

# Enhanced Sampling in Free Energy Calculations: Combining SGLD with the Bennett's Acceptance Ratio and Enveloping Distribution Sampling Methods

Gerhard König,<sup>\*,†</sup> Benjamin T. Miller,<sup>†</sup> Stefan Boresch,<sup>‡</sup> Xiongwu Wu,<sup>†</sup> and Bernard R. Brooks<sup>†</sup>

<sup>†</sup>Laboratory of Computational Biology, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

<sup>‡</sup>Department of Computational Biological Chemistry, Faculty of Chemistry, University of Vienna, Währingerstraße 17, A-1090 Vienna, Austria

**ABSTRACT:** One of the key requirements for the accurate calculation of free energy differences is proper sampling of conformational space. Especially in biological applications, molecular dynamics simulations are often confronted with rugged energy surfaces and high energy barriers, leading to insufficient sampling and, in turn, poor convergence of the free energy results. In this work, we address this problem by employing enhanced sampling methods. We explore the possibility of using self-guided Langevin dynamics (SGLD) to speed up the exploration process in free energy simulations. To obtain improved free energy differences from such simulations, it is necessary to account for the effects of the bias due to the guiding forces. We demonstrate how this can be accomplished for the Bennett's acceptance ratio (BAR) and the enveloping distribution sampling (EDS) methods. While BAR is considered among the most efficient methods available for free energy calculations, the EDS method developed by Christ and van Gunsteren is a promising development that reduces the computational costs of free energy calculations by simulating a single reference state. To evaluate the accuracy of both approaches in connection with enhanced sampling, EDS was implemented in CHARMM. For testing, we employ benchmark systems with analytical reference results and the mutation of alanine to serine. We find that SGLD with reweighting can provide accurate results for BAR and EDS where conventional molecular dynamics simulations fail. In addition, we compare the performance of EDS with other free energy methods. We briefly discuss the implications of our results and provide practical guidelines for conducting free energy simulations with SGLD.

## 1. INTRODUCTION

The determination of free energy differences has become an essential part of the toolbox of computational chemists. Free energy simulations (FES) provide means to calculate relative binding affinities of different ligands to a receptor,<sup>1,2</sup> guide drug development,<sup>3</sup> or study enzymatic reactions<sup>4</sup> and protein synthesis in the ribosome.<sup>5</sup> Despite many successful applications, one has to keep in mind that the accuracy of FES is limited by two factors. First, the quality of the force field has a direct impact on the correctness of the calculated free energy differences.<sup>6</sup> Second, it is necessary to adequately sample the important parts of the phase spaces at the end points of a FES. The recent SAMPL3 binding affinity prediction competition highlights that, even for the binding of relatively small host–guest complexes, sampling is still an issue that accounts for a significant part of the deviations from experimental results.<sup>7–11</sup> An insightful discussion of the importance of sampling in molecular simulation was recently published;<sup>12</sup> furthermore, we refer the reader to an overview of the state of the art of FES, which addresses both of these limitations.<sup>3</sup>

In this work, we explore the utility of a recent method to accelerate sampling, self-guided Langevin dynamics (SGLD),<sup>13</sup> in connection with FES. Originally presented in 2003, SGLD has been successfully applied to a variety of problems.<sup>14–18</sup> However, since the effect of the guiding force on ensemble distributions was initially poorly understood, SGLD has so far been restricted to applications that primarily focus on

enhancing sampling; it was considered unsuitable for FES, where simulating a well-defined statistical mechanical ensemble is essential. Recently, Wu and Brooks have shown how to obtain a canonical ensemble distribution from SGLD,<sup>19</sup> which makes it possible to use SGLD in FES. SGLD certainly is an interesting option to enhance sampling: the overhead above unmodified Langevin dynamics is minimal, and no prior knowledge of the system is required. SGLD enhances the (slow) low frequency degrees of freedom that are usually considered more biologically relevant by redistributing energy into them from high frequency motions; the amount of redistribution is controlled by the user. There is no need to run multiple simulations at different temperatures, and the method is equally applicable to gas phase, implicit solvent, and explicit solvent simulations. Reference 19 demonstrates that SGLD can be considered a simulation in the presence of weighting factors that are derived from the SGLD trajectory and a reference LD simulation. Once the weighting factors  $w^{\text{SGLD}}$  are known for each frame of the trajectory, any (canonical) ensemble average  $\langle X \rangle$  can be obtained from SGLD simulations according to

**Special Issue:** Wilfred F. van Gunsteren Festschrift

**Received:** February 8, 2012

**Published:** June 15, 2012

$$\langle X \rangle = \frac{\langle X w^{\text{SGLD}} \rangle_{\text{SGLD}}}{\langle w^{\text{SGLD}} \rangle_{\text{SGLD}}} \quad (1)$$

by the well-known Torrie–Valleau reweighting formula for simulations in the presence of a biasing potential.<sup>20</sup>

All of the standard methods to compute free energy differences, thermodynamic perturbation (TP),<sup>21</sup> thermodynamic integration (TI),<sup>22</sup> and Bennett's acceptance ratio method (BAR),<sup>23</sup> are based on the calculation of ensemble averages  $\langle X \rangle$ . Therefore, eq 1 makes it possible to use SGLD as the underlying sampling method in FES. One focus of this work is the combination of SGLD with BAR, since BAR and its derivatives are widely considered the computationally most efficient FES techniques.<sup>24–27</sup> However, SGLD could be coupled with thermodynamic perturbation and integration equally well.

A practical limitation of all standard techniques to compute free energy differences is the large amount of computational resources required.<sup>28</sup> In addition to simulations at the end states, i.e., the states corresponding to the physical systems, simulations of several artificial (“alchemical”) intermediate states, typically 10–20, are needed. This makes FES prohibitively expensive in many molecular modeling applications, in particular for high-throughput screening. One way to mitigate the computational cost of FES is to use special techniques to compute multiple free energy differences from a single simulation. A recent promising development in this direction is the enveloping distribution sampling (EDS) method developed by Christ and van Gunsteren.<sup>29–32</sup> Instead of computing individual free energy differences between  $N$  states as in standard FES methods, EDS computes the free energy differences of the  $N$  states relative to a common reference state  $R$ . This can be accomplished from a single simulation, though some exploratory calculations to optimize the reference state may be necessary. Methods such as EDS represent a trade-off between gains in computational efficiency and a loss in accuracy compared to the direct calculation of the free energy differences. The most important factor determining the accuracy of EDS is the choice of parameters characterizing the reference state  $R$ . The simulation of  $R$  not only has to explore the phase space of the hypothetical hybrid state itself but also must sample regions of phase space relevant for each of the real systems of interest. Therefore, sampling may be even more important in EDS than in traditional FES, which suggests testing SGLD in conjunction with both BAR and EDS.

Clearly, any combination of an enhanced sampling method with FES should be tested carefully. In particular, it is prudent to verify the correctness of the approach with simple systems for which free energy differences of interest can be computed either analytically or by direct integration of the partition function. While neither approach is feasible for biological systems, it is well-known how to compute partition functions of polyatomic molecules in the gas phase.<sup>33,34</sup> These types of systems have been successfully employed in earlier methodological work.<sup>35–39</sup> In addition, we compute relative solvation free energy differences of model systems. To test the combination of SGLD with BAR and EDS in the simplest possible manner, we proceed in four steps: First, we test the treatment of bonded terms in FES with SGLD by studying a four-atomic benchmark system (system I), for which analytical reference results are available. Second, we demonstrate that SGLD allows the calculation of free energy differences in cases where normal MD fails. For this purpose, we calculate the free

energy difference between two five-atomic benchmark systems (systems II and III). These were designed to have no significant phase space overlap, and the relevant energy minima are separated by high energy barriers of more than 10 kcal/mol (16 kT).<sup>39</sup> Third, we explore the capability of SGLD to be used with explicit solvent by mutating “ethan-1-ol” to “ethan-2-ol” (i.e., moving the hydroxyl group from carbon 1 to carbon 2) in aqueous solution. This system was selected since the end states are equivalent and the total free energy change is zero, and no reference results are required. Fourth, we present the practicability of our approach for the biologically relevant mutation of Alanine to Serine and study the solvation free energy difference between the two amino acids. As shown recently, this free energy difference of solvation is very sensitive to the “secondary structure”<sup>40</sup> of the blocked amino acids. Thus, it represents a good test for the correct sampling of a nontrivial free energy surface. In addition, we employ this system to compare the performance of EDS with BAR and three different kinds of thermodynamic perturbation.

The remainder of the manuscript is organized as follows. First, we outline the theory of SGLD, BAR, and EDS, including the equations to perform BAR and EDS with reweighted SGLD data. Next, the implementation of EDS in CHARMM version 36 and the methodological details of the simulations are presented. We then proceed to the results for the four benchmark systems as outlined above and conclude with a short discussion concerning the usefulness of these types of calculations. In the Appendix, we lay out the equations for the analytical reference results and a short extract from a CHARMM script that uses EDS.

## 2. THEORY

**2.1. Self-Guided Langevin Dynamics.** **2.1.1. Basics.** The principal idea of SGLD is to selectively enhance the sampling of slow degrees of freedom by using information about the low frequency motions from the history of the simulation.<sup>13</sup> This is done by adding a guiding force  $\mathbf{g}$  to the Langevin equation of motion

$$\frac{\partial \mathbf{p}_i}{\partial t} = \mathbf{f}_i + \mathbf{g}_i - \gamma_i \mathbf{p}_i + \mathbf{R}_i \quad (2)$$

where  $\mathbf{p}$  denotes the momentum,  $\mathbf{f}$  the force,  $\gamma$  the collision frequency, and  $\mathbf{R}$  the random forces applied to maintain a temperature  $T$ ;  $i$  is the atom index. The guiding force depends on three factors

$$\mathbf{g}_i = \lambda_g \gamma_i (\langle \mathbf{p}_i \rangle_L - \xi \mathbf{p}_i) \quad (3)$$

where the guiding factor  $\lambda_g$  sets the strength of the guiding force. For  $\lambda_g = 0$ , eq 2 reduces to normal LD.  $\langle \mathbf{p}_i \rangle_L$  represents a local time average of the momenta over a time period  $t_L$ . The factor  $\xi$  ensures the conservation of energy.

The local (time) average  $\langle \cdot \rangle_L$  serves as a low-pass filter. By an appropriate choice of the averaging time  $t_L$ , it is possible to divide the motion contributions of the fast degrees of freedom (such as bond stretching and angle bending terms, which exhibit a high frequency) from those of the slow degrees of freedom (e.g., dihedral angles that change with low frequency motions). Fast, high frequency fluctuations of the momenta cancel during the averaging process, while slow motions pass through the filter. Thus, the guiding force is automatically applied in the direction of systematic motions and accelerates the conformational change caused by low frequency motions.

The averaging time  $t_L$  defines which motions are considered high or low frequency, i.e., those with a frequency lower than  $(t_L)^{-1}$  are enhanced by the guiding force. For example, bond and angle vibrations usually happen on a subpicosecond timescale, while conformational changes of a protein take nanoseconds or microseconds. Therefore, the optimal  $t_L$  for a specific application has to be estimated by the user (e.g., a  $t_L = 0.2$  ps is used for enhanced sampling of changes of the secondary structure, while a  $t_L = 1$  ps is employed for interdomain motions).

**2.1.2. Reweighting with SGLD.** It is clear that SGLD does not preserve the canonical ensemble, because low frequency motions are accelerated by  $\lambda_g \gamma_i \langle \mathbf{p}_i \rangle_L$ , while the high frequency motions are damped by  $-\lambda_g \gamma_i \mathbf{p}_i$  to conserve energy. Hence, SGLD constantly transfers kinetic energy from the fast to the slow degrees of freedom. At the same time, the slow degrees of freedom lose kinetic energy via the coupling to the fast degrees of freedom. The balance between the two opposing effects leads to two separate “effective” temperatures, one for the high frequency motions and one for the low frequency motions. Recently, Wu and Brooks derived reweighting factors to account for those deviations by employing eq 1. Here, we present a short outline of the derivation found in ref 19.

The conceptional basis of the derivation lies in reformulating the configurational partition function  $\Theta$  in terms of contributions from low and high frequency motions. It starts from the assumption that the local time average of the total potential energy  $\langle \mathbf{U} \rangle_L$  represents the potential energy associated with the slow degrees of freedom

$$U^{\text{lf}} \approx \eta^{\text{lf}} \langle \mathbf{U} \rangle_L \quad (4)$$

where  $\eta^{\text{lf}}$  is a correction factor for the bias introduced to the energy surface of the low frequency motions due to the guiding force. This bias of the potential energies can be determined from the projection of guiding forces  $\mathbf{g}$  on the forces  $\langle \mathbf{f} \rangle_L$  in the SGLD trajectory. If the two variables are uncorrelated,  $\eta^{\text{lf}}$  is 1, and the guiding force can be regarded as a random factor added to the forces (for the selected examples in this paper,  $\eta^{\text{lf}}$  was  $\sim 0.93 \pm 0.03$ ).

Conversely, the bias introduced by the high frequency motions to the energy surface is approximated by the difference between the total potential energy and  $\langle \mathbf{U} \rangle_L$

$$U^{\text{hf}} \approx \eta^{\text{hf}} (U - \langle \mathbf{U} \rangle_L) \quad (5)$$

where  $\eta^{\text{hf}}$  is the correction factor for the bias due to the guiding force. It can be calculated from the covariance of the scalar products of the guiding forces and the internal forces (for the selected examples in this paper,  $\eta^{\text{hf}}$  was  $\sim 0.993 \pm 0.002$ ).

To account for the differences of the effective temperature, the low frequency collision factor  $\chi^{\text{lf}}$  is introduced. It is the ratio of the effective temperature of the low frequency motions  $T^{\text{lf}}$  to the reference temperature of these motions,  $T_0^{\text{lf}}$

$$\chi^{\text{lf}} = \frac{T_0^{\text{lf}}}{T^{\text{lf}}} \quad (6)$$

where the effective low frequency temperature can be calculated from the SGLD trajectory with

$$T^{\text{lf}} = (N_{\text{dof}} \cdot k_B)^{-1} \left\langle \sum_i \frac{\langle \mathbf{p}_i \rangle_L^2}{m_i} \right\rangle_{\text{SGLD}} \quad (7)$$

employing the number of degrees of freedom,  $N_{\text{dof}}$ , the Boltzmann constant,  $k_B$ , and the atomic masses,  $m_i$ . For LD simulations,  $T^{\text{lf}}$  equals  $T_0^{\text{lf}}$ . While  $T_0^{\text{lf}}$  can be estimated on the basis of an SGLD simulation according to ref 19, practical experience shows that the results are more accurate if  $T_0^{\text{lf}}$  is determined from a LD simulation (using the average of the TREPLF output in CHARMM).

The collision factor for the high frequency motions is defined on the basis of the difference between the total temperature of the system  $T$  and the effective temperature of the low frequency motions

$$\chi^{\text{hf}} = \frac{T - T_0^{\text{lf}}}{T - T^{\text{lf}}} \quad (8)$$

thus accounting for the lowered effective temperature of the fast high frequency motions.

With the equations above, it is possible to formulate a partition function for SGLD in terms of low frequency and high frequency contributions

$$\Theta_{\text{SGLD}} = \sum_s e^{-\beta(\chi^{\text{lf}} U_s^{\text{lf}} + \chi^{\text{hf}} U_s^{\text{hf}})} \quad (9)$$

where  $\beta$  has the usual meaning of  $1/k_B T$ . The summation runs over all combinations of microstates and, additionally, possible pathways for the local time average  $\langle \cdot \rangle_L$  to that microstate  $s$ . The microstates, therefore, not only depend on the conformational degrees of freedom but also on their history. However, the main difference from the classical partition function is the use of two different effective temperatures ( $\chi^{\text{lf}} \beta$  and  $\chi^{\text{hf}} \beta$ ) for the low and high frequency motions.

The calculation of the weighting factor for eq 1 requires the unnormalized probabilities of a microstate  $s$  in an LD simulation ( $\tilde{p}_s^{\text{LD}}$ ) and in an SGLD simulation ( $\tilde{p}_s^{\text{SGLD}}$ ), which are given by

$$\tilde{p}_s^{\text{LD}} = e^{-\beta(U_s)} \text{ and } \tilde{p}_s^{\text{SGLD}} = e^{-\beta(\chi^{\text{lf}} U_s^{\text{lf}} + \chi^{\text{hf}} U_s^{\text{hf}})} \quad (10)$$

The weighting factor  $w_s^{\text{SGLD}}$  is calculated from the ratio of these two probabilities

$$w_s^{\text{SGLD}} = \frac{\tilde{p}_s^{\text{LD}}}{\tilde{p}_s^{\text{SGLD}}} = e^{-\beta(U_s - \chi^{\text{lf}} U_s^{\text{lf}} - \chi^{\text{hf}} U_s^{\text{hf}})} \quad (11)$$

To summarize, determining the weighting factors for an SGLD trajectory requires the measurement of two different kinds of contributions. First, the contribution of the guiding force to the energy is accounted for by calculating  $U^{\text{lf}}$  and  $U^{\text{hf}}$  at each time step. Second, the bias due to the two effective temperatures must be calculated. This requires the determination of the effective temperature  $T^{\text{lf}}$  over the whole SGLD trajectory, as well as measuring the reference low frequency temperature  $T_0^{\text{lf}}$  from a reference LD simulation.

**2.2. Free Energy Methods.** **2.2.1. Bennett's Acceptance Ratio Method.** To compute the free energy difference between two states, 0 and 1, BAR utilizes the information obtained from simulations of both states.<sup>23</sup> The free energy difference  $\Delta A$  between states 0 (potential energy function  $U_0$ ) and 1 (potential energy function  $U_1$ ) is given by

$$\Delta A^{0 \rightarrow 1} = \beta^{-1} \left( \ln \frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} \right) + C \quad (12)$$



The subscripts 0 and 1 in eq 12 indicate that the ensemble averages  $\langle \cdot \rangle$  are calculated from the trajectories of the initial (0) and final states (1), respectively. The symbol  $f$  denotes the Fermi function

$$f(x) = \frac{1}{1 + \exp(\beta x)} \quad (13)$$

and

$$C = \beta^{-1} \ln \frac{Q_0 n_1}{Q_1 n_0} \quad (14)$$

where  $Q$  denotes the respective partition function and  $n_0$  and  $n_1$  are the numbers of configurations of states 0 and 1 from which the ensemble averages are evaluated. The unknown constant  $C$ , which corresponds essentially to the free energy difference of interest, is found iteratively. Starting from an initial guess, a value of  $C$  is found so that the argument of the logarithm in eq 12 equals unity, since in this case the free energy difference is given by

$$\Delta A^{0 \rightarrow 1} = -\beta^{-1} \ln \frac{n_1}{n_0} + C \quad (15)$$

As shown in refs 41 and 42, BAR can conveniently be extended to employ weighted averages, leading to

$$\Delta A^{0 \rightarrow 1} = k_B T \ln \left( \frac{\langle f(U_0 - U_1 + C) w_1 \rangle_1 \langle w_0 \rangle_0}{\langle f(U_1 - U_0 - C) w_0 \rangle_0 \langle w_1 \rangle_1} \right) + C \quad (16)$$

where the indices 0 and 1 refer to the respective end state. This method is referred to as non-Boltzmann Bennett (NBB)<sup>41</sup> or weighted Bennett's acceptance ratio (w-BAR).<sup>42</sup>

To use eq 16, it is necessary to evaluate three quantities for each frame of each state's trajectories: For the biased trajectory of state 0,  $U_0$ ,  $U_1$ , and the SGLD weight  $w_0$  must be computed, while for state 1,  $U_0$ ,  $U_1$ , and  $w_1$  are required. Thus, the computational overhead is limited to the determination of the  $w^{\text{SGLD}}$  for both trajectories.

**2.2.2. Enveloping Distributions Sampling.** Enveloping distribution sampling<sup>29–32</sup> relies on the simulation of a single reference state  $R$  to calculate the free energy differences between  $N$  end points. The potential energy of the reference state  $U_R$  mixes the potential energies of all end points according to

$$U_R = -(\beta s)^{-1} \ln \sum_{k=0}^N e^{-\beta s(U_k - E_k^R)} \quad (17)$$

where  $U_k$  denotes the potential energy of end point  $k$ ,  $E_k^R$  is the energy offset parameter that ensures equal sampling of all end states, and  $s$  is the smoothness factor. Note that the EDS equation presented here employs only a single  $s$  for all end points. Thus, it resembles the EDS implementation in ref 30 rather than more recent versions of EDS (e.g., in refs 31 and 32), which are considerably more general. Depending on the choice of  $s$  and  $E_k^R$ , the reference state can be used for different purposes.<sup>30,31</sup> A reference state with  $s = -1$  and  $E_k^R = A_k$  corresponds to the virtual state that is generated by the perturbations in Bennett's acceptance ratio method. However, when this state is actually simulated,  $U_R$  will be dominated by the end state with the highest  $(U_k - E_k^R)$ . For example, if two imaginary end points are used, one with low potential energy ( $U_1 = 1$ ) and one with high potential energy ( $U_2 = 10$ ), using  $s$

$= -1$ ,  $E_1^R = 0$ , and  $E_2^R = 0$  will lead to a reference state potential of  $U_R = 10.001$ , which corresponds to the state with higher energy. This approach will lead to significant sampling problems and, therefore, poor convergence. Therefore, negative values of  $s$  should not be used in EDS.

On the other hand, a reference state with  $s = +1$  and  $E_k^R = A_k$  corresponds to enveloping all important parts of phase space of the end states.  $U_R$  will be dominated by the end state with the smallest  $U_k - E_k^R$ . Returning to our example with two imaginary end points above, using  $s = +1$  will lead to  $U_R = 0.99$ , which corresponds to the state with lower energy. In this case, high energy barriers may remain due to interactions with the environment of the free energy calculation. In practice, the efficiency of EDS can be tuned by the smoothness parameter  $s$ , and it can be further improved by introducing two-dimensional smoothness parameters  $s_{kl}$  for pairs of end states  $k$  and  $l$ .<sup>31,32</sup>

The energy offsets  $E_k^R$  can be optimized by selecting a good estimate for the free energy difference between the reference state and the respective end state  $i$ . This estimate can be obtained from an LD simulation. In connection with SGLD, the same LD simulation can be used to determine  $T_0^{\text{ff}}$  for the reweighting of SGLD and an initial guess of  $E_k^R$ . However, since efficient protocols are available to update  $s$  and  $E_k^R$  in an optimal way,<sup>30–32</sup> we focus on testing the advantages of advanced sampling for the simplified version of EDS described in eq 17. Our results should be regarded as a worst-case scenario of an EDS implementation.

The free energy difference between two arbitrary end states,  $X$  and  $Y$ , is evaluated by their free energy differences from the reference state:

$$\Delta A^{X \rightarrow Y} = -\beta^{-1} \ln \frac{\langle e^{-\beta(U_Y - U_R)} \rangle_R}{\langle e^{-\beta(U_X - U_R)} \rangle_R} \quad (18)$$

Reweighting can be implemented in a straightforward fashion by combining eq 18 with eq 1, which leads to

$$\Delta A^{X \rightarrow Y} = -\beta^{-1} \ln \frac{\langle e^{-\beta(U_Y - U_R)} w_R \rangle_R}{\langle e^{-\beta(U_X - U_R)} w_R \rangle_R} \quad (19)$$

### 3. METHODS

Except for the four-atomic benchmark system and the ethanol mutation, four different approaches were used to calculate the free energy differences: BAR with LD simulations (BAR/LD), EDS with LD simulations (EDS/LD), BAR with SGLD and reweighting based on eq 16 (BAR/SGLD), and EDS with SGLD based on eq 19 (EDS/SGLD). Unless stated otherwise, the BAR calculations were conducted without any intermediate points, relying on simulations of the end points only, and EDS calculations used only a single reference state with an EDS smoothness parameter  $s = 1$ . The energy offset parameters  $E_k^R$  were determined from short LD equilibration simulations with a simulation time of less than 10% of the production run. The same  $E_k^R$  values were used for production of both EDS/LD and EDS/SGLD; i.e.,  $E_k^R$  values were never updated after the initial estimate. To calculate the weights  $w^{\text{SGLD}}$  for BAR/SGLD and EDS/SGLD, the low frequency reference temperature  $T_0^{\text{ff}}$  was determined from reference LD simulations that had equal length and identical energy parameters.

#### 3.1. EDS-Implementation in CHARMM Using MSCALE.

We have implemented the simulation of EDS reference states in CHARMM,<sup>43,44</sup> making use of the MSCALE module.<sup>45</sup>

MSCALE is a communications protocol that has been developed for CHARMM and other programs to facilitate the combination of different simulation methods into a single calculation. The system can be arbitrarily divided into one or more subsystems, allowing for the possibility of overlapping atom selections for the subsystems. For example, it is possible to set up three subsystems, where the first subsystem contains a core region, the second subsystem the environment around the core, and the third subsystem both the core region and the environment.

It is possible to employ a different Hamiltonian for each subsystem with MSCALE. For this purpose, the main process, called the client, spawns a new calculation for each of the subsystems. The spawned processes are called servers and may run any program for which an MSCALE interface has been developed. At each energy and force evaluation, the client sends each server the coordinates that its subsystem requires, and the server returns the energies and forces (as well as other data that might be relevant for the simulation). Therefore, all subsystems are coupled over the main process, and energies and forces from the various subsystems may be combined in an arbitrary manner.

The EDS implementation in CHARMM follows eq 17. It is, therefore, based on a simplified version of EDS with single smoothness parameter  $s$  for all end states, as described in ref 30. Instead of using  $s$  explicitly, the user has to specify a so-called “EDS temperature” ( $T_{\text{EDS}}$ ). The EDS temperature is given by

$$T_{\text{EDS}} = \frac{T}{s} \quad (20)$$

Based on  $T_{\text{EDS}}$ , eq 17 becomes

$$U_{\text{R}} = -k_{\text{B}}T_{\text{EDS}} \ln \sum_{k=0}^N \exp\left(-\frac{U_k - E_k^{\text{R}}}{k_{\text{B}}T_{\text{EDS}}}\right) \quad (21)$$

To perform an EDS simulation, the MSCALE command is invoked by the user, followed by SUBSystem commands that define each end state of the simulation. The use of different atom selections allows both single and multi-topology approaches to free energy calculations while using a common set of coordinates. After the specification of the end states in MSCALE, the EDS command is invoked. It requires the specification of the EDS TEMPerature ( $T_{\text{EDS}}$ ), the number of end states, the names of the MSCALE subsystems that will be used as end states, as well as their corresponding energy offsets  $E_k^{\text{R}}$ . An extract from an example CHARMM script that calls EDS is shown in the Appendix.

The internal communication of the EDS implementation is structured as follows. The main process sends the current atomic coordinates to the subsystem processes. Each subsystem calculates the energies  $U_k$  and atomic forces  $\mathbf{f}_k$  for its corresponding end state and reports them to the main process. Instead of just combining the resulting energies and forces as is normally done in MSCALE, the main process is responsible for calculating the potential energy of the reference state  $U_{\text{R}}$  as well as its corresponding forces (c.f. eq 10 in ref 31), based on the information from the subsystems. The forces of the reference state are then used to conduct a molecular dynamics step that leads to the next set of atomic coordinates. For optimal performance, all interactions that are not affected by the mutation should be calculated in the main process.<sup>46</sup>

**3.2. Simulation Details.** All calculations were conducted with CHARMM version 36, using the CHARMM22 all-atom

force field.<sup>47</sup> Images of the mutations were generated with VMD.<sup>48</sup> Random forces for both LD and SGLD were applied according to a target temperature of 300 K. Except for the four-atomic benchmark system, either all atomic masses were set to at least 10 amu or SHAKE<sup>49</sup> was employed to keep the bonds rigid, allowing for a time step between 1 and 2 fs. A cutoff radius of 998 Å was employed for the gas phase simulations, which is considerably larger than the size of the molecule. Trajectories were written every 10 steps for the four-atomic benchmark system and every 100 steps for all other benchmarks. Further details of the simulations follow when describing the respective system. Tables 1 and 2 give an overview of the simulation details and the EDS parameters. Unless stated otherwise, the reported standard deviations were determined by repeating each free energy simulation four times, starting with different initial random velocities. The energies of the respective states required for BAR and EDS were extracted with the UPEN command of the MSCALE module of CHARMM; the BAR/EDS analysis was carried out by a Perl program.

**3.2.1. Four-Atomic Benchmark Systems.** The calculations on the four-atomic benchmark system test whether BAR in combination with SGLD can reproduce analytically calculable free energy differences of bonded terms. The theoretical background concerning the partition function of small polyatomic molecules is summarized in the Appendix. As demonstrated in refs 35, 36, and 38, free energy differences between small molecules in the gas phase can be computed with high accuracy and precision by MD based FES. The most important practical detail is the choice of the thermostat since (i) such small systems clearly are far from the thermodynamic limit, and (ii) their stiff harmonic nature when represented by a typical force field causes the initial kinetic energy introduced by the assignment of random velocities to not redistribute unless special precautions are taken. Simulations with a Berendsen or a single Nose–Hoover thermostat fail to reproduce the analytical result. In earlier work, we, therefore, used individual Nose–Hoover thermostats for each atom;<sup>35,36</sup> in this work, we employed Langevin dynamics (LD) in the reference simulations. Further, the overall rotation of the molecules must not be removed.<sup>35</sup> The high sensitivity of the free energy results to small deviations from the original ensemble makes them a very useful benchmark for the SGLD reweighting scheme.

The alchemical free energy simulation with LD was carried out with the PERT module of CHARMM,<sup>43</sup> using a single topology approach for the alchemical intermediates.<sup>36</sup> Each free energy difference was calculated using 21  $\lambda$  intermediate states,  $\lambda = 0.00, 0.05, 0.10, \dots, 0.95, 1.00$  with BAR (note that  $\lambda$  intermediate states are not connected with the guiding factor  $\lambda_{\text{g}}$ ). A friction coefficient of 5 ps<sup>−1</sup> was applied to all atoms. For SGLD, an averaging time ( $t_{\text{L}}$ ) of 0.2 ps and a guiding factor  $\lambda_{\text{g}} = 1$  were used. The time step was 1 fs. At each  $\lambda$ , 10 ps were discarded as equilibration, while the free energy data were accumulated over the following 190 ps. Although this is an extremely short protocol, longer simulation lengths had only a minuscule impact on the results (data not shown). The statistical uncertainty was estimated by computing each free energy difference 10 times using independent simulations with different initial random velocities. The results reported in Table 3 are the mean and standard deviations of these 10 individual values.

Table 1. Simulation Types and Details

benchmark	description <sup>a</sup>	$n_s^b$	$t_s^c$	$\lambda_g^d$	$\gamma^e$	$T_0^{f,f}$
four-atomic	BAR gas	880 <sup>g</sup>	0.2	—	5	X
	BAR gas	880 <sup>g</sup>	0.2	—	5	
	BAR gas	880 <sup>g</sup>	0.2	1.0	5	
five-atomic	BAR II gas	4	50	—	5	X
	BAR II gas	4	50	—	300	
	BAR II gas	4	50	1.0	300	
	BAR III gas	4	50	—	5	X
	BAR III gas	4	50	—	300	
	BAR III gas	4	50	1.0	300	
	EDS II/III gas	4	50	—	5	X
	EDS II/III gas	4	50	—	300	
	EDS II/III gas	4	50	1.0	300	
ethanol	BAR solv	44	0.2	—	5	X
	BAR solv	44	0.2	1.0	5	
Ala-Ser	BAR Ala gas	4	0.2	—	5	X
	BAR Ala gas	4	75	—	100	
	BAR Ala gas	4	75	0.75	100	
	BAR Ala GBMV	4	75	—	5	X
	BAR Ala GBMV	4	75	—	100	
	BAR Ala GBMV	4	75	0.75	100	
	BAR Ser gas	4	75	—	5	X
	BAR Ser gas	4	75	—	100	
	BAR Ser gas	4	75	0.75	100	
	BAR Ser GBMV	4	75	—	5	X
	BAR Ser GBMV	4	75	—	100	
	BAR Ser GBMV	4	75	0.75	100	
	EDS Ala/Ser gas	4	75	—	5	X
	EDS Ala/Ser gas	4	75	—	100	
	EDS Ala/Ser gas	4	75	0.75	100	
	EDS Ala/Ser GBMV	4	75	—	5	X
	EDS Ala/Ser GBMV	4	75	—	100	
	EDS Ala/Ser GBMV	4	75	0.75	100	
	TP <sub>0.5</sub> Ala/2+Ser/2 gas	4	75	—	5	X
	TP <sub>0.5</sub> Ala/2+Ser/2 gas	4	75	—	100	
	TP <sub>0.5</sub> Ala/2+Ser/2 gas	4	75	0.75	100	
	TP <sub>0.5</sub> Ala/2+Ser/2 GBMV	4	75	—	5	X
	TP <sub>0.5</sub> Ala/2+Ser/2 GBMV	4	75	—	100	
	TP <sub>0.5</sub> Ala/2+Ser/2 GBMV	4	75	0.75	100	

<sup>a</sup>BAR marks simulations used for BAR calculations. EDS marks simulations using an EDS reference state, followed by a description of the end states. The abbreviation “gas” marks gas phase simulations. “solv” means simulations with explicit solvent, and “GBMV” marks implicit solvent simulations using GBMV. <sup>b</sup>Number of simulations. <sup>c</sup>Simulation time in nanoseconds. <sup>d</sup>SGLD guiding factor, a “—” marks LD simulations. <sup>e</sup>Friction constant in picoseconds<sup>-1</sup>. <sup>f</sup>X marks LD reference calculations to determine  $T_0^{f,f}$  for the reweighting of SGLD trajectories. <sup>g</sup>Benchmark involves eight mutations (see corresponding sections in the Methods), each using the two end states and nine intermediate states, each being repeated 10 times.

An identical initial state was used in all free energy calculations: The equilibrium bond lengths  $r_{12} = r_{23} = r_{34}$  were set to 2 Å; the indices denote the atom numbers. The force constant for all bonds was  $K_{\text{bond}} = 200 \text{ kcal}/(\text{mol } \text{\AA}^2)$ . Similarly, equal equilibrium bond angles  $\theta_{123} = \theta_{234} = 110^\circ$  and angle force constants  $K_{\text{angle}} = 50 \text{ kcal}/(\text{mol } \text{rad}^2)$  were employed. The dihedral force constant was  $k_{1234} = 1 \text{ kcal/}$

Table 2. EDS Details

benchmark	description	$U_1^a$	$U_2^b$	$s$	$E_1^{\text{Rc}}$	$E_2^{\text{Rd}}$	top <sup>e</sup>
2	$\Delta A^{\text{II} \rightarrow \text{III}}$	II	III	1	0	0	S
2	$\Delta A^{\text{II} \rightarrow \text{III}}$	II	III	0.1	0	0	S
2	$\Delta A^{\text{II} \rightarrow \text{III}}$	II	III	0.01	0	0	S
4	$\Delta A_{\text{gas}}^{\text{Ala} \rightarrow \text{Serf}}$	Ala	Ser	1	4.17	13.27	D
	$\Delta A_{\text{GBMV}}^{\text{Ala} \rightarrow \text{Serg}}$	Ala	Ser	1	4.21	11.28	D

<sup>a</sup>First end state. <sup>b</sup>Second end state. <sup>c</sup>Energy offset of first subsystem. <sup>d</sup>Energy offset of second subsystem. <sup>e</sup>Topology of FES. S = single topology, D = dual topology. <sup>f</sup>Free energy difference in gas phase. <sup>g</sup>Free energy difference using the GBMV implicit solvent.

mol with a multiplicity  $n_{1234} = 3$ . We computed the free energy differences to end states with the following changes relative to this initial state analytically, using both regular BAR and BAR in combination with SGLD (see Table 3): (1)  $r_{34} = 3 \text{ \AA}$ ; (2)  $K_{34} = 400 \text{ kcal}/(\text{mol } \text{\AA}^2)$ ; (3)  $r_{34} = 3 \text{ \AA}$ ,  $K_{34} = 400 \text{ kcal}/(\text{mol } \text{\AA}^2)$ ; (4)  $K_{234} = 100 \text{ kcal}/(\text{mol } \text{rad}^2)$ ; (5)  $\theta_{234} = 150^\circ$ ; (6)  $\theta_{234} = 150^\circ$ ,  $K_{234} = 100 \text{ kcal}/(\text{mol } \text{rad}^2)$ ; (7)  $r_{34} = 3 \text{ \AA}$ ,  $K_{34} = 400 \text{ kcal}/(\text{mol } \text{\AA}^2)$ ,  $\theta_{234} = 150^\circ$ ,  $K_{234} = 100 \text{ kcal}/(\text{mol } \text{rad}^2)$ ,  $k_{1234} = 2 \text{ kcal/mol}$ ; (8)  $r_{34} = 3 \text{ \AA}$ ,  $K_{34} = 400 \text{ kcal}/(\text{mol } \text{\AA}^2)$ ,  $\theta_{234} = 150^\circ$ ,  $K_{234} = 100 \text{ kcal}/(\text{mol } \text{rad}^2)$ ,  $k_{1234} = 2 \text{ kcal/mol}$ ,  $n_{1234} = 2$ .

**3.2.2. Five-Atomic Benchmark Systems.** Two five-atomic benchmark systems developed by Leitgeb et al. were employed (systems II and III of reference 39), using a single topology setup for the FES. They are unbranched, nonlinear five-atomic molecules. The two systems are designed to have small phase space overlap and large energy barriers, thus representing a worst case scenario for free energy calculations. Due to the smallness of the systems, quasi-analytical reference results are available (cf. Appendix 1). The equilibrium bond length is 1.53 Å, and all bond angles are 111°. Two dihedral angle terms ( $\phi_1$ ,  $\phi_2$ ) are present.

In system II, the sum of two potentials was applied simultaneously to each dihedral ( $n_{1,1} = 3$ ,  $k_{1,1} = 2 \text{ kcal/mol}$ ,  $n_{1,2} = 1$ ,  $k_{1,2} = 2 \text{ kcal/mol}$ ;  $n_{2,1} = 3$ ,  $k_{2,1} = 2 \text{ kcal/mol}$ ,  $n_{2,2} = 1$ ,  $k_{2,2} = 2 \text{ kcal/mol}$ ). This results in one global minimum and four local minima (cf. Figure 2). The dihedral potentials were identical for system III, but the dihedral force constants were raised ( $n_{1,1} = 3$ ,  $k_{1,1} = 3.5 \text{ kcal/mol}$ ,  $n_{1,2} = 1$ ,  $k_{1,2} = 3 \text{ kcal/mol}$ ;  $n_{2,1} = 3$ ,  $k_{2,1} = 3.5 \text{ kcal/mol}$ ,  $n_{2,2} = 1$ ,  $k_{2,2} = 3 \text{ kcal/mol}$ ). Furthermore, intramolecular electrostatic and Lennard-Jones interactions were added to the system, resulting in two equivalent global minima and one local minimum for system III. However, the three minima are separated by high energy barriers (Figure 2).

For all computations involving the five-atomic benchmark systems, simulation lengths of 50 ns and a time step of 1 fs were employed. Free energy differences were computed with BAR and EDS using three different sets of parameters: LD with a friction coefficient of 5 ps<sup>-1</sup> on all atoms, LD with a friction coefficient of 300 ps<sup>-1</sup> on all atoms to determine the reference low frequency temperature, and SGLD with a friction coefficient of 300 ps<sup>-1</sup> on all atoms and a guiding factor of 1. The different friction constants in LD ( $\gamma = 5 \text{ ps}^{-1}$ ) and SGLD ( $\gamma = 300 \text{ ps}^{-1}$ ) are motivated by the increase of the guiding force with larger values of  $\gamma$  (eq 3).

**3.2.3. Ethanol.** The solvation free energy difference between ethan-1-ol and ethan-2-ol was calculated in explicit water, using the standard thermodynamic cycle to compute solvation free energies.<sup>50</sup> Alchemical free energy differences between the end states were calculated both in the gas phase (yielding  $\Delta A_{\text{gas}}$ )



Table 3. Free Energy Differences of Dummy Atoms in Four-Atomic Benchmark Systems in Kilocalories Per Mole

type of change <sup>a</sup>	ref <sup>b</sup>	LD <sup>c</sup>	SGLD <sup>d</sup>	SGLD <sub>rw</sub> <sup>e</sup>	SGLD <sub>nr</sub> <sup>f</sup>
1. $r_{34}$	−0.483	−0.491	−0.712	−0.747	−0.504
2. $K_{34}$	0.207	0.208	0.178	0.175	0.196
3. $r_{34}K_{34}$	−0.277	−0.268	−0.505	−0.521	−0.307
4. $K_{234}$	0.207	0.206	0.178	0.174	0.196
5. $\theta_{234}$	0.376	0.366	0.231	0.207	0.387
6. $\theta_{234}K_{234}$	0.583	0.573	0.427	0.404	0.595
7. $r_{34}K_{34}\theta_{234}K_{234}k_{1234}$	0.550	0.544	0.156	0.108	0.612
8. $r_{34}K_{34}\theta_{234}K_{234}k_{1234}n_{1234}$	0.550	0.569	0.110	0.026	0.586
RMSD <sup>g</sup>		0.009	0.250	0.288	0.030

<sup>a</sup>The internal parameters changed are indicated. For the abbreviations used and values of the changed parameters, see the Methods. <sup>b</sup>Analytical results (cf. Appendix). <sup>c</sup>Regular BAR results with Langevin dynamics, average standard deviation = 0.033 kcal/mol. <sup>d</sup>BAR/SGLD results without reweighting, average standard deviation = 0.037 kcal/mol. <sup>e</sup>Reweight SGLD/BAR results, average standard deviation = 0.036 kcal/mol. <sup>f</sup>Reweight SGLD/BAR results if net translation and rotation are removed from the guiding force (SGNONET command), average standard deviation = 0.034 kcal/mol. <sup>g</sup>Root mean square deviations from analytical results.

and with explicit solvent (yielding  $\Delta A_{\text{H}_2\text{O}}$ ), employing 11  $\lambda$  points and a dual topology setup. The solvation free energy difference ( $\Delta\Delta A_{\text{solv}}$ ) is given by  $\Delta\Delta A_{\text{solv}} = \Delta A_{\text{H}_2\text{O}} - \Delta A_{\text{gas}}$ . The gas phase simulations were set up as described in ref 40. However, since the free energy results for LD and SGLD in the gas phase were well converged and close to zero ( $<0.0008 \pm 0.011$  kcal/mol), the solvation free energy difference results shown in Table 5 are practically the same as the free energy differences in solution (i.e.,  $\Delta\Delta A_{\text{solv}} \approx \Delta A_{\text{H}_2\text{O}}$ ).

A total of 766 TIP3P water molecules<sup>51,52</sup> were present in the solvent simulations. The simulation box was a truncated octahedron. The side length  $L$  of the cube from which the octahedron was generated was  $L = 35.75$  Å, which was the average box size from a 200 ps constant pressure simulation. For the determination of the free energy difference, we used constant volume simulations with a friction coefficient of  $5 \text{ ps}^{-1}$  on all atoms. The time step was 1 fs, and SHAKE<sup>49</sup> was used to keep the water geometry rigid. Lennard-Jones interactions were switched off between 10 and 12 Å, while electrostatic interactions were computed with the particle mesh Ewald method.<sup>53</sup> Each system was equilibrated for 100 ps at every  $\lambda$  value, followed by 1 ns of production dynamics. For SGLD, two different setups were employed to evaluate the efficiency of the reweighting scheme for large systems. In the first case, the guiding force was applied to all atoms, thus acting on both solute and solvent (SGLD<sup>all</sup>). In the second setup, the application of guiding forces was restricted to the solute by using the ISGSTA and ISGEND commands (SGLD<sup>EtOH</sup>).

**3.2.4. Alanine–Serine.** The solvation free energy differences between Ala and Ser shown in Table 6 were calculated with the same thermodynamic cycle as in the ethanol example, but using a dual topology approach. Free energy differences between Ala and Ser were calculated both in the gas phase (yielding  $\Delta A_{\text{gas}}^{\text{Ala} \rightarrow \text{Ser}}$ ) and with implicit solvent (yielding  $\Delta A_{\text{imp}}^{\text{Ala} \rightarrow \text{Ser}}$ ). The solvation free energy difference is given by  $\Delta\Delta A_{\text{solv}}^{\text{Ala} \rightarrow \text{Ser}} = \Delta A_{\text{imp}}^{\text{Ala} \rightarrow \text{Ser}} - \Delta A_{\text{gas}}^{\text{Ala} \rightarrow \text{Ser}}$ . The implicit solvent simulations utilized GBMV,<sup>54</sup> as described in ref 40. This choice is motivated by the good agreement between GBMV and explicit solvent results. GBMV is able to reproduce the dependency of the solvation free energy difference between Ala and Ser on the backbone conformation with a root-mean-square error (RMSE) of 0.27 kcal/mol compared to explicit solvent. Moreover, GBMV yielded an RMSE of 0.41 for eight solvation free energy

differences between the amino acids Ala, Ser, Val, Thr, Phe, and Tyr (compared to explicit solvent).<sup>40</sup>

For all simulations, a time step of 1.5 fs was chosen. Free energy differences were calculated with EDS and BAR, based on simulations of 75 ns length. The parameters for GBMV were those recommended in gbm.doc in the CHARMM documentation at [www.charmm.org](http://www.charmm.org). The thermodynamic perturbation (TP) results for TP<sub>fw</sub> and TP<sub>bw</sub> in Table 6 were calculated from the same trajectories as BAR. For TP<sub>0.5</sub>, MSCALE generated a reference state with  $U_{\text{R}} = U_{\text{Ala}}/2 + U_{\text{Ser}}/2$ . All other simulation details were identical to those used for BAR. The respective SGLD results for TP were generated on the basis of eq 9 given in ref 39 for non-Boltzmann thermodynamic perturbation.

## 4. RESULTS AND DISCUSSION

**4.1. Four-Atomic Benchmark Systems.** In this example, we test the effect of SGLD on the treatment of bonded terms of dummy atoms in FES (Figure 1). Table 3 summarizes results of

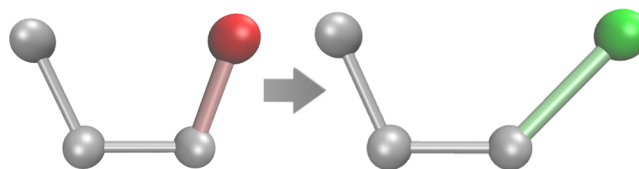


Figure 1. Illustration of the four-atomic benchmark system to test the treatment of bonded terms (for example, changes of bond length and bond angle) in free energy simulations with SGLD.

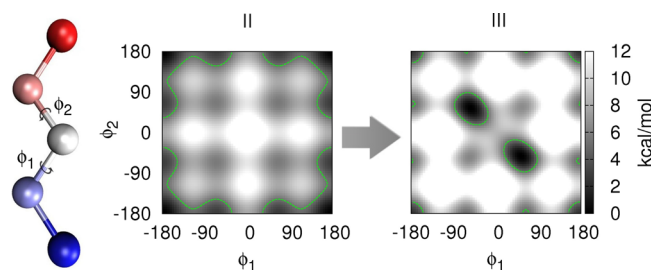
FES for four-atomic model systems without nonbonded interactions. The results of regular LD simulations (third column) are compared with those obtained from SGLD simulations with (fifth column) and without reweighting (fourth column). The column “ref” lists the analytical results, which are calculated as outlined in the Appendix. The agreement between the analytical and LD results is excellent, with the largest absolute deviation being  $\sim 0.02$  kcal/mol. Similarly, the standard deviations of the simulation results are very low (on average,  $\sim 0.03$  kcal/mol). The results of the SGLD simulations with and without reweighting (columns four and five in Table 3) deviate significantly from the reference results (with absolute deviations of up to  $\sim 0.4$  kcal/mol). This finding is not surprising, considering that SGLD is designed to transfer energy from fast degrees of freedom (such as bond and

angle terms) to slow degrees of freedom and, hence, does not correspond to the canonical ensemble. The results illustrate the sensitivity of this simple model system to the details of the simulation setup. Similar deviations can also be found for two- and three-atomic systems (data not shown). Therefore, we investigated further potential sources of error.

If the net translational force and torque for the whole system is removed from the guiding force in SGLD simulations by using the SGNONET command (SGLD<sub>nr</sub>), the results (rightmost column of Table 3) are in agreement with the analytical results, exhibiting a root-mean-square difference of 0.030 kcal/mol. Although the RMSD of SGLD<sub>nr</sub> is slightly higher than that with LD, the deviations are not statistically significant (the average standard deviation is 0.034). The difference between SGLD and SGLD<sub>nr</sub> indicates that, for small systems, the default parameters of SGLD lead to distributions of the kinetic energy that do not correspond to the canonical ensemble. Specifically, SGLD increases the rotational motions of the molecule. This side effect leads to artifacts in the contributions to the free energy difference from the change in moment of inertia (cf. pp. 5150–5151 of ref 35). This essentially kinetic contribution to the free energy difference is presently not accounted for by using the reweighting scheme in eq 11. For the remainder of this paper, only simulations with the SGNONET keyword (i.e., without enhanced translation and rotation) will be employed. Future versions of SGLD in CHARMM will automatically remove all net translation and rotation from the guiding force.

An alternative way to avoid contributions to the free energy difference from the change in moment of inertia is to set up the mutation with a so-called dual topology approach. Dual topology simulations use three separate groups of atoms in the coordinate set: (a) atoms that do not change during the FES (the environment), (b) atoms corresponding to the initial state, and (c) atoms corresponding to the target state. Atoms in groups *b* and *c* do not interact with each other. For example, in a FES of the mutation of Ala to Ser, the backbone, which corresponds to the environment group *a*, is connected to both a CH<sub>3</sub> group (corresponding to Ala) and a CH<sub>2</sub>OH group (corresponding to Ser). A dual topology setup was used in the ethanol and Ala → Ser examples presented in this work, avoiding complications resulting from changes of bonded terms that could potentially arise in single topology.

**4.2. Five-Atomic Benchmark Systems.** This example considers the effect of high energy barriers on FES by employing two five-atomic benchmark systems having both bonded and nonbonded interactions. In contrast to the four-atomic benchmark system, the bond and angle terms are not changed during the mutation, and SHAKE is applied to the bonds. Thus, no contributions to the free energy difference arise from changes of the “hard” bond or angle terms. In addition, the net translation and rotation was removed from the guiding force to avoid unphysical contributions to the free energy. Figure 2 presents the potentials of mean force (PMF) around the two dihedral angles,  $\phi_1$  and  $\phi_2$ , of the two five-atomic benchmark systems II and III. Dark areas signify energy minima, while bright areas represent energy barriers. A dashed green line encircles the energy minima at an energy level of about  $9 k_B T$ , representing the areas of phase space that can usually be accessed by normal LD simulations within a few nanoseconds. It is evident that the two systems are designed to have little phase space overlap. The global energy minima of both states do not overlap; instead, the minima of system II are



**Figure 2.** Potentials of mean force of the five-atomic benchmark systems II and III as a function of the dihedral angles  $\phi_1$  and  $\phi_2$ . Dark areas represent energy minima. A green dashed line encircles the areas with energy levels below  $9 k_B T$ .

located in high energy regions of III and vice versa. In addition, the minima are separated by energy barriers greater than 10 kcal/mol. The five-atomic benchmark systems thus are one of the simplest possible systems for which enhanced sampling in one form or the other is essential to obtaining converged free energy results.

Table 4 lists the free energy difference results between II and III based on LD (second column) and reweighted SGLD (third

**Table 4. Free Energy Difference of Five-Atomic Benchmark Systems in Kilocalories Per Mole**

	LD	SGLD <sub>nr</sub>
reference		−14.28
BAR	no convergence	−14.28 ± 0.10
EDS ( $s = 1$ )	−11.91 ± 0.54	−14.13 ± 0.12
EDS ( $s = 0.1$ )	−12.83 ± 0.76	−14.11 ± 0.20
EDS ( $s = 0.01$ )	−14.05 ± 0.45	−14.08 ± 0.16
EDS ( $s = 0.001$ )	−14.20 ± 0.52	−14.25 ± 0.27

column) simulations. For LD, regular BAR simulations fail to converge, since there is no phase space overlap between the two end points. For EDS with LD, the deviations depend on the chosen  $s$ . For  $s = 1$  and 0.1 (corresponding to  $T_{\text{EDS}} = 300$  and 3000 K), the free energy results deviate from the reference results by 2 and 1.5 kcal/mol. Using thermodynamic integration, Leitgeb et al. found similar deviations for simulations that were trapped in a local energy minimum.<sup>39</sup> This demonstrates that, for values of  $s$  close to 1, EDS can also be afflicted by sampling problems. Starting from an  $s$  lower than 0.01 (which corresponds to an  $T_{\text{EDS}} > 30\,000$  K), EDS with LD is able to reproduce the reference results, but with relatively high standard deviations of greater than 0.45 kcal/mol.

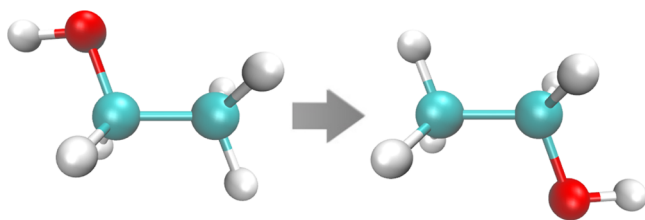
Without reweighting, the SGLD results are  $12.31 \pm 0.06$  for BAR and  $12.08 \pm 0.10$  for EDS with  $s = 1$  (an error of 2 kcal/mol), thus reflecting the bias introduced by SGLD. However, after reweighting, the BAR result is in excellent agreement with the analytical reference, while the deviations of EDS are minuscule (ranging between 0.03 to 0.2 kcal/mol). In addition, all SGLD results exhibit low standard deviations. This demonstrates that reweighting considerably improves the result. Although no energy offsets were used for these EDS calculations, the FES with SGLD still yields accurate results. This feature can be exploited for more advanced versions of EDS that include a parameter update scheme (e.g., ref 32). The key step for the fast convergence of EDS is finding the optimal value of the smoothing parameter  $s$  and the energy offsets  $E_k^R$ . The use of SGLD allows higher values of  $s$  and, thus, increases the efficiency of the calculations. While SGLD is probably not



able to obtain accurate free energy results in all cases, our data demonstrate that it can improve the convergence during the search for the correct  $E_k^R$  parameters, which reduces the time required for EDS.

In the following, we compare the present results with previously reported results by other FES methods with advanced sampling. In ref 39, non-Boltzmann thermodynamic integration with 11 intermediate steps and a total simulation time of 52 ns yielded a free energy difference of  $-14.43 \pm 0.17$  by using adaptive umbrella potentials. In ref 41, non-Boltzmann Bennett with a special umbrella potential and a total simulation time of 84 ns yielded a result of  $-14.38 \pm 0.20$ . This indicates that SGLD is able to improve the sampling compared to LD with an efficiency that is equal or superior to umbrella potentials.

**4.2.1. Ethanol.** In this section, we test the efficiency of the reweighting scheme in explicit solvent. For this purpose, we employ a benchmark system where the two end states are equivalent, and therefore, the total free energy difference is zero. Thus, no reference data are required. In particular, we focus on the alchemical mutation of ethan-1-ol to ethan-2-ol in a box of 766 TIP3P water molecules. In this mutation, the hydroxyl group is transferred from one carbon atom of ethanol to the opposite carbon atom (Figure 3), causing the hydration



**Figure 3.** Mutation of ethan-1-ol to ethan-2-ol in explicit water. Since the end states are equivalent, the free energy difference is zero.

shell to adapt to the new charge distribution. Since this happens on the time scale of a few picoseconds, sampling is not a limiting factor for the accuracy of this FES. However, the primary objective of this benchmark system is to illustrate the limitations of the reweighting scheme in free energy calculations. For this purpose, two different forms of SGLD are employed: In the first case (SGLD<sup>all</sup>), the guiding force is applied to all atoms, thus enhancing the motions of both solute and solvent. In the second setup (SGLD<sup>EtOH</sup>), the application of guiding forces is restricted to the solute, leaving the water unaffected.

The solvation free energy differences ( $\Delta\Delta A_{\text{solv}}$ ) between ethan-1-ol and ethan-2-ol in explicit solvent are listed in Table 5. The first three rows show the results without reweighting. The accuracy and precision of the LD results indicate that the

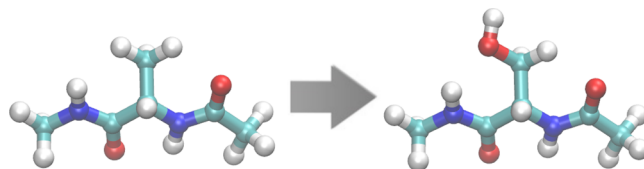
**Table 5. Free Energy Differences ( $\Delta\Delta A_{\text{solv}}$ ) between Ethan-1-ol and Ethan-2-ol in Explicit Solvent in Kilocalories Per Mole**

	$\Delta\Delta A_{\text{solv}}$
LD	$-0.04 \pm 0.05$
SGLD <sup>all</sup>	$-0.03 \pm 0.04$
SGLD <sup>EtOH</sup>	$0.02 \pm 0.03$
SGLD <sup>all</sup> <sub>rw</sub>	$-0.51 \pm 1.24$
SGLD <sup>EtOH</sup> <sub>rw</sub>	$-0.11 \pm 0.05$

simulation times used in this example are long enough to obtain converged results (exhibiting a deviation of  $-0.04$  kcal/mol from the ideal result). The unweighted SGLD results in rows two and three illustrate that the use of SGLD does not introduce any substantial additional errors (yielding deviations  $<0.03$  kcal/mol). However, when the reweighting equation is employed (indicated by the subscript “rw” in rows four and five), both errors and standard deviations increase dramatically. In particular, the application of the reweighting scheme to all atoms of the system (SGLD<sup>all</sup><sub>rw</sub>) leads to an error of  $-0.51$  kcal/mol, which is 10 times higher than that of LD. In addition, the standard deviations amount to 1.24 kcal/mol, which is  $\sim 30$  times higher than in LD.

This truly terrifying result can be explained in terms of the variance of the weighting factors  $w^{\text{SGLD}}$ . The weighting factors for SGLD<sup>all</sup><sub>rw</sub> range between  $10^{-10}$  and  $10^{+40}$ . Since a few data points carry weights that are several orders of magnitude above average, they dominate in the calculation of the reweighted free energy result. For example, the results for SGLD<sup>all</sup><sub>rw</sub> are largely determined from three data points out of the whole trajectory. This finding highlights that SGLD is not size-extensive; i.e., the employment of reweighting becomes more inefficient as the number of atoms of the system increases. Therefore, the use of SGLD should be restricted to the environment directly surrounding the site of interest in the FES. For example, if the application of SGLD is restricted to the solute (SGLD<sup>EtOH</sup><sub>rw</sub>), the weighting factors vary only between 0.1 and 10, leading to a reduction of the deviation to acceptable levels (0.11 kcal/mol) as well as yielding a standard deviation that is comparable to that of LD (0.05 kcal/mol).

**4.3. Alanine–Serine.** In the final example, we compare the performance of BAR and EDS in connection with SGLD for a more biologically relevant system. Here, we mutate Ala into Ser (see Figure 4) both in the gas phase and implicit solvent to



**Figure 4.** Illustration of the mutation of N-acetyl-methylamide alanine to serine.

compute the solvation free energy difference between the two amino acids. Since a dual topology setup is employed, no contributions from changes of the bond angles arise. The choice of this system is motivated by the high sensitivity of the free energy difference between Ala and Ser to the secondary structure of the backbone. Limited or biased sampling of the backbone can lead to free energy differences between  $-1.47$  and  $-4.45$  kcal/mol.<sup>40</sup> The mutation, therefore, provides means to test SGLD for biases in the sampling of secondary structures.

Table 6 lists the solvation free energy differences between Ala and Ser based on LD (second column) and reweighted SGLD (third column) simulations. Focusing on the upper half of Table 6, the solvation free energy differences from BAR and EDS are in excellent agreement. In addition, the results are in fair agreement with explicit and implicit solvent results in refs 40, 41, and 55. However, when comparing the LD results with the SGLD results, there is a subtle, but statistically significant deviation between the two BAR results. Although this deviation

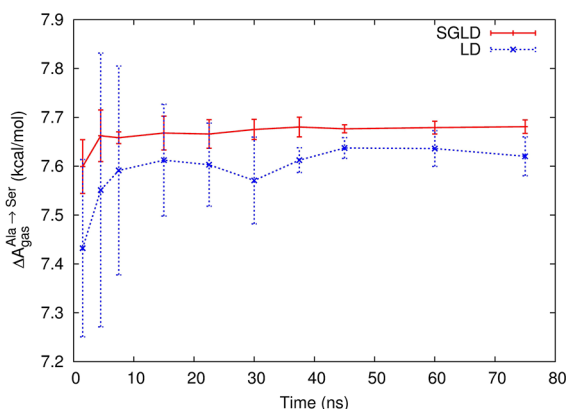
**Table 6.** Solvation Free Energy Differences between Alanine and Serine in Kilocalories Per Mole

	LD	SGLD <sub>rw</sub>
BAR	$-2.98 \pm 0.04$	$-3.12 \pm 0.03$
EDS	$-2.97 \pm 0.04$	$-3.01 \pm 0.03$
TP <sub>fw</sub> <sup>a</sup>	$-2.88 \pm 0.89$	$-2.84 \pm 1.07$
TP <sub>bw</sub> <sup>b</sup>	$-2.50 \pm 0.69$	$-3.10 \pm 0.43$
TP <sub>0.5</sub> <sup>c</sup>	$-2.82 \pm 0.28$	$-2.82 \pm 0.21$

<sup>a</sup>Thermodynamic perturbation based on trajectories of Ala.<sup>b</sup>Thermodynamic perturbation based on trajectories of Ser.<sup>c</sup>Thermodynamic perturbation based on trajectories of  $U_R = U_{Ala}/2 + U_{Ser}/2$ .

amounts to about 5% of the LD results, it should be perfectly acceptable in normal applications, especially considering that errors due to insufficient sampling can be significantly larger.

The effect of SGLD on sampling for this benchmark system is shown in Figure 5. The plot compares the convergence and

**Figure 5.** Comparison of the convergence of free energy simulations based on SGLD (red) and LD (blue). Gas phase free energy differences ( $\Delta A_{\text{gas}}^{\text{Ala} \rightarrow \text{Ser}}$ ) between alanine and serine resulting from BAR are plotted over time. The error bars at each data point are calculated from four different simulations.

standard deviations of LD and SGLD FES. Gas phase free energy differences ( $\Delta A_{\text{gas}}^{\text{Ala} \rightarrow \text{Ser}}$ ) between alanine and serine resulting from BAR are plotted over simulation time. While both LD and SGLD results converge in 10 ns to  $\Delta A_{\text{gas}}$  values that are within 0.05 kcal/mol of the final result, the standard deviations are significantly lower when using SGLD (0.05 compared to 0.28 kcal/mol after ~5 ns). This demonstrates that, even for relatively unproblematic systems like alanine and serine, the use of SGLD can increase the performance of FES.

The other motivation to employ this benchmark system is the capability of BAR to calculate its free energy difference in a single step. As shown previously,<sup>55</sup> the phase space overlap between alanine and serine is not a limiting factor for the accuracy of FES. Our aim here is a fair comparison of the efficiency of several free energy methods without the use of  $\lambda$  intermediate states. In the following, we rely on BAR as a reference. In particular, we compare the performance of EDS with BAR and thermodynamic perturbation based on trajectories of Ala (TP<sub>fw</sub>), Ser (TP<sub>bw</sub>), and a simple half of a d-half mixture of both end states (TP<sub>0.5</sub>). The results for those methods are shown in the lower part of Table 6.

On the basis of the EDS equation to calculate free energies, eq 18, one could wrongly assume that the performance of EDS

resembles TP since the end states are connected to the reference state by two TP steps. However, due to the characteristics of the reference state, the sampling and convergence properties of EDS and TP are different. The data in Table 6 show that EDS is more closely related to BAR than to the three different TP methods, both in terms of absolute value and standard deviations. While EDS is in almost perfect agreement with BAR, the deviations of the TP methods from the reference results range between 0.1 and 0.4 kcal/mol. The standard deviations of TP range between 0.28 and 0.89, while EDS yields the same standard deviations as BAR (0.04 kcal/mol). The close relationship between BAR and EDS was recently derived by theoretical means by Christ and van Gunsteren in ref 30.

## 5. CONCLUSIONS

From the results in this paper, it is possible to provide a few general guidelines for the use of SGLD in FES. The results for the four-atomic benchmark system show that the use of SGLD changes free energy differences of bonded terms of dummy atoms. These errors can be eliminated if the net rotation and translation is removed from the guiding force, reducing the RMSD of SGLD simulations from 0.250 to 0.03 kcal/mol. Therefore, this option will become the default of all future versions of SGLD in CHARMM. Alternatively, errors resulting from changes of the bonded terms can also be avoided by employing dual topology setups.

The results for the five-atomic benchmark system highlight that BAR is not able to yield correct results with LD due to the lack of phase space overlap between the two end points and the high energy barriers of over 10 kcal/mol ( $\sim 16k_B T$ ). Even higher energy barriers are quite common in biological systems, e.g., in the rearrangement of protein side chains.<sup>39</sup> In order to become a standard technique, FES will have to be able to deal with such an energy barrier without human intervention on a case by case basis. Considering that side chain motions or conformation changes in proteins can happen on a scale of microseconds to milliseconds, it becomes evident that sampling problems in MD simulations cannot be addressed by a simple increase of computer power (as was hoped for during past decades). With SGLD, both BAR and EDS lead to the correct results. Considering that SGLD does not require any prior knowledge about the system, it is an attractive choice for (semi)automatic FES protocols. For the examples presented here, the additional effort required for running SGLD with reweighting consisted in equilibrating the system with LD and executing an awk script after the termination of the SGLD simulation.

Notably, EDS simulations with LD yield the correct results for the five-atomic benchmark system if the  $s$  parameter is lower than 0.01. However, this comes at the cost of greater standard deviations than in any corresponding SGLD simulation. This indicates that SGLD can be used to improve the convergence of EDS. In addition, SGLD allows the use of higher  $s$  values, which allows a more efficient use of the EDS update scheme. Another interesting aspect of the results for the five-atomic benchmark systems is that we did not use any energy offsets to balance the sampling between the two end states. This suggests that SGLD can improve the convergence of the more advanced versions of EDS such as used in refs 31 and 32.

The next example, which involves an FES of ethanol in explicit solvent, shows that the performance of SGLD is limited by the size of the system since the variance of the weighting

factors increases with system size. This leads to excessively high standard deviations and inaccuracies of the free energy results, as computed ensemble averages are dominated by the outliers with the highest weight. In extreme cases, the final free energy result will be determined by a single data point. This is a general limitation of FESs that rely on biased sampling followed by reweighting. For optimal performance, it could be necessary to limit the application of SGLD to the solute and, in protein simulations, to atoms close to the site of interest. Alternatively, it might be interesting to explore size-extensive methods.<sup>56</sup>

The last example, involving the mutation of Ala to Ser, is a very sensitive benchmark for the sampling of the backbone by SGLD. As shown recently, the correct generation of the backbone probability distributions is essential for reproducing experimental results.<sup>57</sup> Previous studies demonstrate that the solvation free energies of Ala and Ser are very sensitive to the secondary structure of the backbone, causing deviations of up to 2.57 kcal/mol.<sup>40</sup> For BAR, we find a statistically significant difference of 0.14 kcal/mol between LD and SGLD, which suggests a slight bias due to SGLD. However, this deviation can be considered small in comparison to other sources of error, and the detected loss of accuracy is outweighed by a faster convergence of the results.

In addition, we employ the mutation of Ala to Ser to compare the performance of EDS with BAR and three different kinds of TP. In particular, we focus on TP based on the trajectories of Ala and Ser, as well as on an intermediate state based on  $U_R = U_{\text{Ala}}/2 + U_{\text{Ser}}/2$ . Our data support the hypothesis that the sampling and convergence properties of EDS are different from TP and more closely related to BAR. Even the simplified version of EDS exhibits smaller deviations from the BAR reference results than all versions of TP considered here. The standard deviations from EDS are considerably smaller than those of TP, being equal to the results from BAR. Our results indicate that EDS is a viable method for performing free energy calculations from a single simulation. In particular, EDS is a valuable tool in cases where the simulation of multiple intermediate states would be prohibitively expensive in terms of computational time or human resources. However, it is necessary to keep in mind that the convergence of the EDS results will become poorer with decreasing overlap of the phase spaces of the end points involved. In such cases, the performance of EDS can be improved by optimizing the EDS reference state Hamiltonian (e.g., using an  $s < 1$ ) as well as by increasing the sampling with SGLD.

## ■ APPENDIX 1: ANALYTICAL CALCULATION OF FREE ENERGY DIFFERENCES BETWEEN POLYATOMIC MOLECULES

As reference results for the free energy calculations of the four- and five-atomic benchmark systems, we employ analytical results based on the derivations by Herschbach, Johnston, and Rapp (HJR). In work completely unrelated to computer simulations, HJR introduced an alternative formulation of how to compute the configurational partition function of a polyatomic molecule,<sup>58</sup> which turned out to be extremely useful for computing alchemical free energy differences between such moieties analytically.<sup>35</sup> Summarizing the result of ref 58, the configurational partition function  $Z$  of a polyatomic molecule in the rigid rotator harmonic oscillator (RRHO) approximation is given by

$$Z \approx Z^{\text{RRHO}} = \frac{(2\pi k_B T)^{(3N-6)/2}}{|\mathbf{F}|^{1/2}} J \quad (22)$$

where  $N$  is the number of atoms and  $k_B$  and  $T$  are the Boltzmann constant and the temperature.  $J$  is the Jacobian factor arising from the change of Cartesian to internal coordinates. The crucial element of HJR was to show that  $J$  can be computed as a simple product of factors, one for each atom, i.e.,  $J = \prod_{\alpha=1}^N J_{\alpha}$  (see Table 1 of ref 58).  $\mathbf{F}$  is the force matrix, and in the simple cases considered here (unbranched four- and five-atomic molecules) its determinant is just the product of all force constants present in the system. Equation 22 implicitly assumes dihedral angles to be harmonic as well; however, HJR also show how hindered rotation can be accounted for correctly.

In the RRHO approximation, i.e., assuming that all bonded energy terms are harmonic, the alchemical free energy difference between two molecules ( $i, f$ ) with the same number of atoms is, therefore, given by

$$\Delta A = -k_B T \ln \frac{\prod_{\alpha=1}^N J_{\alpha}^f}{\prod_{\alpha=1}^N J_{\alpha}^i} - k_B T \ln \left( \frac{|\mathbf{F}|^i}{|\mathbf{F}|^f} \right)^{1/2} \quad (23)$$

Keeping in mind that  $|\mathbf{F}|$  in simple situations is just a product of force constants, only those parts of the system that are different between states  $i$  and  $f$  contribute to  $\Delta A$ . Suppose that two unbranched, nonlinear four-atomic molecules are mostly identical but differ by the bond length between atoms 3 and 4 ( $r_{34}$ ) and by the force constants for the bond stretching term between atoms 3 and 4 ( $K_{34}$ ), as well as the angle bending term for atoms 2, 3, and 4 ( $K_{234}$ ), then the free energy difference between the two hypothetical molecules is given by (see refs 35, 36, and 58, for more details and further examples)

$$\Delta A = -k_B T \ln \left( \frac{r_{34}^f}{r_{34}^i} \right)^2 - k_B T \ln \left( \frac{K_{234}^i K_{34}^i}{K_{234}^f K_{34}^f} \right)^{1/2} \quad (24)$$

In the present work, we even went beyond the RRHO approximation and treated the dihedral terms exactly. Following HJR, the terms are taken out of the force matrix, and their contribution to the partition function is obtained by numerical integration, i.e., for a four-atomic molecule with one dihedral angle term

$$Z = Z^{\text{RRHO}} \int d\varphi \exp \left( -\frac{U_{\text{dih}}}{k_B T} \right) \quad (25)$$

Here,  $Z^{\text{RRHO}}$  accounts for the bond stretching and bond angle terms in excellent approximation, whereas the explicit integration over the dihedral angle takes care of the anharmonic dihedral potential. For the five-atomic model systems, there are two dihedral degrees of freedom, and the intramolecular nonbonded interactions were included as well by expressing the distances between atom pairs 1–4, 2–5, and 1–5 in terms of internal coordinates (see ref 39 for further details).

## ■ APPENDIX 2: EXAMPLE SCRIPT TO USE EDS WITH MSCALE

Here, we show an extract from an example script for the use of the EDS implementation in CHARMM. First, two end points called “first” and “second” are set up with MSCALE, spawning two processes that will use one processor each (nproc). The



Hamiltonians of each end point are defined in the input scripts sub1.inp and sub2.inp:

```
mscale nsubs 2
subs first coef 1.0 prog "../exec/gnu/charmm" -
  outp "sub1.out" inpu "sub1.inp" -
  nproc 1 sele all end
subs second coef 1.0 prog "../exec/gnu/charmm" -
  outp "sub2.out" inpu "sub2.inp" -
  nproc 1 sele all end
end
```

After setting up the end points with MSCALE, the EDS command has to be invoked. Here, an EDS temperature  $T_{\text{EDS}} = 300$  is chosen, which is equivalent to an  $s = 1$  at  $T = 300$ . The number of subsystems for EDS (neds) is two, followed by the names of the subsystems and their energy offsets (with  $E_{\text{first}}^{\text{R}} = 1.1$  and  $E_{\text{second}}^{\text{R}} = 2.2$ ).

```
eds temp 300. neds 2 term first 1.1 term second 2.2
```

After the EDS command, the normal CHARMM commands to start a simulation may be used.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: gerhard.koenig@nih.gov.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors would like to thank R. Pastor for critically reading the manuscript. This work was primarily supported by the intramural research program of NHLBI, NIH. The authors would like to dedicate this article to Wilfred van Gunsteren in honor of his 65th birthday.

## REFERENCES

- (1) Oostenbrink, C.; van Gunsteren, W. Free energies of ligand binding for structurally diverse compounds. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6750–6754.
- (2) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **2007**, *371*, 1118–1134.
- (3) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.
- (4) Kästner, J.; Senn, H.; Thiel, S.; Otte, N.; Thiel, W. QM/MM free-energy perturbation compared to thermodynamic integration and umbrella sampling: Application to an enzymatic reaction. *J. Chem. Theory Comput.* **2006**, *2*, 452–461.
- (5) Sund, J.; Ander, M.; Aqvist, J. Principles of stop-codon reading on the ribosome. *Nature* **2010**, *465*, 947–U12.
- (6) Merz, K. M. Limits of Free Energy Computation for Protein-Ligand Interactions. *J. Chem. Theory Comput.* **2010**, *6*, 1018–1027.
- (7) Muddana, H. S.; Varnado, C. D.; Bielawski, C. W.; Urbach, A. R.; Isaacs, L.; Geballe, M. T.; Gilson, M. K. Blind prediction of host–guest binding affinities: a new SAMPL3 challenge. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 475–487.

- (8) Gallicchio, E.; Levy, R. M. Prediction of SAMPL3 host-guest affinities with the binding energy distribution analysis method (BEDAM). *J. Comput.-Aided Mol. Des.* **2012**, *26*, 505–516.
- (9) Lawrenz, M.; Wereszczynski, J.; Ortiz-Sánchez, J. M.; Nichols, S. E.; McCammon, J. A. Thermodynamic integration to predict host-guest binding affinities. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 569–576.
- (10) Mobley, D. L.; Liu, S.; Cerutti, D. S.; Swope, W. C.; Rice, J. E. Alchemical prediction of hydration free energies for SAMPL. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 551–562.
- (11) König, G.; Brooks, B. R. Predicting binding affinities of host-guest systems in the SAMPL3 blind challenge: the performance of relative free energy calculations. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 543–550.
- (12) Mobley, D. Let's get honest about sampling. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 93–95.
- (13) Wu, X.; Brooks, B. R. Self-guided Langevin dynamics simulation method. *Chem. Phys. Lett.* **2003**, *381*, 512–518.
- (14) Damjanović, A.; Miller, B. T.; Wenaus, T. J.; Maksimović, P.; García-Moreno, E.; Brooks, B. R. Open Science Grid Study of the Coupling between Conformation and Water Content in the Interior of a Protein. *J. Chem. Inf. Model.* **2008**, *48*, 2021–2029.
- (15) Damjanović, A.; Wu, X.; García-Moreno, B. E.; Brooks, B. R. Backbone Relaxation Coupled to the Ionization of Internal Groups in Proteins: A Self-Guided Langevin Dynamics Study. *Biophys. J.* **2008**, *95*, 4091–4101.
- (16) Damjanović, A.; García-Moreno, E.; Brooks, B.; Self-guided, B. R. Langevin dynamics study of regulatory interactions in NtrC. *Proteins* **2009**, *76*, 1007–1019.
- (17) Lee, M. S.; Olson, M. A. Protein Folding Simulations Combining Self-Guided Langevin Dynamics and Temperature-Based Replica Exchange. *J. Chem. Theory Comput.* **2010**, *6*, 2477–2487.
- (18) Lee, C.-I.; Chang, N.-Y. Characterizing the denatured state of human prion 121-230. *Biophys. Chem.* **2010**, *151*, 86–90.
- (19) Wu, X.; Brooks, B. R. Toward canonical ensemble distribution from self-guided Langevin dynamics simulation. *J. Chem. Phys.* **2011**, *134*, 134108.
- (20) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187.
- (21) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420.
- (22) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.
- (23) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (24) Shirts, M. R.; Pande, V. S. Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys.* **2005**, *122*, 144107–1–144107–16.
- (25) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*.
- (26) Paliwal, H.; Shirts, M. R. A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *J. Chem. Theory Comput.* **2011**, *7*, 4115–4134.
- (27) Bruckner, S.; Boresch, S. Efficiency of Alchemical Free Energy Simulations I: Practical Comparison of the Exponential Formula, Thermodynamic Integration and Bennett's Acceptance Ratio Method. *J. Comput. Chem.* **2011**, *32*, 1303–1319.
- (28) Oostenbrink, C.; van Gunsteren, W. F. Free energies of ligand binding for structurally diverse compounds. *PNAS* **2005**, *102*, 6750–6754.
- (29) Christ, C.; van Gunsteren, W. Enveloping distribution sampling: A method to calculate free energy differences from a single simulation. *J. Chem. Phys.* **2007**, *126*, 184110.
- (30) Christ, C. D.; van Gunsteren, W. F. Multiple free energies from a single simulation: Extending enveloping distribution sampling to

nonoverlapping phase-space distributions. *J. Chem. Phys.* **2008**, *128*, 174112.

(31) Christ, C. D.; van Gunsteren, W. F. Comparison of Three Enveloping Distribution Sampling Hamiltonians for the Estimation of Multiple Free Energy Differences from a Single Simulation. *J. Comput. Chem.* **2009**, *30*, 1664–1679.

(32) Riniker, S.; Christ, C. D.; Hansen, N.; Mark, A. E.; Nair, P. C.; van Gunsteren, W. F. Comparison of enveloping distribution sampling and thermodynamic integration to calculate binding free energies of phenylethanolamine N-methyltransferase inhibitors. *J. Chem. Phys.* **2011**, *135*, 024105.

(33) McQuarrie, D. A. *Statistical Mechanics*; Harper & Row: New York, 1976.

(34) Knox, J. H. *Molecular Thermodynamics*; John Wiley & Sons: New York, 1978.

(35) Boreesch, S.; Karplus, M. The Jacobian factor in free energy simulations. *J. Chem. Phys.* **1996**, *105*, 5145–5154.

(36) Boreesch, S.; Karplus, M. The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis. *J. Phys. Chem. A* **1999**, *103*, 103–118.

(37) Boreesch, S.; Karplus, M. The Role of Bonded Terms in Free Energy Simulations. 2. Calculation of Their Influence on Free Energy Differences of Solvation. *J. Phys. Chem. A* **1999**, *103*, 119–136.

(38) Boreesch, S. The role of bonded energy terms in free energy simulations — insights from analytical results. *Mol. Simul.* **2002**, *28*, 13–37.

(39) Leitgeb, M.; Schröder, C.; Boreesch, S. Alchemical free energy calculations and multiple conformational substates. *J. Chem. Phys.* **2005**, *122*, 084109.

(40) König, G.; Boreesch, S. Hydration Free Energies of Amino Acids: Why Side Chain Analog Data Are Not Enough. *J. Phys. Chem. B* **2009**, *113*, 8967–8974.

(41) König, G.; Boreesch, S. Non-Boltzmann Sampling and Bennett's Acceptance Ratio Method: How to Profit from Bending the Rules. *J. Comput. Chem.* **2011**, *32*, 1082–1090.

(42) Wereszczynski, J.; McCammon, J. A. Using Selectively Applied Accelerated Molecular Dynamics to Enhance Free Energy Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 3285–3292.

(43) Brooks, B.; Brooks, C., III; Mackerell, A., Jr.; Nilsson, L.; Petrella, R.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boreesch, S.; Caffisch, A.; Caves, L.; Cui, Q.; Dinner, A.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.; Post, C.; Pu, J.; Schaefer, M.; Tidor, B.; Venable, R.; Woodcock, H.; Wu, X.; Yang, W.; York, D.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

(44) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.

(45) Woodcock, H. L.; Miller, B. T.; Hodoscek, M.; Okur, A.; Larkin, J. D.; Ponder, J. W.; Brooks, B. R. MS-SCALE: A General Utility for Multiscale Modeling. *J. Chem. Theory Comput.* **2011**, *7*, 1208–1219.

(46) Christ, C. D.; van Gunsteren, W. F. Simple, Efficient, and Reliable Computation of Multiple Free Energy Differences from a Single Simulation: A Reference Hamiltonian Parameter Update Scheme for Enveloping Distribution Sampling (EDS). *J. Chem. Theory Comput.* **2009**, *5*, 276–286.

(47) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of protein. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(48) Humphrey, W.; Dalke, A.; Schulten, K. VMD — Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, No. 14, 33–38.

(49) Van Gunsteren, W. F.; Berendsen, H. J. C. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* **1977**, *34*, 1311–1327.

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

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

(52) Neria, E.; Fischer, S.; Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **1996**, *105*–1902.

(53) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(54) Lee, M. S.; Feig, M.; Salsbury, F. R.; Brooks, C. L., III. New analytic approximation to the standard molecular volume definition and its application to generalized born calculations. *J. Comput. Chem.* **2003**, *23*, 1348–1356.

(55) König, G.; Bruckner, S.; Boreesch, S. Unorthodox Uses of Bennett's Acceptance Ratio Method. *J. Comput. Chem.* **2009**, *30*, 1712–1718.

(56) Wu, X.; Brooks, B. R. Force-momentum-based self-guided Langevin dynamics: A rapid sampling method that approaches the canonical ensemble. *J. Chem. Phys.* **2011**, *135*.

(57) Best, R. B.; Buchete, N.-V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.* **2008**, *95*, L7–L9.

(58) Herschbach, D. R.; Johnston, H. S.; Rapp, D. Molecular Partition Functions in Terms of Local Properties. *J. Chem. Phys.* **1959**, *31*, 1652–1661.