JCTC Journal of Chemical Theory and Computation

Article

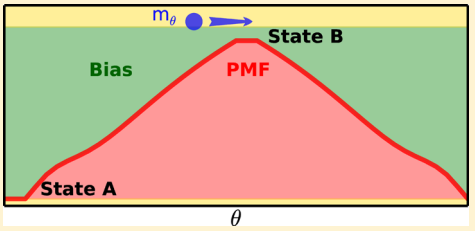
pubs.acs.org/JCTC

Local Elevation Umbrella Sampling Applied to the Calculation of Alchemical Free-Energy Changes via λ-Dynamics: The λ-LEUS Scheme

Noah S. Bieler, Rico Häuselmann, and Philippe H. Hünenberger*

Laboratory of Physical Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland

ABSTRACT: When using molecular dynamics (MD) simulations for the calculation of alchemical free-energy changes, the extended-dynamics method called λ-dynamics (λD) in its currently available implementations does not compare very favorably with thermodynamic integration (TI) in terms of robustness, efficiency, and accuracy, although it is in principle easier to set up, postprocess, and automatize. In the present article, the main shortcomings of the λD approach are carefully analyzed, and possible remedies are proposed. The resulting scheme, called λ-LEUS, involves: (*i*) the use of a simple noninvertible coordinate transformation ensuring a finite sampling of the two physical end states; (*ii*) the application of the local elevation umbrella sampling (LEUS) memory-based biasing scheme to enforce homogeneous sampling and overcome barriers along the alchemical coordinate; (*iii*) recommendations concerning the choice of the mass parameter and of the temperature-coupling scheme for this coordinate; (*iv*) the use of a second-order splines basis set for the memory-based biasing functions. The λ-LEUS scheme is described and tested considering the perturbation of hydroquinone to benzene in water. The results are compared to those of TI calculations and exhibit a superior accuracy-to-efficiency ratio, presumably because dynamic variations in the alchemical coordinate open up pathways to circumvent orthogonal barriers, these pathways being inaccessible when the coordinate is constrained. Therefore, λ-LEUS combines the practical advantages of λD with the robustness of TI, simultaneously affording a slightly enhanced computational efficiency.



1. INTRODUCTION

Atomistic simulations and in particular molecular dynamics (MD) represent nowadays one of the main approaches for calculating relative free energies in condensed-phase (bio)-molecular systems.[1,2] The goal of these calculations is to evaluate, given a classical force-field description, the free-energy difference between two or more well-defined states,[3] either *thermodynamic* (*i.e.* distinct state parameters such as temperature or pressure), *conformational* (*i.e.* distinct molecular conformations), or *alchemical* (*i.e.* distinct molecular topologies).

The distinction between these three types of free-energy differences can be stated as follows. The free-energy difference between two *thermodynamic* states corresponds to a change in free energy upon altering a thermodynamic boundary condition of the system, i.e. a parameter characterizing the influence of the surroundings, the system being otherwise unchanged in terms of degrees of freedom and Hamiltonian. The free-energy difference between two *conformational* states is a logarithmic measure of the relative partition functions corresponding to a common Hamiltonian and thermodynamic state but defined by integrals over different regions in terms of the degrees of freedom of the system. The free-energy difference between two *alchemical* states is a logarithmic measure of the relative partition functions corresponding to a common thermodynamic state and to the entire space in terms of the degrees of freedom of the system but calculated considering two different Hamiltonians that are both functions of these degrees of freedom. Note that the expression alchemical change is used in this article as a synonym for topological change, irrespective whether the change truly involves an unphysical component (e.g., different atom contents of the physical end states or unphysical connecting pathway).

In the ideal situation of infinite sampling from a Boltzmann-weighted distribution, as obtained e.g. using thermostated MD, relative *conformational* free energies could be obtained from a single simulation of the molecular system via direct counting[1,4] (DC) and relative *alchemical* free energies could be obtained from a single simulation of the molecular system in either of the two states (or any other reference state) via one-step free-energy perturbation[5,6] (OSP). In practice, however, these two simple approaches are often inapplicable or numerically inaccurate[7,8] due to the finite-time scale nature of the MD sampling. In both cases, the generic solution to this issue, which is at the core of all practical free-energy calculation methods proposed to date, is to separate the problem of sampling configurations from that of estimating the free-energy difference.

Considering two conformational or alchemical states *A* and *B*, an efficient finite-sampling protocol must satisfy four conditions:[3] (*i*) it must visit the configurations that are the most relevant for states *A* and *B*; (*ii*) it must present a sufficient number of interconversion transitions between these sets of configurations; (*iii*) it must avoid the sampling of an excessive amount of configurations that are relevant neither for *A* nor for *B*, nor for ensuring a sufficient number of transitions; (*iv*) it

Received: March 28, 2014

   dx.doi.org/10.1021/ct5002686 | *J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX

must be associated with a statistical-mechanically sound free-energy estimator.

Numerous free-energy calculation schemes have been proposed to date with the aim of achieving these four goals.[1,2] They can be broadly classified according to the following criteria: (*i*) whether they rely on the definition of a *pathway* connecting *A* with *B* or of an unphysical *reference state* encompassing configurations relevant for both *A* and *B*; (*ii*) whether they rely on a *single* simulation or on *multiple* simulations; (*iii*) for multiple simulations, whether the problem is broken down in a *serial* fashion (each simulation serves to evaluate a small part of the overall free-energy change, obtained by summation) or in a *parallel* fashion (each simulation represents an individual system trajectory, the free energy change being obtained by exponential averaging over initial conditions); (*iv*) for serial simulations, whether the definition of the individual windows relies on *constraints* or on *biasing potentials*, with the possible addition of *replica-exchange* moves;[9] (*v*) whether the method is *a priori* most relevant for *conformational* or *alchemical* calculations; (*vi*) what kind of *free-energy estimator* is applied for the postprocessing of the results. For example, in the context of serial calculations involving two or more states, the Bennet acceptance ratio[10] (BAR) or multiple Bennet acceptance ratio[11] (MBAR) methods, respectively, represent optimal forms of free-energy estimators.

Umbrella sampling[12] (US) is a method relying on biasing potentials and typically applied to conformational problems, either in a single simulation (reference-state method with a single preoptimized biasing potential) or in a serial fashion (pathway method with multiple harmonic potentials). Thermodynamic integration[13−15] (TI) is a pathway method relying on serial simulations involving constraints, applicable to alchemical calculations. Enveloping distribution sampling[16,17] (EDS) is a single-simulation reference-state method using a biasing potential (preoptimized enveloping potential), applicable to both conformational and alchemical problems. Jarzinsky's fast growth[18,19] (FG) method is a pathway method relying on parallel simulations, via exponential averaging of irreversible work estimates over initial conditions.

Often, free-energy methods are applied together with the definition of a coupling *pathway*, associated with a corresponding coupling parameter λ, i.e. a curvilinear coordinate along the selected pathway. In conformational calculations, the pathway will connect the center of the first conformational region *A* (λ = 0) with that of the second region *B* (λ = 1) via a curve in conformational space. In alchemical calculations, the pathway will connect the Hamiltonian of the first alchemical state *A* (λ = 0) with that of the second state *B* (λ = 1) into a hybrid (parametric) Hamiltonian. *A priori*, the nature of the coupling variable λ in the two types of calculations is quite distinct. In conformational calculations, λ is a singled out *generalized coordinate* of the system. In alchemical calculations, λ is a *parameter* of a hybrid Hamiltonian. However, as a key step in the development of free-energy calculation methods, the λ-dynamics (λD) approach[20] (see also its ancestor λ-Monte Carlo[21] and the related Hamiltonian replica-exchange (RE) scheme[22,23]) relies on the introduction of the λ coupling parameter of an alchemical calculation as a dynamical variable within the Hamiltonian of an extended system, thereby making it in effect a conformational coordinate of this extended system. This is achieved by assigning to λ a mass parameter m_λ, a momentum p_λ, and a Newton-like equation of motion (extended Hamiltonian formalism[24−26]). When comparing conformational free-energy calculations with alchemical free-energy calculations relying on λD, it becomes unnecessary to distinguish between the two types of calculations, λ being a degree of freedom of the system (or extended system) in both cases.

Unfortunately, current λD schemes suffer in practice from a number of shortcomings, and the importance of this major methodological step remains somewhat unacknowledged in the literature. When λD is applied "naively", i.e. in the form of a straight inclusion of λ as a dynamical variable within an extended Hamiltonian, these shortcomings are the following:

A. Not only physical and "intermediate" values of λ will be sampled, i.e. with $0 \leq \lambda \leq 1$, but also "out-of-bounds" values with $\lambda < 0$ or $\lambda > 1$.

B. The physical values $\lambda = 0$ and $\lambda = 1$, being points along a curve, only account for an infinitesimal fraction of the sampled extended-system configurations.

C. The mass parameter m_λ and temperature-coupling scheme of the λ degree of freedom must be chosen carefully, so as to guarantee the dynamical coupling of this degree of freedom with the rest of the system and an equilibrium sampling along this coordinate.

D. If appropriate dynamical coupling and equilibrium sampling are achieved, the probability distribution along λ will be Boltzmann-weighted in the corresponding potential of mean force (PMF) $G(\lambda)$, which may result in poor statistics in the neighborhood of either or both of the physical states $\lambda = 0$ or $\lambda = 1$ at finite sampling time.

E. Even in a situation where the physical states $\lambda = 0$ and $\lambda = 1$ have comparable free-energies, *i.e.* $G(0) \approx G(1)$, the presence of high barriers in the PMF may result in an insufficient number of transitions between the two states at finite sampling time.

Note that the combination of problems A and D renders the "naive" implementation of the λD approach fundamentally impossible in (common) situations where the extended-system Hamiltonian diverges to large negative values for $\lambda < 0$ or $\lambda > 1$. The five shortcomings listed above have been addressed with more or less success in the literature pertaining to λD.

In the initial λD article,[20] the sampling of "out-of-bounds" states (problem A) was prevented by the inclusion of a hard-wall mirroring potential energy term acting on the λ variable at 0 and 1, i.e. by inverting p_λ when hitting these boundary values. Alternatively, *coordinate transformations* of the form $\lambda(\theta)$ have been employed,[27−32] where θ is the actual dynamical variable, λ is the Hamiltonian-coupling parameter, and $\lambda(\theta)$ is a continuous and periodic function of θ bounded to the range $\lambda \in [0;1]$ or some slightly extended range.[28,31] The simplest form of periodic coordinate transformation is the zigzag line,[29] in which $\lambda(\theta)$ is defined over the period $\theta \in [0;2[$ by a line segment of slope +1 followed by a line segment of slope −1. This transformation is actually equivalent to the mirroring scheme of the original λD article.[20] The use of nonlinear transformations[27,28,30−32] has also been advocated as a mean to focus the sampling toward the physical states $\lambda = 0$ and $\lambda = 1$ and to smoothen possible free-energy barriers, i.e. as a possible remedy to problems D and E. Clearly, however, the use of a unique nonlinear transformation cannot provide a general solution to the inhomogeneous sampling and transition barrier problems, considering that the form of the PMF is system specific. Furthermore, none of the coordinate transformations proposed to date provides a formally exact remedy to the infinitesimal nature of the physical end points $\lambda = 0$ and $\lambda = 1$,

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

i.e. to problem B. Typically, the ensembles of configurations characterized by $|\lambda| < \varepsilon$ or $|1 - \lambda| < \varepsilon$, where ε is a tolerance parameter (*e.g.* $\varepsilon = 0.1$ or 0.2 in ref 30), are assumed to represent the physical states. However, in (common) situations where the PMF presents a high slope or a negative curvature at $\lambda = 0$ or $\lambda = 1$, this approximation may lead to very large errors due to unphysical states contributing significantly and spuriously to the calculated free-energy difference. In fact, a formal remedy to problem B requires the use of a noninvertible coordinate transformation, i.e. one that maps a full θ segment to the points $\lambda = 0$ and 1 corresponding to the physical states. This possibility was not explored previously and is considered in the present article.

In the original λD article,[20] the solution proposed to diminish the impact of possible free-energy barriers (problem E) is to increase the dimensionality of the λ-space above one, i.e. to assign a different λ variable to each perturbed atom with a holonomic constraint on the corresponding sum. However, although this approach has the merit of enabling relatively flat pathways, it does not address the problem of inhomogeneous sampling (problem D) and extends the one-dimensional range of unphysical λ values to a high-dimensional space of unphysical states, the size of which grows exponentially with the number of dimensions, i.e. of perturbed atoms. To address problems D and E in a general fashion would require a method to optimize the coordinate transformation for each specific PMF. Alternatively one may prefer to rely on a biasing potential optimized for the specific alchemical transformation, i.e. to use a memory-based biasing potential.

In the recent years, memory-based (sometimes called history-based[33,34] or adaptive[35]) biasing approaches have proven extremely successful in the context of conformational free-energy calculations and progressively supplanted multiple-windows approaches in these calculations, under many different names and flavours (see Introduction of ref 3 for an overview). A practical example for these kinds of methods is the local elevation umbrella sampling (LEUS) approach,[36] which relies on a combination of the local elevation method[37] (LE; also popularized under the name metadynamics[33]) and the umbrella sampling method[12] (US). The principle underlying LEUS is to construct an optimal biasing potential along a specific reaction coordinate as a sum of local grid-based functions, the weights of which are preoptimized dynamically for homogeneous sampling along this coordinate (LE phase) and then frozen for appropriate equilibrium sampling (US phase). The PMF corresponding to the physical (unbiased) ensemble is then reconstructed by an appropriate reweighting. The main advantages of memory-based methods such as LEUS over multiple-windows (serial) methods in the context of conformational free-energy calculations are: (*i*) shorter overall equilibration time (one simulation to equilibrate vs a set thereof for serial methods); (*ii*) ease of automatization (simple protocol, two simulations only, simple postprocessing).

To our knowledge, the only attempts to date for combining λD with memory-based biasing methods are the λ-metadynamics approach of Wu et al.[29] and the double-integration orthogonal space tempering (DI-OST) scheme of Zheng and Yang.[32] In spite of its novelty, λ-metadynamics still suffers from a number of shortcomings in terms of practical applications. In particular, it does not rely on an equilibrium sampling phase (the negative of the preoptimized biasing potential is directly used as an estimate for the PMF), does not address appropriately problems B and C, and requires a very large number of local basis functions to model the negative of the PMF (1800 or 4200 in ref 29). The DI-OST scheme addresses the need to not only enhance the sampling along θ, for which the adaptive biasing force[38,39] (ABF) scheme is used, but also in the space of orthogonal degrees of freedom. Inspired by Marcus' theory, the Hamiltonian derivative is assumed to be a suitable order parameter for enhancing the orthogonal sampling, by coupling its value to another virtual variable called ϕ. However, although ABF is used to improve the sampling along ϕ as well, this sampling is simultaneously restricted by introducing an effective sampling temperature (higher than the system temperature) along with the ϕ restraint. Hence, the faster convergence of DI-OST may result at least in part from limiting the extent of the orthogonal phase space (see e.g. Figures 7c and 8c in ref 32).

Last, the choices of the mass parameter m_λ and of the temperature-coupling scheme for the λ variable are expected to have a significant impact on the accuracy of λD calculations (problem C). They should be chosen appropriately, so as to guarantee the dynamical coupling of the λ degree of freedom to the rest of the system and an equilibrium sampling along this coordinate. Either a too low mass or a too high mass may induce the decoupling of the λ variable from the conformational degrees of freedom of the system (adiabatic decoupling[40−42]), also depending on the thermostatting scheme applied to the two sets of degrees of freedom (e.g., joint or separate, deterministic or stochastic, coupling strength). A too low mass may also result in fast λ variations, leading to integration errors and numerical instabilities. These effects are at present poorly characterized and the mass parameter m_λ either not reported[32] or chosen based on relatively obscure prescriptions. For example, the original λD article[20] reports the use of a mass of 20 or 40 atomic mass units[43] (abbreviated as u = g mol^{-1}), which is inconsistent with the actual units of m_λ, namely mass times square length. Similarly, although reporting m_λ in the appropriate units, the authors of ref 44 still suggest as a rough prescription that this mass should be of the same order of magnitude as that of the heaviest atom in the system, which is again unit-inconsistent. To underline the difference in units, m_λ is called a mass parameter rather than a mass in this article.

The goal of the present article is to remedy problems A−E simultaneously, leading to a robust, general, and efficient protocol for the calculation of free-energy changes along two-states (one-dimensional) alchemical processes. The key components of this scheme, termed λ-LEUS, are: (*i*) the use of a simple noninvertible coordinate transformation $\lambda(\theta)$, a zigzag curve with plateaus instead of tips, that is independent of the underlying PMF and ensures a finite sampling of the two physical end states; (*ii*) the application of the LEUS sampling enhancement scheme along λ, to enforce homogeneous sampling and overcome barriers; (*iii*) recommendations concerning the choice of m_λ and of the temperature-coupling scheme. As a small technical addition, the first-order spline basis functions used in the current LEUS implementation[45] (see Appendix A therein) are replaced here by second-order splines, permitting the representation of steep PMFs without spurious "staircase" effects.

The λ-LEUS scheme is described and tested considering a simple model system, the perturbation of hydroquinone to benzene in water. The results are compared to those of TI calculations in terms of accuracy and efficiency. The influence of different values for m_λ as well as of the width w of the plateaus in $\lambda(\theta)$ is also investigated.

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

2. THEORY

2.1. Thermodynamic Integration and λ-Dynamics. Consider two alchemical states A and B, corresponding to the Hamiltonians $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. The pathway-based calculation of the alchemical free-energy difference between these two states relies on the definition of a hybrid Hamiltonian $\tilde{\mathcal{H}}(\lambda)$ which depends parametrically on a coupling parameter λ in such a way that $\tilde{\mathcal{H}}(0) = \mathcal{H}_A$ and $\tilde{\mathcal{H}}(1) = \mathcal{H}_B$, i.e., the system is in state A for $\lambda = 0$ and in state B for $\lambda = 1$. For simplicity, the explicit dependence of $\tilde{\mathcal{H}}$, $\mathcal{H}_A$ and $\mathcal{H}_B$ on the coordinates $\mathbf{r}$ and momenta $\mathbf{p}$ of all atoms in the system is omitted from the notation.

In the TI method,[13−15] the free-energy change between the states A and B is written

$$\Delta G_{AB} = \int_0^1 \left\langle \frac{\partial \tilde{\mathcal{H}}(\lambda)}{\partial \lambda} \right\rangle_{\lambda'} \mathrm{d}\lambda' \tag{1}$$

where $\langle\cdot\rangle_\lambda$ denotes the Boltzmann-weighted ensemble averaging over $\mathbf{r}$ and $\mathbf{p}$ at a given value of λ. In practice, the integrand in eq 1 is calculated for a finite set of fixed λ values by direct averaging over thermostated MD simulations of finite durations and integrated by numerical quadrature.[46,47]

Even though TI has proven over many years to be one of the most robust tools for the determination of alchemical free-energy differences, it is not necessarily the most simple and efficient method to solve this problem. In particular the design of a TI protocol requires the manual selection of many parameters, including the number and positioning of the λ-points, the specification of the initial configurations as well as of the equilibration and sampling times at each λ-point, and the choice of the quadrature scheme.[46,47] The computational efficiency will depend on the selection of these parameters, but their optimization, if undertaken at all, represents a nontrivial and time-consuming task. Furthermore, an equilibration period is required for each λ-point (i.e., for each simulation), corresponding to nonproductive simulation time.

In the λD approach,[20] the coupling parameter λ is treated as a dynamical variable of an extended Hamiltonian $\mathcal{H}(\lambda)$ defined as

$$\mathcal{H}(\lambda) = \tilde{\mathcal{H}}(\lambda) + \frac{p_\lambda^2}{2m_\lambda} \tag{2}$$

which obeys a Newtonian-like equation of motion of the form

$$\ddot{\lambda} = \frac{\dot{p}_\lambda}{m_\lambda} = -\frac{1}{m_\lambda}\frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \tag{3}$$

where p_λ is the momentum of the λ variable, and m_λ its mass parameter, with actual units of mass times square length. Based on a single simulation of the extended ensemble, the free-energy difference between the two states can formally be calculated as

$$\Delta G_{AB} = -\frac{1}{\beta} \ln \frac{\langle \delta(\lambda - 1)\rangle}{\langle \delta(\lambda)\rangle} \tag{4}$$

where δ is the Dirac delta function and $\beta = (k_B T)^{-1}$, k_B being the Boltzmann constant and T the absolute temperature. Clearly, however, the ensemble averages involved in eq 4 cannot be evaluated exactly based on a finite simulation due to the infinitesimal-range nature of the δ function. This function must therefore be approximated by some finite-ranged binning function with an arbitrary half width ε. The error incurred by such an approximation on the results of the calculation can be very large, especially when the PMF features a large slope or negative curvature at $\lambda = 0$ or $\lambda = 1$, as illustrated numerically in Appendix A.

The "naive" approach corresponding to the direct application of eqs 2−4 results in the five shortcomings A−E listed in the Introduction section. In the following sections, solutions for these issues are presented. The resulting scheme is called λ-LEUS and includes: (*i*) a coordinate transformation; (*ii*) a memory-based biasing procedure; (*iii*) a free-energy estimator; (*iv*) recommendations for the choices of m_λ and of the thermostatting scheme.

2.2. Coordinate Transformation. The principle of the coordinate transformation[27−32] is to dissociate the Hamiltonian-coupling variable λ from the corresponding dynamical variable, relabeled θ, in such a way that the function $\lambda(\theta)$ is continuous and periodic in θ, its value being bounded to the range $\lambda \in [0;1]$. With this change, and renaming m_λ to m_θ as well as p_λ to p_θ, eqs 2 and 3 become

$$\mathcal{H}(\theta) = \tilde{\mathcal{H}}(\lambda(\theta)) + \frac{p_\theta^2}{2m_\theta} + \mathcal{B}(\theta) \tag{5}$$

and

$$\ddot{\theta} = \frac{\dot{p}_\theta}{m_\theta} = -\frac{1}{m_\theta}\left(\frac{\partial \tilde{\mathcal{H}}(\lambda)}{\partial \lambda}\frac{\mathrm{d}\lambda(\theta)}{\mathrm{d}\theta} + \frac{\mathrm{d}\mathcal{B}(\theta)}{\mathrm{d}\theta}\right) \tag{6}$$

The inclusion of a biasing potential-energy term $\mathcal{B}(\theta)$ will be discussed in the next section. For the moment, it is set to $\mathcal{B}(\theta) = 0$. The advantage of such a transformation is to enable the use of an unbounded variable θ in the dynamics, while restricting the extreme values of the coupling parameter λ to the physical values 0 and 1.

The λ-LEUS scheme involves a simple coordinate transformation which also circumvents the impossibility of evaluating exactly the ensemble averages of eq 4 from finite simulations. This function, illustrated in Figure 1 (top panel, green line), corresponds to a periodic zigzag line[29] interrupted by plateaus at $\lambda = 0$ and $\lambda = 1$. It is defined over the period $\theta \in [0;P[$ by the equation

$$\lambda(\theta) = \begin{cases} 0 & \text{if } 0 \le \theta \le w \\ \dfrac{\theta - w}{W} & \text{if } w < \theta < \dfrac{P}{2} \\ 1 & \text{if } \dfrac{P}{2} \le \theta \le \dfrac{P}{2} + w \\ \dfrac{P - \theta}{W} & \text{if } \dfrac{P}{2} + w < \theta < P \end{cases} \tag{7}$$

where w is the width of the plateaus, W is the inverse slope of the zigzag segments (in this work always set to 1), and $P = 2(w + W)$ is the period. The corresponding derivative, required for the application of eq 6, is given by

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

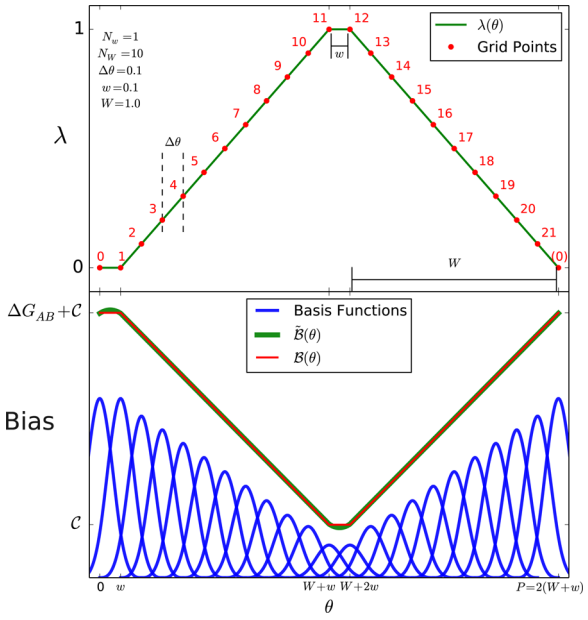


Figure 1. Schematic illustration of the coordinate transformation and of the memory-based biasing potential in the λ-LEUS scheme. The upper panel shows the coordinate transformation $\lambda(\theta)$ of eq 7 and the grid discretization of θ for the standard parameter settings used in the present calculations. Only one period $[0;P[$ of the function is shown, where $P = 2(W+w)$. For $\theta \in [0, w]$ and $[W+w, W+2w]$ the value of λ is 0 or 1, corresponding to one of the two physical end states. The parameters N_w and N_W represent the number of grid spacings $\Delta\theta$ in one plateau and one slope, respectively. The lower panel shows the representation of the biasing potential energy as a memory-weighted sum of grid-centered basis functions Φ (cubic B-splines, Appendix B), either without ($\tilde{\mathcal{B}}$, eq 10, green) or with linear ($\mathcal{B}$, eq 11, red) interpolation. The parameters are the same as for the upper panel and the biasing potential would be the one appropriate for a linear PMF with $\Delta G_{AB} > 0$. The linear interpolation only affects the plateaus and is essential to remove sampling inhomogeneities and ringing effects at the plateau edges.

$$\frac{d\lambda(\theta)}{d\theta} = \begin{cases} +\dfrac{1}{W} & \text{if } w < \theta < \dfrac{P}{2} \\ -\dfrac{1}{W} & \text{if } \dfrac{P}{2} + w < \theta < P \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Although the value of θ is formally unbounded during the dynamics, it can always be refolded by periodicity into the interval $[0;P[$. In the following, it will be assumed that such a refolding is applied, i.e. that $\theta \in [0;P[$, which simplifies the notation without affecting the results in any way.

Applying the above coordinate transformation in the situation where $\mathcal{B}(\theta) = 0$ in eqs 5 and 6, eq 4 can be rewritten as

$$\Delta G_{AB} = -\frac{1}{\beta} \ln \frac{\langle \delta(\lambda(\theta) - 1) \rangle}{\langle \delta(\lambda(\theta)) \rangle} \tag{9}$$

Considering this equation, the advantage of the plateaus in $\lambda(\theta)$ becomes evident. Although the δ functions in eq 4 account for an infinitesimal point along the λ coordinate, the δ functions in eq 9 now account for a finite segment along the θ coordinate (length fraction w/P for each of the two states). As a result, the ensemble averages in eq 9 can be evaluated exactly based on a finite simulation, i.e. without approximation of the δ functions by finite-ranged binning functions of arbitrary half-width ε.

2.3. Biasing Procedure. In addition to the coordinate transformation, the LEUS method is used to bias the simulation in such a way that homogeneous sampling along the λ coordinate is guaranteed. The memory-based biasing potential-energy term corresponds to the term $\mathcal{B}(\theta)$ in eqs 5 and 6, where the dependence of $\mathcal{B}$ on the memory $\mathbf{M}$ is omitted from the notation for simplicity. The LEUS method[36] consists of two phases, a local-elevation[37] (LE) phase and an umbrella-sampling[12] (US) phase. In the LE phase, the biasing potential $\mathcal{B}$ is constructed by systematically driving the system out of previously visited regions. In this nonequilibrium optimization phase, the memory $\mathbf{M}$ is time-dependent. The biasing potential $\mathcal{B}$ is then frozen and used for the subsequent US phase. In this equilibrium sampling phase, the memory $\mathbf{M}$ is thus time-independent.

The biasing potential is supported by a grid of N_P equally spaced points $k = 0,...,N_P - 1$ along θ and spans the reference period $[0;P[$ with a grid spacing $\Delta\theta = P/N_P$ as illustrated in Figure 1 (top panel, red points). The inverse slope W and plateau width w are required to be integer multiples of $\Delta\theta$, *i.e.* $W = N_W\Delta\theta$ and $w = N_w\Delta\theta$. The notation $\tilde{\theta} = \theta/\Delta\theta$ is introduced to represent θ in units of the grid spacing, so that in particular the coordinates of the grid points are $\tilde{\theta}_k = k$. Finally, the memory $\mathbf{M}$ is a N_P-dimensional vector specifying the contribution of each grid point to the biasing potential $\mathcal{B}$.

A priori, one would like to define $\mathcal{B}$ as a memory-weighted sum of local basis functions Φ centered at the successive grid points, i.e. as

$$\tilde{\mathcal{B}}(\theta) = \sum_{k=0}^{N_P-1} M_k \Phi(min(\tilde{\theta} - k)) \tag{10}$$

where the function *min* returns the minimal periodic distance, i.e. refolds its argument to the interval $[-N_P/2;\ N_P/2[$. Different choices are possible for the local basis function Φ, e.g. Gaussian[36,37] or first-order spline.[45] In λ-LEUS, second-order splines[48] (cubic B-splines) are employed, following the example of ref 34. The corresponding expression for Φ (eq B.5) as well as the reasons motivating this choice are detailed in Appendix B. Note that these functions satisfy in particular $\Phi(0) = 1$, $\Phi(\pm 1) = 1/4$, and $\Phi(t) = 0$ for $|t| \geq 2$, i.e. the function range extends up to the 4 neighboring grid points. Thus, in particular, the biasing potential at a grid point k is the sum of M_k plus one-quarter of the memory values at the two neighboring grid points (taking periodicity into account). In addition, Φ is twice continuously differentiable. The biasing potential of eq 10 is denoted $\tilde{\mathcal{B}}$ because it is not the one directly employed in λ-LEUS (see below).

The coordinate transformation $\lambda(\theta)$ and, consequently, the PMF $G(\theta)$ are not continuously differentiable at the points where the plateaus begin and end. Since the representation $\tilde{\mathcal{B}}$ of the biasing potential is by construction continuously differentiable, it follows that the optimality condition $\tilde{\mathcal{B}}(\theta) \approx -G(\theta)$ cannot be satisfied exactly. In practice, the optimization of a biasing potential of the form $\tilde{\mathcal{B}}$ during the LE phase will result in nonhomogeneous values over the plateaus, as illustrated in Figure 1 (bottom panel, green line) for the case of a linear PMF $G(\lambda)$ between states A and B (with $\Delta G_{AB} > 0$) and a plateau width corresponding to one grid unit ($w = \Delta\theta$). For plateau widths larger than $\Delta\theta$, this

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

may lead to pronounced oscillations at the plateau edges, analog to the ringing effect observed when representing a discontinuous curve with a truncated Fourier series (also called the Gibbs phenomenon). To remove these oscillations and enable free diffusion of the θ variable along the plateaus, a linear interpolation scheme is implemented. This scheme is active between the grid points m and n defining the edges of the plateaus, namely $(m, n) \in \{(0, N_w), (N_W + N_w, N_W + 2N_w)\}$. With this modification, the biasing potential $\mathcal{B}$ actually employed in the λ-LEUS scheme reads

$$\mathcal{B}(\theta) = \begin{cases} (\tilde{\theta} - m)\dfrac{\tilde{\mathcal{B}}_n - \tilde{\mathcal{B}}_m}{n - m} + \tilde{\mathcal{B}}_m & \text{if } m < \tilde{\theta} < n \\ \tilde{\mathcal{B}}(\theta) & \text{otherwise} \end{cases} \tag{11}$$

where $\tilde{\mathcal{B}}_k = \tilde{\mathcal{B}}(k\Delta\theta)$ is the biasing potential at grid point k.

The memory build-up procedure in the LE phase is identical to that formulated for the ball-and-stick LEUS (B&S-LEUS) method.[3] At each simulation time step, the grid point k closest to $\tilde{\theta}$ is determined (taking periodicity into account). The memory element M_k is then increased per time step by a force constant increment k_{LE}, weighted by a reduction factor, as

$$M_k(t + \Delta t) = M_k(t) + k_{\mathrm{LE}} \cdot f_{red}^R \tag{12}$$

The integer exponent or reduction counter R of f_{red} counts the number of double-sweeps of the θ range since the start of the LE phase, i.e. the counter is increased by one every time all the grid points have been visited at least once and then revisited at least once again since the last increment.

2.4. Free-Energy Estimator. In the absence of any biasing potential $\mathcal{B}$, the PMF $G(\theta)$ along θ could be obtained from an extended-ensemble simulation via

$$G(\theta') = -\frac{1}{\beta} \ln\langle \delta(\theta - \theta') \rangle + C \tag{13}$$

where C is an unknown constant. In the presence of a biasing potential, eq 13 must be replaced by the corresponding reweighted expression[12]

$$G(\theta') = -\frac{1}{\beta} \ln\langle \delta(\theta - \theta') \exp(+\beta\mathcal{B}(\theta)) \rangle_b + C \tag{14}$$

where $\langle\cdot\rangle_b$ denotes ensemble averaging over the biased ensemble. Similarly, the nonperiodic free-energy profile along λ is given by

$$G(\lambda') = -\frac{1}{\beta} \ln\langle \delta(\lambda(\theta) - \lambda') \exp(+\beta\mathcal{B}(\theta)) \rangle_b + C' \tag{15}$$

In the following, the convention that $G(0) = 0$ will be used to set the value of C'. The free-energy difference between the states A and B is then obtained via

$$\Delta G_{AB} = -\frac{1}{\beta} \ln \frac{\langle \delta(\lambda(\theta) - 1) \exp(+\beta\mathcal{B}(\theta)) \rangle_b}{\langle \delta(\lambda(\theta)) \exp(+\beta\mathcal{B}(\theta)) \rangle_b} \tag{16}$$

This equation generalizes eq 9 to the case of a biased simulation.

In practice, i.e. for a finite set of trajectory frames i, the ensemble averages in eq 16 are replaced by finite sums, so that the free-energy difference is calculated as

$$\Delta G_{AB} = -\frac{1}{\beta} \ln \frac{\sum_{\forall i.\lambda_i=1} \exp(\beta(\mathcal{B}_i - \mathcal{B}_{\max}))}{\sum_{\forall i.\lambda_i=0} \exp(\beta(\mathcal{B}_i - \mathcal{B}_{\max}))} \tag{17}$$

where $\mathcal{B}_i$ s the value of the biasing potential in frame i and $\mathcal{B}_{\max}$ its maximal value over all frames. The subtraction of $\mathcal{B}_{\max}$ in eq 17 does not affect the results but improves the accuracy of the calculations at finite machine precision by keeping the highest precision for the configurations with the highest weights.

2.5. Mass Parameter and Thermostatting Scheme. In the limit of infinite sampling, the choice of the mass parameter m_θ and of a thermostat coupling scheme for the conformational and alchemical sets of degrees of freedom of the extended system (e.g., joint or separate, deterministic of stochastic, coupling strength) should have no influence on the results, provided that the numerical integration errors (finite time step) are negligible and that the employed thermostat is statistical-mechanically correct,[49] e.g. Andersen[24] or Nosé–Hoover[25,26,50] thermostat. For simulations of finite durations, however, these choices must be made carefully to ensure a strong coupling between the two sets of degrees of freedom and between these and the thermostat, i.e. an exchange of kinetic energy that is significantly faster than the simulation time scale.

Either a too low or a too high mass parameter m_θ may lead to a situation of adiabatic decoupling[40,41] between the two sets of degrees of freedom, i.e. to a violation of equipartition between them. Note that methods like adiabatic free-energy decoupling[40,41] (AFED) and temperature accelerated MD (TAMD)[51] rely on the adiabatic decoupling of different sets of degrees of freedom at different temperatures. However, they only provide an approximation for the PMF,[52] so that even if the two sets of degrees of freedom exchange no kinetic energy, the PMF may still be wrong when the relaxation times of the orthogonal degrees of freedom is not much smaller than the time scale of the θ-motion. In contrast, the λ-LEUS method is an equilibrium method where the θ degree of freedom is meant to act like an additional (but otherwise "normal") configurational degree of freedom, kinetically coupled with the other ones.

On the one hand, if the mass parameter is too low, the motion along θ may become too fast for the thermostat (at a given coupling time) to maintain its temperature appropriately. On the other hand, if the mass parameter is too high, θ follows its course inertially, the influence of the underlying PMF becoming weak, i.e. only significant over long times (possibly longer than the simulation time). Furthermore, even with an appropriate choice of m_θ, the joint coupling of the two sets of degrees of freedom to a common thermostat may still lead to a steady state rather than an equilibrium situation, depending on the relative rates of heat transfer between alchemical and conformational subsystems and between them and the thermostat. This could result in constant but different temperatures for these subsystems.[53−55] Initial attempts with joint coupling revealed a tendency of the θ variable to be "hotter" than the conformational temperature for the process considered in the present article (data not shown). Based on these considerations, the λ-LEUS scheme relies on: (*i*) a choice of m_θ preoptimized (see below) for strong coupling between the θ variable and the conformational variables; (*ii*) a Nosé–Hoover chain thermostat[56] with strong heat-bath coupling for the θ variable; (*iii*) an arbitrary thermostat (e.g., weak coupling[57]) for the conformational degrees of freedom.

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

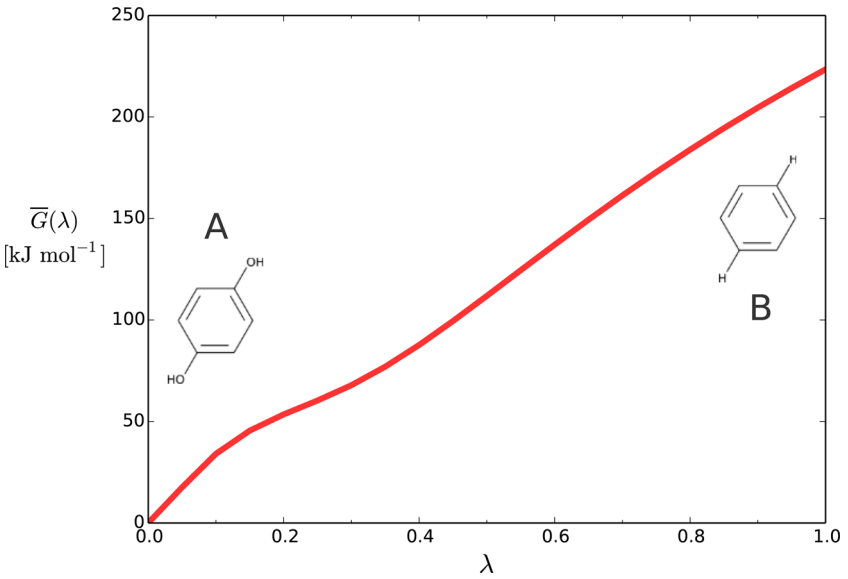


Figure 2. The alchemical process used as a test case in the present study. A hydroquinone molecule (state A) transforms into a benzene molecule (state B) in water. The red curve represents the potential of mean force $\overline{G}(\lambda)$ along the alchemical coupling parameter λ as averaged over the 10 TI calculations and anchored with $\overline{G}(0) = 0$. The corresponding average free-energy difference is $\overline{\Delta G}_{AB} = 223.49 \pm 0.08$ kJ mol^{-1}, the error estimate corresponding to the standard error on the mean.

To facilitate the optimization of m_θ so as to achieve strong coupling, it is useful to introduce characteristic times τ_θ as measures for the diffusivity of the system along θ. Two such measures are considered here. Based on a simple dimensional analysis, the thermal coupling time τ_θ^c is defined as

$$\tau_\theta^c = \sqrt{\beta m_\theta} \tag{18}$$

and satisfies the intuitive expectation that θ motions are slower, i.e. take more time, when m_θ is increased or T decreased. The thermal coupling time can also be interpreted as the inverse spread of the Maxwell−Boltzmann velocity distribution $p(\dot\theta)$

$$p(\dot\theta) = \sqrt{\frac{\beta m_\theta}{2\pi}}\, e^{-\beta m_\theta \dot\theta^2/2} = \frac{\tau_\theta^c}{\sqrt{2\pi}} e^{-(\tau_\theta^c \dot\theta)^2/2} \tag{19}$$

As an alternative measure, the diffusion time τ_θ^d is defined as the inverse diffusion constant D_θ as obtained from the Einstein equation[58]

$$(\tau_\theta^d)^{-1} = D_\theta = \lim_{t\to\infty} \frac{1}{2t} \langle [\theta(\tau + t) - \theta(\tau)]^2 \rangle_t \tag{20}$$

where $\langle\cdot\rangle_t$ denotes averaging over τ at constant t. Note that θ must be unfolded, i.e. allowed to diffuse across periodic intervals, for the calculation of D_θ using this equation. The analysis of the different τ_θ for different m_θ permits to assess the coupling strength. In practice, a broad range of m_θ values was found to actually give results of comparable accuracies for the process considered in the present article.

3. COMPUTATIONAL DETAILS

All simulations were carried out using a modified version of the GROMOS MD++ program[59−61] together with the 53A6 force field[62] and the simple point charges (SPC) water model.[63] The model alchemical transformation considered involved aqueous hydroquinone (*p*-benzenediol, state A) transforming into benzene (including two dummy atoms, state B), as illustrated in Figure 2. Soft-core interactions[64] were used for the oxygen to hydrogen and the hydroxyl hydrogen to dummy mutations, to avoid the singularity at $\lambda = 1$, with softening parameters $\alpha_{LJ} = 0.5$ and $\alpha_{CRF} = 0.5$ nm^2 for Lennard-Jones and electrostatic interactions, respectively.

The system consisted of a single solute molecule solvated in 1091 water molecules within a cubic periodic box of edge 3.233 nm (density of 970 kg m^{-3}). The simulations were performed at constant volume and temperature. The temperature of all atomic degrees of freedom was kept close to 300 K using weak coupling[57] to a single heat bath with a coupling constant of 0.1 ps. The equations of motion were integrated using the leapfrog algorithm[65] with a time step of 2 fs. All bond lengths as well as the water hydrogen−hydrogen distances were constrained by application of the SHAKE procedure[66] with a relative geometric tolerance of 10^{-4}. The translational center of mass motion was removed every 10000 steps. The nonbonded interaction were calculated using a twin-range cutoff scheme[67] with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively, and an update frequency of 5 timesteps for the short-range pairlist and intermediate range interactions. The mean effect of the electrostatic interactions beyond the long-range cutoff distance was accounted for by means of a reaction-field correction[68,69] using a relative dielectric permittivity of 61 as appropriate[70] for the SPC water model.

The calculations performed included: (*i*) a set of 10 independent TI calculations of ΔG_{AB} with different initial velocities; (*ii*) a set of 10 independent λ-LEUS calculations of ΔG_{AB} with different initial velocities, using an inverse slope $W = 1$, a plateau width $w = 0.1$, and a mass parameter $m_\theta = 75$ u nm^2; (*iii*) a series of 5 independent λ-LEUS simulations with different plateau widths w at $W = 1$ and $m_\theta = 75$ u nm^2; (*iv*) a series of 11 independent λ-LEUS simulations with different m_θ at $W = 1$ and $w = 0.1$.

For all TI calculations 41 equispaced λ-points were used. Each λ-point was sampled during 0.9 ns after 0.1 ns equilibration, resulting in a total equilibration time $t_{EQ} = 4.1$ ns and a total sampling time $t_{TI} = 36.9$ ns for each ΔG_{AB} estimate. The integration was carried out using Simpson's rule.[47]

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

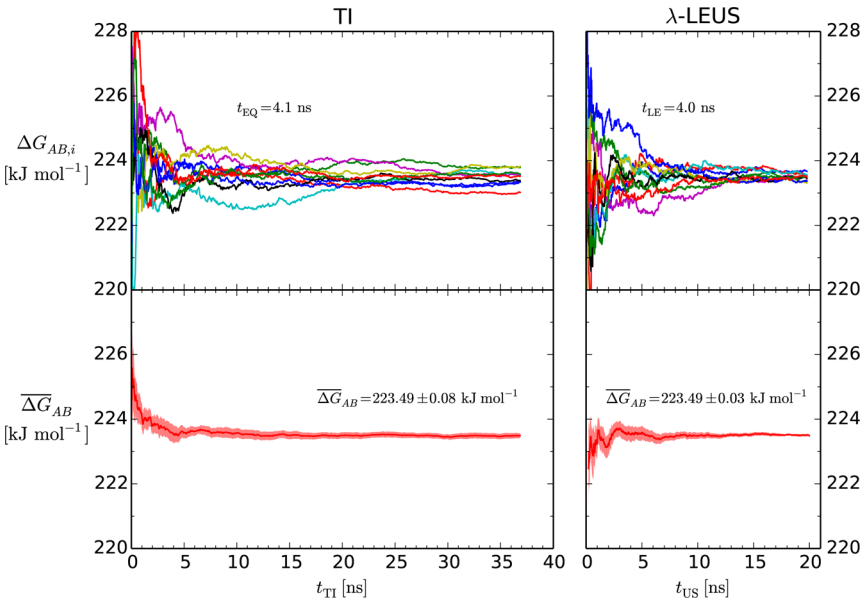


Figure 3. Convergence properties of the TI and λ-LEUS calculations. The left panels show the results of 10 independent TI calculations. The right panels show the results of 10 independent λ-LEUS calculations. All of them were started with different initial velocity random seeds. The top panels show the value of $\Delta G_{AB,i}$ for each simulation i as a function of the total sampling time, t_{TI} for TI and t_{US} for λ-LEUS. The preceding total equilibration (TI) or LE build-up (λ-LEUS) times are t_{EQ} = 4.1 ns and t_{LE} = 4.0 ns. The bottom panels show the mean value of ΔG_{AB} as a function of the total sampling time. The shaded area accounts for the standard error on the mean. The final average values of $\overline{\Delta G}_{AB}$ were 223.49 ± 0.08 kJ mol^{-1} for TI and 223.49 ± 0.03 kJ mol^{-1} for λ-LEUS.

For all λ-LEUS simulations, the grid spacing was set to $\Delta\theta$ = 0.1 and the inverse slope to W = 1 leading to N_W = 10 grid spacings per slope segment. For the LE phase, the force-constant increment k_{LE} was normally set to 10^{-3} kJ mol^{-1} per time step (see exceptions below) and the force-reduction factor f_{red} to 0.95. The θ variable was thermostated using a Nosé–Hoover chain[56] of 10 thermostat variables, each with a relaxation time of 0.1 ps.

For the set of 10 reference λ-LEUS calculations to be compared with the TI results, the mass parameter m_θ was set to 75 u nm^2 and the plateau width to w = 0.1, leading to N_w = 1 grid spacing per plateau and N_P = 22 grid points in total (situation illustrated in Figure 1). The duration of the LE phase was t_{LE} = 4 ns (approximately equal to the total equilibration time t_{EQ} of the TI simulations) and that of the US phase t_{US} = 20 ns (approximately half of the total sampling time t_{TI} of the TI simulations).

In the 5 λ-LEUS simulations probing the effect of the plateau width, the mass parameter m_θ was set to 75 u nm^2, the force-constant increment k_{LE} was changed to 5 × 10^{-4} kJ mol^{-1} per time step, and the plateau width was varied from 0.0 to 0.4 in steps of 0.1. The duration of the LE phase was 10 ns and that of the US phase 20 ns. In the case of w = 0.0, the half-width of the binning function (approximating the δ function of eq 16) was set to either ε = 0.01 or 0.001

In the 11 λ-LEUS simulations probing the effect of the mass parameter, the plateau width w was set to 0.1 and the mass parameter m_θ was varied from 1 to 200 u nm^2 in 11 steps ($m_\theta \in$ {1, 25, 50, 60, 75, 90, 100, 125, 150, 175, 200} u nm^2). The durations of the LE and the US phases were both 20 ns.

The diffusion time τ_θ^d of eq 20 was calculated via a linear least-squares fit over the interval $t\in$[0,300] ps. All plots were created with Python (www.python.org) and the Matplotlib library.[71]

4. RESULTS

4.1. Calculations Using TI. The TI calculations resulted in 10 independent estimates for the PMF $G(\lambda)$, each requiring t_{EQ} = 4.1 ns and t_{TI} = 36.9 ns total equilibration and sampling times, respectively. The resulting average PMF $\overline{G}(\lambda)$ is shown in Figure 2. The free-energy differences $\Delta G_{AB,i}$ obtained from each calculation i are shown in the top left panel of Figure 3 as a function of the total sampling time t_{TI} invested. Interestingly, the spread of the different $\Delta G_{AB,i}$ curves remains approximately constant after about 20 ns, i.e. roughly 0.5 ns sampling per λ-point. The corresponding average estimate $\overline{\Delta G}_{AB}$ as well as the corresponding estimated error (standard error on the mean) are shown as a function of the total sampling time t_{TI} per calculation in the bottom left panel. The final value of $\overline{\Delta G}_{AB}$, considering the 10 simulations after the full sampling period of 36.9 ns per calculation, is 223.49 ± 0.08 kJ mol^{-1}. The analogous estimate using trapezoidal instead of Simpson quadrature yields 223.37 ± 0.07 kJ mol^{-1} with convergence curves nearly indistinguishable from those in Figure 3 (data not shown).

In order to understand why significant differences persist between the individual $\Delta G_{AB,i}$ estimates even at relatively long sampling times, the mean values $\bar{X}$ and mean absolute deviations MAD(X) were calculated for $G(\lambda)$ and $\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda$ as

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{21}$$

and

$$\mathrm{MAD}(X) = \frac{1}{N}\sum_{i=1}^{N} |X_i - \bar{X}| \tag{22}$$

with N = 10. The MAD was preferred here over the root-mean-square deviation (RMSD) to avoid giving excessive weights to

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

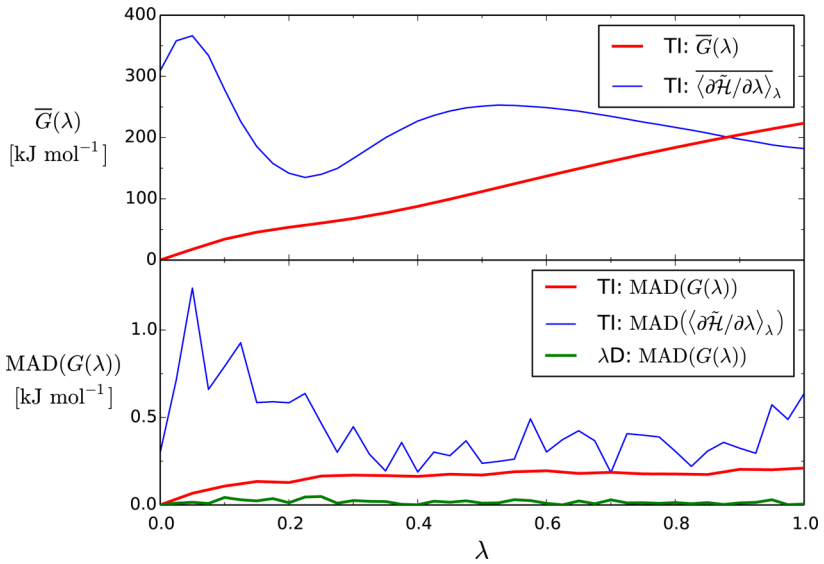


Figure 4. Distribution of sampling errors along the λ coordinate. The top panel shows the average PMF $\overline{G}(\lambda)$ calculated based on the 10 independent TI calculations (red curve; same as in Figure 2) and the corresponding average of $\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda$ (blue curve). The bottom panel shows the associated error, namely the MAD of $G(\lambda)$ (red curve) and the MAD of $\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda$ (blue curve). The MAD of $G(\lambda)$ calculated based on the 10 independent λ-LEUS simulations is also shown for comparison (green curve).

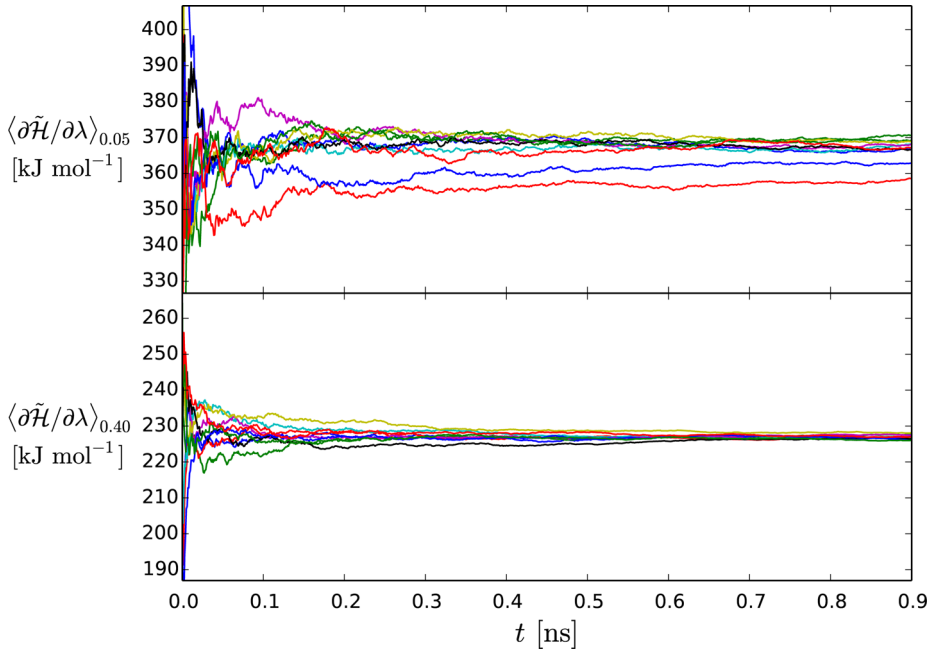


Figure 5. Convergence properties of $\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda$ for two illustrative values of λ. The running averages of $\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda$ from 10 independent TI calculations is shown as a function of the TI sampling time t at these λ points. The top panel shows slow convergence for λ = 0.05 and the bottom panel fast convergence for λ = 0.40.

large deviations.[72] The curves of the mean derivative $\overline{\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda}$ and the mean PMF $\overline{G}(\lambda)$ are shown in the top panel of Figure 4, while the corresponding MADs are shown in the bottom panel. The MAD of $\langle \partial\tilde{\mathcal{H}}/\partial\lambda \rangle_\lambda$ is largest between λ = 0.0 and 0.2, i.e. close to the hydroquinone state. As a result the MAD of $G(\lambda)$ increases rapidly up to $\lambda \approx 0.2$ and remains constant thereafter.

As illustrated in Figure 5, which compares the running averages of $\partial\tilde{\mathcal{H}}/\partial\lambda$ at the λ-points 0.05 and 0.40 for the 10 independent simulations, the increased error between λ = 0.0 and 0.2 stems from the apparent convergence of the different simulations to different average values on the 0.9 ns time scale. This suggests the presence of significant barriers along the degrees of freedom orthogonal to θ, which cannot easily be crossed even on the ns-time scale. These barriers, occurring in the region close to the hydroquinone state, are presumably related to the progressive loss of solute−solvent hydrogen-bonding interactions, transiently leading to conformation sets in slow interconversion.

4.2. Calculations Using λ-LEUS. The λ-LEUS calculations meant for the comparison with TI also resulted in 10

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

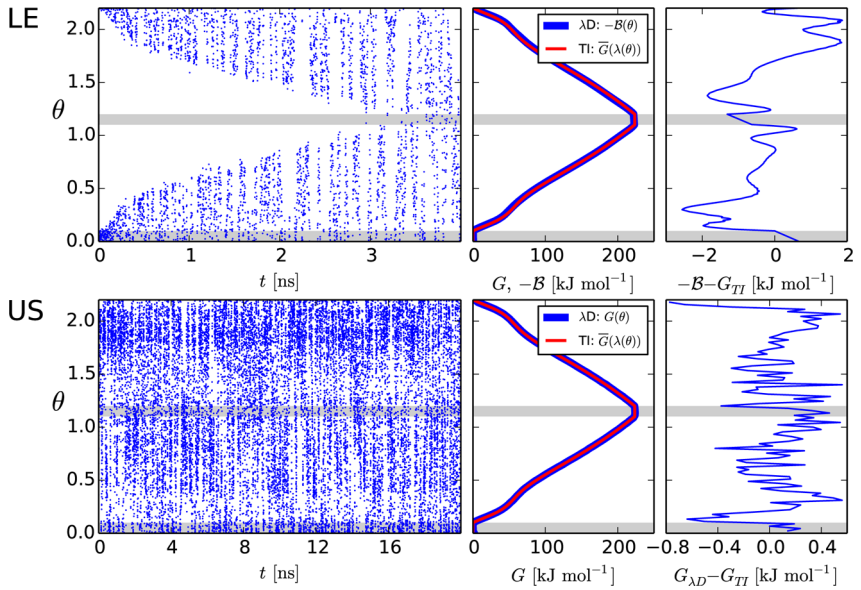


Figure 6. Illustrative example of the convergence properties of a λ-LEUS simulations. The top panel shows the data for the LE phase, the bottom panel the data for the US phase. The gray shaded areas represent the plateaus. The left panels show the time series of the θ variable. The top center panel shows the TI reference curve $\bar{G}(\theta(\lambda))$ (red) and the negative biasing potential $-\mathcal{B}(\theta)$ at the end of the LE phase (blue). The bottom center panel shows the same TI reference curve (red) and the free-energy profile $G(\theta)$ calculated from the US phase of the λ-LEUS simulation (blue). The right panels show the difference between the two curves in the center panels. The corresponding time series of the velocity $\dot{\theta}$ is shown in Figure 9.

independent estimates for the PMF $G(\lambda)$, each requiring t_{LE} = 4 ns LE build-up and t_{US} = 20 ns US sampling time, respectively. The convergence properties for one of these 10 calculations are illustrated in Figure 6. During the LE phase (top left panel), the biasing potential is constructed, and the system progressively evolves from sampling θ values close to 0 (equivalent to 2.2 by periodicity), i.e. in the neighborhood of the hydroquinone state of lowest free energy, to sampling the entire range [0;2.2[of θ values. At the end of this period, the negative $-\mathcal{B}(\theta)$ of the biasing potential is as expected close to the average PMF $\bar{G}(\lambda(\theta))$ evaluated using TI (top center panel). However, a closer look at the quantity $-\mathcal{B}(\theta) - \bar{G}(\lambda(\theta))$ (top right panel) suggests that using $-\mathcal{B}(\theta)$ directly as an estimate for the PMF, as generally done in metadynamics[33] and in λ-metadynamics,[29] would lead to relatively large errors, as already pointed out previously.[36,73] In particular, the magnitude of the error would be as large as about 2 kJ mol^{-1} close to θ = 0.2 (negative) and θ = 2.0 (positive), a significant negative error also being observed close to θ = 1.4. The two former values both correspond to λ = 0.2, i.e. close to the hydroquinone state, in the region where the PMFs evaluated using TI presented the largest variations (Figure 4). Arguably, this difference could be reduced by increasing the LE build-up time above 4 ns, considering that sweeping of the full θ range from about 3.5 ns onward enables the build-up rate to progressively decrease to zero. At 4 ns, the force-reduction counter R of eq 12 is about 7, see Table 1, i.e. the build-up rate has already been decreased by a factor $0.95^7 \approx 0.7$. However, the use of a subsequent US sampling phase enables a much better control of the error.[3,36,73] Following this strategy, the build-up time was limited here to 4 ns so as to be approximately equal to the corresponding total TI equilibration time t_{EQ} for comparison purposes.

During the US phase (bottom left panel of Figure 6), the biasing potential is frozen at the stage reached by the end of the LE phase, and the system keeps sampling the entire range [0;2.2[of θ values, now under equilibrium conditions. As expected from the discrepancy between $-\mathcal{B}(\theta)$ and $\bar{G}(\lambda(\theta))$ (see above), the regions close to θ = 0.2 and 1.4 are somewhat undersampled, and those close to θ = 2.0 somewhat oversampled. At the end of this period, the PMF $G(\lambda)$ estimated from the λ-LEUS simulation is also very close to the average PMF $\bar{G}(\lambda(\theta))$ evaluated using TI (bottom center panel). However, a closer look at the difference (bottom right panel) shows that the error is significantly reduced compared to the direct use of $-\mathcal{B}(\theta)$ as an estimate for the PMF, being now of at most 0.6 kJ mol^{-1}.

The free-energy differences $\Delta G_{AB,i}$ obtained from each of the 10 calculations i are shown as a function of the total sampling time t_{US} in the top right panel of Figure 3. Compared to the corresponding TI curves (top left panel; note the twice shorter time scale for λ-LEUS), the convergence properties at short

Table 1. Reduction Counter at the End of the Build-up Phase[a]

seed	R	w	R	m_θ [u nm^2]	R
1	5	0.0	23	1	348
2	8	0.1	24	25	125
3	7	0.2	17	50	86
4	5	0.3	13	60	86
5	7	0.4	9	75	85
6	6			90	71
7	7			100	70
8	7			125	74
9	7			150	64
10	8			175	58
				200	55

[a]In this table the values of the reduction counters R at the end of the LE phase (t_{LE} = 4, 10, or 20 ns from left to right) for the three different sets of λ-LEUS calculations are listed; namely for the 10 calculations with different initial velocity random seeds (left), for the 5 calculations with different plateau widths w (center) and for the 11 simulations with different mass parameters m_θ (right).

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

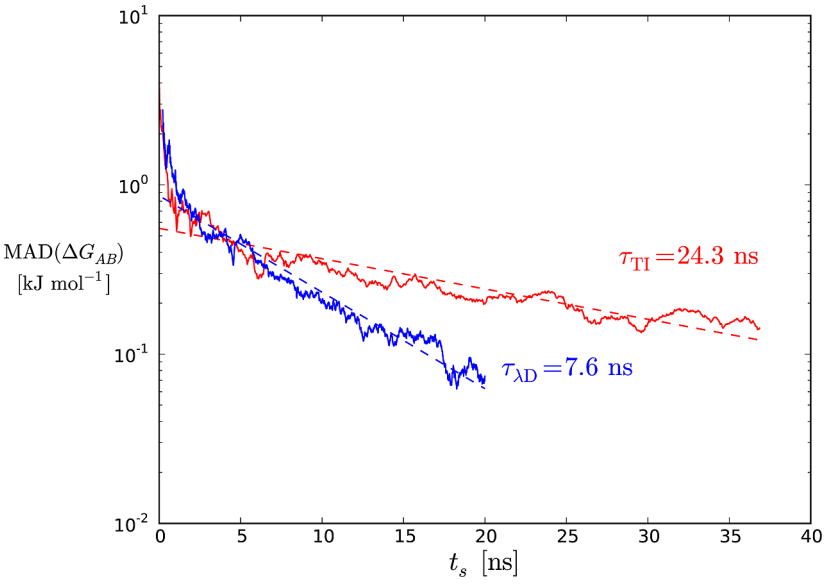


Figure 7. Comparison of the convergence properties of TI and λ-LEUS. The MAD of ΔG_{AB} from the 10 TI calculations (red) is compared to that from the 10 λ-LEUS calculations (blue) on a logarithmic scale, as a function of the total sampling time t_s (t_{EQ} and t_{US}, respectively). The dashed curves show an exponential fit of the form $A \exp(-t_s/\tau_{MAD})$, where $\tau_{MAD,TI}$ = 24.3 ns and $\tau_{MAD,\lambda D}$ = 7.1 ns.

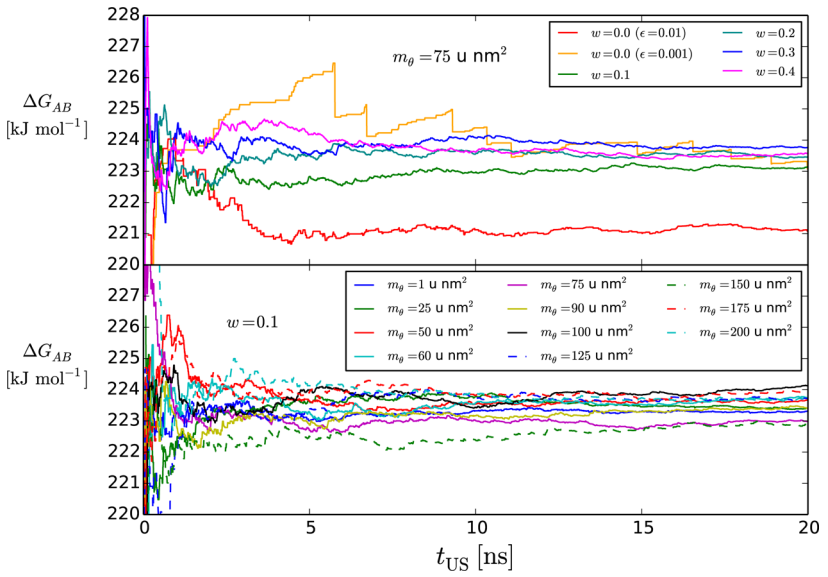


Figure 8. Sensitivity of the λ-LEUS results to plateau width and mass parameter. The top panel shows ΔG_{AB} as a function of the sampling time t_{US} for different plateau widths w (with W = 1 and m_θ = 75 u nm^2). For the width w = 0.0 (no plateau), two curves are shown corresponding to definitions of the physical states using binning functions with half-widths of either ε = 0.01 or ε = 0.001. The bottom panel shows the dependence of ΔG_{AB} as a function of the sampling time t_{US} for different mass parameters m_θ (with W = 1 and w = 0.1).

time ($\leq$10 ns) are comparable, but the λ-LEUS curves evidence less spread at 20 ns than the corresponding TI curves at the same time or even after 39.6 ns. The corresponding average estimate $\overline{\Delta G_{AB}}$ as well as the corresponding estimated error (standard error on the mean) are shown as a function of the total sampling time t_{US} per calculation in the bottom right panel of Figure 3. The final value of $\overline{\Delta G_{AB}}$ considering the 10 calculations after the full sampling period of 20 ns per calculation is 223.49 $\pm$ 0.03 kJ mol^{-1}, to be compared with 223.49 $\pm$ 0.08 kJ mol^{-1} for TI. The two results are in perfect agreement, but the λ-LEUS result is about two to three times more precise in spite of the halved sampling time. The MAD of $G(\lambda)$ across the λ-LEUS calculations, which is shown against that of TI in the bottom panel of Figure 4, also reveals a better convergence along the entire PMF, including the region 0.0–0.2 where TI displayed the most pronounced variations.

To make the above comparison of the convergence properties of TI and λ-LEUS more quantitative, the MAD of ΔG_{AB} across the 10 calculations is shown for both methods on a logarithmic scale in Figure 7 as a function of the total sampling time (t_{TI} or t_{US}, respectively). Beyond about 5 ns, both curves become approximately linear, corresponding to an exponential decrease of the MAD. The corresponding exponential time constants τ_{MAD} are 24.3 ns for TI, indeed significantly longer than that for λ-LEUS, namely 7.1 ns.

A likely explanation for the faster convergence (higher sampling efficiency) of λ-LEUS compared to TI arises from the observation that the latter method constrains λ in separate

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

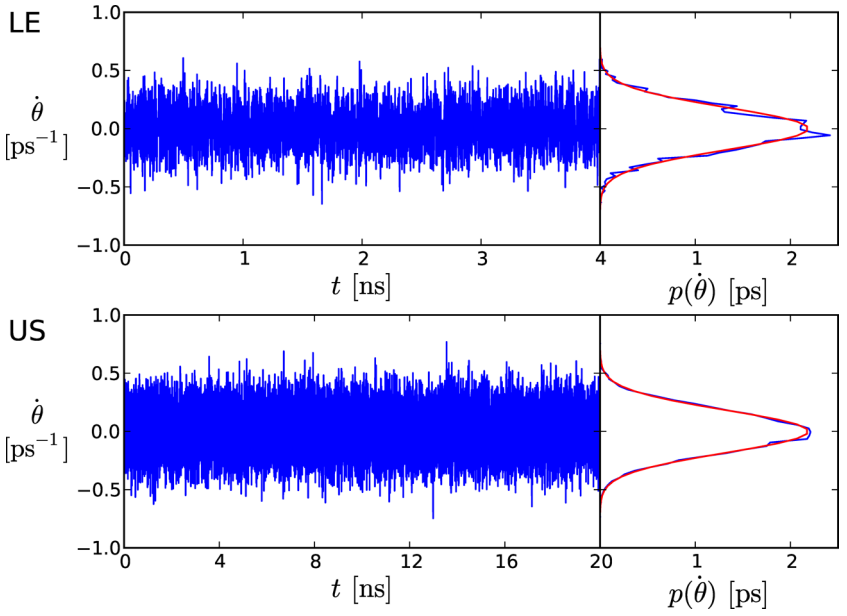


Figure 9. Illustrative example of the θ velocity properties in a λ-LEUS simulation. The top panel shows the LE phase, the bottom one the US phase. The left panels show the time series of $\dot{\theta}$. The right panels show the corresponding normalized probability distribution (blue) and the expected analytical Maxwell−Boltzmann distribution (eq 19; red curve). The corresponding time series of the θ variable is shown in Figure 6.

simulations, whereas the former method lets this variable evolve freely in a single simulation. As seen in the previous section, the range between λ = 0.0 and 0.2 appears to involve orthogonal barriers that are then not easily crossed on the ns-time scale for fixed values of λ. However, the resulting parallel valleys join close to λ = 0.0 or 0.2. As a result, λ-LEUS can more easily overcome such orthogonal barriers by enabling variable λ pathways circumventing them. Note, finally, that the TI protocol employed here for comparison purposes has not been optimized for efficiency.[46,47] In particular, it may be argued that the use of 41 λ-points represents an overkill and that the application of the MBAR estimator[11] should be preferred over that of Simpson quadrature. Nevertheless, the ability of λ-LEUS to circumvent orthogonal barriers via variable-λ pathways will still remain inaccessible even to a highly optimized TI protocol.

4.3. Influence of the Plateau Width. To examine the influence of the plateau width w on the calculated results, 5 additional λ-LEUS simulations were undertaken, involving plateau widths ranging from 0.0 to 0.4 in steps of 0.1. The other parameters were the same as used above, except for a longer build-up time t_{LE} = 10 ns and a halved force-constant increment k_{LE} = 5 × 10^{-4} kJ mol^{-1} per time step. The free-energy differences $\Delta G_{AB,i}$ obtained from each of the 5 calculations i are shown as a function of the total sampling time t_{US} in Figure 8 (top panel).

As soon as a plateau is introduced ($w \geq 0.1$) its width does not affect the results significantly over the range considered, with a convergence rate and a spread in the final values comparable to those observed for the previous calculations with w = 0.1 (Figure 3, top right panel). In principle one would expect broader plateaus to increase the statistical efficiency (more time spent in the physical states) and decrease the convergence rate (fewer transitions between these states). The latter effect is visible indirectly in the values of the reduction counters R at the end of the 10 ns LE phases of these simulations, listed in Table 1, which decrease with increasing plateau width.

The situation is different if the width is set to zero. In this case, the physical states correspond to infinitesimal points at θ = 0 and 1 in the θ range [0;2[, and the δ functions in eq 16 have to be replaced by finite-ranged binning functions. This leads to an error proportional to the half-width ε and the mean of the slopes of the PMF at the end states as explained in Appendix A. The $\Delta G_{AB,i}$ value obtained using w = 0.0 and a half-width ε = 0.01 is well converged, but, as expected, lower by about 2.5 kJ mol^{-1} compared to the other estimates due to the spurious contribution of unphysical states. The binning error principally affects state B which is a free-energy maximum along θ. In contrast, using w = 0.0 and a half-width ε = 0.001 results in a more accurate value for $\Delta G_{AB,i}$ after a long sampling time but leads to a more erratic convergence due to the lack of statistics in the end-state bins. Another issue related to the absence of plateaus is that the interpolation scheme of eq 11 is inactive, i.e. the memory-based biasing function is continuously differentiable and cannot provide an exact representation of the negative of the PMF $G(\theta)$ which has a discontinuous first derivative at the tips (θ = 0 or 1).

4.4. Influence of the Mass Parameter. To examine the influence of the mass parameter m_θ on the calculated results, 11 additional λ-LEUS simulations were undertaken, involving mass parameters ranging from 1 to 200 u nm^2. The other parameters were the same as used above, except for a longer build-up time t_{LE} = 20 ns along with the normal force-constant increment of k_{LE} = 10^{-3} kJ mol^{-1} per time step. The free-energy differences $\Delta G_{AB,i}$ obtained from each of the 11 calculations i are shown as a function of the total sampling time t_{US} in Figure 8 (bottom panel).

The value of the mass parameter does not affect the results significantly over the range considered, with a convergence rate and a spread in the final values comparable to those observed for the previous calculations with m_θ = 75 u nm^2 (Figure 3, top right panel). In principle one would expect the mass parameter to have little effect on the calculated values as long as it is appropriate to achieve: (i) strong coupling (fast exchange of kinetic energy on the sampling time scale) between the

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

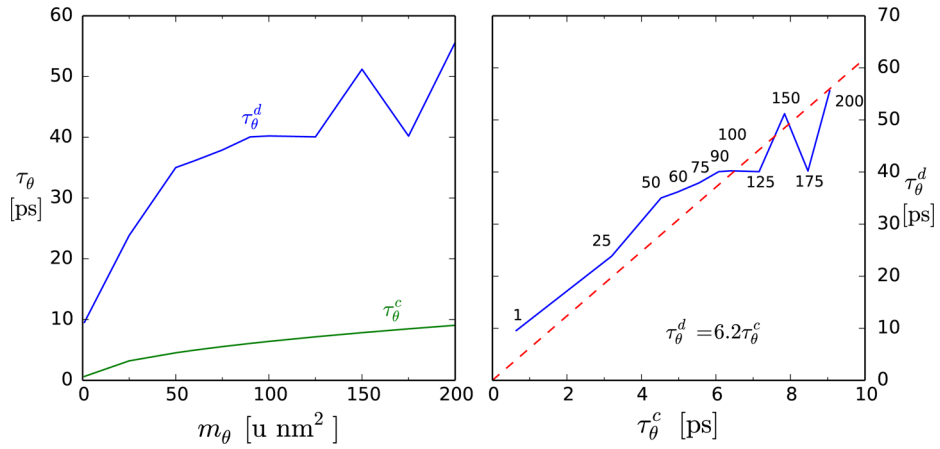


Figure 10. Characteristic times associated with the diffusion along the θ variable. The left panel shows the thermal time τ_θ^c (eq 18; green curve) and the diffusive time τ_θ^d (eq 20; blue curve) for different values of the mass parameter m_θ. The right panel shows τ_θ^d as a function of τ_θ^c (blue curve) along with a fit (eq 23; red dashed line), where k_θ = 6.2. The numbers indicate the different values of m_θ.

alchemical variable θ and the conformational degrees of freedom of the system; (*ii*) strong coupling of each of the two sets of degrees of freedom to its separate thermostat, avoiding a steady-state situation involving distinct temperatures (determined by the different rates of exchange of kinetic energy). In such a "healthy" situation, one expects the velocity $\dot{\theta}$ of the θ variable to fluctuate rapidly (including changes of sign) and to display a Maxwell distribution at the selected simulation temperature. As shown in Figure 9 for an illustrative λ-LEUS simulation with m_θ = 75 u nm^2, this situation is achieved in the present simulations.

The dependence of the diffusivity along θ on the value of the mass parameter can be made more explicit by considering the characteristic times τ_θ^c and τ_θ^d of eqs 18 and 20. These characteristic times are displayed as a function of m_θ in Figure 10 (left panel). As expected, increasing m_θ reduces the diffusivity, i.e. slows down the motions of θ. This effect is also visible indirectly in the values of the reduction counters R at the end of the 20 ns LE phase of these simulations, listed in Table 1, which decrease with increasing mass parameter. The kinetic theory of mass-dependent diffusion in liquids[74] suggests that the inverse of the diffusion constant is proportional to the square of the mass, thus suggesting a simple linear dependence between τ_θ^c and τ_θ^d, namely

$$\tau_\theta^d = k_\theta \tau_\theta^c \tag{23}$$

As illustrated in the right panel of Figure 10, such a relationship indeed approximately holds, with $k_\theta \approx 6.2$.

The distribution of $\partial\tilde{\mathcal{H}}/\partial\lambda$ as a function of λ along the sampling is displayed in Figure 11 for one TI and two λ-LEUS sample simulations, in the latter case considering the two extreme values of the mass parameter m_θ. Expectedly, the point distributions define the same average $G(\lambda)$. They also display comparable spreads as seen by their MAD of $\partial\tilde{\mathcal{H}}/\partial\lambda$. Comparing the two values of the mass parameter, a too rapid rate of change of the θ-variable at low m_θ would cause a noticeable difference between the values of $\partial\tilde{\mathcal{H}}/\partial\lambda$ for TI and λD, since the orthogonal degrees of freedom would not have enough time to relax. Since no such change occurs, it appears that any choice of the mass parameter m_θ between 1 and 200 u nm,^{2} is appropriate for the present system.

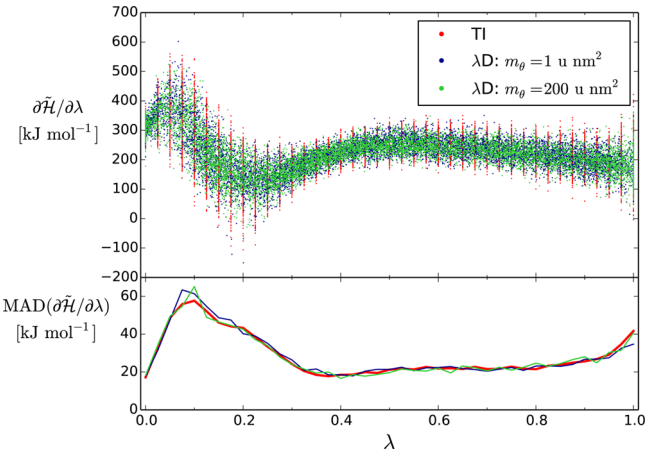


Figure 11. Values of the Hamiltonian derivative in illustrative TI and λ-LEUS calculations. The top panel shows the values of $\partial\tilde{\mathcal{H}}/\partial\lambda$ for a sample TI simulation (red points), along with those for two different λ-LEUS simulations with mass parameters of 1 and 200 u nm^2, respectively (blue and green points). The bottom panel shows the corresponding MADs of $\partial\tilde{\mathcal{H}}/\partial\lambda$.

5. CONCLUSION

In the present study, a new scheme termed λ-LEUS is proposed to alleviate the five shortcomings of λD (points A–E in the Introduction section) for calculating free-energy differences along two-states (one-dimensional) alchemical processes. The main ingredients of this scheme are: (*i*) the use of a simple noninvertible coordinate transformation $\lambda(\theta)$ involving plateaus of width w that ensure a finite sampling of the two physical end states; (*ii*) the application of the LEUS sampling enhancement method along θ, to enforce homogeneous sampling and overcome barriers; (*iii*) recommendations concerning the choice of the mass parameter m_θ and of the temperature-coupling scheme; (*iv*) the use of second-order splines basis functions for the memory-based biasing potential, permitting the representation of a steep PMF without spurious "staircase" effects.

The λ-LEUS scheme was tested considering the perturbation of hydroquinone to benzene in water as a model system. The results are comparable to those of reference TI calculations but with a superior accuracy-to-efficiency ratio in the sampling.

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

This improved sampling performance is interpreted as a consequence of allowing λ-relaxation processes, which help overcome barriers along the orthogonal set of degrees of freedom and thus achieve more thorough sampling of the phase space in less simulation time. The influence of the plateau width w (between 0.1 and 0.4) and of the mass parameter m_θ (between 1 and 200 u nm^2) on the results is found to be limited. However, the use of sharp tips (w = 0.0) is not recommended. It leads to the necessity of approximating the δ functions defining the physical end states using finite-width binning functions. This corresponds to including contributions from unphysical states in the calculated free-energy difference. The precision and the accuracy of the results then depends on the choice of an arbitrary half-width ε, the optimal compromise at finite-sampling time being generally still affected by significant errors.

From a practical perspective, the design of a λ-LEUS protocol requires fewer human decisions and its automatization is easier compared to a TI protocol, considering that the design of a TI protocol requires the manual and possibly iterative selection of many parameters (number and positioning of the λ-points, specification of initial configurations as well as equilibration and simulation times at each λ-point, choice of a quadrature scheme). The input parameters of a λ-LEUS calculation are the inverse slope W, the plateau width w, the grid spacing $\Delta\theta$, the mass parameter m_θ, the durations t_{LE} and t_{US} of the LE and US phase, respectively, the force-constant increment k_{LE} per time step, and the reduction factor f_{red}. The monitoring of the force-reduction counter R offers a simple way to stop the LE phase at an appropriate time point, while the monitoring of the convergence of ΔG_{AB} permits to stop the subsequent US phase at an appropriate time point.

The choices of W and w are correlated, their common scaling leading (along with that of other parameters) to an effectively equivalent scheme. For this reason, the choice W = 1 as a reference is recommended. With this choice, a large plateau width w enhances the sampling efficiency of the physical states but may slow down the convergence by reducing the number of transitions between these states. Large masses are less coupled to the conformational degrees of freedom, so that noise in the PMF is effectively integrated out. On the other hand, small masses sample more transitions, due to faster motion along θ.

The higher accuracy-to-efficiency ratio of λ-LEUS compared to TI observed in the sampling (see above) may also be accompanied by an improvement at the level of equilibration. In a TI calculation, the total equilibration time is proportional to the number of λ-points used, because each point must be equilibrated prior to sampling. In λ-LEUS, the analog of the equilibration of TI is the LE build-up phase. Considering that the biasing potential in the LE phase has to fill up a certain total area above the PMF, it makes sense to vary $\Delta\theta$ and k_{LE} is such a way that $\Delta\theta k_{LE}$ is constant. In this case, comparable results for a given t_{LE} should be obtained. However, although a smaller $\Delta\theta$ permits a more accurate representation of the negative of the PMF by the biasing potential, increasing k_{LE} to arbitrarily large values is not recommended, since it may promote "hill surfing".[3] Note also that special approaches could be considered to accelerate the build-up phase by using an initial approximate setting of the biasing potential (instead of starting from zero). For example, the memory may be initialized by using the result of an initial slow-growth sweep, a biasing potential optimized for a related alchemical transformation or one for the same transformation in a slightly different environment.

One possible shortcoming of the λ-LEUS scheme is that, contrary to TI, it does not present itself in a straightforwardly parallelizable format. Still, in principle, one could envision schemes in which the results of multiple parallel US sampling simulations are combined into a single PMF. Similarly, schemes can be envisioned in which the LE build-up phase is performed using multiple systems that simultaneously contribute to the construction of the biasing potential.

Finally, a major shortcoming common to all free-energy calculation methods is also shared by both TI and λ-LEUS, namely that accurate free-energy differences can only be obtained when the overall sampling time scale is significantly longer than the time scale associated with all relaxation processes orthogonal to the λ-variable.[32,75,76] Clearly, the test case considered here should be viewed as a relatively easy case from this point of view. The mutation of hydroquinone to benzene in water involves changes in local dipole and van der Waals cavity size in the solute. The orthogonal relaxation time for local electrostatic changes in water is on the order of the single-molecule rotational correlation time (about 1 ps) or the dielectric (Debye) relaxation time (about 10 ps).[77] The corresponding orthogonal relaxation time for local cavity-size changes in this solvent can be estimated[78] to be on the order of 200 ps or less, depending on the cavity size. This is obviously a very favorable situation for simulations on the nanosecond time scale.

The present article shows that λD can be used to provide, within the λ-LEUS scheme, a robust, general, and efficient protocol for the calculation of free-energy changes along two-states (one-dimensional) processes. Future work on the λ-LEUS scheme will follow two main lines. First toward the investigation of the performance of the method in the context of more challenging systems, where orthogonal relaxation occurs on longer time scales or involves barriers. In this case, one expects difficulties arising from Hamiltonian lag problems, which will have to be characterized and remedied. Second toward an extension of the method to multiple-state calculations, i.e. the obtension of relative free energies for a set of $N > 2$ compounds from a single simulation. Instead of introducing $N-1$ coupling parameters, this will be achieved using a pseudo-one-dimensional framework, where the evolution of the θ variable will successively map the N states (with associated plateaus) in a predefined order, before returning to the first state by periodicity.

A. BINNING ERROR

In the following, a rough estimate is provided for the error introduced when calculating the free-energy difference ΔG_{AB} by approximating the δ function in eq 4 (or eqs 9 or 13 when w = 0.0) by a finite binning function. Only the λ-dependent Hamiltonian $\tilde{\mathcal{H}}(\lambda)$ of the TI method is considered, but the tilde is omitted for simplicity. Using the Boltzmann probability

$$p(\lambda) = \mathcal{N} \exp[-\beta \mathcal{H}(\lambda)] \tag{A.1}$$

where $\mathcal{N}$ is a normalization constant, the free energy G_k of a state k at λ_k is given by

$$G_k = -\frac{1}{\beta} \ln \int_{\lambda_k-\varepsilon}^{\lambda_k+\varepsilon} p(\lambda) \mathrm{d}\lambda + C \tag{A.2}$$

$$= -\frac{1}{\beta} \ln \int_{\lambda_k-\varepsilon}^{\lambda_k+\varepsilon} \mathcal{N} \exp(-\beta \mathcal{H}(\lambda)) \mathrm{d}\lambda + C \tag{A.3}$$

dx.doi.org/10.1021/ct5002686 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

where ε is the half-width of the binning function, and C is an arbitrary constant. Using a first-order Taylor expansion for $\mathcal{H}(\lambda)$ around λ_k, one has

$$\mathcal{H}(\lambda) \approx \mathcal{H}(\lambda_k) + \mathcal{H}'_{\lambda_k}(\lambda - \lambda_k) \tag{A.4}$$

with the short-hand notation $\mathcal{H}'_{\lambda_k} = \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_{\lambda_k}$ The free energy can be calculated as

$$\begin{aligned}\Delta G_{AB} = -\frac{1}{\beta}\ln \cdot & \left[\frac{\mathcal{N}}{-\beta\mathcal{H}'_1}\exp\{-\beta(\mathcal{H}(1) + \mathcal{H}'_1(\lambda - 1))\}\right]_{1-\varepsilon}^{1} \\ & + \frac{1}{\beta}\ln\left[\frac{\mathcal{N}}{-\beta\mathcal{H}'_0}\exp\{-\beta(\mathcal{H}(0) + \mathcal{H}'_0\lambda)\}\right]_0^{\varepsilon} \\ & = \mathcal{H}(1) - \mathcal{H}(0) - \frac{1}{\beta}\ln\left(\frac{-\mathcal{H}'_0}{+\mathcal{H}'_1}\cdot\frac{1 - \exp[+\beta\mathcal{H}'_1\varepsilon]}{1 - \exp[-\beta\mathcal{H}'_0\varepsilon]}\right)\end{aligned} \tag{A.5}$$

Setting $\mathcal{H}'_0 = a$, $\mathcal{H}'_1 = b$ and $\beta\varepsilon = x$, the numerator and denominator of the logarithmic term in eq A.5 may be expanded in a first-order Taylor series as a function of x

$$\ln\left(\frac{-a}{b}\cdot\frac{1 - \exp[bx]}{1 - \exp[-ax]}\right) \approx \ln\left(\frac{-a}{b}\cdot \frac{1 - \left(1 + bx + \frac{b^2}{2}x^2\right)}{1 - \left(1 - ax - \frac{a^2}{2}x^2\right)}\right) \tag{A.6}$$

$$= \ln\left(\frac{1 + \frac{b}{2}x}{1 - \frac{a}{2}x)}\right) \approx \ln\left(\frac{\exp\left[\frac{b}{2}x\right]}{\exp\left[\frac{-a}{2}x\right]}\right) = \frac{a + b}{2}x \tag{A.7}$$

where the third step involves approximating the numerator and denominator by exponential functions. Applying this simplification to eq A.5 yields

$$\Delta G_{AB} \approx \mathcal{H}(1) - \mathcal{H}(0) - \frac{\mathcal{H}'_0 + \mathcal{H}'_1}{2}\varepsilon \tag{A.8}$$

In case of an infinitesimal half-width ($\varepsilon \to 0$), the last term in the right-hand side vanishes, and the correct free-energy is recovered. Otherwise, the error is proportional to the half-width ε and depends on the slope of $\mathcal{H}$ at $\lambda = 0$ and 1. Using a simple linear Hamiltonian $\mathcal{H}(\lambda) = \lambda\Delta\mathcal{H}$ as an example, where $\Delta\mathcal{H} = 200$ kJ mol^{-1}, the resulting error would be -2 kJ mol^{-1} and -0.2 kJ mol^{-1} for a half-width of $\varepsilon = 0.01$ or 0.001, respectively.

B. SECOND-ORDER SPLINES

Cubic splines are a powerful tool to interpolate smoothly between a set of N+1 discrete points (x_i,y_i), with $i = 0,...,N$ and $x_i < x_{i+1}$.[48] A cubic spline curve s is defined in terms of a piecewise union of N cubic functions s_i, each defined between x_i and x_{i+1}, and satisfying the following conditions

$$s_i(x_i) = y_i \tag{B.1}$$

$$s_i(x_i) = s_{i-1}(x_i) \tag{B.2}$$

$$s'_i(x_i) = s'_{i-1}(x_i) \tag{B.3}$$

$$s''_i(x_i) = s''_{i-1}(x_i) \tag{B.4}$$

where s' and s'' represent the first and second derivatives with respect to x, respectively.

This leads to a system of N+2 linear equation, which can be solved with linear complexity. In the case of a constant spacing $\Delta x = x_{i+1} - x_i$, the resulting solution can be expressed in terms of basis functions Φ, called second-order splines (cubic B-splines) and defined as

$$\Phi(t) = \begin{cases} \frac{(2 - |t|)^3}{4} & 1 \le |t| \le 2 \\ \frac{4 - 6|t|^2 + 3|t|^3}{4} & |t| \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{B.5}$$

The spline curve s can then be expressed as a weighted sum of the basis functions, i.e.

$$s = \sum_{i=0}^{N-1} w_i \Phi\left(\frac{x - x_i}{\Delta x}\right) \tag{B.6}$$

where w_i is the weight of the basis function i. The advantage of the second-order splines employed here over the first-order splines used in the standard LEUS implementation[45] (see Appendix A therein) is that, although the former can be used to represent exactly a straight horizontal line, only the latter can represent exactly a line of finite slope. In the presence of a slope, the first-order splines lead to "staircase" effects. These effects are not problematic when the slope is moderate (as typical for conformational free-energy calculations). They become, however, a major nuisance for high slopes as typical for alchemical free-energy calculations. The price to pay is that second-order splines are slightly less local, with a 5-point instead of a 3-point support.

AUTHOR INFORMATION

Corresponding Author

*Phone: 41 44 632 55 03. Fax: 41 44 632 10 39. E-mail: phil@igc.phys.chem.ethz.ch.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

N.S.B. thanks Niels Hansen for supplying the starting files for the hydroquinone to benzene mutation and Zhixiong Lin for insightful discussions. Financial support from the Swiss National Foundation (Grants 21-132739 and 21-138020) is gratefully acknowledged.

REFERENCES

(1) van Gunsteren, W. F.; Daura, X.; Mark, A. E. Computation of free energy. *Helv. Chim. Acta* **2002**, *85*, 3113.

(2) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. Basic ingredients of free energy calculations: A review. *J. Comput. Chem.* **2010**, *31*, 1569.

(3) Hansen, H. S.; Hünenberger, P. H. Ball-and-stick local elevation umbrella sampling: molecular simulations involving enhanced

sampling within conformational or alchemical subspaces of low internal dimensionalities, minimal irrelevant volume and problem-adapted geometries. *J. Chem. Theory Comput.* **2010**, *6*, 2622.
(4) Hünenberger, P. H.; Granwehr, J. K.; Aebischer, J.-N.; Ghoneim, N.; Haselbach, E.; van Gunsteren, W. F. Experimental and theoretical approaches to hydrogen-bonded diastereomeric interactions in a model complex. *J. Am. Chem. Soc.* **1997**, *119*, 7533.
(5) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420.
(6) Liu, H.; Mark, A. E.; van Gunsteren, W. F. Estimating the relative free energy of different molecular states with respect to a single reference state. *J. Phys. Chem.* **1996**, *100*, 9485.
(7) Lin, Z.; Timmerscheidt, T. A.; van Gunsteren, W. F. Using enveloping distribution sampling (EDS) to compute the free enthalpy difference between right- and left-handed helices of a β-peptide in solution. *J. Chem. Phys.* **2012**, *137*, 064108/1.
(8) Lin, Z.; van Gunsteren, W. F. Influence of variation of a side chain on the folding equilibrium of a β-peptide: limitations of one-step perturbatio. *J. Comput. Chem.* **2013**, *34*, 1899.
(9) Swendsen, R. H.; Wang, J.-S. Replica Monte Carlo simulation of spin-glasses. *Phys. Rev. Lett.* **1986**, *57*, 2607.
(10) Bennet, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245.
(11) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105/1.
(12) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187.
(13) Kirkwood, J. G. Quantum statistics of almost classical assemblies. *Phys. Rev.* **1933**, *44*, 31.
(14) Kirkwood, J. G. Quantum statistics of almost classical assemblies. *Phys. Rev.* **1934**, *45*, 116.
(15) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300.
(16) Chen, J.; Im, W.; Brooks, C. L., III Balancing solvation and intramolecular interactions: Toward a consistent generalized Born force field. *J. Am. Chem. Soc.* **2006**, *128*, 3728.
(17) 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*, 174112/1.
(18) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690.
(19) Jarzynski, C. Equilibrium free-energy differences from non-equilibrium measurements: A master-equation approach. *Phys. Rev. E* **1997**, *56*, 5018.
(20) Kong, X.; Brooks, C. L., III λ-dynamics: A new approach to free energy calculations. *J. Chem. Phys.* **1996**, *105*, 2414.
(21) Tidor, B. Simulated annealing on free energy surfaces by a combined molecular dynamics and Monte Carlo approach. *J. Phys. Chem.* **1993**, *97*, 1069.
(22) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *134*, 141.
(23) Comer, J.; Roux, B.; Chipot, C. Achieving ergodic sampling using replica-exchange free-energy calculations. *Mol. Sim.* **2014**, *40*, 218.
(24) Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384.
(25) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255.
(26) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511.
(27) Knight, J. L.; Brooks, C. L., III Multisite λ-dynamics for simulated structure-activity relationship studies. *J. Chem. Theory Comput.* **2011**, *7*, 2728.
(28) Donnini, S.; Tegeler, F.; Groenhof, G.; Grubmüller, H. Constant pH molecular dynamics in explicit solvent with λ-dynamics. *J. Chem. Theory Comput.* **2011**, *7*, 1962.
(29) Wu, P.; Hu, X.; Yang, W. λ-metadynamics approach to compute absolute solvation free energy. *J. Phys. Chem. Lett.* **2011**, *2*, 2099.
(30) Knight, J. L.; Brooks, C. L., III Applying efficient implicit nongeometric constraints in alchemical free energy simulations. *J. Comput. Chem.* **2011**, *32*, 3423.
(31) Donnini, S.; Tegeler, F.; Groenhof, G.; Grubmüller, H. Correction to constant pH molecular dynamics in explicit solvent with λ-dynamics. *J. Chem. Theory Comput.* **2013**, *9*, 3261.
(32) Zheng, L.; Yang, W. Practically efficient and robust free energy calculations: Double-integration orthogonal space tempering. *J. Chem. Theory Comput.* **2012**, *8*, 810.
(33) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562.
(34) Babin, V.; Roland, C.; Sagui, C. Adaptively biased molecular dynamics for free energy calculations. *J. Chem. Phys.* **2008**, *128*, 134101/1.
(35) Paine, G. H.; Scheraga, H. A. Prediction of the native conformation of a polypeptide by a statistical-mechanical procedure. I. Backbone structure of enkephalin. *Biopolymers* **1985**, *24*, 1391.
(36) Hansen, H. S.; Hünenberger, P. H. Using the local elevation method to construct optimized umbrella sampling potentials: Calculation of the relative free energies and interconversion barriers of glucopyranose ring conformers in water. *J. Comput. Chem.* **2010**, *31*, 1.
(37) Huber, T.; Torda, A. E.; van Gunsteren, W. F. Local elevation: A method for improving the searching properties of molecular dynamics simulation. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695.
(38) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169.
(39) Darve, E.; Rodriguez-Gomez, D.; Pohorille, A. Adaptive biasing force method for scalar and vector free energy calculation. *J. Chem. Phys.* **2008**, *128*, 144120/1.
(40) Rosso, L.; Mináry, P.; Zhu, Z.; Tuckerman, M. E. On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. *J. Chem. Phys.* **2002**, *116*, 4389.
(41) Rosso, L.; Tuckerman, M. E. An adiabatic molecular dynamics method for the calculation of free energy profiles. *Mol. Simul.* **2002**, *28*, 91.
(42) Abrams, J. B.; Rosso, L.; Tuckerman, M. E. Efficient and precise solvation free energies via alchemical adiabatic molecular dynamics. *J. Chem. Phys.* **2006**, *125*, 074115/1.
(43) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 5th ed.; W.H. Freeman & Co. Ltd.: 2007.
(44) Guo, Z.; Durkin, J.; Fischermann, T.; Ingram, R.; Prongay, A.; Zhang, R.; Madison, V. Application of the λ-dynamics method to evaluate the relative binding free energies of inhibitors to HCV protease. *J. Med. Chem.* **2003**, *46*, 5360.
(45) Hansen, H. S.; Daura, X.; Hünenberger, P. H. Enhanced conformational sampling in molecular dynamics simulations of solvated peptides: fragment-based local elevation umbrella sampling. *J. Chem. Theory Comput.* **2010**, *6*, 2598.
(46) Bruckner, S.; Boresch, S. Efficiency of alchemical free energy simulations. I. A practical comparison of the exponential formula, thermodynamic integration, and Bennett's acceptance ratio method. *J. Comput. Chem.* **2010**, *32*, 1303.
(47) Bruckner, S.; Boresch, S. Efficiency of alchemical free energy simulations. II. Improvements for thermodynamic integration. *J. Comput. Chem.* **2010**, *32*, 1320.
(48) Habermann, C.; Kindermann, F. Multidimensional spline interpolation: Theory and applications. *Comput. Econ.* **2007**, *30*, 153.
(49) Hünenberger, P. H. Thermostat algorithms for molecular dynamics simulations. *Adv. Polym. Sci.* **2005**, *173*, 105.
(50) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695.
(51) Maragliano, L.; Vanden-Eijnden, E. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.* **2006**, *426*, 168.
(52) Kunz, A.-P. E.; Liu, H.; van Gunsteren, W. F. Enhanced sampling of particular degrees of freedom in molecular systems based

on adiabatic decoupling and temperature or force scaling. *J. Chem. Phys.* **2011**, *135*, 104106/1.
(53) Harvey, S. C.; Tan, R. K.-Z.; Cheatham, T. E., III The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comput. Chem.* **1998**, *19*, 726.
(54) Chiu, S.-W.; Clark, M.; Subramaniam, S.; Jakobsson, E. Collective motion artifacts arising in long-duration molecular dynamics simulations. *J. Comput. Chem.* **2000**, *21*, 121.
(55) Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. The "hot-solvent/cold-solute" problem revisited. *J. Chem. Theory Comput.* **2008**, *4*, 1293.
(56) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635.
(57) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; di Nola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684.
(58) Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **1905**, *322*, 549.
(59) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Comput. Phys. Commun.* **2012**, *183*, 890.
(60) Kunz, A.-P. E.; Allison, J. R.; Geerke, D. P.; Horta, B. A. C.; Hünenberger, P. H.; Riniker, S.; Schmid, N.; van Gunsteren, W. F. New functionalities in the GROMOS biomolecular simulation software. *J. Comput. Chem.* **2012**, *33*, 340.
(61) van Gunsteren, W. F. *The GROMOS software for biomolecular simulation*. Available at http://www.gromos.net (accessed 05/05/2011).
(62) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656.
(63) 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, The Netherlands, 1981; pp 331−342.
(64) Beutler, T. C.; Mark, A. E.; van Schaik, R.; Gerber, P. R.; van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, *222*, 529.
(65) Hockney, R. W. The potential calculation and some applications. *Methods Comput. Phys.* **1970**, *9*, 136.
(66) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327.
(67) Berendsen, H. J. C.; van Gunsteren, W. F.; Zwinderman, H. R. J.; Geurtsen, R. G. Simulations of proteins in water. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 269.
(68) Barker, J. A.; Watts, R. O. Monte Carlo studies of the dielectric properties of water-like models. *Mol. Phys.* **1973**, *26*, 789.
(69) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451.
(70) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* **2001**, *115*, 1125.
(71) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90.
(72) Gorard, S. Revisiting a 90-year old debate: The advantages of the mean absolute deviation. *Br. J. Educ. Stud.* **2005**, *53*, 417.
(73) Perić-Hassler, L.; Hansen, H. S.; Baron, R.; Hünenberger, P. H. Conformational properties of glucose-based disaccharides using molecular dynamics simulations with local elevation umbrella sampling. *Carbohydr. Res.* **2010**, *345*, 1781.
(74) Ehmler, H.; Heesemann, K.; Rätzke, K.; Fauppel, F. Mass dependence of diffusion in a supercooled metallic melt. *Phys. Rev. Lett.* **1998**, *80*, 4919.
(75) Zheng, L.; Chen, M.; Yang, W. Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20227.
(76) Zheng, L.; Yang, W. Essential energy space random walks to accelerate molecular dynamics simulations: Convergence improvements via an adaptive-length self-healing strategy. *J. Chem. Phys.* **2008**, *129*, 014105/1.
(77) Buchner, R.; Barthel, J.; Stauber, J. The dielectric relaxation of water between 0°C and 35°C. *Chem. Phys. Lett.* **1999**, *306*, 57.
(78) Beutler, T. C.; Béguelin, D. R.; van Gunsteren, W. F. Free energy of cavity formation in solvent: Computational, methodological, and physical aspects. *J. Chem. Phys.* **1995**, *102*, 3787.