JCTC Journal of Chemical Theory and Computation

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)

Clara D. Christ† and Wilfred F. van Gunsteren*,†

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zürich, Switzerland

Received October 7, 2008

Abstract: We present an automatic adaptive scheme which allows fast optimization of the reference Hamiltonian parameters in enveloping distribution sampling (EDS). Six different variants of the update scheme have been tested on a condensed phase test system which included the recurrent deletion and creation of complete water molecules in water. All six schemes gave accurate free energy estimates with absolute errors of up to 1 kJ/mol for the worst scheme and up to 0.1 kJ/mol for the best scheme. Configurational sampling is focused on the regions where the end state energy difference distributions intersect, explaining the high accuracy and precision of the free energy estimates. The new update scheme makes the application of EDS to other systems, e.g. in ligand binding studies, easy as no reference state Hamiltonian parameters have to be chosen by the user. The only necessary input are the Hamiltonians of the various end states involved.

1. Introduction

Estimation of free energies from molecular simulation has been an active field of research for several decades and many review articles and books are devoted to the topic.[1−18] Although the basic equations used for free energy calculations have been proposed decades ago,[19,20] it still remains challenging to accurately calculate free energies. This is due to two main challenges that have to be met when estimating free energies from molecular simulation. First, the system of interest, e.g. different ligands binding to a common receptor, has to be described by an appropriate model. Classical models are often used to this end as they allow computationally cheap evaluation of the energy and the forces. For the model to be appropriate it must describe the thermodynamics of the system correctly. References 21 and 22 compare the accuracy of several classical models used for free energy prediction. In the current work, we will focus on the second challenge which consists of finding an efficient evaluation scheme for the free energy. Estimation of free energies involves calculation of the partition function which is for most cases not accessible analytically. Fortunately, in many applications it suffices to calculate relative free energies, which reduces the problem to the estimation of partition function ratios, i.e. the estimation of relative probabilities. As an example, let us consider the free energy of folding of a peptide. Let us assume that we have a clear measure of when the peptide is in the folded state and when in the unfolded. Further assume that we have a sampling scheme (e.g., molecular dynamics) that efficiently samples configuration space. We can then estimate the free energy of folding from the ratio of number of folded configurations to unfolded configurations we have encountered. This simple example illustrates in an intuitive fashion (see section 2 for more rigorous arguments) that an efficient simultaneous sampling of the configuration space of all end states involved allows the calculation of the free energy differences. In this example, the two states were defined by the same Hamiltonian plus a set of criteria or constraints that distinguish the folded from the unfolded state. If the two or more end states correspond, however, e.g. to two chemically distinct species a combined Hamiltonian has to be constructed. This is often

* Corresponding author. e-mail: wfvgn@igc.phys.chem.ethz.ch.
† Swiss Federal Institute of Technology.

10.1021/ct800424v CCC: $40.75 © 2009 American Chemical Society
Published on Web 01/09/2009

done by the so-called coupling parameter approach[23] where the Hamiltonian H is now a function of the coupling parameter λ, and λ is chosen such that $H(\lambda = 0)$ corresponds to state A and $H(\lambda = 1)$ to state B. This basic form of the Hamiltonian is used in many methods; however, there are many variants: first, the functional dependence on λ differs in various methods, and second, the way the configuration space of this newly constructed combined state is sampled may differ: either λ is a parameter and the sampling over the whole λ range is achieved by performing multiple simulations at fixed λ-values (as e.g. in thermodynamic integration (TI),[19] free energy perturbation (FEP),[20] the Bennett acceptance ratio method (BAR),[24] overlap-sampling,[25] or also in Hamiltonian-replica-exchange),[26−28] or λ is allowed to change during the simulation either by Metropolis Monte Carlo moves like in chemical Monte-Carlo molecular dynamics (CMC/MD),[29] or because it is treated as a dynamic variable as in λ-dynamics.[30] If λ is treated dynamically, a biasing scheme[31] has to be used in order to ensure sampling over the whole λ range. Alternatively, free energy estimates can also be obtained from multiple irreversible, independent simulations in which λ is changed so fast that the system is driven out of equilibrium.[32−34]

If the important configuration space of the λ combined Hamiltonian is well sampled during the simulation it contains as a subset the important configuration space of the endstates A and B, allowing for an accurate estimation of the free energy difference.[35] However, the coupling parameter way of combining the end state Hamiltonians has two major drawbacks: first, a lot of computer time is spent on the simulation of regions of configurational space at intermediate λ values and, second, an extension of the approach to multiple end states rapidly makes the calculation unfeasible.

As traditional methods to calculate free energy differences are not easily extended to multiple end states, we have developed enveloping distribution sampling (EDS).[36,37] EDS is an implementation of the umbrella sampling method[31] and belongs to the class of importance sampling[38] methods. EDS is designed to allow for sampling of the important configuration space of multiple end states during a single simulation of a so-called reference state. This goal is also pursued in single-step perturbation.[39,40] Unlike the "hand-made" reference states used in single-step perturbation, EDS uses the strategy of expanding the sampled ensemble[41−46] using the following Hamiltonian $H_R = -(\beta s)^{-1} \ln \sum_i^N \exp[-\beta s(H_i - E_i^R)]$[37,42,47,48] (see section 2) where the energy offset parameters E_i^R ensure equal sampling of all N end states and the smoothness parameter s is chosen such that transitions between regions of phase space important to the different end states can occur, i.e. as low as necessary to overcome barriers and as high as possible in order to avoid unnecessary widening of the configuration space that has to be sampled during the simulation. As we have shown in previous work,[36] the accuracy of free energy differences, which can be calculated "on the fly" from the reference state simulation, strongly depends on the chosen parameters. In this work we, therefore, develop and compare schemes that allow an automatic update of the reference Hamiltonian parameters.

In section 2, the EDS working equations are derived together with the equations used for parameter update. Section 3 states the simulation protocols and the followed parameter update schemes. As a test system, we chose annihilation and creation of five water molecules in liquid water (see section 3). The system is designed such that all free energy differences are zero, i.e. the exact result is known. However, it is a highly challenging system as a water molecule is annihilated at one point in space and created at another point in space when moving from the important configuration space of one end state to that of another. In section 4 the results of the different update schemes are presented and discussed.

2. Theory

The free energy F of a state X is defined as

$$F_X = -\beta^{-1} \ln Q_X \tag{1}$$

where $\beta^{-1} = k_B T$, k_B is Boltzmann's constant, T is the absolute temperature, and Q_X is the partition function of state X

$$Q_X = (h^{3N_p} N_p!)^{-1} \int\int \exp[-\beta H_X(\mathbf{p}, \mathbf{r})]\, d\mathbf{p}\, d\mathbf{r} \tag{2}$$

Here, h is Planck's constant, N_p is the number of particles, $\mathbf{r}$ and $\mathbf{p}$ are the $3N$-dimensional vectors of the particle positions and conjugate momenta, respectively, and H_X is the Hamiltonian of state X. The factor $(N_p!)^{-1}$ only occurs for indistinguishable particles. If the Hamiltonian can be split into a kinetic $K_X(\mathbf{p})$ and a potential energy part $V_X(\mathbf{r})$,

$$H_X(\mathbf{p}, \mathbf{r}) = K_X(\mathbf{p}) + V_X(\mathbf{r}) \tag{3}$$

a separation of the corresponding free energies is allowed,

$$F_X = -\beta^{-1} \ln\{\int \exp[-\beta K_X(\mathbf{p})]\, d\mathbf{p}\} - \beta^{-1} \ln\{\int \exp[-\beta V_X(\mathbf{r})]\, d\mathbf{r}\} + \beta^{-1} \ln(h^{3N_p} N_p!) \tag{4}$$

In molecular simulations, the kinetic part of the Hamiltonian typically reads $K(\mathbf{p}) = \sum_{i=1}^N \mathbf{p}_i^2/(2m_i)$ and, therefore, the first integral in eq 4 can be solved analytically. In the following we will omit the kinetic part and the constant term for simplicity. Unlike the kinetic term the potential energy part of the Hamiltonian usually involves interaction terms that cannot be separated, making it impossible to solve the second integral (the configurational integral) in eq 4 analytically. Therefore, the free energy of a multidimensional system such as e.g. a molecule in solution cannot be calculated using eq 4. Rather than calculating the absolute free energy by evaluation of the configurational integral, one can calculate the free energy difference between two states

$$\Delta F_{BA} = F_B - F_A = -\beta^{-1} \ln \frac{Q_B}{Q_A} \tag{5}$$

where one now has to calculate a ratio of partition functions rather than the partition functions themselves. Inserting the expressions for the partition functions one finds[20]

$$\Delta F_{BA} = -\beta^{-1} \ln\langle \exp[-\beta(V_B - V_A)] \rangle_A \tag{6}$$

where $\langle\rangle_A$ indicates an average over an ensemble sampled at state A. Unfortunately, this expression will only yield reasonable free energy estimates if the important phase space of state B is a subset of the important phase space of state A.[35] Moreover, the expected error is minimal only for $A = B$.[24,37]

Introducing the energy difference distributions

$$\rho_A(\Delta V;\Delta V_{BA}) = \langle\delta[\Delta V - (V_B - V_A)]\rangle_A$$
$$\rho_B(\Delta V;\Delta V_{BA}) = \langle\delta[\Delta V - (V_B - V_A)]\rangle_B \quad (7)$$

we obtain[49−51]

$$\rho_B(\Delta V;\Delta V_{BA})\exp[-\beta\Delta F_{BA}] = \rho_A(\Delta V;\Delta V_{BA})\exp[-\beta\Delta V] \quad (8)$$

or

$$\ln\frac{\rho_A(\Delta V;\Delta V_{BA})}{\rho_B(\Delta V;\Delta V_{BA})} = +\beta\Delta V - \beta\Delta F_{BA} \quad (9)$$

That is, an alternative way to calculate the free energy difference ΔF_{BA} is to calculate the $\rho_A(\Delta V; \Delta V_{BA})$ and $\rho_B(\Delta V; \Delta V_{BA})$ distributions, e.g. from two simulations at states A and B. Equation 8 indicates that the free energy difference is the energy difference ΔV where these two distributions intersect or, alternatively, one can use eq 9 to do a linear regression and obtain $-\beta\Delta F_{BA}$ as the ordinate intercept. These equations show that in order to obtain reasonable relative free energy estimates the energy difference distributions must overlap. If the important phase space regions of states A and B lie far apart, this will, however, not be the case.

In order to sample the $\rho_A(\Delta V; \Delta V_{BA})$ and $\rho_B(\Delta V; \Delta V_{BA})$ distributions over the necessary range and to ensure that the important phase space of all end states is properly sampled, one can estimate the free energy difference between states A and B via the simulation of a reference state R,

$$\Delta F_{BA} = -\beta^{-1}\ln\frac{\langle\exp[-\beta(V_B - V_R)]\rangle_R}{\langle\exp[-\beta(V_A - V_R)]\rangle_R} \quad (10)$$

A reference state Hamiltonian which does sample the important phase space of both A and B reads[37,42,47,48]

$$V_R(r) = -(\beta s)^{-1}\ln\left\{\sum_{i=1}^{N}\exp[-\beta s(V_i(r) - E_i^R)]\right\} \quad (11)$$

where N denotes the number of EDS states, e.g. in the case of two states A and B, $N = 2$. E_i^R are energy offset parameters, and $s > 0$ is the dimensionless smoothness parameter. For $E_i^R = F_i$ and $s = 1$, eq 11 is the Hamiltonian that minimizes the expected error of eq 10.[37] Figure 1 shows a pictorial representation of the reference state of three end states V_A, V_B, and V_C (shown as black lines). V_{R1} corresponds to an unoptimized reference state Hamiltonian i.e. $E_i^R = 0$. The middle and the lower panel show that this leads to uneven sampling of the important configuration space of the three states. For example, configurations important to the B state are hardly important for V_{R1}. Equal sampling of all end states can be obtained by setting the energy offsets E_i^R to the corresponding free energies F_i. These are, however, not known in the beginning and have to be updated iteratively.

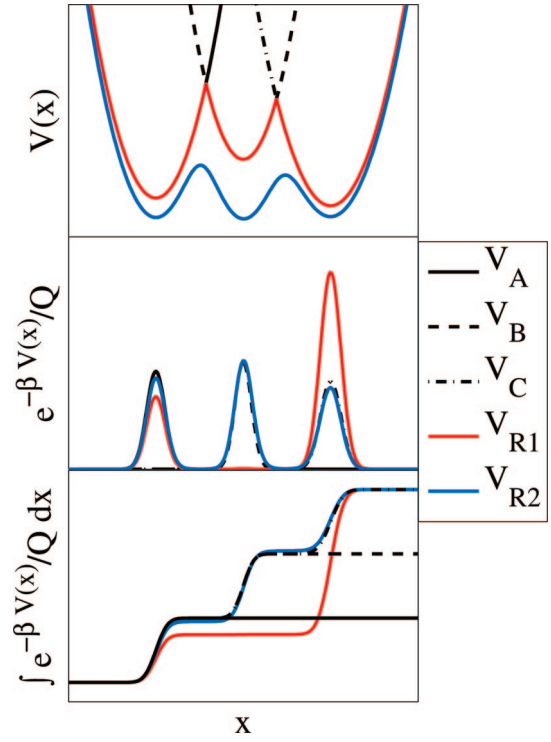

Figure 1. Pictorial representation of the influence of the parameters s and E_i^R of the reference Hamiltonian. V_{R1} is an unoptimized reference state Hamiltonian which does not lead to equal sampling of all three end states (V_A, V_B, and V_C). V_{R2} is optimized such that barriers are reduced and all states are sampled evenly. (Note that for better comparison, the end state distributions are divided by three in the central plot. In the bottom plot, the integrals over these distributions are summed up.)

Adjusting the energy offset parameters is however not sufficient as high barriers on the reference state potential energy surface between regions of configuration space important to one state and to another one may prevent efficient sampling. The smoothness parameter s can be lowered in order to decrease these barriers. Adjustment of energy offsets and smoothness parameters has been done for V_{R2} which now ensures that the important configuration space of all end states is sampled.

When using the reference state Hamiltonian eq 11 in a molecular dynamics implementation, one has to integrate the equations of motion,

$$\begin{aligned}\dot{\mathbf{r}}_k(t) &= m^{-1}\mathbf{p}_k(t)\\ \dot{\mathbf{p}}_k(t) &= \mathbf{f}_k(t) = \left(-\frac{\partial V_R(\mathbf{r})}{\partial \mathbf{r}_k}\right)\\ &= \sum_{i=1}^{N}\left\{\frac{\exp[-\beta s(V_i - E_i^R)]}{\sum_{j=1}^{N}\exp[-\beta s(V_j - E_j^R)]}\left(-\frac{\partial V_i(\mathbf{r})}{\partial \mathbf{r}_k}\right)\right\}\\ &= \sum_{i=1}^{N}\left\{\left[\sum_{\substack{j=1\\ j\neq i}}^{N}\exp[-\beta s(\Delta V_{ji} - \Delta E_{ji}^R)] + 1\right]^{-1}\left(-\frac{\partial V_i(\mathbf{r})}{\partial \mathbf{r}_k}\right)\right\}\end{aligned} \quad (12)$$

where $\Delta V_{ji} = V_j - V_i$. That is, the force on an atom is a sum over force contributions of all different states. All force

contributions are multiplied with a prefactor and all these prefactors add up to one. The parameters in the prefactor determine how often a state is visited (E_i^R) and whether transitions from configuration space important to one state to that important to another state are possible (s).

In order to update the energy offset parameters, one has to "count" how often a state is visited. This can be done by evaluating

$$E_i^R(\text{new}) = -\beta^{-1} \ln\left\langle \left(\sum_{\substack{j=1 \\ j\neq i}}^{N} \exp[-\beta(\Delta V_{ji} - \Delta E_{ji}^R)] + 1 \right)^{-1} \right\rangle_R + E_i^R \quad (13)$$

for all states i. If a configuration is important to state i and not important to the other states j then $\Delta V_{ji} = V_j - V_i$ is a big positive number and all exponential functions in eq 13 are approximately zero, i.e. the "counting function" (inside $\langle\rangle$) returns a value close to one. If a configuration is not important to i but to a another state k, then for $j = k$ the exponential function will return a big number, i.e. the counting function returns a value close to zero. So, if a state i is insufficiently visited, the energy offset E_i^R is raised in order to increase the number of visits of state i in the next iteration. Note that the energy offset is $E_i^R(\text{new}) = \Delta F_{iR}$ if $s = 1$, which can be verified by inserting the reference state Hamiltonian (eq 11) into the equation for ΔF_{iR}.

If the important parts of phase space of two states i and j lie far apart then ΔV_{ji} (see eq 12) is likely to be big and no transitions between these regions of phase space will occur. By lowering s ($s > 0$), one can compensate for this. An optimized s parameter is obtained by solving

$$\ln \sum_{\substack{j=1 \\ j\neq i}}^{N} \left\{ \left[\langle \exp\left[-\beta\left(|\Delta V_{ji}| - \Delta E_{ji}^R\right)\right] \rangle_i \right]^s \right\} = \ln(N-1) - 1 \quad (14)$$

numerically for s for all states i and taking the lowest s. The reasoning behind this heuristic optimization equation is the following. Assume we are sampling configurations which are of importance to state i but unfavorable for all the other $N - 1$ states. Then all ΔV_{ji} will be in general big positive numbers ($V_j \gg V_i$). Therefore, the force prefactor in front of $(-\partial V_i(\mathbf{r})/\partial \mathbf{r}_k)$ (see eq 12) will be approximately one whereas all the other force prefactors ($i' \neq i$) will be approximately zero and no transitions from state i to any of the other states i' will be observed. However, if s ($s > 0$) is lowered, transitions to the other states can occur. How much s needs to be lowered depends on the magnitude of $\Delta V_{ji} - \Delta E_{ji}^R$ when sampling configurations that are of importance to state i. An obvious choice to estimate the average magnitude of this quantity would be to evaluate $\langle \Delta V_{ji} - \Delta E_{ji}^R \rangle_i$. As the distribution of ΔV_{ji} values might, however, be very broad and have tails at very high energies, we chose to use

$$-\beta^{-1} \ln\langle \exp\left[-\beta(|\Delta V_{ji}| - \Delta E_{ji}^R)\right] \rangle_i \quad (15)$$

instead. This is an estimate of the smallest $\Delta V_{ji} - \Delta E_{ji}^R$ values observed when sampling configurations of importance to state i, i.e. the configurations which have most overlap with state j and where a transition to this state is most likely. Taking the absolute value is done for numerical reasons. In general ΔV_{ji} will be positive if a configuration is of importance to state i and negative if it is of importance to state j. In the latter case, $\exp[-\beta(\Delta V_{ji} - \Delta E_{ji}^R)]$ can be very large and might contribute to the calculated ensemble average (eq 15) although the configuration has negligible importance for the i ensemble. The basic (heuristic) Ansatz is now to enforce $\beta^{-1} = s(\Delta V_{ji} - \Delta E_{ji}^R)$ (see eq 12). Substituting this and eq 15 into the equation for the force prefactor (see eq 12), we obtain

$$\sum_{\substack{j=1 \\ j\neq i}}^{N} \exp\left[-\beta s\left(-\beta^{-1} \ln\langle \exp\left[-\beta(|\Delta V_{ji}| - \Delta E_{ji}^R)\right] \rangle_i\right)\right] = \sum_{\substack{j=1 \\ j\neq i}}^{N} \exp[-\beta\beta^{-1}] \quad (16)$$

which, after some rearrangement, leads to eq 14.

3. Simulation Protocols

The test system consisted of five (solute) water molecules in a cubic box of (solvent) water (1175 simple point charge (SPC)[52] water molecules in total, box length: 3.31 nm). In each of the five states, one of the five (solute) water molecules was interacting with the solvent water molecules while the other four were noninteracting ("dummy molecules"), i.e. their nonbonded interactions with the other water molecules were set to zero. Therefore, all states consist of one interacting solute water molecule and four noninteracting water molecules in a box of solvent water. This implies that all states have the same free energy.

All simulations were performed under *NVT* conditions using the weak-coupling method[53] ($T = 300$ K, $\tau_T = 0.1$ ps, and solute and solvent coupled to separate temperature baths). The translational motion of and the rotational motion around the center of mass was removed every 10 000 steps. Energies of the reference state and the end states were saved every 0.1 ps. The SHAKE algorithm[54] (tolerance: 0.0001) was used to constrain all bond lengths and angles to their ideal values. Nonbonded interactions were calculated using the triple range method[55,56] ($R_{\text{CUTP}} = 0.8$ nm, $R_{\text{CUTL}} = R_{\text{rf}} = 1.4$ nm, $\varepsilon_{\text{rf}} = 61$, $\kappa = 0$). The pairlist was updated every fifth step. The leapfrog algorithm was used to integrate Newton's equations of motion ($\Delta t = 0.002$ ps).

In order to test the efficiency of the automatic parameter update using eqs 13 and 14, we chose initial values for the parameters of the reference state Hamiltonian which were far from the optimal ones. The five initial energy offsets were chosen to be $E_i^R = \{0, 50, 100, 150, 200\}$ kJ/mol and the initial smoothness parameter was $s = 1$. Note that after each optimization of parameters the energy offsets were made relative to E_1^R. This has no influence on the trajectories or the calculated free energies as only differences of energy offset parameters occur in eq 12.

We tested six different schemes to update the parameters of the reference state Hamiltonian. These schemes, which are listed below, differ in three main points:

• Whether eqs 13 and 14 are iterated until convergence in the new parameters is obtained (denoted as "reweighting" in the listing below) or whether the calculation is stopped after one iteration (no reweighting). As eq 13 involves an ensemble average over the reference state ensemble, we calculate this average by reweighting the configurations, i.e.

$$\langle X\rangle_{R_{\mathrm{new}}} = \langle X \exp[-\beta(V_{R_{\mathrm{new}}} - V_R)]\rangle_R / \langle \exp[-\beta(V_{R_{\mathrm{new}}} - V_R)]\rangle_R \quad (17)$$

Although this numerical iteration of eqs 13 and 14 has negligible cost compared to the time spent in generating the configurations of the reference state ensemble and should in principle ensure much faster convergence of the energy offset and s parameters, it might still not be useful to iterate eqs 13 and 14 at early stages of the optimization procedure where the important phase space of the end states is not well sampled yet. Therefore, we have tested both strategies.

• When the parameters should be optimized, i.e. after how much simulation time new parameters are calculated. This can be at fixed positions in time, e.g. after 150 ps, $150 + 2 \times 150 = 450$ ps, and $150 + 2 \times 150 + 4 \times 150 = 1050$ ps, which corresponds to an update after the 1st, 3rd, and 7th run if each run is 150 ps (denoted "update 1, 3, 7,..." below). However, this need not be the most efficient interval of parameter update. Another strategy we tested was to update once the sum of the statistical errors of the ΔF_{iR} values calculated from the runs with the current parameters is below the sum of errors calculated from the runs that were performed with the previous set of parameters. This scheme is denoted "update when error in ΔF is smaller than with previous parameters" below.

• Which runs should be taken into account when calculating the new parameters. If the E_i^R parameters of previous runs differ only slightly ($<k_BT$) from the current ones, it would be a waste of computing effort not to take these runs also into account when calculating the new energy offsets E_i^R and smoothness parameter s. This strategy has been pursued in the last two update schemes. Combining the previously mentioned points led to the following update schemes:

1. no reweighting, update 1, 3, 7,...
2. reweighting, update 1, 3, 7,...
3. no reweighting, update when error in ΔF is smaller than with previous parameters.
4. reweighting, update when error in ΔF is smaller than with previous parameters.
5. reweighting, update when error in ΔF is smaller than with previous parameters. When the E_i^R values differ less than k_BT from current parameters for the first time, take all subsequent runs into account when calculating optimized parameters. That is, compare the energy offsets E_i^R of the current run with the energy offsets of all the previous runs (starting from run 1). If all E_i^R differ less than k_BT from the current E_i^R, then take the trajectories from this previous run and from all subsequent runs into account when calculating new energy offsets and a new s parameter.
6. reweighting, update when error in ΔF is smaller than with previous parameters or after 1, 3, 7,... When the E_i^R differ less than k_BT from current parameters for the first time, take all subsequent runs into account when calculating optimized parameters (see also 5). Each update scheme consisted of 127 subsequent simulations ("runs") of 150 ps leading to a total simulation time of 19.05 ns for each of the above-mentioned schemes. When calculating ΔF_{iR}, E_i^R, and s values, the first 50 ps were discarded for equilibration.

The update schemes should optimize the energy offsets to $E_i^R = \{0, 0, 0, 0, 0\}$ kJ/mol. However, it is not clear which s parameter is the optimal one. In order to see whether the update schemes converge to the optimal s parameter, we ran 13 simulations at fixed parameters. All energy offsets were set to zero. The s values were 0.0041, 0.0082, 0.0164, 0.0328, 0.0657, 0.0821, 0.0903, 0.0985, 0.1149, 0.1313, 0.2627, 0.5254, and 1.0508. The simulation time for each of the 13 simulations was 7.5 ns. The rest of the protocol was as for the simulations with parameter update. Furthermore, a 7.5 ns non-EDS simulation of one interacting solute water molecule and four "dummy" solute water molecules in solvent water was run using the same settings as for the EDS simulations at fixed parameters.

All simulations were performed using a modified version of the GROMOS05[57] molecular simulation package. The method has been implemented such that only the perturbed interactions, i.e. the Hamiltonian terms that differ in the various end states, are calculated for every end state. All other interactions, i.e. in general the vast majority of the interactions, are calculated only once per time step. Moreover, the same list of atom pairs that interact with each other via nonbonded interactions is used for all end states. This results in an EDS simulation of N states being computationally much cheaper than N independent simulations.

4. Results and Discussion

In this study we have tested various algorithms that allow for an automatic updating of the reference Hamiltonian parameters needed for efficient EDS simulation. As we have shown in earlier work[37] and will show in the following, the convergence of the free energy estimates strongly depends on the chosen parameters for the reference Hamiltonian (eq 11) making it mandatory to have an efficient algorithm to automatically determine these parameters. Therefore, we have proposed several automatic update schemes (see section 3), whose ability to update the parameters efficiently is reported and discussed in this section.

The convergence of the energy offsets is shown in Figure 2. As all states have the same free energy, the expected optimized energy offsets are $E_i^R = \{0, 0, 0, 0, 0\}$ kJ/mol. Update scheme 6 was fastest in optimizing the energy offsets from $E_i^R = \{0, 50, 100, 150, 200\}$ kJ/mol down to zero. Schemes 1 and 2 both calculate new parameters at fixed points during the simulation (after the 1st, 3rd, and 7th, runs). Whereas scheme 2 iterates eqs 13 and 14 until convergence, scheme 1 stops after one iteration. Although reweighting might not be reasonable at the very beginning of the simulation due to insufficient sampling, it is clear that scheme 2 which does use reweighting optimized the energy offsets faster. This shows that reweighting speeds up the convergence of the parameter estimation considerably as soon as the parameters are in a range that allows reasonable sampling

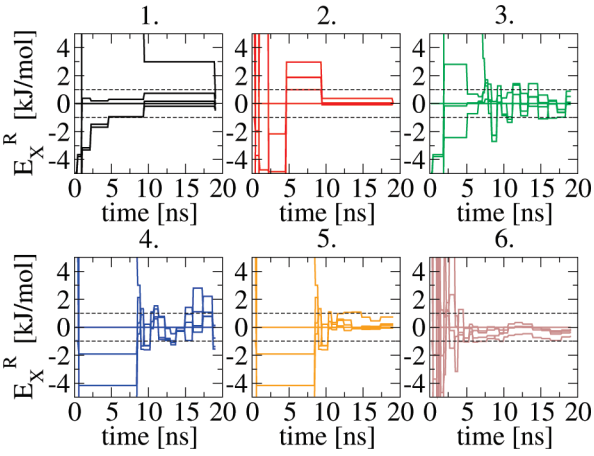

Figure 2. Convergence of the four ($X = 2,..., 5$) energy offsets $E_X^R - E_1^R$ (in kilojoules per mole) for the six parameter update schemes (see Simulation Protocols). (The dashed lines are to help guide the eye.)

of configuration space for all end states. Scheme 2 reliably optimized the energy offsets, however rather slowly. Schemes 3 and 4 were the first attempts to make convergence faster changing the criteria of when to do a calculation of new parameters. Instead of at fixed positions during simulation time the parameters were now optimized if the sum of the statistical errors of the ΔF_{iR} values calculated from the runs with the current parameters were below the sum of errors calculated from the runs which had been performed with the previous set of parameters. Especially at the very beginning of the optimization process, this does not seem to be a good criterion as can be seen (Figure 2) from the long time it takes until the energy offsets come down to the ±1 kJ/mol band. Furthermore, once reasonable energy offsets are found, the update criterion is fulfilled frequently leading to calculation of new energy offsets from very short pieces of trajectory. This leads to fluctuations of the energy offsets around zero. This problem has been solved in scheme 5. It differs from scheme 4 as it takes trajectories obtained with previous parameters also into account when calculating new parameters, if the previous energy offsets differ less than kT. As can be seen from Figure 2, this prevented the fluctuations around the optimal energy offset. Scheme 6 combines the findings of schemes 1−5: it uses reweighting and the information from previous runs once the energy offsets are within kT of the current energy offsets, and it updates the parameters either after runs 1, 3, 7,... or once the statistical error in ΔF_{iR} becomes lower. Figure 2 shows that this strategy allowed the fastest optimization of the energy offsets.

A similar discussion holds for the estimation of the smoothness parameter s (see Figure 3). Also the smoothness parameter s is optimized fastest by scheme 6. Much faster than with all other schemes the s parameter is close to the final optimized s. Whether this smoothness parameter s leads to the most accurate free energy estimate was investigated by performing 13 simulations at different s values and fixed energy offsets ($E_i^R = \{0, 0, 0, 0, 0\}$ kJ/mol). That the optimized s parameter is the best choice for s can be seen from Figure 4. It shows all $N(N-1)/2 = 10$ $|\Delta F_{XY}|$ estimates (thin solid lines), where X and Y denote two end states, and

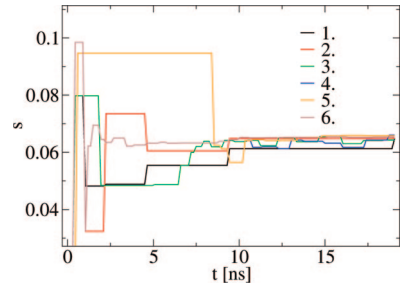

Figure 3. Convergence of the smoothness parameter s for the six parameter update schemes (see Simulation Protocols).

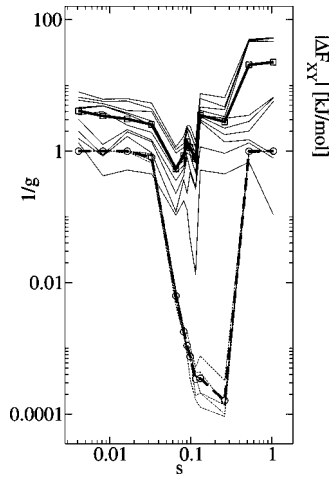

Figure 4. Absolute error of the ten ΔF_{XY} estimates between the five end states (X, $Y = 1,..., 5$) as a function of the smoothness parameter s for the 13 s values mentioned in Simulation Protocols (thin solid lines). The thick solid line (with squares) denotes the mean over these 10 error estimates. The thin dashed lines indicate $1/g$ obtained from the five ΔF_{XR} ($X = 1,..., 5$) estimates. Here g is the statistical inefficiency,[60] i.e. given a time series of M data points M/g gives the number of uncorrelated data points. The thick dashed line (with circles) indicates the mean over the five $1/g$ estimates.

their average (thick solid line) as a function of s. As all states have the same free energy, all ΔF_{XY} estimates should be zero and $|\Delta F_{XY}|$ is equal to the absolute error of the free energy estimate. Minimal errors were obtained for s values between 0.06 and 0.1. The lowest error was obtained for $s = 0.0657$ which agrees well with the final optimized s value of update scheme 6 (0.0653). Figure 4 furthermore shows that the number of uncorrelated data points M/g, where g is the statistical inefficiency,[60] contained within a time series of M data points strongly varies with s (dashed lines).

Figure 5 shows free energy differences ΔF_{XY} between the end states which were calculated from a 4 ns moving window (moving in 1 ns steps). Again update scheme 6 performed best as the calculated ΔF_{XY} values lie quickly within 1 kJ/mol from the correct result (0 kJ/mol). The calculated free energy estimates obtained from the six simulations with parameter updates and the 13 simulations at fixed parameter choices are shown in Tables 1 and 2. With an appropriate equilibration time, all six update schemes give good estimates for the free energy differences. For the simulations at fixed s values, reasonable free energy estimates were only obtained

Table 1. Free Energy Differences and Statistical Uncertainties[37,60] (in kilojoules per mole) between the Five (1−5) End States and the Reference (*R*) State Obtained from Six Simulations Performed with Six Different Update Schemes (See *Simulation Protocols*)[a]

	1	2	3	4	5	6
ΔF_{1R}	7.8 ± 0.3	5.6 ± 0.3	5.8 ± 0.3	6.5 ± 0.3	5.4 ± 0.3	5.7 ± 0.3
ΔF_{2R}	7.5 ± 0.2	5.7 ± 0.3	6.0 ± 0.3	6.8 ± 0.3	6.1 ± 0.3	5.3 ± 0.3
ΔF_{3R}	7.2 ± 0.3	5.7 ± 0.3	5.9 ± 0.3	6.2 ± 0.3	6.7 ± 0.3	5.3 ± 0.3
ΔF_{4R}	7.3 ± 0.3	5.8 ± 0.3	5.7 ± 0.3	6.5 ± 0.3	5.7 ± 0.3	5.5 ± 0.3
ΔF_{5R}	7.5 ± 0.2	5.7 ± 0.3	5.9 ± 0.3	6.2 ± 0.3	5.9 ± 0.3	5.3 ± 0.3
ΔF_{21}	−0.3 ± 0.4	0.1 ± 0.4	0.2 ± 0.4	0.3 ± 0.5	0.7 ± 0.5	−0.4 ± 0.4
ΔF_{31}	−0.6 ± 0.5	0.1 ± 0.5	0.1 ± 0.5	−0.3 ± 0.5	1.2 ± 0.5	−0.4 ± 0.4
ΔF_{41}	−0.5 ± 0.5	0.1 ± 0.5	−0.1 ± 0.4	−0.1 ± 0.4	0.2 ± 0.5	−0.2 ± 0.4
ΔF_{51}	−0.3 ± 0.4	0.0 ± 0.4	0.1 ± 0.4	−0.4 ± 0.5	0.4 ± 0.4	−0.3 ± 0.4
ΔF_{32}	−0.3 ± 0.4	0.0 ± 0.5	0.0 ± 0.5	−0.6 ± 0.5	0.6 ± 0.4	0.0 ± 0.4
ΔF_{42}	−0.2 ± 0.4	0.0 ± 0.4	−0.3 ± 0.4	−0.4 ± 0.5	−0.5 ± 0.5	0.2 ± 0.4
ΔF_{52}	0.0 ± 0.4	−0.1 ± 0.4	−0.1 ± 0.4	−0.6 ± 0.5	−0.3 ± 0.4	0.1 ± 0.5
ΔF_{43}	0.1 ± 0.4	0.0 ± 0.5	−0.2 ± 0.4	0.2 ± 0.5	−1.0 ± 0.5	0.2 ± 0.4
ΔF_{53}	0.3 ± 0.4	−0.1 ± 0.5	0.0 ± 0.4	−0.1 ± 0.5	−0.8 ± 0.5	0.1 ± 0.4
ΔF_{54}	0.3 ± 0.4	−0.1 ± 0.4	0.2 ± 0.4	−0.3 ± 0.5	0.2 ± 0.5	−0.2 ± 0.5

[a] The averaging is performed over the last 9.6 ns of the 19.05 ns simulation time.

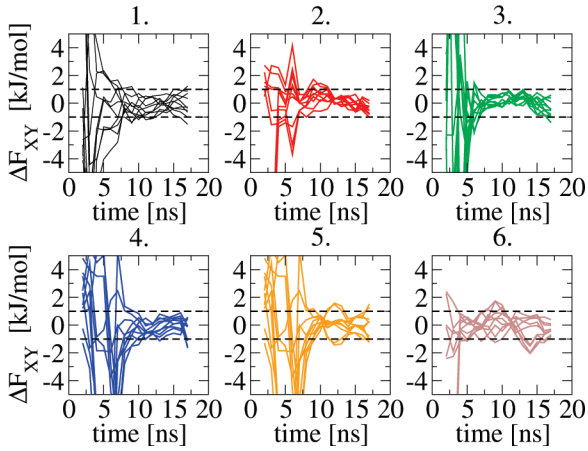

Figure 5. Ten free energy differences ΔF_{XY} (in kilojoules per mole) between the end states (X, $Y = 1,..., 5$) calculated from a 4 ns moving window (moving in 1 ns steps) for the six parameter update schemes (see Simulation Protocols). (The dashed lines are to help guide the eye.)

for a small range around the optimal s, as could already be seen from Figure 4.

The high sensitivity of the free energy estimate to the s parameter can be further explained using energy difference distributions. Figure 6 shows the energy difference distributions (eq 7) $\rho_X(\Delta V; \Delta V_{XY})$, $\rho_Y(\Delta V; \Delta V_{XY})$, and $\rho_R(\Delta V; \Delta V_{XY})$ for all ten $X-Y$ pairs and all 13 s values. Starting at $s = 0.0657$, we see that the $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ are well formed and that they match the distributions obtained from an independent, non-EDS simulation. Recall that the free energy difference between two states X and Y is the energy difference where the $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ intersect (see section 2, eq 8). From Figure 6, we see that the highest probability of $\rho_R(\Delta V; \Delta V_{XY})$ is where $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ distributions intersect, i.e. during the reference state simulation the sampling is focused on the important crossing region. The importance of focusing the sampling on the relevant energy difference space has recently been pointed out by Min and Yang.[58] Their approach was to improve sampling by adding a biasing potential on ΔV in order to increase the sampled ΔV range. A different strategy was pursued by Wu,[59] who uses a modified Metropolis acceptance rule to constrain the sampling to a given ΔV range and obtains the energy-difference distribution from overlapping umbrella windows. In EDS, sampling is automatically focused only on the crucial ΔV range once the parameters of the reference state Hamiltonian have been optimized.

The effect of suboptimal s values can be studied in the other plots of Figure 6. Increasing the s value leads to broader sampling over the ΔV range, leading to worse sampling of the crossing region and longer convergence times (see also Table 2). Starting with $s = 0.1313$ not all five states are sampled equally anymore—the ten $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ distributions, respectively, start to differ. For $s = 0.5254$ and higher s values, only the state from which the simulation started is sampled and no meaningful free energy estimates can be obtained. A decrease of s very quickly distorts the potential energy surface of the reference state Hamiltonian such that the important regions of the configuration spaces of the end states are not longer minima on this surface. This leads to very narrow $\rho_R(\Delta V; \Delta V_{XY})$ distributions around $\Delta V = 0$ kJ/mol and badly formed $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ distributions. This suboptimal sampling of the end states is also reflected in the rather inaccurate free energy estimates for these s values (see Table 2).

Figure 7 shows in another way how the accuracy of the free energy estimates depends on the chosen s parameter. It shows $\ln(\rho_X(\Delta V; \Delta V_{XY})/\rho_Y(\Delta V; \Delta V_{XY}))$ as a function of ΔV for the 13 s values. From eq 9, we see that the ordinate intercept is $-\beta\Delta F_{XY}$. For $s = 0.0657$, the lines nicely cross the ordinate at $-\beta\Delta F_{XY} = 0$. Moving to smaller or larger s values, the variance increases. For $s = 0.5254$ and larger, no reasonable free energy estimates can be obtained, because only a few ΔV values near the region where the end state energy difference distributions intersect ($\Delta V = 0$) are sampled in the trajectories. The average temperatures calculated from the slopes of a linear regression of the data shown in Figure 7 (see eq 9) are 330, 319, 302, 302, 303, 303, 303, 303, 303, 303, 303, 301, and 301 K for the 13 s values. For $s = 0.0657-0.2627$, where $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ are well formed (see Figure 6), the extracted

Table 2. Free Energy Differences and Statistical Uncertainties[37,60] (in kilojoules per mole) between the Five (1−5) End States and the Reference (*R*) State Obtained from 13 Simulations Performed with Different *s* Parameters[a]

s =	0.0041	0.0082	0.0164	0.0328	0.0657	0.0821	0.0903	0.0985	0.1149	0.1313	0.2627	0.5254	1.0508
ΔF_{1R}	937.4	450.9 ± 1.2	205.7 ± 1.3	63.7 ± 1.6	5.9 ± 0.3	4.1 ± 0.4	4.0 ± 0.6	3.8 ± 0.6	5.0 ± 1.0	1.4 ± 0.7	1.3 ± 0.8	46.8 ± 2.0	53.2 ± 1.6
ΔF_{2R}	942.0	446.9 ± 1.4	202.1 ± 1.6	62.2 ± 2.3	5.7 ± 0.4	3.6 ± 0.4	5.5 ± 0.8	4.6 ± 0.7	3.5 ± 0.8	8.9 ± 1.9	5.6 ± 2.3	50.3 ± 1.7	52.4 ± 1.1
ΔF_{3R}	940.4	452.1 ± 1.3	200.0 ± 2.4	66.2 ± 1.1	4.8 ± 0.3	4.2 ± 0.5	3.3 ± 0.5	5.8 ± 0.9	3.8 ± 0.9	4.2 ± 1.3	4.7 ± 2.1	48.8 ± 2.4	46.7 ± 1.8
ΔF_{4R}	944.0	446.0 ± 2.3	205.9 ± 0.9	61.8 ± 2.0	5.4 ± 0.3	5.4 ± 0.5	5.7 ± 0.6	3.3 ± 0.5	4.5 ± 1.6	6.8 ± 1.7	7.8 ± 2.4	47.5 ± 2.3	53.3 ± 1.3
ΔF_{5R}	936.0	451.7 ± 1.2	203.8 ± 1.3	60.8 ± 2.1	5.9 ± 0.3	4.7 ± 0.5	3.4 ± 0.5	3.8 ± 0.9	3.8 ± 1.2	4.8 ± 1.5	5.5 ± 2.2	0.0 ± 0.0	0.0 ± 0.0
ΔF_{21}	4.6	−4.0 ± 2.3	−3.6 ± 2.0	−1.5 ± 3.8	−0.2 ± 0.6	−0.5 ± 0.7	1.4 ± 1.1	0.8 ± 1.0	−1.5 ± 1.5	7.5 ± 2.3	4.3 ± 2.7	3.5 ± 3.6	−0.8 ± 2.0
ΔF_{31}	3.1	1.3 ± 1.8	−5.8 ± 2.7	2.5 ± 1.9	−1.1 ± 0.5	0.2 ± 0.7	−0.7 ± 1.0	1.9 ± 1.2	−1.1 ± 1.4	2.8 ± 1.8	3.3 ± 2.5	2.0 ± 4.3	−6.5 ± 3.4
ΔF_{41}	6.6	−4.9 ± 2.7	0.1 ± 1.6	−1.9 ± 2.5	−0.5 ± 0.5	1.4 ± 0.7	1.7 ± 0.9	−0.5 ± 1.0	−0.5 ± 2.2	5.4 ± 2.2	6.5 ± 2.7	0.7 ± 4.3	0.1 ± 2.1
ΔF_{51}	−1.3	0.9 ± 1.9	−2.0 ± 1.8	−2.9 ± 3.7	0.1 ± 0.5	0.6 ± 0.8	−0.6 ± 0.9	−0.1 ± 1.4	−1.2 ± 1.6	3.3 ± 1.9	4.2 ± 3.1	−46.8 ± 2.0	−53.2 ± 1.6
ΔF_{32}	−1.5	5.3 ± 2.0	−2.1 ± 3.0	4.0 ± 2.5	−0.9 ± 0.6	0.6 ± 0.7	−2.2 ± 1.0	1.2 ± 1.3	0.4 ± 1.3	−4.7 ± 2.4	−0.9 ± 3.9	−1.5 ± 4.1	−5.7 ± 2.1
ΔF_{42}	2.0	−0.9 ± 3.7	3.8 ± 1.9	−0.4 ± 3.4	−0.3 ± 0.5	1.8 ± 0.8	0.3 ± 1.0	−1.2 ± 1.0	1.1 ± 1.9	−2.1 ± 3.6	2.2 ± 4.7	−2.8 ± 4.0	0.9 ± 2.4
ΔF_{52}	−5.9	4.9 ± 2.5	1.7 ± 2.0	−1.4 ± 3.1	0.2 ± 0.6	1.1 ± 0.7	−2.0 ± 1.1	−0.8 ± 1.3	0.3 ± 1.8	−4.2 ± 3.5	−0.1 ± 3.2	−50.3 ± 1.7	−52.4 ± 1.1
ΔF_{43}	3.5	−6.2 ± 2.7	5.9 ± 2.6	−4.4 ± 2.4	0.6 ± 0.5	1.2 ± 0.8	2.4 ± 1.0	−2.4 ± 1.1	0.7 ± 2.3	2.6 ± 3.0	3.2 ± 4.2	−1.3 ± 4.7	6.6 ± 2.5
ΔF_{53}	−4.4	−0.4 ± 1.9	3.8 ± 2.7	−5.4 ± 2.4	1.1 ± 0.5	0.5 ± 0.7	0.1 ± 0.9	−2.0 ± 1.3	−0.1 ± 1.8	0.5 ± 2.9	0.9 ± 4.1	−48.8 ± 2.4	−46.7 ± 1.8
ΔF_{54}	−7.9	5.7 ± 2.6	−2.1 ± 1.8	−0.9 ± 3.7	0.5 ± 0.5	−0.7 ± 0.8	−2.3 ± 0.9	0.4 ± 1.3	−0.8 ± 2.4	−2.1 ± 2.4	−2.3 ± 4.7	−47.5 ± 2.3	−53.3 ± 1.3

[a] For *s* = 0.0041, the variance is zero and no error estimate could be calculated. Simulation time: 7.5 ns (minus 150 ps for equilibration).

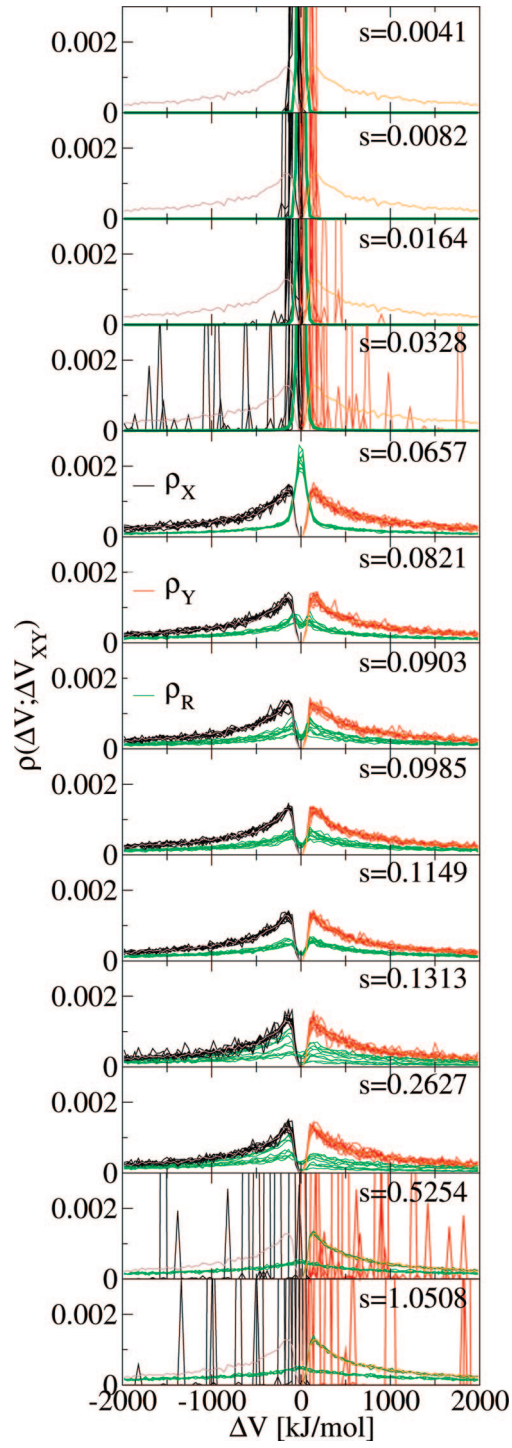

Figure 6. Energy difference distributions (eq 7) $\rho_X(\Delta V; \Delta V_{XY})$ (black), $\rho_Y(\Delta V; \Delta V_{XY})$ (red), and $\rho_R(\Delta V; \Delta V_{XY})$ (green) for all 10 *X*−*Y* pairs and all 13 *s* values (see Simulation Protocols). The distributions obtained from an independent non-EDS simulation are shown in brown ($\rho_1(\Delta V; \Delta V_{12})$) and orange ($\rho_1(\Delta V; \Delta V_{21})$).

temperatures agree with the average temperature which was of 303 K for all 13 simulations.

It should be stressed that for this test system $\Delta V = 0$ kJ/mol can be observed for three different kind of configurations. First, for low *s* values ($s = 0.0041-0.0328$) regions of configuration space important to the reference state are sampled which are equally unfavorable for the end states,

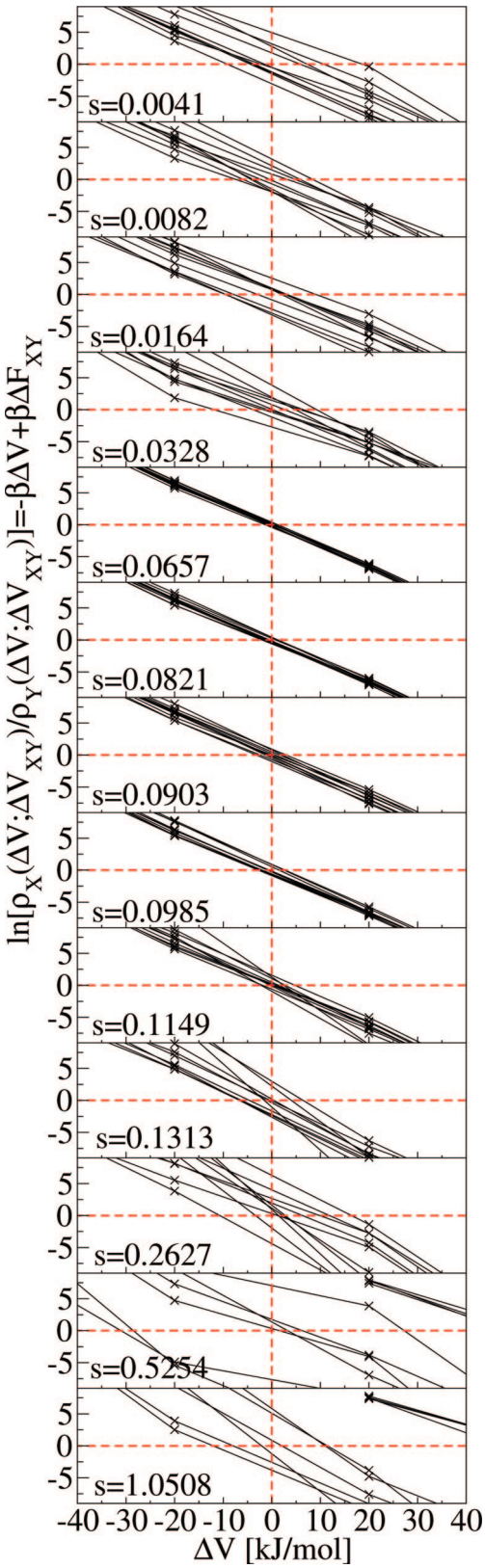

Figure 7. Linearized representation (eq 9) of energy difference distributions for all 10 $X-Y$ pairs: $\ln(\rho_X(\Delta V; \Delta V_{XY})/\rho_Y(\Delta V; \Delta V_{XY}))$ as a function of ΔV for all 13 s values (see Simulation Protocols). (The dashed lines are to help guide the eye.)

leading to energy differences ΔV_{XY} of zero ($V_X \approx V_Y \gg V_R$). This can be observed in Figure 6 for the lowest four s values. The $\rho_R(\Delta V; \Delta V_{XY})$ distribution is centered around zero, yet the important phase space of the end states is not well sampled (badly formed $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ distributions) which is a necessary (but not sufficient) condition for reasonable free energy difference estimates. If parts of the $\rho_X(\Delta V; \Delta V_{XY})$ and $\rho_Y(\Delta V; \Delta V_{XY})$ distributions are not sampled, this will shift the intersection point and therefore the free energy estimate unless the neglected areas are of the same size and do not lie in the intersection region. Second, as in the current simulation setup the five solute water molecules are free to move, a zero energy difference can also be observed if two water molecules happen to occupy the same position ($V_X = V_Y = V_R$). If this happened for longer periods during the simulation, the chosen test system would be trivial. It does indeed happen occasionally during the simulation that two water molecules occupy the same position, they however separate again after some picoseconds. In order to check whether these occasional encounters influence the presented results, we have recalculated the free energy differences using only frames where the distance between any two of the five solute water molecules is larger than 0.5 nm. The difference in the calculated ΔF_{XR} values was found to be smaller than the statistical uncertainty of these values. The third type of configuration where a zero energy difference occurs is when during the reference state simulation a transition from the important configuration space of one end state to that of another end state occurs ($V_X = V_Y > V_R$). For $s = 0.0657$, many of these transitions occur, which explains the focusing of $\rho_R(\Delta V; \Delta V_{XY})$ around $\Delta V = 0$. With increasing s the number of these transitions decreases which is reflected in the $\rho_R(\Delta V; \Delta V_{XY})$ distributions (i.e., the density around $\Delta V = 0$ decreases), in the decrease of the number of uncorrelated data points (see Figure 4, $s = 0.0657-0.2627$), and consequently in the statistical uncertainties of the free energy estimates (see Table 2, $s = 0.0657-0.2627$). It is due to this decrease in the number of transitions (down to zero for $s = 0.5254-1.0508$) that the nominally optimal smoothness parameter $s = 1$ is in practice not optimal. For the two smoothness parameter values $s = 0.06568$ and $s = 0.1149$ we have recalculated the free energy differences using the same number of uncorrelated data points, i.e. for the higher s value the complete time series was used in the analysis and for the lower s value only $M_{s=0.06568} = M_{s=0.1149} g_{s=0.06568}/g_{s=0.1149}$ data points were used. Interestingly but not surprisingly, the obtained mean absolute error of the 10 ΔF_{XY} estimates is somewhat higher for $s = 0.06568$ ($\approx$2 kJ/mol) than for $s = 0.1149$ ($\approx$0.8 kJ/mol). This is in line with the expected nominally optimal smoothness parameter being $s = 1$. In practice, we seek the lowest s value that still ensures sampling of the complete end state energy difference distributions as this will ensure a precise and accurate free energy estimate, where the precision depends on sufficient transitions between regions of configuration space important to the end states, and the accuracy depends on complete sampling of these parts of configuration space. The presented schemes for the iterative update of the reference Hamiltonian parameters have been shown to be able to find this optimal value for the smoothness parameter.

5. Conclusions

We have successfully tested different schemes that allow for an automatic updating of the reference Hamiltonian parameters in enveloping distribution sampling (EDS). As a test system, we chose liquid water in which particular molecules were created and deleted. We selected five water molecules to define five states. Each state consisted of one interacting solute water molecule and four noninteracting water molecules in a box of solvent water. Starting from a set of reference Hamiltonian parameters which was far from optimal, all schemes optimized the parameters to the expected energy offset values (0 kJ/mol) and to a unique smoothness parameter. One scheme (scheme 6) was fastest in optimizing and will be used in further applications. The simulations we performed at fixed smoothness parameter s showed that the optimized s parameter gives rise to the most accurate free energy estimate (see Figure 4). The automatic update scheme is a big step toward application of EDS by nonexpert users. No parameters have to be chosen at the beginning of the simulation. The only input to the initial reference Hamiltonian are the Hamiltonians of the various end states.

Future work will imply testing of the method on flexible molecules where an optimal balance between sampling within the important configuration space of one end state and transitions between parts of configuration space important to different end states might pose a problem. A further challenge are systems where the important configuration space of some end states lie close together and those of others are far apart. For such systems a single smoothness parameter s approach might break down. Therefore, we are currently testing also multiple s approaches.

Acknowledgment. The authors would like to thank Chris Oostenbrink for reading the manuscript and for helpful discussions. Financial support by the National Center of Competence in Research (NCCR) Structural Biology and by Grant No. 200021-109227 of the Swiss National Science Foundation (SNSF) is gratefully acknowledged.

References

(1) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.

(2) van Gunsteren, W. F. Methods for calculation of free energies and binding constants: Successes and problems. In *Computer simulation of biomolecular systems: theoretical and experimental applications*; van Gunsteren, W. F., Weiner, P. K., Eds.; ESCOM Science publishers B. V.: Leiden, 1989.

(3) Reynolds, C. A.; King, P. M.; Richards, W. G. *Mol. Phys.* **1992**, *76*, 251–275.

(4) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407–435.

(5) van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. Computation of free energy in practice: Choice of approximations and accuracy limiting factors. In *Computer simulation of biomolecular systems: theoretical and experimental applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM Science publishers B. V.: Leiden, 1993; Vol. 2.

(6) Kofke, D. A.; Cummings, P. T. *Mol. Phys.* **1997**, *92*, 973–996.

(7) Gelman, A.; Meng, X. L. *Statist. Sci.* **1998**, *13*, 163–185.

(8) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211–243.

(9) Chipot, C.; Pearlman, D. A. *Mol. Simul.* **2002**, *28*, 1–12.

(10) van Gunsteren, W. F.; Daura, X.; Mark, A. E. *Helv. Chim. Acta* **2002**, *85*, 3113–3129.

(11) Brandsdal, B. O.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V. B.; Aqvist, J. *Adv. Protein Chem.* **2003**, *66*, 123–158.

(12) Kofke, D. A. *Fluid Phase Equilib.* **2005**, *228*, 41–48.

(13) Rodinger, T.; Pomes, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 164–170.

(14) Meirovitch, H. *Curr. Opin. Struct. Biol.* **2007**, *17*, 181–186.

(15) Chipot, C.; Pohorille, A. *Free energy calculations: Theory and applications in chemistry and biology*; Springer: Berlin, 2007.

(16) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.

(17) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. *Annu. Rep. Comput. Chem* **2007**, *3*, 41–59.

(18) Jorgensen, W. L.; Thomas, L. L. *J. Chem. Theory Comput.* **2008**, *4*, 869–876.

(19) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.

(20) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(21) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.

(22) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2007**, *111*, 2242–2254.

(23) Squire, D. R.; Hoover, W. G. *J. Chem. Phys.* **1969**, *50*, 701–706.

(24) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.

(25) Lu, N. D.; Wu, D.; Woolf, T. B.; Kofke, D. A. *Phys. Rev. E* **2004**, *69*, 057702.

(26) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042–6051.

(27) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058–9067.

(28) Affentranger, R.; Tavernelli, I.; Di Iorio, E. E. *J. Chem. Theory Comput.* **2006**, *2*, 217–228.

(29) Pitera, J.; Kollman, P. *J. Am. Chem. Soc.* **1998**, *120*, 7557–7567.

(30) Kong, X.; Brooks, C. L. *J. Chem. Phys.* **1996**, *6*, 2414–2423.

(31) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.

(32) Jarzynski, C. *Phys. Rev. E* **1997**, *56*, 5018–5035.

(33) Crooks, G. E. *Phys. Rev. E* **2000**, *61*, 2361–2366.

(34) Jarzynski, C. *Phys. Rev. E* **2006**, *73*, 046105.

(35) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.

(36) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Phys.* **2007**, *126*, 184110.

(37) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Phys.* **2008**, *128*, 174112.

(38) Srinivasan, R. *Importance Sampling: Applications in Communications and Detection*; Springer: Berlin, Heidelberg, NY, 2002.

(39) Liu, H.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *100*, 9485–9494.

(40) Oostenbrink, C.; van Gunsteren, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6750–6754.

(41) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9–12.

(42) Han, K. K. *Phys. Lett. A* **1992**, *165*, 28–32.

(43) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776–1783.

(44) Escobedo, F. A.; de Pablo, J. J. *J. Chem. Phys.* **1995**, *103*, 2703–2710.

(45) Smith, G. R.; Bruce, A. D. *J. Phys. A: Math. Gen.* **1995**, *28*, 6623–6643.

(46) Engkvist, O.; Karlstrom, G. *Chem. Phys.* **1996**, *213*, 63–76.

(47) Han, K. K. *Phys. Rev. E* **1996**, *54*, 6906–6910.

(48) Chen, Y. G.; Hummer, G. *J. Am. Chem. Soc.* **2007**, *129*, 2414–2415.

(49) Shing, K. S.; Gubbins, K. E. *Mol. Phys.* **1982**, *46*, 1109–1128.

(50) Powles, J. G.; Evans, W. A. B.; Quirke, N. *Mol. Phys.* **1982**, *46*, 1347–1370.

(51) Jacucci, G.; Quirke, N. *Lect. Notes Phys.* **1982**, *166*, 38–57.

(52) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981; pp 331−342.

(53) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(54) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comp. Phys.* **1977**, *23*, 327–341.

(55) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular simulation: The GROMOS96 manual and user guide*; Vdf Hochschulverlag AG an der ETH Zürich; Zürich, 1996.

(56) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.

(57) Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F. *J. Comput. Chem.* **2005**, *26*, 1720–1751.

(58) Min, D. H.; Yang, W. *J. Chem. Phys.* **2008**, *128*, 191102.

(59) Wu, D. *J. Chem. Phys.* **2008**, *128*, 224105.

(60) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 26–41.

CT800424V