Satisfying the fluctuation theorem in free-energy calculations with Hamiltonian replica exchange

Matthew A. Wyczalkowski and Rohit V. Pappu
Department of Biomedical Engineering and Center for Computational Biology, Washington University in St. Louis, St. Louis, Missouri 63130, USA

(Received 20 November 2007; published 11 February 2008; corrected 20 February 2008)

An error measure, referred to as the hysteresis error, is developed from the Crooks fluctuation theorem to evaluate the sampling quality in free-energy calculations. Theory and the numerical free energy of hydration calculations are used to show that Hamiltonian replica exchange provides a direct route for minimizing the hysteresis error. Replica exchange swap probabilities yield the rate at which the hysteresis error falls with the simulation length, and this result can be used to decrease bias and statistical errors associated with free-energy calculations based on multicanonical simulations.

DOI: 10.1103/PhysRevE.77.026104 PACS number(s): 82.20.Wt

I. INTRODUCTION AND OVERVIEW

The free energies of solvation provide quantitative assessments of the driving forces for spontaneous processes such as protein folding, binding, self-assembly, and solubility. Formally, the free energy of solvation in the canonical ensemble is the free-energy change ΔF associated with the transfer of a solute from the gas phase to a fixed position in the solvent [1]. Operationally, one has access to a range of techniques to obtain estimates for ΔF [2,3]. Kirkwood [4] showed that one could introduce arbitrary parameters into potential functions and continuously vary the degree of coupling between specific molecules in a dense fluid. The device of coupling parameters leads to simple expressions for the chemical potentials of any component of the fluid. If the component is the solute molecule, which is transferred from the gas phase into the solvent, then a single coupling parameter λ, where $0 \leq \lambda \leq 1$, modulates solute-solvent interactions in the system's potential function. The limits $\lambda=0$ and $\lambda=1$ correspond to the pure solvent and solvent plus fully grown solute, respectively. Intermediate values of λ correspond to potential functions that include only a part of the solute-solvent interactions. The Kirkwood coupling parameter plays a central role in equilibrium methods for calculating ΔF. One carries out a series of independent canonical simulations where each simulation is associated with a distinct potential function, characterized by a specific λ value. As it samples the equilibrium ensemble, each simulation generates a series of work values, which are then used to estimate the free-energy change across the entire λ schedule.

The multicanonical approach described above takes advantage of the simple formalism developed by Kirkwood for calculating ΔF. However, in practice, standard free-energy calculations based on multicanonical simulations are plagued by slow convergence and inaccurate estimates of ΔF [5]. Errors may be divided into statistical and bias (or finite sampling) errors [6]. The former stem from the fluctuations of the free-energy estimator and can be estimated by block averaging or bootstrap methods [7,8]. Since the statistical error decreases as the inverse square root of the simulation length, it is frequently used as an indicator of the convergence of the multicanonical simulation. While statistical errors are random fluctuations of short simulation results about some mean value, the bias error is an error of the mean value itself and it changes with simulation length. As discussed by Zuckerman and Woolf [9], bias errors have two causes: The free-energy estimates are nonlinear averages, and the work distributions on which such estimates are based will typically have long tails which are rarely sampled, and yet these are important to the average. The latter point is important: Rare events dominate free-energy estimates, and one seldom observes these events in short simulations. As a result, the average drifts with simulation length, resulting in inaccurate estimates for ΔF from bias error even when the statistical error is small. The magnitude of the bias error is difficult to quantify directly, as it requires knowledge of the actual free-energy difference, the very quantity we wish to determine. Furthermore, small fluctuations in the estimate for ΔF may not be indicative of convergence, but rather of inadequate sampling of the rare but important configurations. To address these problems, we develop an alternate measure of free-energy error, one based on deviations from equilibrium distributions.

Crooks [10] derived a fluctuation theorem (Appendix A 1) valid for stochastic, microscopically reversible dynamics, which relates the distribution of dissipated work values along a forward and reverse path as

$$\exp(\beta W_D) = \frac{P_F(\beta W_D)}{P_R(-\beta W_D)}. \tag{1}$$

Here, $\beta=(k_B T)^{-1}$, $P_F(\beta W_D)$ is the probability distribution for dissipated work associated with switching λ from λ_0 to λ_1, and $P_R(-\beta W_D)$ is the corresponding distribution for the reverse process. If the canonical simulations for each value of λ sample the equilibrium ensemble adequately, then the distributions of dissipated work obtained over the course of free-energy calculations will satisfy Eq. (1).

In this work, we develop a readily measured error estimate, the hysteresis error ϵ_H, which quantifies the degree to which the observed work distributions obey the Crooks fluctuation theorem. Hamiltonian replica exchange, a multicanonical equilibration technique, effectively reduces the hysteresis error. We relate the average replica exchange swap probability to the degree of overlap between equilibrium ensembles, as well as to the rate at which ϵ_H falls. Based on

this, we may construct an optimized λ schedule to further minimize the hysteresis error for an entire simulation.

The remainder of this presentation is organized as follows: The theory section introduces the hysteresis error in the context of the Crooks fluctuation theorem followed by a formal illustration of how Hamiltonian replica exchange minimizes ϵ_H, the definition of swap probability as a measure of the overlap between different equilibrium ensembles, and a connection between the amount of overlap and minimization of ϵ_H. We calculate the free energy of hydration for acetamide to demonstrate how to estimate ϵ_H and minimize this error using replica exchange coupled to standard multicanonical simulations. We conclude with a summary and a discussion of the features of our methodology.

II. THEORY

A. Background

The free energy of replica i in the canonical ensemble at temperature T, whose potential $U_i(\Gamma)=U(\Gamma,\lambda_i)$ is a function of system configuration Γ and the parameter λ_i, is formally given as [11]

$$F_i = -\beta^{-1} \ln\left\{\int d\Gamma \exp[-\beta U_i(\Gamma)]\right\}. \tag{2}$$

At equilibrium, the probability of observing configuration Γ is given as

$$\rho_i(\Gamma) = \exp\{\beta[F_i - U_i(\Gamma)]\}. \tag{3}$$

To calculate the free-energy change δF associated with switching the Hamiltonian from U_0 to U_1 we perform simulations at λ_0 and λ_1 and calculate the forward and reverse work as

$$W^F(\Gamma) = U_1(\Gamma) - U_0(\Gamma), \tag{4a}$$

$$W^R(\Gamma) = U_0(\Gamma) - U_1(\Gamma). \tag{4b}$$

For the forward and reverse work values the configuration Γ is typically drawn from the equilibrium ensemble of U_0 and U_1, respectively. The free-energy perturbation (FEP) method [12] utilizes forward and reverse work distributions to provide two independent estimators for δF,

$$\delta F^F_{FEP} = -\beta^{-1} \ln\langle\exp(-\beta W^F)\rangle_0, \tag{5a}$$

$$\delta F^R_{FEP} = +\beta^{-1} \ln\langle\exp(-\beta W^R)\rangle_1, \tag{5b}$$

where the forward estimator δF^F_{FEP} utilizes forward work values from the simulation at U_0 and the reverse estimator the reverse work from U_1. Note that in both cases δF is associated with the process of switching $\lambda_0 \rightarrow \lambda_1$. These two estimators have different convergence rates [6]. Therefore, while in practice the two estimates should be equal, in simulations with finite sampling they are generally different.

Another free-energy estimator, the Bennett acceptance ratio [13], uses both the W^F and W^R distributions to obtain a free-energy estimate. It is generally more accurate [14] and is employed later in this paper for numerical free-energy estimates, but will not be considered for theoretical development.

B. Hysteresis error

The hysteresis error ϵ_H is defined as the difference between the forward and reverse δF_{FEP} estimates,

$$\epsilon_H \equiv \delta F^F_{FEP} - \delta F^R_{FEP}. \tag{6}$$

ϵ_H has contributions from both the statistical and bias error of the FEP estimators [6,9]. The bias error of the two estimators is typically in the opposite direction. While the statistical error may dominate the ϵ_H for a given simulation, in averages over multiple short simulations the dominant contribution to the average hysteresis error is the sum of the forward and reverse FEP bias.

We take ϵ_H as a measure of sampling quality and aim to minimize its magnitude between all pairs of neighboring replicas. The validity of using ϵ_H as a general sampling error is based on a relationship between it and the fluctuation theorem of Crooks, Eq. (1), derived below.

Switching the parameter $\lambda_0 \rightarrow \lambda_1$ (and vice versa) is equivalent to performing nonequilibrium work; the difference between the work performed and the free-energy change of the system is the dissipated work, defined in the forward and reverse directions as

$$W^F_D(\Gamma) = W^F(\Gamma) - \delta F, \tag{7a}$$

$$W^R_D(\Gamma) = W^R(\Gamma) + \delta F. \tag{7b}$$

Crooks [10] equates W^F_D and W^R_D to the entropy production caused by changing $\lambda_0 \rightarrow \lambda_1$ and $\lambda_1 \rightarrow \lambda_0$, respectively, for the given configuration.

The distributions $P_F(W_D)$ and $P_R(W_D)$ give the probability of realizing a specific value for the dissipated work in the forward and reverse directions, respectively. The distributions are related to each other by the fluctuation theorem shown in Eq. (1), which we have rederived in Appendix A 1 for the specific case of instantaneous switching between configurations with different λ values. In practice, Eq. (1) will not be satisfied exactly because of errors due to finite sampling. To take simulation errors into account, we rewrite Eq. (1) with an arbitrary error term ϵ^*_{FT} and with observed (rather than ideal) dissipated work distributions P^*_F and P^*_R,

$$\exp[\beta W_D + \beta\epsilon^*_{FT}(W_D)] = \frac{P^*_F(\beta W_D)}{P^*_R(-\beta W_D)}. \tag{8}$$

Equation (8) is constructed such that the Crooks fluctuation theorem is recovered and $\epsilon^*_{FT}=0$ when the observed work distributions match the correct distributions. The hysteresis error ϵ_H and the fluctuation error ϵ^*_{FT} are related to each other as (see Appendix A 2)

$$\epsilon_H = -\beta^{-1} \ln\langle\exp(-\beta\epsilon^*_{FT})\rangle^*_0, \tag{9}$$

where $\langle\cdots\rangle^*$ is defined as the average obtained from a finite simulation. The more closely a simulation obeys the relationship (1), the smaller the hysteresis error ϵ_H and vice versa. In the next section, we will discuss methods to reduce ϵ_H,

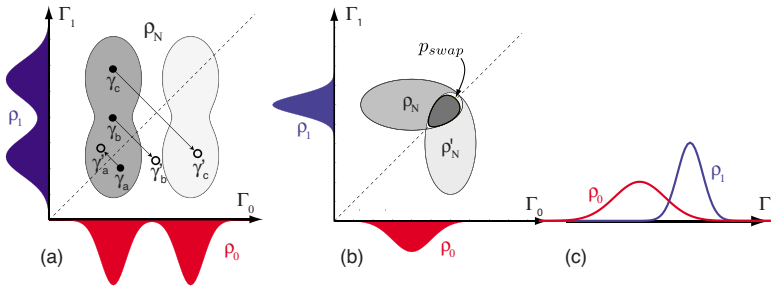

FIG. 1. (Color online) A graphical representation of replica exchange. (a) The independent (high-dimensional) configuration spaces Γ_0 and Γ_1 have probability distributions ρ_0 and ρ_1, respectively, and the joint equilibrium ensemble ρ_N is drawn over this domain. The ρ_0 system has a kinetic barrier (represented by the two disconnected lobes), and with no replica exchange the system explores only the configurations of the shaded domain. A replica exchange swap is a reflection of the configuration pair γ about the $\Gamma_0=\Gamma_1$ diagonal axis, and three swap attempts are shown: The configuration pair γ_a swaps successfully and becomes γ'_a, but it does not sample otherwise inaccessible regions; a swap of γ_b fails because γ'_b is not in the equilibrium ensemble; and the swap of γ_c succeeds and allows the system to explore otherwise inaccessible regions of phase space. (b) The equilibrium domain ρ_N and its swapped image ρ'_N are drawn. Swaps are feasible only for configuration pairs which belong to both ρ_N and ρ'_N. This overlap region, labeled p_{swap}, is the domain where the integrand of Eq. (17b) is large and its integral corresponds to the average swap probability $\langle p_{swap}\rangle$. (c) The overlap of the ρ_0 and ρ_1 distributions along the common configuration $\Gamma_0=\Gamma_1$. For the hysteresis error to converge, the λ_0 simulation must observe configurations where $\rho_1>\rho_0$ and the λ_1 simulation must adequately sample the region $\rho_0>\rho_1$. The frequency with which this occurs is approximately given by $\langle p_{swap}\rangle$.

which in turn leads to the satisfaction of the Crooks fluctuation theorem.

C. Replica exchange

In a Hamiltonian replica exchange [15,16] simulation, Monte Carlo moves are employed to exchange configurations Γ (or, equivalently, parameters λ) between two replicas with probability

$$P_{swap} = \min[1, \exp(-\beta \Delta U_{swap})], \tag{10}$$

where

$$\Delta U_{swap} = U_0(\Gamma_1) + U_1(\Gamma_0) - U_0(\Gamma_0) - U_1(\Gamma_1) \tag{11a}$$

$$= W^F + W^R \tag{11b}$$

$$= W_D^F + W_D^R. \tag{11c}$$

Γ_0 and Γ_1 denote configurations drawn at random from the equilibrium ensembles of U_0 and U_1, respectively. For convenience, we write $\gamma=(\Gamma_0,\Gamma_1)$ as a pair of such configurations and $\gamma'=(\Gamma_1,\Gamma_0)$ is the swapped configuration pair.

Since Γ_0 and Γ_1 are independent configurations, we can consider the probability of sampling Γ_0 in the equilibrium ensemble of U_0 *and* sampling Γ_1 in the equilibrium ensemble of U_1; this is the native probability $\rho_N(\gamma)$. Analogously, the joint probability of sampling the swapped configurations, Γ_1 from ρ_0 and Γ_0 from ρ_1 is given as $\rho'_N(\gamma)$

$$\rho_N(\gamma) = \rho_0(\Gamma_0)\rho_1(\Gamma_1), \tag{12a}$$

$$\rho'_N(\gamma) = \rho_0(\Gamma_1)\rho_1(\Gamma_0) = \rho_N(\gamma'). \tag{12b}$$

Replica exchange swaps are conveniently visualized by plotting the independent configurations Γ_0 and Γ_1 along orthogonal axes and the equilibrium ensemble of the system as an isocontour of ρ_N, illustrated in Fig. 1(a).

At equilibrium, the relative probability of observing a pair of replicas in their swapped versus native configurations is

$$\frac{\rho'_N}{\rho_N} = \exp(-\beta \Delta U_{swap}), \tag{13}$$

which is derived from definitions (12), (3), and (11a). We will refer to this as an interreplica equilibrium relationship.

In an infinitely long simulation, (13) will be satisfied exactly, but this will generally not be the case for finite simulations, where inadequate sampling of configuration space will result in inaccurate probability estimates. However, in simulations with replica exchange we expect the interreplica equilibrium relationship to be satisfied more closely than in simulations without replica exchange, because the swap move distributes configuration pairs in such a way as to satisfy Eq. (13). To illustrate, consider the system in Fig. 1(a) where the U_0 replica is presumed to be stuck in the left lobe of the ρ_0 distribution because of a kinetic barrier. Without replica exchange, only the heavily shaded region of ρ_N will be sampled accurately. The simulation will not have a correct estimate for $\rho'_N(\gamma_c)=\rho_N(\gamma'_c)$, since ρ_0 for the swapped configuration, never having been observed, will be inaccurate. Consequently, Eq. (13) will not hold. Replica exchange directly populates swapped configurations (e.g., γ'_c), thereby improving the statistics of ρ'_N and allowing interreplica equilibrium to be achieved more quickly for all configurations in ρ_N.

The degree to which Eq. (13) is satisfied determines the magnitude of the hysteresis error. To illustrate this, suppose that the distribution ρ'_N has some small error $\rho_\epsilon(\Gamma_0,\Gamma_1)$ due to finite sampling, so that we write $(\rho'_N+\rho_\epsilon)$ as the numerator in Eq. (13). In Appendix A 3 we show, by integrating over all configuration pairs, that the relationship between the hysteresis error and the error of sampling the swapped distribution, ρ_ϵ is

$$\epsilon_H \simeq -\beta^{-1}\int d\Gamma_0 d\Gamma_1 \rho_\epsilon. \tag{14}$$

The hysteresis error, then, will be minimized when the estimated swapped configuration probabilities ρ'_N are consistent with the equilibrium distribution. Since replica exchange populates the swapped configurations directly, it provides an efficient route to minimizing ϵ_H.

D. Swap probability

Analysis of the average replica exchange swap probability is complicated by the fact that the Metropolis function [Eq. (10)] is not analytical. For the purposes of interpreting this quantity, we will instead consider the Fermi swap probability

$$p_{swap} = f(\beta \Delta U_{swap}),$$

where $f(x)$ is defined as

$$f(x) = 1/[1+\exp(x)]. \tag{15}$$

(See [13] for a discussion.) We use p_{swap} to denote the Fermi swap probability and P_{swap} for the Metropolis swap probability; while the theoretical development uses p_{swap}, replica exchange moves are accepted or rejected using P_{swap}. A simulation with either the Metropolis or Fermi swap probability will yield a Boltzmann distribution of swapped and unswapped configurations [Eq. (13)]. While the exact numerical values of the Fermi and Metropolis swap probabilities will differ somewhat, their qualitative behavior and the conclusions drawn here will hold for both.

The average Fermi swap probability for two systems evolving independently is

$$\langle p_{swap}\rangle \equiv \langle\langle f(\beta\Delta U_{swap})\rangle_0\rangle_1 \tag{16a}$$

$$= \int d\Gamma_0 d\Gamma_1 \rho_N f(\beta \Delta U_{swap}), \tag{16b}$$

which can be written as

$$\langle p_{swap}\rangle = \left\langle \left\langle \frac{\rho'_N}{\rho_N+\rho'_N}\right\rangle_0\right\rangle_1 \tag{17a}$$

$$= \int d\Gamma_0 d\Gamma_1 \frac{\rho_N \rho'_N}{\rho_N+\rho'_N}. \tag{17b}$$

The integrand of (17b) is a normalized probability of observing a given configuration pair, and the average swap probability is then the overlap of ρ_N and ρ'_N. See Fig. 1(b) for a graphical interpretation. Thus, a large average swap probability implies a large overlap between the equilibrium distributions of the two replicas and a low $\langle p_{swap}\rangle$ indicates that the configurations these replicas adopt are distinct.

We can expand (16a) in a Taylor series about $\lambda = \lambda_0 + \delta_\lambda$. To leading order in δ_λ, we find that in the neighborhood of λ_0 the average swap probability is (see Appendix A 4)

$$\langle p_{swap}\rangle \simeq \frac{1}{2} - \frac{\beta^2\delta_\lambda^2}{4} C_\lambda, \tag{18}$$

where

$$C_\lambda \equiv \mathrm{var}\left(\frac{\partial U}{\partial \lambda}\right) = \langle(\partial U/\partial\lambda)^2\rangle_0 - \langle \partial U/\partial\lambda\rangle_0^2.$$

C_λ, then, determines the rate at which the average swap probability declines as the difference in λ between the two replicas, δ_λ, increases, although this linear analysis is accurate only for small δ_λ.

E. Swap probability and the hysteresis error convergence rate

We now demonstrate that the average swap probability between two replicas gives a measure of how quickly the hysteresis error decreases, on average, over the course of a simulation. The hysteresis error is the difference between the forward and reverse δF_{FEP}, and since the forward and reverse FEP estimators do not converge at equal rates [6], it is the slower of these which governs the convergence of ϵ_H.

We may rewrite Eq. (5a) as

$$\langle \exp(-\beta W_D^F)\rangle_0 = 1. \tag{19}$$

For this to hold, we must sample configurations where $W_D^F < 0$; since the dissipated work is on average greater than zero by the second law of thermodynamics, such configurations tend to be rare [17]. As a result, the convergence rate of δF_{FEP}^F is governed by the probability of observing negative dissipated forward work values. Likewise, the convergence of δF_{FEP}^R is dictated by observations of $W_D^R < 0$. We can understand this criterion graphically with the relationships (see Appendix A 1)

$$\frac{\rho_0(\Gamma_0)}{\rho_1(\Gamma_0)} = \exp[\beta W_D^F(\Gamma_0)], \tag{20a}$$

$$\frac{\rho_1(\Gamma_1)}{\rho_0(\Gamma_1)} = \exp[\beta W_D^R(\Gamma_1)]. \tag{20b}$$

In the context of Fig. 1(c), observing $W_D^F < 0$ corresponds to sampling configurations from the ρ_0 distribution where $\rho_1 > \rho_0$, and for $W_D^R < 0$ we require $\rho_0 > \rho_1$ when sampled from the ρ_1 distribution.

Turning our attention to the average swap probability, we note that ΔU_{swap}, which is the sum of W_D^F and W_D^R, is negative whenever $\rho'_N > \rho_N$ [by Eq. (13)]. Configurations for which this is the case are sampled by a simulation only in the lower-right half of the domain labeled p_{swap} in Fig. 1(b). The larger this domain, whose size is given by the average swap probability, the more frequently negative values of W_D^F and W_D^R are observed and the more quickly the hysteresis error converges. A numerical confirmation of this argument—that low swap probabilities correspond to large hysteresis errors and vice versa—is demonstrated in the results section.

III. METHODS

The computational system consists of 21 replicas, each with a different λ, which are simulated independently to obtain equilibrium statistics. The parameter λ controls the nonbonded interactions between an acetamide (ACE) solute and the water molecules. Two independent sets of simulations

were performed, with and without replica exchange, in order to investigate the effect of this technique.

The Lennard-Jones and Coulomb interactions between the water and ACE molecules are scaled by λ_{LJ} and λ_C, respectively. We scaled both parameters simultaneously, such that $\lambda_{LJ}=\lambda_C$; the single parameter λ then refers to both terms. This choice, while not commonplace in free-energy calculations, was made to simplify the replica exchange implementation, and since the free energy is a state function, any path through (λ_{LJ},λ_C) space is valid [18]. The specific way in which the Lennard-Jones and Coulomb terms scale with λ is described in Appendix B. λ varies across the 21 replicas from 0 to 1 in increments of 0.05.

Each replica consists of 343 water molecules and one ACE molecule, which is rigid and whose position is fixed in the central box. All simulations were performed at constant temperature (298 K) and volume (21.8 Å cubic box) using Metropolis Monte Carlo sampling. Parameters from the OPLS-AA force field [19] and four-site TIP4P water model [20] were used to model the solute and solvent, respectively. Minimum image boundary conditions and spherical cutoffs were employed for the Coulomb and Lennard-Jones potentials. The cutoff radius was 10.5 Å for electrostatic interactions and 10 Å for van der Waals interactions. Cutoffs were group based for the former and atom based for the latter. No long-range corrections were employed. All simulations were carried out using the MCCCS Towhee [21] Monte Carlo simulation package [36].

The initial configurations for all replicas were identical and correspond to the end point of a preequilibration run with ACE in water. For each replica, simulations consisted of 2×10^6 cycles, where a cycle corresponds to 343 Monte Carlo moves; each move combines rotations and translations of a randomly chosen individual water molecule. The initial 10^5 cycles were discarded for equilibration. The average acceptance rate for all replicas was 31%.

The replica exchange simulation consists of a number of simulation rounds, where each replica evolves independently, separated by swap rounds, when a number of swap attempts take place. The length of the simulation round was drawn from a normal distribution with a mean of 500 and standard deviation of 50 cycles. Five hundred cycles is the approximate energy autocorrelation "time." The swap round consists of 21^2 swap attempts between randomly selected replica pairs. Allowing swaps beyond neighboring replicas increases the efficiency of replica exchange by allowing a replica to traverse the entire range of λ from 0 to 1 more quickly than if only neighbor swaps were permitted [22].

During the course of the simulation, the native $[U_i(\Gamma_i)]$ and foreign $[U_{j\neq i}(\Gamma_i)]$ potential energies, as well as values for $dU/d\lambda_C$ and $dU/d\lambda_{LJ}$ (where $dU/d\lambda=dU/d\lambda_{LJ}+dU/d\lambda_C$), were saved every ten cycles. These were then post-processed to obtain the free energies, the hysteresis error, swap probabilities, and C_λ, regardless of whether actual replica exchange swaps took place. The total free energy of hydration, ΔF, is the sum of all free-energy changes $(\delta F)_i$ between neighboring replicas i and $i+1$, calculated using the Bennett acceptance ratio method [13],

TABLE I. The hydration free energy of acetamide. (a) The Helmholtz hydration free energy ΔF for the current work, as calculated by the Bennett acceptance ratio, and the root-mean-square hysteresis error. The ΔF statistical errors are calculated by the bootstrap method. (b) Published values of the Gibbs free energy ΔG, obtained both computationally and experimentally. All computational results utilize the OPLS-AA force field for the solute acetamide. Also noted are the water model and free-energy estimator (TI, thermodynamic integration; FEP, free-energy perturbation; BAR, Bennett acceptance ratio).

(a) Acetamide free energy of hydration: current work

	ΔF (kcal/mol)	$\epsilon_{\rm rms}$ (kcal/mol)
No replica exchange	-8.35 ± 0.051	0.120
Replica exchange	-8.14 ± 0.053	0.023

(b) Acetamide free energy of hydration: literature

	ΔG (kcal/mol)	Details
MacCallum and Tieleman [26]	-8.25 ± 0.26	TIP4P, TI
Shirts *et al.* [25]	-8.20 ± 0.03 [a]	TIP3P, TI
Chang *et al.* [27]	$-8.54\pm0.1-0.3$	TIP4P, BAR
Udier-Blagović *et al.* [28]	$-9.65\pm0.3-0.5$	TIP4P, FEP
Experimental [29]	−9.54	

[a]No long-range van der Waals corrections.

$$\Delta F \equiv \sum_i^{M-1} (\delta F)_i,$$

where M is the total number of replicas. Similarly, $\epsilon_{\rm rms}$ is the root mean square of the hysteresis error $(\epsilon_H)_i$ between neighboring replicas,

$$\epsilon_{\rm rms} \equiv \sqrt{\sum_i^{M-1} (\epsilon_H)_i^2/M}.$$

Statistical errors for ΔF were estimated using the bootstrap method [8]. With the simulation data set consisting of N observations, we drew n^* observations at random and with replacement to create one bootstrap estimate ΔF^*. This process was repeated 10 000 times, and the standard deviation among all the ΔF^* is the estimated error of ΔF. n^* is the expected number of independent observations in the data set; here, $n^*=1900$ with the assumption that there is one independent observation per two internal energy autocorrelation "times" [23].

IV. RESULTS

A. Acetamide free energy of hydration

The hydration free energies we calculate for acetamide are in line with results obtained by other researchers, as shown in Table I. All numerical results differ somewhat from experimental values due to differences in force field parameters. Our calculations were carried out in the canonical ensemble. Therefore, we obtain estimates for the Helmholtz free energy ΔF, whereas the experimental and other computational values obtain estimates for the Gibbs free energy

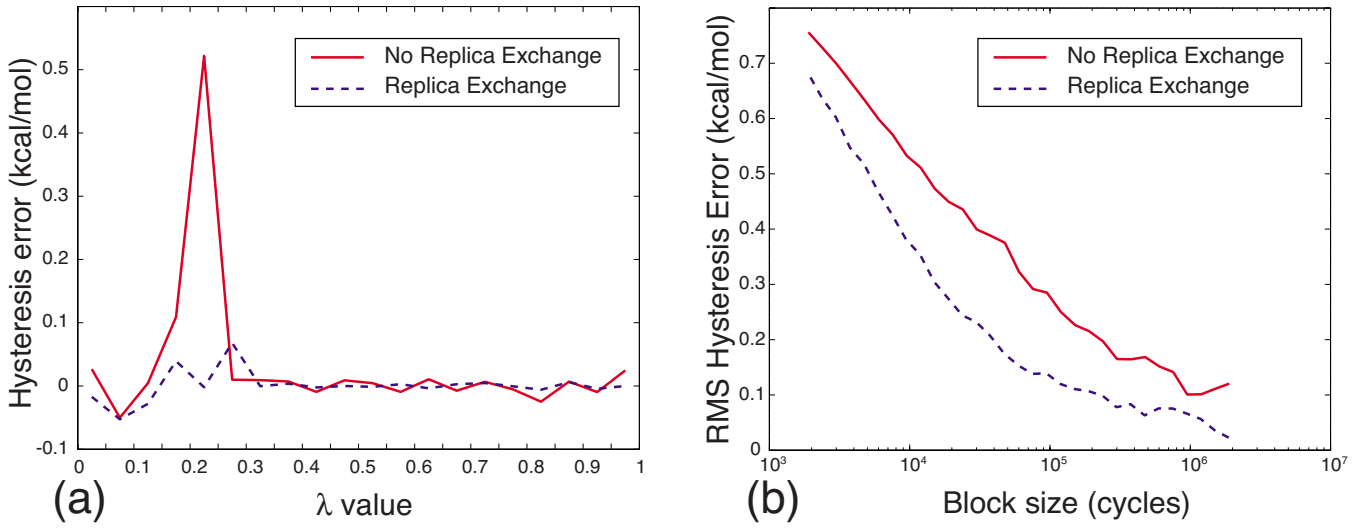

FIG. 2. (Color online) (a) The hysteresis error between neighboring replicas. Replica exchange effectively reduces the hysteresis error for replica pairs. (b) Block averages of the root-mean-square hysteresis error, showing that the hysteresis error falls with increasing block size. Replica exchange increases the rate at which hysteresis error is lowered, thereby achieving the same magnitude error with simulations which are on average 4–8 times shorter.

ΔG. Since the box volume at $\lambda=0$ was adjusted to correspond to 1 atm, the distinction between these two values should be negligible [24] even with some "pressurizing" due to insertion of the acetamide [25]; test calculations of ΔG in the *N-P-T* ensemble confirm this assertion (data not shown). The general consistency between our results and those of others serves to verify our implementation and sampling technique.

The methods to calculate ΔG in the *N-P-T* ensemble do not differ from those for calculating ΔF in the *N-V-T* ensemble; in particular, the replica exchange swap probability (10) does not change, since the pressure-volume work is reversible and does not contribute to the dissipated work.

Table I shows the differences between results obtained with and without replica exchange. As expected from our theoretical considerations, we find that the root-mean-square hysteresis error is lowered by an order of magnitude when replica exchange is coupled to the multicanonical sampling protocol. However, it should be noted that the statistical error estimated using the bootstrap method remains unaffected. This is not an artifact of the bootstrap method used to estimate statistical errors. Instead, fluctuations in estimates for δF originate in fluctuations of the underlying work distribution, shown in Eq. (1). As long as both simulations sample the work distribution adequately, they will have similar statistical error associated with them. As a cautionary note, low statistical errors can also be caused by inadequate sampling of the appropriate work distributions. The statistical error between two replicas can be reduced by decreasing the λ distance between them, and an optimal λ schedule can reduce it for an entire simulation.

B. Hysteresis error and replica exchange

For a fixed λ schedule, the hysteresis error may be reduced with either an improved sampling methodology like replica exchange or longer simulations per replica. The effects of both approaches are illustrated in Fig. 2.

Panel (a) shows ϵ_H for each neighboring replica pair. The hysteresis error is not uniform across all pairs, with spikes in the region $\lambda=0.1-0.3$. Replica exchange systematically reduces the hysteresis error for all pairs of replicas.

Panel (b) illustrates how both longer sampling and replica exchange affect the hysteresis error. Block averaging shows that the average root-mean-square hysteresis error declines consistently with longer simulations. This reduction can be improved with replica exchange; in fact, a simulation with replica exchange will achieve the same magnitude of ϵ_{rms} about five times more quickly than one without replica exchange.

C. Average swap probability

Figure 3 shows downward spikes in the swap probability for values of λ where the hysteresis error is large in Fig. 2(a). These results are consistent with the proposal that swap probability between two replicas is an indicator of the rate at which ϵ_H is minimized. The same region is characterized by a positive spike in C_λ, which is expected based on the relationship between the swap probability and C_λ in Eq. (18). However, while the swap probability calculation requires two separate simulations, estimates of C_λ can be obtained from just one. Moreover, $\langle p_{swap} \rangle$ varies as the distance between the replicas changes, complicating the interpretation if the λ schedule is not uniform. Evaluation of C_λ as a function of λ using a preliminary, coarse λ schedule can identify regions where the swap probability is expected to be low, and can be used to construct optimal λ schedules, as discussed in Sec. V B.

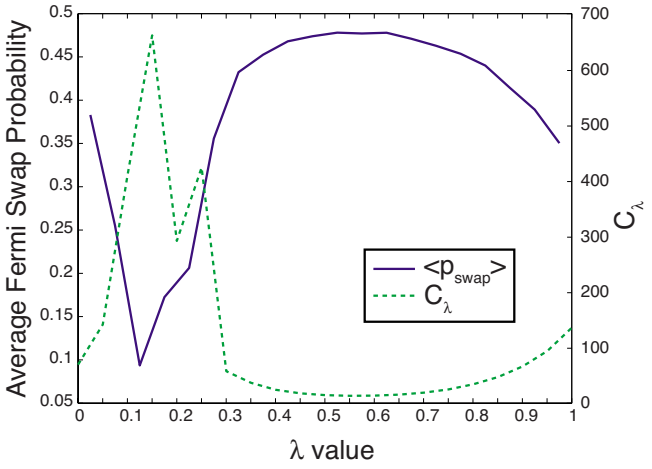

FIG. 3. (Color online) The average swap probability between adjacent replicas and $C_\lambda = \mathrm{var}(\partial U/\partial\lambda)$ evaluated for each replica (from the replica exchange simulation; simulation with no replica exchange is not significantly different). Spikes in C_λ indicate regions of low swap probability.

V. DISCUSSION

A. Physical interpretation of the C_λ profile

To gain a physical interpretation of the profile for C_λ shown in Fig. 3, we plot in Fig. 4 the average water density in a 2.5-Å sphere surrounding the carbonyl carbon of acetamide. The plot shows that water occupancy around the growing solute decreases rapidly in the range of $\lambda \sim 0.15$. The expulsion and rearrangement of water molecules during cavitation leads to a large shift in the equilibrium ensemble, giving rise to a pronounced spike in C_λ. (Smaller shifts in C_λ near $\lambda = 1$ reflect electrostatic effects and are not observed for simulations where $\lambda_C = 0$, data not shown.) Thus, C_λ profiles

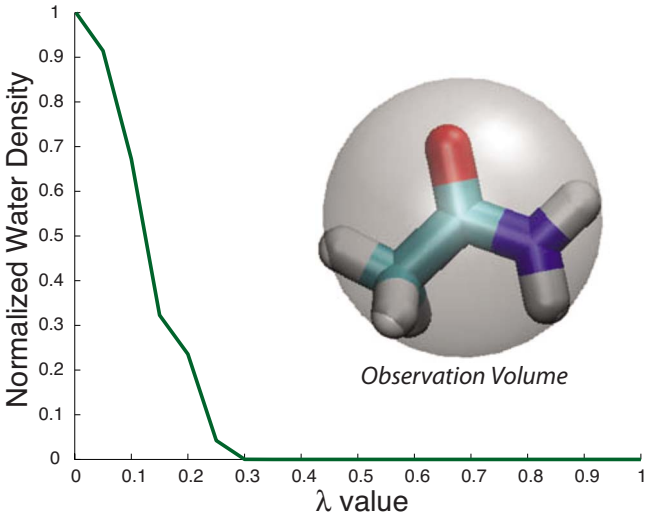

FIG. 4. (Color online) Water density within 2.5 Å of the acetamide carbonyl carbon as λ varies. The inset illustrates the position and size of the observation volume with respect to an acetamide molecule. Density is normalized by the bulk density. As λ increases, waters are expelled by the growing cavity.

serve as useful probes for detecting large shifts in equilibrium ensembles. Regions where the equilibrium ensembles change most rapidly are the regions that contribute the largest errors in free-energy calculations.

B. Optimal λ schedule for free-energy calculations

For given computational resources, with the number of replicas and the simulation length fixed, the root-mean-square hysteresis error of a simulation may be decreased by optimizing the λ schedule or the distribution of λ across the replicas. The swap probability gives the rate at which the average hysteresis error falls between two replicas, and in an optimized simulation it would be uniform across all replica pairs. In practice, it is difficult to obtain the λ schedule which makes the swap probability exactly uniform, but reasonable approximations can be made by using the linearized swap probability, given by Eq. (18).

First, it is necessary to perform some number of preliminary simulations to obtain C_λ along a coarse λ schedule. These initial simulations need not be as long as the final production runs, since C_λ converges more quickly than δF and is more tolerant of error. With a rough estimate of $C_\lambda(\lambda)$ in hand, the λ schedule can be adjusted to ensure that the linear swap probability is uniform between all replicas. Alternatively, one might simply shift replicas from where C_λ is small to where it is large. Both approaches are only approximate and break down when the linear response assumption in Eq. (18) ceases to be valid. They may be applied iteratively as C_λ is evaluated for new λ schedules.

The aim of an optimal λ schedule is to place replicas close together in regions where the C_λ profile shows spikes. This ensures reasonable swap probabilities and minimal hysteresis errors in regions that are problematic. Preliminary investigations show that even when the schedule is improved in an *ad hoc* manner, hysteresis as well as statistical errors decrease.

C. Replica exchange

Replica exchange provides a Monte Carlo move which may allow a replica to access a distant part of its equilibrium ensemble in one step. It is no substitute for conformational exploration within a replica. This point, while obvious, must be emphasized in the context of the hysteresis error, which does not report on the quality of intrareplica sampling. As an extreme but illustrative case, consider a system of some number of frozen replicas, each with a different configuration, which undergo replica exchange moves but no conformational changes. With just a modest number of swaps, these configurations attain the probability distribution described by Eq. (13) and the hysteresis error is zero. The system has achieved interreplica equilibrium, but the intrareplica probability distribution has not been obtained. In practice, the majority of Monte Carlo moves must be within a replica. The optimal frequency of swap moves remains an open question, although preliminary simulations suggest that more frequent swaps reduce the hysteresis error more quickly.

It is worthwhile to relate the replica exchange-based free-energy calculations presented here to other generalized en-

semble techniques, particularly λ-dynamical methods [30–32]. There, λ is a dynamical variable which evolves in time according to the conjugate force $\partial U/\partial\lambda$, with additional terms typically added to the Hamiltonian to associate with λ a momentum and to restrict its range. Since λ fluctuates, equilibrium distributions associated with a fixed λ cannot be calculated and measures like the swap probability and the hysteresis error are no longer defined. Nonetheless, both classes of techniques attempt to distribute atomic configurations across a range of λ values according to a Boltzmann distribution; λ-dynamical generalized ensemble techniques generate this distribution dynamically, whereas replica exchange techniques utilize a Markov chain to the same effect. Sampling difficulties in both cases are associated with phase changes and a large variance of $\partial U/\partial\lambda$, or C_λ. This results in low swap probabilities (for replica exchange) or regions in λ space not easily traversed (for λ dynamics), and these problems may be overcome with improved λ schedules or modified biasing potentials, respectively. One practical advantage of replica exchange techniques is that they are readily parallelizable across a number of computers, a trait not shared by all λ-dynamical methods.

VI. SUMMARY AND CONCLUSION

In a simulation of multiple replicas, each sampling the equilibrium ensemble of a different Hamiltonian, swapping configurations between replicas is a nonequilibrium work process. Accordingly, the work needed to perform such swaps has a distribution of values, as described by the Crooks fluctuation theorem. The hysteresis error ϵ_H developed here measures how closely a given simulation reproduces these work distributions between a pair of replicas.

The hysteresis error is particularly useful in the context of free-energy calculations. It reports on the combined bias of the forward and reverse free-energy perturbation techniques, and it measures how completely individual replicas sample their equilibrium ensemble. The root-mean-square hysteresis error, which reports on ϵ_H for the whole λ schedule, may be decreased by running a longer simulation, employing replica exchange, utilizing an improved λ schedule, or all of these approaches.

The average swap probability is another useful measure and can be calculated whether or not replica exchange is employed. Since it determines the rate at which the hysteresis error decreases with simulation length, the swap probability can be used to optimize the λ schedule. With a uniform average swap probability the hysteresis error falls evenly between all replica pairs. This maximizes the efficiency of simulations with fixed computational resources, avoiding unnecessary replicas where the hysteresis is low and preventing excessive errors from regions where the hysteresis error is large.

The swap probability, along with a related measure C_λ, yields insight into the microscopic behavior of a system. The swap probability is low and C_λ is large when the equilibrium ensemble changes rapidly with λ—for instance, during phase changes. Slow convergence and bias errors in free-energy calculations arise when there are spikes in the C_λ profile along the λ schedule, resulting in large hysteresis errors.

ACKNOWLEDGMENTS

We are grateful to Hoang Tran, Albert Mao, Xiaoling Wang, Nicholas Lyle, and in particular Alan Chen and Andreas Vitalis for helpful discussions and insightful comments. This work was supported by the National Science Foundation through Grants Nos. MCB-0416766 and MCB-0718924.

APPENDIX A: DERIVATIONS

1. Fluctuation theorem derivation

We derive the Crooks fluctuation theorem (1) in the context of instantaneously switching $\lambda_0 \rightarrow \lambda_1$ (forward) and $\lambda_1 \rightarrow \lambda_0$ (reverse). Expanding the ratio ρ_0/ρ_1 with (3) for an arbitrary configuration Γ,

$$\frac{\rho_0(\Gamma)}{\rho_1(\Gamma)} = \exp[\beta(F_0 - F_1) - \beta(U_0 - U_1)]$$
$$= \exp(-\beta\delta F + \beta W^F)$$
$$= \exp(\beta W_D^F), \tag{A1a}$$

and similarly,

$$\frac{\rho_1(\Gamma)}{\rho_0(\Gamma)} = \exp[\beta W_D^R(\Gamma)], \tag{A1b}$$

where the definitions of work (4a) and (4b) and dissipated work (7a) and (7b) were used.

We integrate ρ_1 from (A1a) over all configurations, but consider contributions only from those Γ for which the forward dissipated work value takes on a specific value, W_D:

$$\int d\Gamma\ \rho_0(\Gamma)\exp[-\beta W_D^F(\Gamma)]\delta[\beta W_D - \beta W_D^F(\Gamma)]$$
$$= \int d\Gamma\ \rho_1(\Gamma)\delta[\beta W_D - \beta W_D^F(\Gamma)]. \tag{A2}$$

Since, from (A1a) and (A1b),

$$W_D^F(\Gamma) = -W_D^R(\Gamma),$$

(A2) becomes

$$\int d\Gamma\ \rho_0(\Gamma)\exp[-\beta W_D^F(\Gamma)]\delta[\beta W_D - \beta W_D^F(\Gamma)]$$
$$= \int d\Gamma\ \rho_1(\Gamma)\delta[\beta W_D + \beta W_D^R(\Gamma)]. \tag{A3}$$

We define $P^F(W_D)$ as the probability of observing a given dissipated work value in the forward switching process, and it can be expressed as an integral over all configurations which yield this value,

$$P^F(W_D) = \int d\Gamma\ \rho_0(\Gamma)\delta[\beta W_D - \beta W_D^F(\Gamma)]. \tag{A4a}$$

Likewise, the probability of observing a given dissipated work value in the reverse switching process is

$$P^R(W_D) = \int d\Gamma\ \rho_1(\Gamma)\delta[\beta W_D - \beta W_D^R(\Gamma)]. \quad \text{(A4b)}$$

With these definitions, (A3) may be written as

$$\exp(-\beta W_D)P^F(\beta W_D) = P^R(-\beta W_D),$$

which is equivalent to (1).

2. Fluctuation theorem and hysteresis error

The relationship between some arbitrary deviation of a simulation from the Crooks fluctuation theorem and the hysteresis error is derived by first rewriting Eq. (8) as

$$P_R^*(-\beta W_D)\exp(\beta W_D) = P_F^*(\beta W_D)\exp(-\beta\epsilon_{FT}^*). \quad \text{(A5)}$$

Inserting the δF_{FEP}^R definition (5b) into the definition of the hysteresis error (6), expanding the reverse work with (7b) and using the δF_{FEP}^F estimate for δF,

$$\begin{aligned}\epsilon_H &= \delta F_{FEP}^F - \beta^{-1}\ln\langle\exp(-\beta W^R)\rangle_1^* \\ &= \delta F_{FEP}^F - \beta^{-1}\ln[\langle\exp(-\beta W_D^R)\rangle_1^* \\ &\quad \times\exp(\beta\delta F_{FEP}^F)] \\ &= -\beta^{-1}\ln[\langle\exp(-\beta W_D^R)\rangle_1^*].\end{aligned}$$

We now expand the estimated ensemble average as an integral over all values of βW_D^F, with P_R^* the normalized histogram of βW_D^R obtained from a simulation,

$$\epsilon_H = -\beta^{-1}\ln\left[\int_{-\infty}^{+\infty} d[\beta W_D^R]P_R^*(\beta W_D^R)\exp(-\beta W_D^R)\right].$$

As βW_D^R is a dummy variable, we change it to $-\beta W_D$

$$\epsilon_H = -\beta^{-1}\ln\left[\int_{-\infty}^{+\infty} d[\beta W_D]P_R^*(-\beta W_D)\exp(\beta W_D)\right],$$

where we implicitly multiplied the integrand by −1 to preserve the limits of integration. With (A5) the above can be written as

$$\epsilon_H = -\beta^{-1}\ln\left[\int_{-\infty}^{+\infty} d(\beta W_D)P_F^*(\beta W_D)\exp(-\beta\epsilon_{FT}^*)\right],$$

which reduces to (9).

3. Interreplica equilibrium and hysteresis error

We can relate an small arbitrary error in the calculated distribution ρ_N' to the hysteresis error by considering a small error $\rho_\epsilon(\Gamma_0,\Gamma_1)$ in the otherwise correctly estimated ρ_N'. Rewriting (13),

$$\rho_N' + \rho_\epsilon = \rho_N\exp(-\beta\Delta U_{swap}),$$

we integrate over all configuration pairs and rewrite ΔU_{swap} with (11b),

$$\begin{aligned}&\int d\Gamma_0 d\Gamma_1\rho_N' + \int d\Gamma_0 d\Gamma_1\rho_\epsilon \\ &= \int d\Gamma_0 d\Gamma_1\rho_0(\Gamma_0)\rho_1(\Gamma_1)\exp[-\beta W^F(\Gamma_0)]\int d\Gamma_1\rho_1(\Gamma_1) \\ &\quad\times\exp[-\beta W^R(\Gamma_1)]. \quad \text{(A6)}\end{aligned}$$

With the sampling error contained in ρ_ϵ, the ρ_N' term [expanded with (12b)] is identically 1. Taking the logarithm and dividing by β, (A6) becomes

$$-\beta^{-1}\ln\left[1 + \int d\Gamma_0 d\Gamma_1\rho_\epsilon\right] = \delta F_{FEP}^R - \delta F_{FEP}^F, \quad \text{(A7)}$$

where we have used the δF_{FEP} definitions (5a) and (5b). With the approximation $\ln(1+x)\simeq x$ for small x and the definition of ϵ_H, Eq. (6), we obtain Eq. (14).

4. Linearized average swap probability

Here we consider the average Fermi swap probability between two replicas whose λ parameters differ by a small amount δ (written as δ_λ in the text). For convenience we define

$$\begin{aligned}\mu &\equiv \beta\Delta U_{swap} \\ &= \beta[U_\delta(\Gamma_0) - U_0(\Gamma_0) + U_0(\Gamma_\delta) - U_\delta(\Gamma_\delta)],\end{aligned}$$

where Γ_0 and Γ_δ are configurations drawn from the equilibrium distributions U_0 and U_δ parametrized by λ_0 and $\lambda_0+\delta$, respectively. We expand U_δ as a Taylor series about λ_0,

$$U_\delta(\Gamma) = U_0(\Gamma) + \delta V_0(\Gamma) + \frac{\delta^2}{2}W_0(\Gamma) + O(\delta^3),$$

with

$$V_0 \equiv \left.\frac{\partial U}{\partial\lambda}\right|_{\lambda=\lambda_0},$$

$$W_0 \equiv \left.\frac{\partial^2 U}{\partial\lambda^2}\right|_{\lambda=\lambda_0}.$$

μ can then be written as

$$\mu = \beta\delta[V_0(\Gamma_0) - V_0(\Gamma_\delta)] + \frac{\beta\delta^2}{2}[W_0(\Gamma_0) - W_0(\Gamma_\delta)].$$

Note that μ is small $[O(\delta)]$; thus, with the identities

$$\exp(x) = 1 + x + x^2/2 + \cdots, \quad \text{(A8a)}$$

$$\frac{1}{1+x} = 1 - x + x^2 - \cdots, \quad \text{(A8b)}$$

we may write the Fermi swap probability between configurations Γ_0 and Γ_δ as

$$p_{swap} = \frac{1}{1+\exp\mu}$$

$$= \frac{1}{2}\left(\frac{1}{1+\mu/2+\mu^2/4+O(\mu^3)}\right)$$

$$= \frac{1}{2}[1-(\mu/2+\mu^2/4)+(\mu/2+\mu^2/4)^2+O(\mu^3)]$$

$$= \frac{1}{2}-\frac{1}{4}\mu+O(\mu^3).$$

The average swap probability is the ensemble average over all configuration pairs,

$$\langle\langle p_{swap}\rangle_0\rangle_\delta = \frac{1}{2}-\frac{1}{4}\langle\langle\mu\rangle_0\rangle_\delta$$

$$= \frac{1}{2}-\frac{1}{4}\left(\beta\delta\langle V_0\rangle_0+\frac{\beta\delta^2}{2}\langle W_0\rangle_0-\beta\delta\langle V_0\rangle_\delta - \frac{\beta\delta^2}{2}\langle W_0\rangle_\delta\right)+O(\delta^3). \quad \text{(A9)}$$

To evaluate $\langle\cdots\rangle_\delta$, we first obtain Q_δ, the partition function at $(\lambda_0+\delta)$,

$$Q_\delta \equiv \int d\Gamma \exp(-\beta U_\delta)$$

$$= \int d\Gamma \exp(-\beta U_0)[1-\beta\delta V_0+O(\delta^2)]$$

$$= Q_0[1-\beta\delta\langle V_0\rangle_0+O(\delta^2)]$$

and its reciprocal

$$Q_\delta^{-1} = Q_0^{-1}[1+\beta\delta\langle V_0\rangle_0+O(\delta^2)].$$

We can now evaluate $\langle V_0\rangle_\delta$ and $\langle W_0\rangle_\delta$, retaining only terms which will remain $O(\delta^2)$ or larger in (A9):

$$\langle V_0\rangle_\delta \equiv Q_\delta^{-1}\int d\Gamma \exp(-\beta U_\delta)V_0$$

$$= Q_0^{-1}(1+\beta\delta\langle V_0\rangle_0)\int d\Gamma(1-\beta\delta V_0)\exp(-\beta U_0)V_0$$

$$= (1+\beta\delta\langle V_0\rangle_0)(\langle V_0\rangle_0-\beta\delta\langle V_0^2\rangle_0)$$

$$= \langle V_0\rangle_0+\beta\delta(\langle V_0\rangle_0^2-\langle V_0^2\rangle_0)$$

and

$$\langle W_0\rangle_\delta \equiv Q_\delta^{-1}\int d\Gamma \exp(-\beta U_\delta)W_0$$

$$= Q_0^{-1}[1+O(\delta)]\int d\Gamma \exp(-\beta U_0)W_0[1-O(\delta)]$$

$$= \langle W_0\rangle_0+O(\delta).$$

Finally, (A9) becomes

$$\langle\langle p_{swap}\rangle_0\rangle_\delta = \frac{1}{2}-\frac{\beta^2\delta^2}{4}(\langle V_0^2\rangle_0-\langle V_0\rangle_0^2)+O(\delta^3), \quad \text{(A10)}$$

equivalent to Eq. (18), which is valid for small δ.

APPENDIX B: U_{LJ} AND U_C FUNCTIONAL FORMS

The functional forms of both the Coulomb and Lennard-Jones potentials were developed for this work based on three criteria

(i) Configurations where the solute and solvent overlap may be observed for $\lambda=0$. For such configurations, we require (a) that swaps be permitted with reasonable frequency for small λ (e.g., $\lambda=0.1$) and (b) that the swap probabilities fall off quickly thereafter; in particular, we wish to avoid the situation where the swap probability declines only very near $\lambda=1.0$.

(ii) We require that $\partial U/\partial\lambda$ not always be zero for $\lambda=0$ to avoid complications with the thermodynamic integration (TI) estimator. While we do not report results using TI in this work, we wish to construct a λ schedule that works with all estimators.

(iii) In this work, $\lambda_{LJ}=\lambda_C$. Therefore, Lennard-Jones repulsion must dominate Coulombic attraction at very small atomic separations.

While various ways to scale the potential have been discussed in the literature [33–35], none of these satisfied all of our requirements. It should be noted that condition (iii) is somewhat arbitrary and more common scaled potentials may be used if the insertion process scales the Lennard-Jones prior to the Coulomb potential. The specific profiles of Figs. 2(a) are dependent on the choice of Coulomb and Lennard Jones functional forms, as well as the relationship between λ_C and λ_{LJ}.

Coulomb scaling. We employ a modified version of the linear soft-core scaling [34]; for two atoms of charges q_i and q_j distance r apart, the potential energy is λ_C as

$$U_C(r,\lambda_C) = \lambda_C\frac{q_iq_j}{\alpha_C(1-\lambda_C)+r}, \quad \text{(B1)}$$

where α_C controls the "soft core" term and for small λ_C imposes a minimum effective atomic separation. $\alpha_C=1.5$ Å for all simulations in this work.

Lennard-Jones scaling. The Lennard-Jones potential between two particles may be written generally as

$$U_{LJ}(r,\lambda_{LJ}) = BA(A-1), \quad \text{(B2)}$$

where, for the unscaled Lennard-Jones potential,

$$A(r) = \left(\frac{\sigma}{r}\right)^6, \quad B = 4\epsilon.$$

Simple linear scaling by λ_{LJ} of the Lennard-Jones potential is

known to be unsatisfactory, and a number of alternate forms have been introduced. We developed the exponential soft core

$$A(r,\lambda_{LJ}) = 1 \Big/ \left[\alpha_{LJ}(1-\lambda_{LJ})^b + \left(\frac{r}{\sigma}\right)^6 \right], \qquad \text{(B3a)}$$

$$B(\lambda_{LJ}) = 4\epsilon \frac{1-e^{-k\lambda_{LJ}}}{1-e^{-k}}, \qquad \text{(B3b)}$$

with $a=4$, $k=1$, and $\alpha_{LJ}=0.5$ Å. The precise position along the λ coordinate of the swap probability trough (see Fig. 3) is specific to this Lennard-Jones potential.

[1] A. Ben-Naim, *Solvation Thermodynamics* (Plenum Press, New York, 1987).
[2] F. M. Ytreberg, R. H. Swendsen, and D. M. Zuckerman, J. Chem. Phys. **125**, 184114 (2006).
[3] R. M. Levy and E. Gallicchio, Annu. Rev. Phys. Chem. **49**, 531 (1998).
[4] J. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[5] D. L. Mobley, E. Dumont, J. D. Chodera, and K. A. Dill, J. Phys. Chem. B **111**, 2242 (2007).
[6] N. Lu and D. A. Kofke, J. Chem. Phys. **114**, 7303 (2001).
[7] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, New York, 2002).
[8] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall, New York, 1993).
[9] D. M. Zuckerman and T. B. Woolf, J. Stat. Phys. **114**, 1303 (2004).
[10] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[11] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
[12] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[13] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[14] M. R. Shirts and V. S. Pande, J. Chem. Phys. **122**, 144107 (2005).
[15] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[16] H. Fukunishi, O. Watanabe, and S. Takada, J. Chem. Phys. **116**, 9058 (2002).
[17] C. Jarzynski, Phys. Rev. E **73**, 046105 (2006).
[18] A. A. Chialvo and J. M. Haile, Fluid Phase Equilib. **37**, 293 (1987).
[19] W. Jorgensen, D. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).
[20] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein, J. Chem. Phys. **79**, 926 (1983).
[21] M. G. Martin and J. I. Siepmann, J. Phys. Chem. B **103**, 4508 (1999).
[22] C. Predescu, M. Predescu, and C. V. Ciobanu, J. Phys. Chem. B **109**, 4189 (2005).
[23] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon Press, Oxford, 1999).
[24] H. Qian and J. J. Hopfield, J. Chem. Phys. **105**, 9292 (1996).
[25] M. R. Shirts, J. W. Pitera, W. C. Swope, and V. S. Pande, J. Chem. Phys. **119**, 5740 (2003).
[26] J. L. MacCallum and D. P. Tieleman, J. Comput. Chem. **24**, 1930 (2003).
[27] J. Chang, A. M. Lenhoff, and S. I. Sandler, J. Phys. Chem. B **111**, 2098 (2007).
[28] M. Udier-Blagović, P. M. D. Tirado, S. A. Pearlman, and W. L. Jorgensen, J. Comput. Chem. **25**, 1322 (2004).
[29] R. Wolfenden, Biochemistry **17**, 201 (1978).
[30] X. Kong and C. L. Brooks III, J. Chem. Phys. **105**, 2414 (1996).
[31] R. Bitetti-Putzer, W. Yang, and M. Karplus, Chem. Phys. Lett. **377**, 633 (2003).
[32] J. B. Abrams, L. Rosso, and M. E. Tuckerman, J. Chem. Phys. **125**, 074115 (2006).
[33] T. C. Beutler, A. E. Mark, R. C. van Scheik, P. R. Gerber, and W. F. van Gunsteren, Chem. Phys. Lett. **222**, 529 (1994).
[34] J. W. Pitera and W. F. van Gunsteren, Mol. Simul. **28**, 45 (2002).
[35] M. R. Shirts and V. S. Pande, J. Chem. Phys. **122**, 134508 (2005).
[36] See http://towhee.sourceforge.net. Version maw4_17_4 (maw-dev branch) was used for this work.