JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

A Toolkit for the Analysis of Free-Energy Perturbation Calculations

Peng Liu,[†] François Dehez,[‡] Wensheng Cai,[†] and Christophe Chipot*[,‡,§]

[†]College of Chemistry, Nankai University, Tianjin, 300071, People's Republic of China

[‡]Équipe de dynamique des assemblages membranaires, UMR 7565, Nancy Université, BP 239, 54506 Vandoeuvre-lès-nancy cedex, France

[§]Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana−Champaign, 405 North Mathews, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: As computational power inexorably continues to grow, harnessing the capabilities of novel processing units and architectures, free-energy calculations are progressively brought to the level of routine modeling tools for exploring the thermodynamic properties of increasingly larger molecular assemblies. Within these premises, free-energy perturbation (FEP) arguably represents the most commonly chosen approach for tackling transformations of a chemical nature between thermodynamic states. To augment the accuracy, the precision, and, hence, the reliability of these calculations, a number of good practices have been established. In the present contribution, a new toolkit, coined ParseFEP, is proposed to follow these prescriptions in a user-friendly environment. Written as a Tcl plugin, it allows FEP calculations carried out using the popular molecular-dynamics package NAMD to be analyzed seamlessly within the visualization platform VMD. The potential of the toolkit is probed through a number of illustrative examples, which demonstrate cogently how pathological cases, often related to convergence issues, can be detected and remedied by means of a pictorial approach.

INTRODUCTION

To understand and capture the thermodynamic detail of most chemical processes, in particular, those of biological relevance, an accurate determination of the underlying free-energy change is generally required.[1−5] Such is the case, for instance, of the permeation of chemical compounds, peptides, and ions through the cell membrane, or the association of small molecules like hormones and neurotransmitters to proteins. These chemical processes, which are of paramount importance for the communication of the cell with its environment, cannot be predicted reliably without prior knowledge of the associated free energy.

Arguably enough, methods targeted at the computation of free-energy differences can be clustered into four classes: (i) perturbation theory[6,7] and free-energy perturbation (FEP) calculations,[8] (ii) determination of gradients and integration thereof, i.e., thermodynamic integration-like approaches,[9] (iii) histograms,[10,11] and (iv) nonequilibrium work (NEW).[12,13] Though NEW shares common theoretical foundations with FEP, the latter can be viewed as a limiting case of the former.[12]

Perturbation theory, in which the Hamiltonian of the target state is described as the sum of the reference Hamiltonian and a perturbation term,[7] is arguably the oldest route toward the estimation of free-energy differences. Originally devised to understand how the physical properties of hard-core molecules are altered by the introduction of a rudimentary, attractive potential[8]—many years before computer simulations became a standard tool of the theoretician—FEP calculations have greatly benefited over the past decades from significant advances on the computational front,[14,15] which have helped expand their field of applications to continuously larger molecular assemblies and increasingly more complex transformations.[3,5]

While harnessing the power of the computer architectures developed in recent years has undeniably contributed to promoting FEP to the level of other predictive numerical approaches, it may appear somewhat paradoxical that these simulations are still performed predominantly in academic environments. To a certain extent, the difficulty of macromolecular, multipurpose force fields to reproduce in a systematic fashion and with an acceptable accuracy[4,16−23] key thermodynamic observables of biological relevance, like binding affinities,[24−36] explains why such dogmas like FEP calculations being about as unreliable as they are prohibitively expensive have pervaded beyond the walls of nonacademic research and development facilities.

The adequacy of molecular mechanical potential energy functions to address the challenging enterprise of measuring reliably free-energy differences, however, only represents one aspect of an intricate puzzle.[16] Armed with an hypothetically very accurate multipurpose force field, the theoretician is still left with the human tasks of setting up and analyzing the simulations.[37−39] These tasks have proven to be sufficiently time-consuming to preclude the blind, possibly automated use of free-energy calculations in nonacademic environments, in particular in the context of lead identification and high-throughput virtual screening. While a significant effort has been invested in visualization platforms to facilitate the design and setup of the simulations by means of user-friendly graphical user interfaces, appraisal of convergence as well as management

Received: March 23, 2012
Published: July 18, 2012

   dx.doi.org/10.1021/ct300242f | J. Chem. Theory Comput. 2012, 8, 2606−2616

of errors appear to constitute the main stumbling stone that has prevented FEP from achieving the status of a routine method shared by other modeling tools, like molecular docking.

It may be contended that post-treatment of a large amount of data to assess whether or not the ensemble average has converged,[40] measure the statistical and the systematic errors associated with the free-energy difference, and possibly go beyond the simple FEP estimator represents the most arduous, human-involved component of free-energy calculations. Although the theory underlying the a posteriori analysis of these calculations was developed many years ago, its application as a series of precepts or good practices, in particular in the framework of complex transformations of chemical or biological relevance, has remained hitherto scarce.

It has been recently reasserted as a compendium of such good practices[41] that free-energy estimates can be significantly improved by combining the data accrued from bidirectional transformations, that error bars ought to accompany systematically published free-energy differences, and that a strategy for the stratification of the reaction path be sought to reduce the variance associated with the computed estimates. Central to these prescriptions is a graphical consistency check of several observables to assess the reliability of the computed free-energy estimate.

In the present contribution, a new toolkit, coined ParseFEP and aimed at facilitating the task of the theoretician in following the aforementioned precepts, is proposed as a plugin within the user-friendly environment of the popular visualization platform VMD.[42] This toolkit has been conceived to analyze and visualize the results of FEP calculations carried out with the molecular-dynamics (MD) package NAMD,[43] albeit its modular design guarantees a straightforward adaptation to other similar, fashionable MD codes.

After summarizing the theoretical underpinnings that form the scaffold of the plugin, the latter is utilized in a series of examples that convincingly illustrate on the basis of a pictorial approach how pathological simulations can be detected and how the variance of free-energy calculations can be measured and reduced. The subsequent section describes an extension of perturbation theory to the estimation of internal energies, or enthalpies, and entropies. Next, an alternate route to assessing the convergence of FEP calculations is proposed by modeling probability distributions using Gram−Charlier expansions. The article closes with an overview of the main features of the toolkit and how the latter ought to be utilized to improve the reliability of free-energy calculations, alongside suggestions of possible directions to enhance the plugin.

THEORETICAL BACKDROP AND IMPLEMENTATION

In the present section, the theoretical underpinnings of FEP calculations are outlined, alongside the good practices recommended to perform and analyze them, and which form the methodological scaffold of the toolkit described herein. For all intents and purposes, this theoretical background is presented in the context of sampling the Boltzmann measure in the canonical ensemble.

The Free-Energy Perturbation Framework. The quantity of interest that we endeavor to estimate is the free-energy difference between the reference state, 0, and the target state, 1, of some molecular assembly:

$$\Delta A = -\frac{1}{\beta}\ln\frac{Q_1}{Q_0} \tag{1}$$

where Q_0 and Q_1 are the partition functions representative of these states, and $\beta = 1/k_B T$, with k_B being the Boltzmann constant and T, the temperature.

At a given value of a general extent parameter, λ, introduced to transform reversibly between the reference and the target states, the partition function can be related to the Boltzmann distribution that describes the equilibrium state of the molecular assembly:

$$P(\mathbf{x}, \mathbf{p}_x, \lambda) = \frac{\exp[-\beta\mathcal{H}(\mathbf{x}, \mathbf{p}_x, \lambda)]}{\int d\mathbf{x}\, d\mathbf{p}_x \exp[-\beta\mathcal{H}(\mathbf{x}, \mathbf{p}_x, \lambda)]} \equiv \frac{1}{Q_\lambda}\exp[-\beta\mathcal{H}(\mathbf{x}, \mathbf{p}_x, \lambda)] \tag{2}$$

where $\mathcal{H}(\mathbf{x},\mathbf{p}_x,\lambda)$ is the Hamiltonian characterizing the intermediate state λ; $\mathbf{x}$ represents the Cartesian coordinates of the molecular assembly and $\mathbf{p}_x$, its conjugated momenta.

In the FEP framework, the free-energy difference between the reference, 0, and the target, 1, states is expressed as an equilibrium average:

$$\exp(-\beta\Delta A) = \langle\exp(-\beta\Delta U)\rangle_0 \tag{3}$$

Remembering that the kinetic term of the partition function cancels out if the mass remains unchanged in the course of the transformation, it will be henceforth omitted. Here, ΔU stands for the difference in potential energies, $U(\mathbf{x},\lambda = 1) - U(\mathbf{x},\lambda = 0)$. $\langle\cdots\rangle_0$ is an equilibrium configurational average over the microstates representative of the reference state. The present transformation is often referred to as forward transformation.

Bidirectional Transformations. Brute-force application of eq 3 is often referred to as unidirectional transformation. A homologous, backward transformation can be imagined, wherein sampling is performed from the target state, 1, in lieu of the reference state, 0. Under these premises, the FEP identity now writes

$$\exp(+\beta\Delta A) = \langle\exp(+\beta\Delta U)\rangle_1 \tag{4}$$

Among the suggested good practices for FEP calculations, it is strongly recommended to estimate free-energy differences using a bidirectional approach, which incorporates samples from both forward and backward distributions

$$\frac{P_0(\Delta U)}{P_1(\Delta U)} = \exp[-\beta(\Delta A - \Delta U)] \tag{5}$$

Here, $P_0(\Delta U)$ is the probability distribution characterizing the reference state and $P_1(\Delta U)$, that of the target state. It is convenient to restate eq 3 on the basis of the probability distribution for the forward equilibrium ensemble:[44−47]

$$\exp(-\beta\Delta A) = \int_{-\infty}^{+\infty} d\Delta U\, P_0(\Delta U)\exp(-\beta\Delta U) \tag{6}$$

FEP, like most free-energy methods in general, is burdened by convergence issues.[40] Abrupt changes in the free energy, in particular in the early stage of sample accrual, are a manifestation that the average is dominated by rare events.[40,41] The latter are rooted in the overlap of the equilibrium distributions $P_0(\Delta U)$ and $P_1(\Delta U)$. Convergence of the calculation based upon eq 3 imposes that the predominant contribution of the average arise from microstates of the

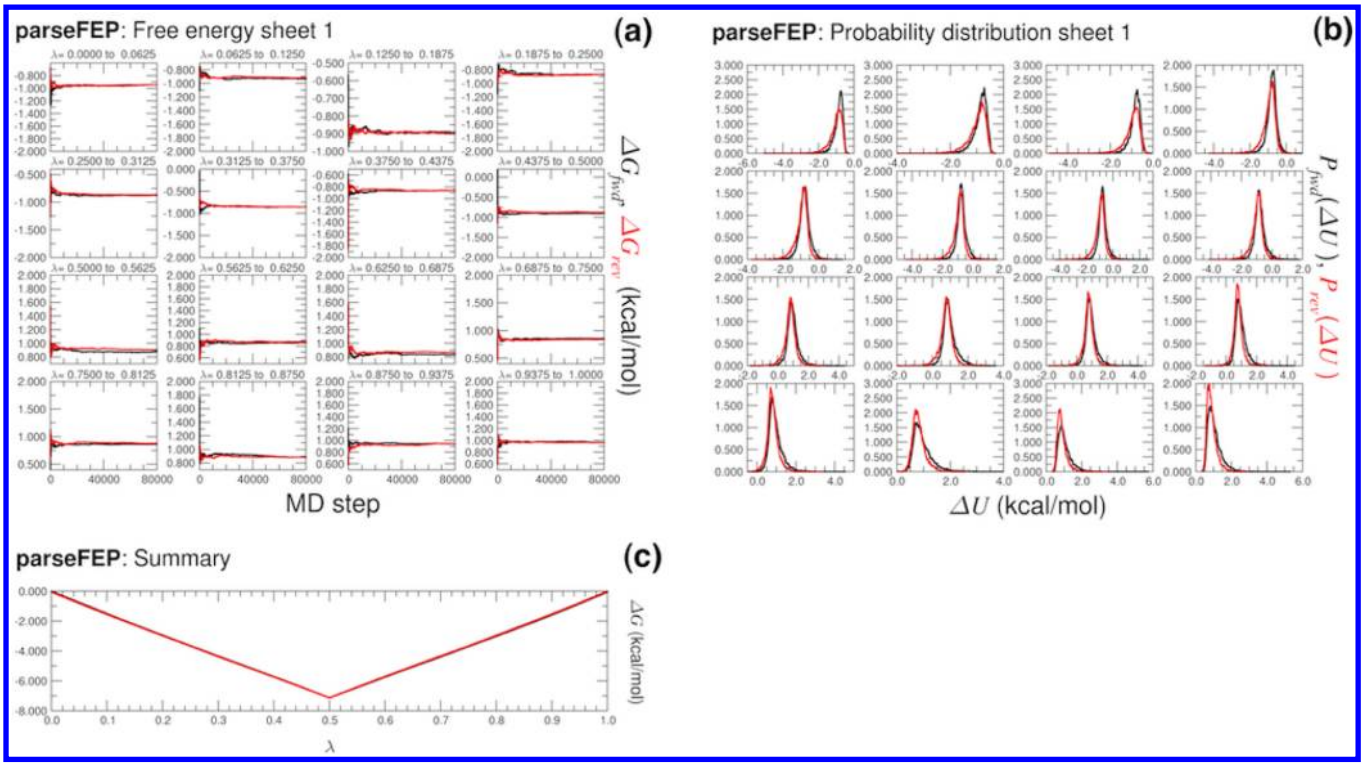

Figure 1. Typical graphical output of ParseFEP obtained in the case of the reversible zero-sum ethane-to-ethane transformation carried out, employing a 16-window stratification strategy. Time-evolution of the free energy (a), histograms of the probability distributions (b), and overall free-energy change (c) for the forward (black solid line) and backward (light solid line) transformations. In practice, a pseudo-propane molecule is utilized, wherein one methyl group is annihilated as the other one is created, resulting in a force-field independent zero net free-energy change.[48] Computational details of the simulations are gathered in the Supporting Information (SI).

reference ensemble that are also low-energy states of the target ensemble. When the overlap between $P_0(\Delta U)$ and $P_1(\Delta U)$ is limited, the important microstates are only marginally sampled, thereby impeding convergence.

Aside from the mere time evolution of the free energy, the toolkit proposed herein allows the user to assess the degree of overlap between the reference and the target states of the transformation (see Figure 1). Pathological scenarios, wherein the overlap is mediocre, can be readily detected from a pictorial representation of the underlying one-dimensional probability distributions. Under such circumstances, stratification of the transformation ought to be optimized.

Among the established good practices,[41] it has been emphasized that it is always preferable to determine a free-energy difference by combining the data sets from both the reference and the target equilibrium ensembles, that is, by combining the results of the forward and the backward transformations, which can be done a posteriori.

A solution that satisfies the problem of bidirectional sampling can be found by considering an intermediate state, i, between the reference and the target states, such that

$$\exp(-\beta\Delta A) = \frac{Q_1}{Q_0} = \frac{Q_i/Q_0}{Q_i/Q_1} = \frac{\langle \exp[-\beta(U_i - U_0)]\rangle_0}{\langle \exp[-\beta(U_i - U_1)]\rangle_1} \tag{7}$$

An obvious choice for i would obey $U_i = (U_0 + U_1)/2$, often referred to as "half umbrella" or "simple overlap sampling" (SOS).[49] In spirit, this scheme is closely related to the seminal idea of "double-wide sampling". The corresponding free-energy estimator then writes

$$\exp(-\beta\hat{\Delta}A^{\mathrm{SOS}}) = \frac{\langle \exp(-\beta\Delta U/2)\rangle_0}{\langle \exp(+\beta\Delta U/2)\rangle_1} \tag{8}$$

The definition of i, however, bears some arbitrariness. Its choice can be optimized through the introduction of a weighting function, $w = \exp\{-\beta[U_i - (U_0 + U_1)/2]\}$, so that the free-energy estimator now writes

$$\exp(-\beta\hat{\Delta}A) = \frac{\langle w \exp(-\beta\Delta U/2\rangle_0}{\langle w \exp(+\beta\Delta U/2\rangle_1} \tag{9}$$

A proper route to optimizing the intermediate state consists in weighting the contributions from the equilibrium ensembles 0 and 1 that minimize the variance, σ^2, of $\hat{\Delta}A$. In a nutshell, this constitutes the basic idea of the Bennett Acceptance Ratio (BAR) method.[10] In the first-order approximation, the optimal choice for w is a hyperbolic secant function:

$$w = \mathrm{sech}[+\beta(\Delta U - c)/2] = \frac{2}{\exp[+\beta(\Delta U - c)/2] + \exp[-\beta(\Delta U - c)/2]} \tag{10}$$

where constant $c = \Delta\hat{A} + (1/\beta)\ln(N_1/N_0)$.

N_0 and N_1 are the numbers of configurations sampled from the equilibrium reference and target states. In practice, since the BAR analysis is performed as a post-treatment, N_0 and N_1 are identical, although they need not be. Put together, the BAR estimator of the free-energy difference writes

$$\exp(\beta\Delta\hat{A}^{\mathrm{BAR}}) = \frac{\langle f[-\beta(\Delta U - c)]\rangle_1}{\langle f[+\beta(\Delta U - c)]\rangle_0}\exp(+\beta c) \tag{11}$$

where $f(x) = 1/[1 + \exp(x)]$ is the so-called Fermi function. Constant c ought to be solved in a self-consistent fashion. In the ParseFEP toolkit, the FEP estimator is utilized as the initial guess of the iterative process.

Error Associated to Free-Energy Perturbation Calculations. Central to the good practices proposed for performing, analyzing, and publishing free-energy calculations is the provision of an error estimate associated with the computed free-energy difference.[41] As it would be inconceivable to report an experimental measurement bereft of an error bar, free-energy estimates ought to be published with the error attached to it. One of the reasons that explain why such an error bar is seldom provided is rooted in the complex nature of this quantity, which includes a statistical component, namely, the variance of the free-energy calculation—i.e., its precision—as well as a systematic one—i.e., its accuracy.[50] The latter is more difficult to apprehend, as it embraces many contributions of different origins, chief among which is the finite length of the simulation. The parameters of the algorithms utilized to generate the equilibrium ensemble alongside with the potential energy function constitute other sources of systematic error.

In ParseFEP, the focus is on the statistical component of the error.[40] The variance for the free-energy estimate is provided for unidirectional transformations:

$$\sigma_{\Delta A}^2 = \frac{1}{N_0\beta^2}\frac{\langle \exp(-2\beta\Delta U)\rangle_0}{\langle \exp(-\beta\Delta U)\rangle_0^2} - \frac{1}{N_0\beta^2} \tag{12}$$

as well as for bidirectional transformations combined by means of the SOS estimator:

$$\sigma_{\Delta A}^{2\ \mathrm{SOS}} = \frac{1}{N_0\beta^2}\frac{\sigma^2_{\langle \exp(-\beta\Delta U/2)\rangle_0}}{\langle \exp(-\beta\Delta U/2)\rangle_0^2} + \frac{1}{N_1\beta^2}\frac{\sigma^2_{\langle \exp(+\beta\Delta U/2)\rangle_1}}{\langle \exp(+\beta\Delta U/2)\rangle_1^2} \tag{13}$$

or the BAR estimator:[41,51,52]

$$\sigma_{\Delta A}^{2\ \mathrm{BAR}} = \frac{1}{N_0\beta^2}\left[\frac{\langle f^2(x)\rangle_0}{\langle f(x)\rangle_0^2} - 1\right] + \frac{1}{N_1\beta^2}\left[\frac{\langle f^2(-x)\rangle_1}{\langle f(-x)\rangle_1^2} - 1\right] \tag{14}$$

where $x = \beta(\Delta U - c)$.

Under most circumstances, configurations of both the reference and the target states are expected to be correlated.[40] N_0 and N_1 ought to be divided by the corresponding sampling ratios, $(1 + 2\tau_0)$ and $(1 + 2\tau_1)$, where τ_0 and τ_1 are the correlation lengths for the forward and backward transformations, respectively. The sampling ratio can be viewed as the minimum number of samples found between two independent observables. Correlation lengths are estimated in an automated fashion based upon the group renormalization algorithm of Flyvbjerg and Petersen.[53]

Accurate determination of the systematic contribution to the overall error associated with the free-energy change is eminently challenging, as this component is rooted in a number of sources of different nature, which are usually difficult to apprehend.[40,41,54−56] Should a bidirectional simulation be run, the hysteresis between forward and backward transformations is often utilized as a proxy for the systematic error. On the other hand, the theoretician is tempted in many instances to compare directly the computational estimate of the free-energy change with its experimental counterpart. The discrepancy between the two values arises from several origins, chief among which are possible inaccuracies in the potential energy function,[21] should ergodic sampling be assumed. Periodicity-induced artifacts constitute yet another possible cause of error, especially in the case of ionic solvation.[57] Aside from force-field and simulation-protocol issues, the primary source of systematic errors is the finite length of the free-energy calculation. In particular, small statistical errors, inferred from estimators 12 and 14 do not necessarily guarantee equally small systematic errors. To understand this apparent nonsequitur, solvation of a flexible molecule will serve as an illustration. Here, the chemical compound can adopt two distinct conformations, which are separated by a free-energy barrier of sufficient height to never be crossed during a simulation of limited length. From a statistical standpoint, the free-energy calculation may seem to have converged, as suggested by the sampling of all important configurations of the solvent surrounding the molecule. Yet, the inaccessibility of the alternate conformation—a rare event from the perspective of Boltzmann sampling—implies an overall incomplete sampling due to the finite length of the simulation and, hence, an appreciable systematic error. Pictorial approaches have proven valuable to detect large systematic errors reflected by poor overlap in the equilibrium distributions $P_0(\Delta U)$ and $P_1(\Delta U)$, in spite of noticeably small variances. Integration of those regions where the histograms do not overlap has been shown to constitute a route for estimating quantitatively the systematic error associated with the free-energy change,[54,55] albeit more elaborate schemes have been devised to address the concept of accuracy[50] in free-energy calculations.[41,58,59]

Perturbative Estimation of Internal Energies and Entropies. The FEP identity can be extended to the estimation of internal energy, or enthalpy, and entropy differences.[5,60] Remembering that these quantities are connected through $\Delta A = \Delta u - T\Delta S$, it follows that the internal energy writes

$$\Delta u = \frac{\langle U_1 \exp(-\beta\Delta U)\rangle_0}{\langle \exp(-\beta\Delta U)\rangle_0} - \langle U_0\rangle_0 \tag{15}$$

from whence the entropy can be readily inferred:

$$\Delta S = \frac{1}{T}\left(\frac{\langle U_1 \exp(-\beta\Delta U)\rangle_0}{\langle \exp(-\beta\Delta U)\rangle_0} - \langle U_0\rangle_0\right) + k_B \ln\langle \exp(-\beta\Delta U)\rangle_0 \tag{16}$$

Modeling the Underlying Probability Distributions. As has been discussed previously, appraising the convergence of FEP calculations can be achieved by measuring the degree of overlap between the underlying probability distributions, $P_0(\Delta U)$ and $P_1(\Delta U)$, characterizing the equilibrium ensembles for the forward and the backward transformations. In unidirectional simulations, measuring the shift between the peak of $P_0(\Delta U)$ and the product $P_0(\Delta U)\ \exp(-\beta\Delta U)$ constitutes a reasonable indicator of convergence. This can be understood by considering second-order perturbation theory, for which the free energy writes $\Delta A = \langle \Delta U\rangle_0 - \beta\sigma^2/2$, where σ^2 is the variance of $P_0(\Delta U)$, which, in this particular instance, is a Gaussian distribution. It follows from the latter that $P_0(\Delta U)\ \exp(-\beta\Delta U)$ is shifted from $P_0(\Delta U)$ toward lower energies by $\beta\sigma^2$.

Under most circumstances, ΔU does not obey a normal law, to the extent that $P_0(\Delta U)$ can markedly depart from a Gaussian distribution. This deviation from the ideal Gaussian case may be expressed in the form of a Gram−Charlier expansion:[61]

$$P_0(\Delta U) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp(-\Delta U^2/2\sigma^2) \times h(\Delta U) \quad (17)$$

where

$$h(x) = \sum_n \frac{c_n}{n!} H_n(x) \quad (18)$$

Here, $H_n(x)$ is the Hermite polynomial of rank n and c_n, the corresponding coefficient. ParseFEP features an interpolation of the underlying probability distributions and the evaluation of integral 6 at the desired order of the Gram−Charlier expansion. Derivation of the c_n coefficients and their relationship with the free-energy change is detailed in the SI.

IMPLEMENTATION DETAIL

ParseFEP has been implemented as a Tcl plugin. It is directly integrated in the extensions/analysis menu of VMD, from whence it opens as a graphical user interface. It requires as an input the output file of an FEP calculation performed with NAMD,[43] i.e., the alchOutFile file—or the alchOutFile files generated in a forward and a backward transformation in the event of a bidirectional free-energy calculation was carried out. After loading the relevant files, users can perform the following analysis:

(a) Free-energy estimates: For unidirectional FEP calculation, ParseFEP computes the free-energy difference and provides an estimate of the statistical error. For bidirectional FEP calculations, the results of the forward and the backward simulations can be combined using either the SOS estimator or the BAR estimator. For each approach, an estimation of the statistical error is supplied.

(b) Pictorial representation: For unidirectional FEP calculations, ParseFEP plots the underlying probability distribution, the Boltzmann weight, the product thereof, and the evolution of the free energy as a function of the number of molecular-dynamics steps. For bidirectional FEP calculations, ParseFEP provides a graphical representation of the probability distribution characterizing the forward and the backward transformations, alongside the evolution of the free energy. The resulting graphic sheets are drawn by XMGrace and displayed by ImageMagick on a Unix platform and saved as PNG files for further usage. On Windows platforms, the Multiplot plugin of VMD is utilized to generate the graphs.

(c) Gram−Charlier interpolation: The probability distribution for each window or stratum can be approximated by means of a Gram−Charlier expansion. Convergence of this expansion can be seen as an additional diagnostic tool to assess the convergence of FEP calculations.

(d) Enthalpies and entropies: The free-energy perturbation formalism can be extended to determine enthalpy, or internal energy, and entropy differences. Caution should be exercised because the latter estimates are inherently less accurate than that of the free energy.

APPLICATIONS

Optimizing the Stratification Strategy. Probably the simplest applicable strategy for reducing the systematic error is stratification[62]—what is the optimum number of stages that guarantees a converged free-energy change? The pictorial representation of ParseFEP can be utilized to detect pathological cases, wherein the probability distributions for the forward and backward transformations, $P_0(\Delta U)$ and $P_1(\Delta U)$, overlap poorly. This symptom is generally encountered when the stratification strategy is either suboptimal or simply inadequate.

It ought to be emphasized again that the integrand in eq 6 is proportional to the probability distribution of ΔU sampled from state 1 through:

$$\exp(-\beta\Delta U)\, P_0(\Delta U) = \exp(-\beta\Delta A)\, P_1(\Delta U) \quad (19)$$

It follows from the latter that ΔA can be estimated with an appreciable reliability under the condition that microstates of the target equilibrium ensemble have been sampled from the reference ensemble.

To probe the influence of the number of intermediate states on the free-energy estimates, a series of reversible annihilations of 4-methylimidazole, i.e., the side chain of ε-L-histidine, in water using between 2 and 128 intermediate states was performed. The overall length of each simulation was kept constant, equal to 12.8 ns. The free energies are gathered in Figure 2. As the number of strata increases from 2 to 128, the calculated free energy appears to reach an asymptotic limit, noteworthily close to the experimental value—though agreement with experimental results is not central to this study. The rudimentary estimate of the systematic error based on the hysteresis between forward and backward transformations decreases as intermediate states are added.

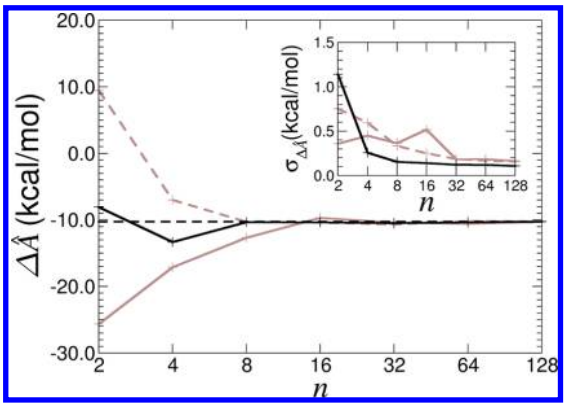

Figure 2. Hydration free energy of 4-methylimidazole—i.e., the side chain of ε-L-histidine, obtained using different stratification strategies, but keeping the overall number of steps constant. n denotes the number of intermediate states. The free energy for the forward, annihilation, and the backward, creation, transformations are shown as light, solid and dashed lines, respectively. The BAR estimate is shown as a dark, solid line. The experimental hydration free energy of −10.3 kcal/mol[63] is displayed as a dashed horizontal line. Inset: Statistical error associated with the hydration free energy. Color coding is identical to that above. Computational details of the simulations are gathered in the SI.

As the number of stages increases from 2 to 128, the calculated free energy seems to attain a plateau, while the statistical error decreases gradually. The present results suggest

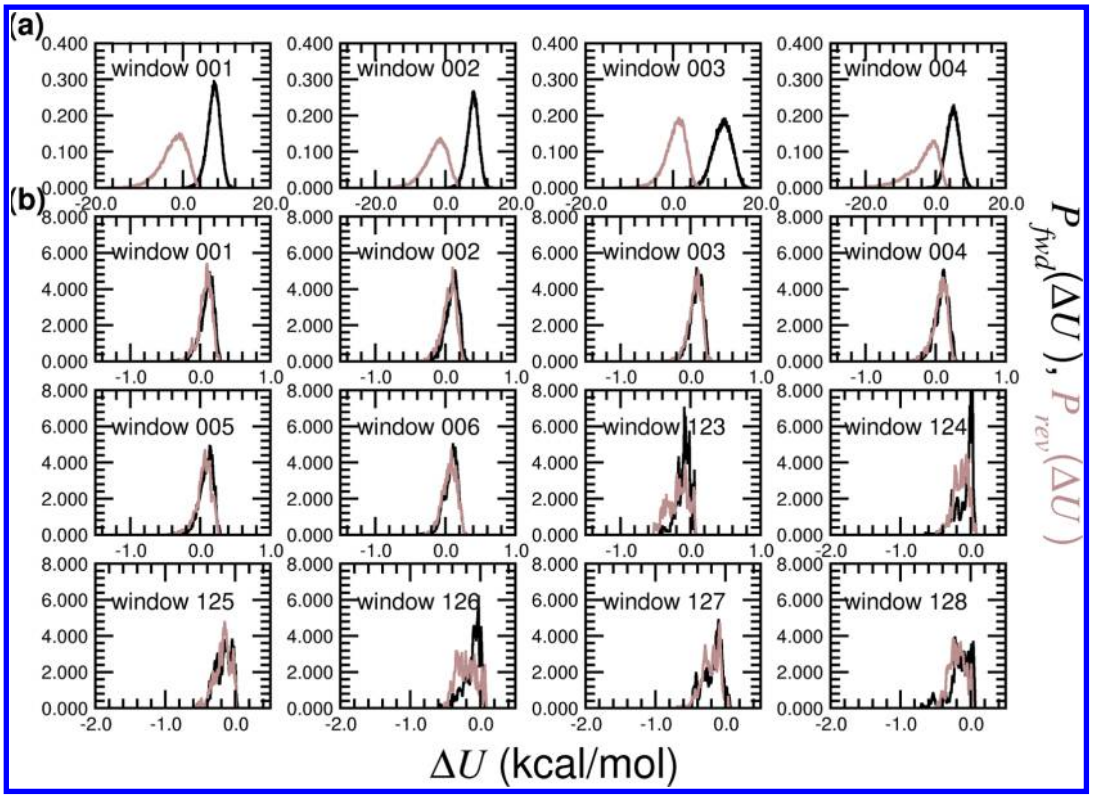

Figure 3. Probability distribution function underlying the immersion in bulk water of 4-methylimidazole. Shown in this figure are the distributions characterizing the forward, $P_0(\Delta U)$ (dark lines) and the backward, $P_1(\Delta U)$ (light lines) transformations for (a) all four intermediate states in a four-window stratification strategy and (b) the first and the last six intermediate states in a 128-window stratification strategy. Note that as the number of strata increases from 4 to 128, the probability distributions become significantly narrower at the same time as the overlap between $P_0(\Delta U)$ and $P_1(\Delta U)$ improves significantly.

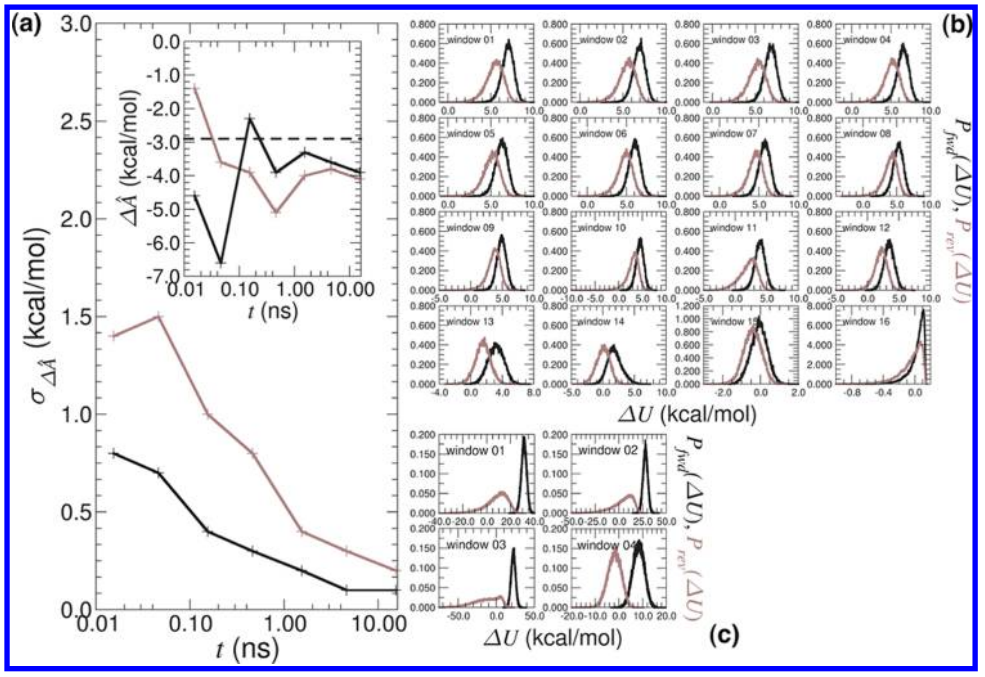

Figure 4. Binding of a potassium ion to the 18-crown-6 crown ether. The free energy of binding is determined bidirectionally by means of a reversible annihilation of the cation in its bound and free, unbound states, employing two stratification strategies consisting of four and 16 windows, and increasing progressively the sampling in each one of the latter. Statistical errors and binding free energies determined using the BAR estimator (a). Results for the four- and 16-window strategy are shown as light and black lines. Histograms of the underlying probability distributions obtained for the longest sampling, namely, 15.36 ns, with a 16- (b) and four-window (c) scheme. Computational details of the simulations are gathered in the SI.

that augmenting the number of strata causes not only the statistical error but also the systematic error to decrease, as may be observed in Figure 3. Here, the underlying probability distributions determined from a four-window and a 128-window free-energy calculation are compared.

For the four-window stratification strategy, the histograms are noticeably wide and suggestive of an appreciable error on the free-energy estimate. Furthermore, the overlap between probability distributions $P_0(\Delta U)$ and $P_1(\Delta U)$ characterizing the forward and the backward transformations is rather poor in each stratum. In sharp contrast, the histograms for the 128-window stratification strategy have a much smaller variance, which, in turn, improves the overlap between $P_0(\Delta U)$ and $P_1(\Delta U)$ significantly.

Put together, the above paradigmatic example illustrates the necessity to stratify the reaction path when undertaking alchemical transformations and to perform the free-energy calculation bidirectionally as a consistency check. Monitoring the underlying probability distributions allows pathological, poorly converged simulations to be detected and subsequently corrected by optimizing the number of intermediate stages.

Reducing the Variance of Free-Energy Calculations. In general, free-energy calculations are burdened by convergence issues,[40] which imposes for efficacy reasons that the computational cost of the simulation be optimized. Here, the balance between the number of intermediate states forming the stratification strategy and the sampling per stratum is discussed in light of the prototypical free-energy calculation of a potassium ion binding crown ether 18-crown-6. Two staging schemes were utilized, consisting of four and 16 intermediate λ states. In each case, seven sampling strategies were considered, with total simulation times ranging from 0.01536 to 15.36 ns. To prevent the cation from drifting away from the complex in the course of its annihilation, a geometrical restraint was enforced, guaranteeing that its position coincide with that of the center of mass of 18-crown-6. The contribution of this restraint can be evaluated analytically as $-1/\beta \ln(c^0\Delta v)$, where c^0 is the standard concentration and Δv, the volume element of the crown ether explored by the potassium ion. In the present instance, the free-energy contribution due to the positional restraint amounts to +4.8 kcal/mol.

As illustrated in Figure 4, irrespective of the number of strata, increasing the sampling reduces the variance associated with the computed binding free energy, which, to a large extent, is a statement of the obvious, considering, for instance in eq 14, the dependence of ${\sigma_{\Delta A}}^2$ on N_0, the number of samples. Of greater interest is the effect of the stratification strategy on the measured free-energy change. In the event where only four intermediate states are used to annihilate reversibly the potassium ion, either in its unbound or bound state, the statistical error remains appreciably larger than with 16 strata, even for the longest simulation, namely, 15.36 ns. A glimpse at the probability distributions of Figure 4 is sufficient to appreciate why a four-window stratification strategy is inadequate to tackle the proposed alchemical transformation. The histograms are widespread, suggestive that when multiplied by the relevant Boltzmann weight, the integrand of 6 will be heavily shifted toward low-energy regions, which have evidently not been sampled in the course of the free-energy calculation. In sharp contrast, the 16-window stratification scheme yields appreciably smaller variances and, hence, more precise free-energy estimates.

In order to compare the results of the FEP calculations with the experimental value, a distinction ought to be made at this stage between the concepts of precision and accuracy.[50] The variance associated with the BAR estimator only represents the statistical component of the overall error, which encompasses several other important contributions. Of particular concern is the adequacy of the macromolecular potential energy function to handle the problem of interest—e.g., accounting for induction phenomena, albeit this question goes beyond the scope of the present article. As has been discussed previously, force-field inaccuracies constitute only one aspect of the discrepant free energies measured computationally and experimentally. The pictorial representation of $P_0(\Delta U)$ and $P_1(\Delta U)$ in Figure 4 reveals that even for appreciably long simulation times, using 16 stages, the overlap is still suboptimal, which is a reflection of the finite-length nature of the simulation. A cheap and simple approach to measure the associated systematic error is undoubtedly the hysteresis between forward and backward calculations. On the basis of the histograms of the probability distributions, a somewhat more sophisticated route can, however, be considered, wherein the relative inaccuracy for the forward transformation of a bidirectional free-energy calculation is determined from the area of the section of $P_1(\Delta U)$ that does not overlap with $P_0(\Delta U)$, and symmetrically, for the backward transformation, from the area of the section of $P_0(\Delta U)$ that does not overlap with $P_1(\Delta U)$.[55]

Measuring Free-Energy Contributions. ParseFEP allows the user to estimate enthalpies, or internal energies, and entropies based on an extension of the FEP identity 15 and 16. Contrary to free-energy calculations, which focus on the interaction of the perturbed part of the molecular assembly with its environment, estimation of internal energies and entropies involves the averaging of all interactions, including that of the environment. The evaluation of Δu and ΔS is, thus, burdened by large numbers, rendering their convergence difficult. It may be argued, however, that since the two terms of eq 15 are obtained from the same simulation—as opposed to the mere difference $\langle U_1\rangle_1 - \langle U_0\rangle_0$—partial error cancellation may reduce large inaccuracies.

The prototypical case of ethanol in water was considered to illustrate the capacity of ParseFEP to provide an estimate of Δu and ΔS, alongside the free-energy difference associated with a given alchemical transformation.

The hydration free energy of ethanol and the corresponding enthalpy and entropy contributions were computed from a double annihilation of the solute in water and in the gas phase. The reaction pathway connecting the initial and the final states of the annihilation and creation transformations was stratified in a series of 16 strata of equal width. Independent trajectories of 16, 32, and 64 ns were produced in both directions and used to evaluate ΔA_{hydr}, Δu_{hydr}, and ΔS_{hydr} after various sampling strategies, ranging from 8 to 64 ns at a step of 8 ns.

The evolution of the thermodynamic properties as a function of the sampling time is reported in Figure 5. The hydration free energy was already converged after 8 ns of sampling. The estimated value -4.4 ± 0.0 kcal/mol, determined with the BAR method, is consistent with the experimental reference of -5.1 kcal/mol[64] and with theoretical estimates published in the literature.[21,65,66]

Shirts et al.,[21] in order to increase the accuracy of the computed hydration free energies, used a long-range correction (LRC) to account for solute−solvent van der Waals

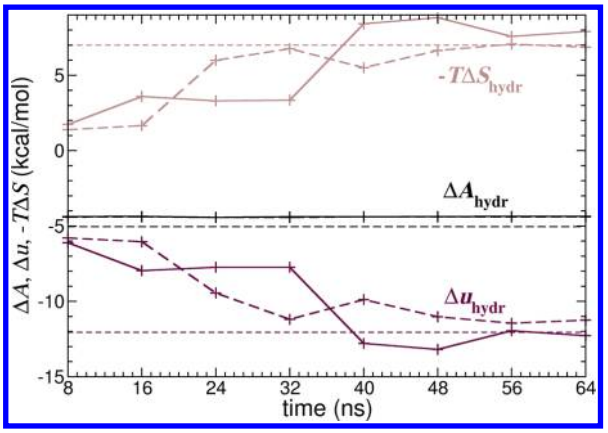

Figure 5. Evolution of the free energy (black) and its enthalpic (dark) and entropic (light) contribution associated with the hydration of ethanol as a function of the simulation time. Solid lines and long-dashed lines denote the variation of the thermodynamic quantities computed from the annihilation and the creation transformations, respectively. Horizontal, dashed lines correspond to the values determined experimentally at 298.15 K.[64] Computational details of the simulations are gathered in the SI.

interactions extended beyond the switching distance (r_{switch}). The correction to the hydration free energy of ethanol, with r_{switch} = 8.0 Å, was estimated to −0.3 kcal/mol, with an overall ΔA_{hydr} = −4.2 kcal/mol. Employing, respectively, a polarizable force field based on the classical Drude oscillator and the charge equilibration formalism, Anisimov et al.[65] and Zhong and Patel[66] predicted a more favorable hydration, namely, −5.3 kcal/mol (LRC = −0.4 kcal/mol) and −4.9 kcal/mol (LRC = −0.5 kcal/mol), respectively.

The discrepancy of about −0.7 kcal/mol between experiment and theory resides to a large extent in the force field description. CHARMM nonbonded parameters account for the mean polarization of the condensed phase and therefore are not appropriated to estimate precisely the gas-phase contribution to the hydration free energy. When seeking for accuracy, the energy of distortion accounting for the change of polarity of the surroundings can be estimated based on atomic charges optimized from high-level ab initio calculation in the gas phase.[67]

Determination of both enthalpy and entropy of hydration necessitate much longer trajectories, and convergence is witnessed after approximately 40 ns of sampling. The estimated values for Δu_{hydr} = −11.7 kcal/mol and $-T\Delta S_{hydr}$ = +7.3 kcal/mol compare nicely with the corresponding quantities, −12.1 kcal/mol and +7.0 kcal/mol, obtained experimentally. Interestingly enough, convergence of Δu_{hydr} and ΔS_{hydr} occur over a much shorter time scale for the creation process, compared to annihilation. This trend, similar to what is observed for ΔA_{hydr}, is a consequence of a larger dissipation energy produced along the creation pathway, the number of time steps required to achieve convergence of the free energy being linked to dissipation.

Even though the proposed strategy allows the experimental enthalpy and entropy contributions to the hydration free energy of ethanol to be reproduced, slow convergence issues make it a costly enterprise and, hence, a rather inefficient one. Other avenues can be pursued to access enthalpies and entropies. One alternative relies upon the following thermodynamic relationships:

$$\Delta u = -\left(\frac{\partial \beta \Delta A}{\partial \beta}\right)_{N,V} \quad \text{and} \quad \Delta S = -\left(\frac{\partial \Delta A}{\partial T}\right)_{N,V} \tag{20}$$

which require the estimation of ΔA at both $T - \Delta T$ and $T + \Delta T$.[68,69] An extensive discussion of possible routes for computing free-energy contributions is provided by Chipot and Pohorille.[5]

Modeling the Underlying Probability Distributions. Modeling probability distributions at a given order of a Gram−Charlier expansion and evaluating the integral of eq 6 arguably constitute a possible route to probe the convergence of FEP calculations. The underlying idea here is that poor sampling strategies, in particular from the standpoint of stratification, generally result in probability distributions endowed with a substantial variance. As has been discussed previously, in the ideal, Gaussian case, when multiplied by the corresponding Boltzmann weight, such distributions yield an integrand of expression 6, the peak of which can be heavily shifted toward low energies, hence, falling in regions where no statistical information has been accrued. Under such circumstances, $P_0(\Delta U)$ cannot be modeled within any given order of expansion of eq 17. In fact, although the free-energy change determined on the basis of the latter expression might appear to have leveled off at a certain, sufficiently high order of the Gram−Charlier series, it does not coincide with the value provided by the FEP estimator of eq 3.

As illustrated in Figure 6, in the prototypical case of 4-methylimidazole hydration, i.e., the side chain of ε- l-histidine, a stratification strategy limited to only two intermediate states

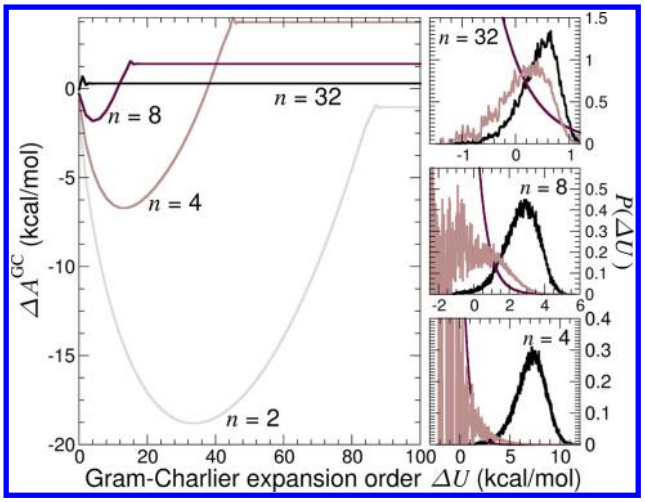

Figure 6. Convergence of the free energy (left panel) obtained using a Gram−Charlier expansion to describe the probability distribution (right panel) characterizing the first stage, or window, of a stratified annihilation of 4-methylimidazole, i.e., the side chain of ε-L-histidine), in bulk water. Here, n denotes the number of strata utilized to simulate the overall transformation, from which the negative of the hydration free energy may be inferred. Bidirectional free-energy calculations relying on a four-, eight-, and 32-window stratification strategy yield a BAR estimator of the net annihilation free energy, i.e., the negative of the hydration free energy, equal to 9.8 ± 0.3, 10.4 ± 0.2, and 10.5 ± 0.1 kcal/mol, respectively, to be compared to an experimental value of −10.27 kcal/mol. Shown in the right panel are the probability distribution, $P_0(\Delta U)$ (black line), the Boltzmann weight (dark line), and the product thereof, i.e., the integrand of eq 6 (light line), for three different stratification strategies. Computational details of the simulations are gathered in the SI.

yields for the first window a free-energy difference that varies significantly with the order of the Gram−Charlier expansion, before eventually reaching a plateau at −1.4 kcal/mol and departing conspicuously from the FEP estimate, equal to +12.7 kcal/mol. Doubling the number of windows leads to an improved agreement between the two estimators for the first stratum of the annihilation of the solute in bulk water. The Gram−Charlier expansion converges faster, and the free-energy change levels off at a value of +3.7 kcal/mol, virtually identical to that supplied by the FEP estimator. Yet, not too unexpectedly, the integrand of eq 6 overlaps only marginally with $P_0(\Delta U)$, from which it is markedly shifted away. Increasing the number of windows to eight confirms the agreement between the alternate estimators, the free-energy change inferred from the Gram−Charlier expansion reaching a plateau at a perceptibly lower order of the latter. On the other hand, the integrand of eq 6 falls in a range of potential energy values that has not been explored in the course of the simulation, thus, casting doubt on the accuracy of the measured free-energy difference. Resorting to a 32-window stratification strategy, it can be observed that convergence of the free energy determined using expansion of eq 17 is almost instantaneous. The quickly converging free-energy difference at a low order of the Gram−Charlier expansion is mirrored in the overlapping probability distribution, $P_0(\Delta U)$, and integrand of eq 6, suggestive of an appropriate sampling strategy.

Put together, while interpolation of probability distributions by means of Gram−Charlier expansions might appear to constitute only a minute improvement over a simple pictorial approach, in which the degree of overlap between $P_0(\Delta U)$ and $P_0(\Delta U)\ \exp(-\beta\Delta U)$, or between $P_0(\Delta U)$ and $P_1(\Delta U)$ in bidirectional transformations, are monitored, it can also be seen as a complementary option to assess the convergence of FEP calculations. Whereas ParseFEP currently only offers the possibility to compute free-energy differences on the basis of eq 17, up to a given order of the expansion, the theoretical scaffold is readily available to extend the use of Gram−Charlier expansions to extrapolate probability distributions.

CONCLUSION AND PERSPECTIVE

When conducting free-energy calculations, in particular those of a perturbative nature, following a number of precepts or "good practices" is strongly recommended.[41] Among the latter, monitoring the overlap between the underlying probability distributions arguably constitutes the key prescription. Overlap of the configurational ensembles—specifically that low-energy configurations of the target state are also low-energy configurations of the reference state—is a *sine qua non* condition for the reliable estimation of free-energy differences. Moreover, combining the statistical information accrued in forward and backward transformations improves the reliability and the efficiency of free-energy calculations. Last, stratification[62] reduces the systematic error of the simulation, albeit not necessarily the statistical error. Finding the proper balance between the number of stages and the optimum amount of sampling per stratum remains a challenging task, for which no a priori criterion applies.

In this contribution, a toolkit coined ParseFEP is proposed as a Tcl plugin within the popular molecular graphics platform VMD[42] to facilitate the application of the suggested "good practices" in free-energy calculations. This plugin provides the end user with a number of features designed for the analysis of FEP calculations carried out with the MD program NAMD,[43] which includes (i) the display of the histograms of the probability distributions inferred from uni- and bidirectional simulations, i.e., $P_0(\Delta U)$ and $P_1(\Delta U)$, (ii) SOS and BAR estimators to combine the results of bidirectional simulations, (iii) the estimate of entropies and enthalpies, and (iv) the modeling of probability distributions by means of Gram−Charlier expansions. The implementation of the latter features has been probed in a variety of molecular systems. As can be cogently shown in the case of 4-methylimidazole, i.e., the side chain of ε-L-histidine, hydration, a pictorial approach, wherein the histograms of the probability distributions are plotted, can be utilized advantageously to detect pathological simulations, in particular those employing a suboptimal stratification strategy. In the instance of a potassium ion binding crown ether 18-crown-6, the balance between the number of intermediate states in the stratification strategy and the amount of sampling per stage is discussed in light of both the results obtained with the BAR estimator and the visualization of the underlying probability distributions. Hydration of ethanol provides a convincing example that, in a number of favorable circumstances, perturbation theory can be extended to the estimation of enthalpies and entropies.[60] Far from asserting that this route constitutes the method of choice, we suggest that, provided appropriate sampling, it can be utilized at no additional cost as an a posteriori treatment of FEP calculations. Last, returning to the hydration of 4-methylimidazole, we show that Gram−Charlier expansions utilized as model probability distributions represent an interesting alternate option to probe the convergence of free-energy calculations.

Although ParseFEP responds in large measure to the needs expressed by the community of theoreticians interested in the routine computation of free-energy differences by facilitating the analysis and interpretation of FEP calculations, ample room is still available for improvement of the toolkit. Among the avenues for further development, one that holds much promise is the generalization of the BAR estimator to combine data collected from multiple simulations, i.e., the so-called mBAR algorithm.[70] Another possible direction is an extension of the use of Gram−Charlier expansions to probe convergence of free-energy calculations not only by interpolating probability distributions but also by extrapolating them.[71] Along this line, exploration of alternate expansions, like the asymptotic Edgeworth expansion, might be desirable on account of the alleged convergence issues of the Gram−Charlier series.[72] Finally, while the end user often resorts to the hysteresis of a bidirectional simulation to estimate the systematic error associated with the free-energy change, implementation of a more rigorous framework for accessing the accuracy of the calculation would be desirable. A number of these ideas are currently examined and tested in ParseFEP by us and by others.

ASSOCIATED CONTENT

Supporting Information

Derivation of the formalism based on Gram−Charlier expansions and computational details of the simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: chipot@ks.uiuc.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are indebted to the Centre Informatique National de l'Enseignement Supérieur in Montpellier, France, for provision of a generous amount of computer time. The Agence Nationale de la Recherche is gratefully acknowledged (grant MEGAS).

■ REFERENCES

(1) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395−2417.
(2) Simonson, T. Free energy calculations. In *Computational Biochemistry and Biophysics*; Becker, O. M., MacKerell, A. D., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker Inc.: New York, 2001; pp 169−197.
(3) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430−437.
(4) Chipot, C.; Pearlman, D. A. *Mol. Sim.* **2002**, *28*, 1−12.
(5) *Free Energy Calculations. Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer Verlag: New York, 2007.
(6) Born, M. *Z. Phys.* **1920**, *1*, 45−48.
(7) Landau, L. D. *Statistical Physics*; The Clarendon Press: Oxford, U. K., 1938.
(8) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(9) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300−313.
(10) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245−268.
(11) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635−2638.
(12) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(13) Jarzynski, C. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **1997**, *56*, 5018−5035.
(14) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050−3054.
(15) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564−568.
(16) van Gunsteren, W.; Beutler, T.; Fraternali, F.; King, P.; Mark, A.; Smith, P. Computation of free energy in practice: Choice of approximations and accuracy limiting factors. In *Computer Simulation of Biomolecular Systems*; van Gunsteren, W., Weiner, P., Wilkinson, A., Eds.; Escom Science Publishers: Leiden, The Netherlands, 1993; pp 315−348.
(17) Chipot, C.; Millot, C.; Maigret, B.; Kollman, P. A. *J. Phys. Chem.* **1994**, *98*, 11362−11372.
(18) Ding, Y. B.; Bernardo, D. N.; Kroghjespersen, K.; Levy, R. M. *J. Phys. Chem.* **1995**, *99*, 11575−11583.
(19) Morgantini, P. Y.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 6057−6063.
(20) Rick, S. W.; Berne, B. J. *J. Phys. Chem. B* **1997**, *101*, 10488−10493.
(21) Shirts, M.; Pitera, J.; Swope, W.; Pande, V. *J. Chem. Phys.* **2003**, *119*, 5740−5761.
(22) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(23) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(24) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742−3746.
(25) van Gunsteren, W. F. Methods for calculation of free energies and binding constants: Successes and problems. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; Van Gunsteren, W. F., Weiner, P. K., Eds.; Escom: Leiden, The Netherlands, 1989; pp 27−59.
(26) Miyamoto, S.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402−8406.
(27) Åqvist, J. *J. Comput. Chem.* **1996**, *17*, 1587−1597.
(28) Åqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. *Acc. Chem. Res.* **2002**, *35*, 358−365.
(29) Archontis, G.; Simonson, T.; Karplus, M. *J. Mol. Biol.* **2001**, *306*, 307−327.
(30) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535−9551.
(31) Oostenbrink, C.; van Gunsteren, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6750−6754.
(32) Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6825−6830.
(33) Wang, J.; Deng, Y.; Roux, B. *Biophys. J.* **2006**, *91*, 2798−2814.
(34) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118−1134.
(35) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234−2246.
(36) Jiang, W.; Roux, B. *J. Chem. Theory Comput.* **2010**, *6*, 2559−2565.
(37) Pearlman, D. A.; Charifson, P. S. *J. Med. Chem.* **2001**, *44*, 3417−3423.
(38) Chipot, C.; Rozanska, X.; Dixit, S. B. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 765−770.
(39) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.
(40) Lelièvre, T.; Stoltz, G.; Rousset, M. *Free Energy Computations: A Mathematical Perspective*; Imperial College Press: London, 2010.
(41) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(42) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(43) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, L.; Kalé, R. D.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(44) McDonald, I. R.; Singer, K. *J. Chem. Phys.* **1967**, *47*, 4766−4772.
(45) McDonald, I. R.; Singer, K. *Discuss. Faraday Soc.* **1967**, *43*, 40−49.
(46) Valleau, J. P.; Torrie, G. M. A Guide to Monte Carlo for Statistical Mechanics: Byways. In *Modern Theoretical Chemistry*; Berne, B. J., Ed.; Plenum: New York, 1977; pp 169−194.
(47) Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M. *J. Chem. Phys.* **1986**, *85*, 6720.
(48) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532−4545.
(49) Lu, N.; Kofke, D. A.; Woolf, T. B. *J. Comput. Chem.* **2004**, *25*, 28−39.
(50) Kofke, D.; Cummings, P. *Fluid Phase Equilib.* **1998**, *150*, 41−49.
(51) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.
(52) Hahn, A. M.; Then, H. *Phys. Rev. E: Stat. Nonlinear, Soft Matter Phys.* **2009**, *80*, 031111.
(53) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461−466.
(54) Lu, N.; Kofke, D. A. *J. Chem. Phys.* **2001**, *114*, 7303−7312.
(55) Lu, N.; Kofke, D. A. *J. Chem. Phys.* **2001**, *115*, 6866−6875.
(56) Lu, N.; Woolf, T. B. Understanding and improving free energy calculations in molecular simulations: Error analysis and reduction methods. In *Free Energy Calculations. Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer Verlag: Berlin, 2007.
(57) Hünenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856−1872.
(58) Zuckerman, D. M.; Woolf, T. B. *Phys. Rev. Lett.* **2002**, *89*, 180602.
(59) Zuckerman, D. M.; Woolf, T. B. *J. Stat. Phys.* **2004**, *114*, 1303−1323.
(60) Wan, S.; Stote, R. H.; Karplus, M. *J. Chem. Phys.* **2004**, *121*, 9539−9548.
(61) Cramèr, H. *Mathematical Methods of Statistics*; Princeton University Press: Princeton, NJ, 1957.
(62) Valleau, J. P.; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457−5462.

(63) Wolfenden, R.; Andersson, L.; Cullins, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849−855.
(64) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016−2027.
(65) Anisimov, V. M.; Vorobyov, I. V.; Roux, B.; Mackerell, A. D. *J. Chem. Theory Comput.* **2007**, *3*, 1927−1946.
(66) Zhong, Y.; Patel, S. *J. Phys. Chem. B* **2010**, *114*, 11076−11092.
(67) Chipot, C. *J. Comput. Chem.* **2003**, *24*, 409−415.
(68) Kubo, M. M.; Gallicchio, E.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 10527−10534.
(69) Levy, R. M.; Gallicchio, E. *Annu. Rev. Phys. Chem.* **1998**, *49*, 531−567.
(70) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(71) Pohorille, A.; Darve, E. A Bayesian approach to calculating free energies in chemical and biological systems. *Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Aip Conference Proceedings*, Melville, NY, 2006; pp 23−30.
(72) Blinnikov, S.; Moessner, R. *Astron. Astrophys. Suppl. Ser.* **1998**, *130*, 193−205.