. Rev. Lett.* 86 (2001) 2050.
- [12] Q. Yan, J.J. de Pablo, *Phys. Rev. Lett.* 90 (2003) 0357011.
- [13] A.M. Ferrenberg, D.P. Landau, R.H. Swendsen, *Phys. Rev. E* 51 (1995) 5092.
- [14] A.M. Ferrenberg, R.H. Swendsen, *Phys. Rev. Lett.* 63 (1989) 1195.
- [15] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, *J. Comp. Chem.* 13 (1992) 1011.
- [16] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, *J. Comp. Chem.* 16 (1995) 1339.
- [17] J.R. Errington, *J. Chem. Phys.* 118 (2003) 9915.
- [18] J.R. Errington, *Phys. Rev. E* 67 (2003) 012102.
- [19] M. Fitzgerald, R.R. Picard, R.N. Silver, *Europhys. Lett.* 46 (1999) 282.
- [20] M. Fitzgerald, R.R. Picard, R.N. Silver, *J. Stat. Phys.* 98 (2000) 321.
- [21] J.-S. Wang, T.K. Tay, R.H. Swendsen, *Phys. Rev. Lett.* 82 (1999) 476.
- [22] M.S. Shell, P.G. Debenedetti, A.Z. Panagiotopoulos, *J. Chem. Phys.* 119 (2003) 9406.
- [23] M.K. Fenwick, F.A. Escobedo, *J. Chem. Phys.* 120 (2004) 3066.
- [24] G.E. Crooks, *J. Stat. Phys.* 90 (1998) 1481.
- [25] G.E. Crooks, *Phys. Rev. E* 61 (2000) 2361.
- [26] D.J. Evans, *Mol. Phys.* 101 (2003) 1551.
- [27] C. Jarzynski, *Phys. Rev. E* 56 (1997) 5018.
- [28] C. Jarzynski, *Phys. Rev. Lett.* 78 (1997) 2690.
- [29] M. Watanabe, W.P. Reinhardt, *Phys. Rev. Lett.* 65 (1990) 3301.
- [30] R.W. Zwanzig, *J. Chem. Phys.* 22 (1954) 1420.
- [31] D.A. Kofke, P.T. Cummings, *Mol. Phys.* 92 (1997) 973.
- [32] D.A. Kofke, P.T. Cummings, *Fluid Phase Equilib.* 150 (1998) 41.
- [33] G. Hummer, A. Szabo, *Proc. Natl. Acad. Sci. U.S.A.* 98 (2001) 3658.
- [34] J. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante, *Science* 296 (2002) 1832.
- [35] J.T. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante, *Biophys. J.* 82 (2002) 193a.
- [36] B. Onoa, J.T. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C.J. Bustamante, *Biophys. J.* 82 (2002) 49a.
- [37] C.H. Bennett, *J. Comp. Phys.* 22 (1976) 245.
- [38] K.S. Shing, K.E. Gubbins, *Mol. Phys.* 46 (1982) 1109.
- [39] N.D. Lu, D.A. Kofke, *J. Chem. Phys.* 114 (2001) 7303.
- [40] D. Wu, D.A. Kofke, *J. Chem. Phys.*, in press.
- [41] D. Wu, D.A. Kofke, *Phys. Rev. E*, in press.
- [42] D.A. Kofke, *Mol. Phys.* 102 (2004) 405.
- [43] R.M. Gray, *Entropy and Information Theory*, Springer-Verlag, New York, 1990.
- [44] N.D. Lu, D.A. Kofke, T.B. Woolf, *J. Comp. Chem.* 25 (2004) 28.
- [45] N.D. Lu, J.K. Singh, D.A. Kofke, *J. Chem. Phys.* 118 (2003) 2977.
- [46] N.D. Lu, D. Wu, T.B. Woolf, D.A. Kofke, *Phys. Rev. E* 69 (2004) 0577021.
- [47] S.X. Sun, *J. Chem. Phys.* 118 (2003) 5769.