A Generalized Free Energy Perturbation Theory Accounting for End States with Differing Configuration Space Volume

R. Thomas Ullmann* and G. Matthias Ullmann

Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstrasse 30, BGI,
95447 Bayreuth, Germany

We present a generalized free energy perturbation theory that is inspired by Monte Carlo techniques and based on a microstate description of a transformation between two states of a physical system. It is shown that the present free energy perturbation theory stated by the Zwanzig equation follows as a special case of our theory. Our method uses a stochastic mapping of the end states that associates a given microstate from one ensemble with a microstate from the adjacent ensemble according to a probability distribution. In contrast, previous free energy perturbation methods use a static, deterministic mapping that associates fixed pairs of microstates from the two ensembles. The advantages of our approach are that end states of differing configuration space volume can be treated easily also in the case of discrete configuration spaces and that the method does not require the potentially cumbersome search for an optimal deterministic mapping. The application of our theory is illustrated by some example problems. We discuss practical applications for which our findings could be relevant and point out perspectives for further development of the free energy perturbation theory.

1. Introduction

Free energy is the central quantity of thermodynamics and thus of outstanding importance in physics, chemistry, and biology. Free energy differences provide the driving force for any physical process and ultimately of life. Free energy calculations have contributed to our understanding of the reactivity and function of complex systems encountered in physics, chemistry, and biology and especially in the interesting problems emerging at their interface. These calculations supplement and guide experiment by modeling details not easily accessible to experiment. Continuous improvements have been achieved over the years both in theory and in practical application of free energy calculation methods in such diverse fields as biochemistry, nanotechnology, materials research, pharmacy, fluid phase theory, bioenergetics, and biotechnology.[1–8]

A problem not worked out so far occurs in free energy perturbation (FEP) when dealing with transformations between two systems for which no unequivocal one-to-one mapping of initial and final microstates exists. The purpose of the present paper is to derive a universally applicable free energy perturbation theory based on a microscopic description of the system.

Motivating Problem. Let us consider a simple system consisting of three microstates with equal energy $E = 0$. The Helmholtz free energy difference between an ensemble consisting only of the first microstate and an ensemble consisting of the remaining microstates two and three is $-\beta^{-1} \ln 2$, where $\beta^{-1} = k_{\mathrm{B}}T$, as can easily be seen from elementary statistical mechanics. Analogous problems from our own work occur, for example, in the calculation of pK_a values and protonation free energies for titratable protein residues or protein cofactors.[9] Prototypical realizations of the example system are carboxylic acids, where the first microstate corresponds to the deprotonated state while the remaining two microstates correspond to the two

possible positions of the proton at either of the two oxygen atoms of the carboxy group in the protonated state. Such problems can in general not be solved by simple symmetry corrections, because the microstates might not be equivalent from the onset or because they become nonequivalent due to differing interactions in an anisotropic environment. Equivalent protonation forms of amino acid residues, for example, lose their equivalency in the anisotropic protein environment.

Suppose we want to calculate this free energy difference from a FEP simulation. Traditional FEP is based on the following equation derived by Zwanzig in 1954[10]

$$\exp(-\beta \Delta F_{0 \to 1}) = \int \rho_0 \exp[-\beta(E_1(\mathbf{q}_1) - E_0(\mathbf{q}_0))] \, \mathrm{d}\mathbf{q}_0$$
$$\text{with } \mathbf{q}_1 = \mathscr{M}(\mathbf{q}_0), \mathscr{M}: \mathbf{q}_0 \to \mathbf{q}_1 \text{ bijective and } |\mathrm{d}\mathbf{q}_0| = |\mathrm{d}\mathbf{q}_1|$$
$$(1)$$

This equation tells us to sample configurations $\mathbf{q}_0$ of the initial ensemble according to its equilibrium probability distribution ρ_0 and to collect state energy difference samples $E_1(\mathbf{q}_1) - E_0(\mathbf{q}_0)$. Each of these energy differences describes a forced transition from a sampled configuration $\mathrm{d}\mathbf{q}_0$ to one from the final ensemble $\mathrm{d}\mathbf{q}_1$. This transition or FEP move can be considered as perturbation of the system.[11] In choosing the term FEP move we seek to carry on and extend an existing interpretation of the perturbation as a special kind of Monte Carlo trial move[11] or step.[12] As we will show in the Theory section, the Zwanzig equation implicitly assumes the existence of an unequivocal one-to-one mapping $\mathscr{M}$ of each configuration of the initial ensemble to exactly one configuration of the final ensemble and vice versa. More precisely, the mapping is bijective and preserves the volume of the differential volume elements mapped onto each other. Due to this one-to-one mapping, the configurations of each pair $\mathbf{q}_0, \mathbf{q}_1$ can be perceived as one configuration $\mathbf{q}$ subject to different energy functions E_0, E_1 in the initial and the final

* To whom correspondence should be addressed. E-mail: thomas.ullmann@uni-bayreuth.de.
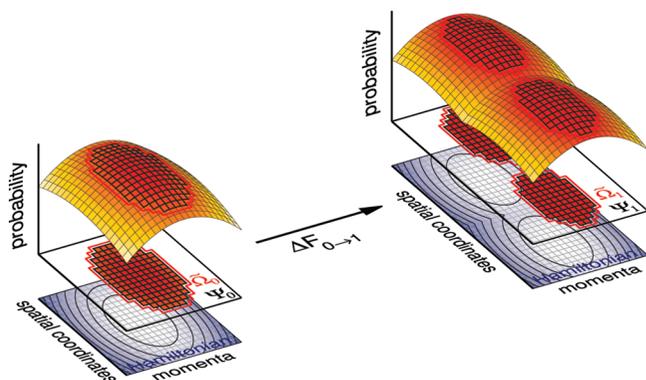
Figure 1. Schematic view of the problem to solve. Sought is the free energy change $\Delta F_{0\rightarrow 1}$ for the transformation of a system from state 0 to state 1 (end states). The plots for the two end states show the occupation probability distributions and the underlying energy surfaces in phase space. The transformation describes the traveling of the system between two regions of phase space with differing extent and energy landscape. The phase space is spanned by the spatial coordinates and the momenta of the particles, depicted here exemplary for a single particle in one spatial dimension. The contour plots at the bottom show projections of the energy surfaces into the phase space plane (constant contour interval, darker shade means higher energy). Overlaid are plots of the probability surfaces and their projections into the phase space plane. In the projections, only the significantly populated regions of phase space are shown. Each square on the surfaces represents an infinitesimal volume element in phase space. The integral over all volume elements of the phase space regardless of their occupation probability gives the *total phase space volume* Ψ. The phase space elements with a significant occupation probability are indicated by their dark shade and bold frame. These elements constitute the thermally accessible regions of phase space enclosed in the red contours. The extent of these regions in phase space is termed *effective phase space volume* and denoted by the symbol $\tilde{\Omega}$. The example for the plot is taken from the classical particle in the box models with one or two harmonic energy wells presented in the example section.

state. Within this picture, the transition between the initial and final state can be interpreted as a perturbation or switching of the energy function.[10,11] Accordingly, eq 1 is written in most cases without indices attached to the configuration **q** or with the index 0 for both end states

$$\exp(-\beta\Delta F_{0\rightarrow 1}) = \int \rho_0 \exp[-\beta(E_1(\mathbf{q}) - E_0(\mathbf{q}))]\, \mathrm{d}\mathbf{q} \tag{2}$$

This perception does, however, not necessarily correspond to the real physical situation modeled in a free energy simulation. In fact, the configurations $\mathbf{q}_0, \mathbf{q}_1$ and the systems described by them can be quite different.[12–14] The required one-to-one mapping is impossible to define in our example system, since the number of possible configurations differs between initial and final ensemble. An apparent solution seems to be found in a random selection of one of the two possible final configurations. It is easily seen, however, that this approach will also not lead to the correct result, because the calculated energy difference for each of the two possible FEP moves is zero. Thus, already with our simple model system we were not able to accomplish the calculation of the free energy difference in a FEP simulation with the Zwanzig equation.

Generalized Statement of the Problem and Our Proposed Solution. A generalized formulation of the problem which this article attempts to solve is schematically depicted in Figure 1. The aim is the calculation of the free energy difference between two states of a system corresponding to two regions of phase

space whose extent can differ. In most cases, the initial and final states of a transformation will be defined on the basis of the system configuration, while the range of the momenta is not restricted. For a classical system all momentum-dependent terms cancel in this case from free energy differences and can thus be neglected in their calculation (see Supporting Information). Consequently, the calculation can be carried out entirely in the configuration space of the system which is spanned by the coordinate axes of all particles residing in the system.

We will derive a novel formalism for FEP calculations which can be applied to arbitrary transformations in which the initial and the final states may possess a differing configuration space volume. The formalism is applicable to systems defined on discrete or continuous configuration spaces. The Zwanzig equation[10] will be recovered as special case of our formalism. Our formalism describes the transformation by a statistical mechanical average over energy differences for microscopic transitions between the two ensembles of microstates. Inspired by Monte Carlo methods, these transitions can be interpreted as special kind of Monte Carlo trial moves termed FEP move here.

The principal question to be answered by our derivations is, what is to be done if there are multiple possible mappings of the end states onto each other and accordingly multiple possible FEP moves? The derivations will result in three different simulation schemes. The first simulation scheme uses energy differences from all possible FEP moves and leads directly to the desired transformation free energy. The two remaining simulation schemes use single FEP moves chosen according to probability distributions. That is, these simulation schemes use a stochastic mapping of the end states onto each other instead of a static, deterministic mapping used by previous FEP methods. The single-move simulation schemes need to be augmented by a correction term that accounts for the relative configuration space volume of the end states. As an interesting side product, our theory offers insights into general statistical thermodynamics regarding the relation of configurational entropy and volume of the configuration space regions effectively explored by a system.

Recently, Jarzynski presented an alternative generalization of the Zwanzig equation named targeted FEP.[14] Targeted FEP uses a deterministic mapping to optimize the overlap[11,15] between the end states resulting in a more efficient simulation (see section 3.2 for an example). This mapping is bijective but not necessarily volume preserving. That is, the coordinates of one end state can be mapped to the coordinates of the other end state by an arbitrary coordinate transformation. Thus, targeted FEP is also applicable if the configuration space volume of the end states differs. Targeted FEP has, however, two drawbacks. First, a suitable mapping has to be found which may be cumbersome in the general case. Second and more important for our own applications, the method is not generally applicable to systems defined on a discrete configuration space. This is because a bijective mapping can only be defined if the number of discrete microstates is the same in both ensembles.

In the next section, we will formally derive our novel FEP formalism. The subsequent section illustrates the theoretical findings with some example problems. The example section is followed by a discussion of applications for which we believe our findings could be of importance and possible future prospects of the FEP theory opened by our theoretical findings. We close with a summary of the main findings of this paper and the conclusions thereof.

Generalized Free Energy Perturbation Theory

J. Phys. Chem. B, Vol. 115, No. 3, 2011 **509**

2. Theory

2.1. Definitions. Before we begin our derivation, it is necessary to define some terms and quantities to avoid ambiguities and to provide a formal basis for the derivation of our FEP theory. Mathematically, our physical system is described by a set of variables defining the momenta $\mathbf{p}$ (not to be confused with the occupation probability p) and a second set of variables $\mathbf{q}$ defining the spatial coordinates. A given set of particular values for the coordinates $\mathbf{q}$ defines a configuration of the system. The configuration can also include further parameters, as for example partial charges. A microstate of the system is defined by particular values for all variables describing the system, that is by a particular configuration of the system and particular momenta of all particles of the system. The microstate is fully described by a particular configuration of the system if the momenta are not considered explicitly.

The free energy difference $\Delta F_{0 \to 1}$ refers to a transformation between two states 0 and 1. The initial and final states of a transformation are often termed end states in free energy calculations. It is important to notice that in general the initial and final states are ensembles of microstates of the system. Sometimes the term macrostate is used to indicate the ensemble nature of these states. The transformation can be partitioned into multiple stages, where initial and final state of a stage can also correspond to a fictitious (but mathematically well-defined) intermediate state constituting a chimera of the end states.[16] These intermediate states can be used to improve the efficiency of a free energy calculation, as briefly sketched in section B1 of the Supporting Information. The ensembles which constitute initial and final state encompass subsets of the entire phase or configuration space Γ, which may or may not overlap. In our case, initial and final state will be defined on the basis of the configuration. We denote sets of configurations constituting the initial ensemble and final ensemble as configuration subspaces Γ_0 and Γ_1, respectively. The relation between the total configuration space Γ and the configuration subspaces encompassed by the initial and final states Γ_0, Γ_1 can be formally defined as

$$\Gamma \supseteq \Gamma_0 \cup \Gamma_1 \tag{3}$$

indicating that the union of the initial configuration subspace and the final configuration subspace is a subset of, or equal to, the total configuration space of the system. The initial and the final configuration subspaces may or may not overlap.

When taking the integral of unity over all configurations constituting a configuration space or configuration subspace, we obtain its total configuration space volume[17]

$$\Psi = \int_{\Gamma} 1 \, d\mathbf{q} \tag{4}$$

The symbol Ψ is chosen to distinguish the total configuration space volume from a spatial volume. Analogously, the total phase space volume would be obtained by integrating over the momentum variables also. Here, total means that Ψ corresponds to the entire volume of configuration or phase space theoretically accessible to the system within the respective configuration space regardless of the energetics and actual probability distribution. The total phase or configuration space volume Ψ is not to be confused with the corresponding effective phase or configuration space volume actually populated by the system denoted by $\tilde{\Omega}$. The difference between Ψ and $\tilde{\Omega}$ is depicted schematically in Figure 1. Note that the actual effective phase

space volume does not necessarily correspond to a region of phase space with sharp boundaries, since the occupation probability density is nonzero throughout the whole phase space on which the model is defined wherever the energy function has a finite value. A more theoretical definition of the effective phase space volume is given below. The DN-dimensional continuous phase or configuration space volume of a system of N particles in a D-dimensional spatial region in a Cartesian coordinate system has units of action $[(\text{J s})^{DN}]$, or units of length $[\text{m}^{DN}]$ (in SI units). See the Supporting Information for an example calculation of the phase and configuration space volumes of a system formulated in continuous, Cartesian coordinates using SI units.

Without loss of generality for the formalism, we will in the following use a description of the system based on its configuration. For a system formulated in a full phase space description, the microstate energy E becomes the Hamiltonian explicitly accounting for the momenta of the system particles and the integrations must be performed over the momenta also (see the Supporting Information for an explicit example calculation). For classical systems, however, the momentum-dependent part of the integral constituting the partition function can be separated and cancels in free energy differences (see for example refs 17 and 18 and the Supporting Information).

2.2. Derivation of the Multimove FEP Equation. We start our derivation with the expression of the Helmholtz free energy difference in terms of the configurational partition functions Q_0 and Q_1 of initial ensemble and final ensemble, respectively

$$
\begin{aligned}
\Delta F_{0 \to 1} &= -\beta^{-1} \ln \frac{Q_1}{Q_0} \\
&= -\beta^{-1} \ln \frac{\int_{\Gamma_1} \exp[-\beta E_1] \, d\mathbf{q_1}}{\int_{\Gamma_0} \exp[-\beta E_0] \, d\mathbf{q_0}}
\end{aligned}
\tag{5}
$$

where the integrations are performed over all $d\mathbf{q}_0$ and $d\mathbf{q}_1$ configurations of the respective configuration subspaces. All the integrations in the following parts of this article are to be understood as total integrals over the whole domain on which the integration variables are defined unless otherwise stated (explicit example calculations can be found in the Supporting Information).

The microstate of the system is described by a set of variables. On the basis of these variables, we partition the system in two subsystems: the reactive subsystem and the environment subsystem. The variables with a different range of possible values in the initial configuration subspace and the final configuration subspace define the reactive subsystem, whereas the variables with the same range of possible values in both subspaces define the environment subsystem. This partitioning is, however, just a mathematical device to make the calculations more efficient. The environment subsystem can freely respond to changes in the reactive subsystem. Accordingly, we can express the total configuration $\mathbf{q}$ of the system as combination of the configurations of these subsystems

$$\mathbf{q} = \{\mathbf{q}^{\text{e}}, \mathbf{q}^{\text{r}}\} \tag{6}$$

where the superscripts e and r indicate environment and reactive subsystem, respectively. The total configuration space volume is equal to the product of the total configuration space volumes of the environment subsystem and the reactive subsystem

$$\Psi = \Psi^e \Psi^r \tag{7}$$

The microstate energy E is a function of the configuration of the environment and of the reactive subsystem. The total state energy can thus be written as sum of terms of the energy function depending on the configuration of both subsystems $E^r(\mathbf{q}^e, \mathbf{q}^r)$ and terms depending only on that of the environment subsystem $E^e(\mathbf{q}^e)$

$$E_i = E^e(\mathbf{q}^e) + E_i^r(\mathbf{q}^e, \mathbf{q}^r) \tag{8}$$

where the subscript $i = \{0,1\}$ indicates the initial or final state, respectively. The term $E^e(\mathbf{q}^e)$ carries no subscript because it is identical for the initial and final state. In the following, the explicit statement of the dependencies in the parentheses will be omitted to keep the notation concise.

Partitioning the energy function and the configuration in the numerator of eq 5 according to this concept transforms eq 5 into

$$\Delta F_{0\to 1} = -\beta^{-1} \ln \frac{\int \exp[-\beta E^e] \exp[-\beta E_1^r] \, d\mathbf{q}^e \, d\mathbf{q}_1^r}{\int \exp[-\beta E_0] \, d\mathbf{q}_0} \tag{9}$$

where the integrations are again performed over all $d\mathbf{q}_0 = d\mathbf{q}^e \, d\mathbf{q}_0^r$ and $d\mathbf{q}_1 = d\mathbf{q}^e \, d\mathbf{q}_1^r$ configurations of the initial and final configuration subspaces, expressed now in terms of the configurations of the environment and the reactive subsystem. The next step constitutes the major difference from the derivation of Zwanzig.[10] We account for multiple possible transitions of a microstate from the initial ensemble to a microstate from the final ensemble by multiplying the integrand in the enumerator of eq 9 by unity expressed as

$$1 = \frac{1}{\Psi_0^r} \int \exp[-\beta(E^e - E^e)] \exp[-\beta(E_0^r - E_0^r)] \, d\mathbf{q}_0^r \tag{10}$$

and obtain

$$\Delta F_{0\to 1} = -\beta^{-1} \ln \left[\frac{\int \exp[-\beta(E_1^r - E_0^r + E^e + E_0^r)] \, d\mathbf{q}^e \, d\mathbf{q}_0^r \, d\mathbf{q}_1^r}{\Psi_0^r \int \exp[-\beta E_0] \, d\mathbf{q}_0} \right] \tag{11}$$

Before we proceed with our derivation, it seems worthwhile commenting on why the right-hand side of eq 10 is equal to 1. The product of the exponential energy terms in eq 10 is equal to 1 for each configuration $\mathbf{q}_0^r$ of the reactive subsystem of the initial state. When taking the integral of unity over all configurations of the reactive subsystem, we obtain the total configuration space volume $\Psi_0^r = \int 1 \, d\mathbf{q}_0^r$ theoretically accessible to the reactive subsystem within the configuration subspace of the initial state. To obtain unity on the left-hand side of eq 10, we thus have to divide the right-hand side by Ψ_0^r. In the case of a discrete configuration space, this volume becomes simply the total number of possible configurations of the reactive subsystem within the initial configuration subspace.

In eq 11 the canonical occupation probability density of a configuration in the initial configuration subspace given by

$$\rho_0 = \frac{\exp[-\beta(E^e + E_0^r)]}{\int \exp[-\beta E_0] \, d\mathbf{q}_0} \tag{12}$$

can be recognized. We thus obtain from eq 11

$$\Delta F_{0\to 1} = -\beta^{-1} \ln \left[\frac{1}{\Psi_0^r} \int \rho_0 \exp[-\beta(E_1^r - E_0^r)] \, d\mathbf{q}^e \, d\mathbf{q}_0^r \, d\mathbf{q}_1^r \right] \tag{13}$$

Equation 13 shows the benefit of partitioning the configuration into reactive and environment subsystem for a simulation. During the calculation of the energy difference, only energy terms depending on the configuration of the reactive subsystem need to be computed while terms depending solely on the configuration of the environment subsystem cancel from the difference. By symmetry, the corresponding equation for the reverse transformation is simply obtained by exchanging source and target indices. Equation 13 can be written in a more compact form as

$$\exp(-\beta \Delta F_{0\to 1}) = \frac{1}{\Psi_0^r} \langle \int \exp[-\beta(E_1^r - E_0^r)] \, d\mathbf{q}_1^r \rangle_0 \tag{14}$$

where $\langle \cdots \rangle_0$ denotes the ensemble average over the initial state.

A special case occurs if the reactive subsystem has just one possible configuration in each end state. In this case, the variables constituting $\mathbf{q}_0^r$ and $\mathbf{q}_1^r$ have only one possible value specific to either end state; i.e., they become constants. The dimensionality of the configuration subspaces is then reduced relative to that of the entire configuration space, because the constants $\mathbf{q}^r$ are omitted from the integration in eq 4. Consequently, the total configuration space volume of the reactive subsystem Ψ^r is unity for both end states. The total volumes of initial and final configuration subspaces $\Psi_0 = \Psi^e \Psi_0^r$ and $\Psi_1 = \Psi^e \Psi_1^r$ are equal in this special case. Furthermore, the mapping of initial to final configurations $\mathcal{M}: \mathbf{q}_0 \to \mathbf{q}_1$ is, in this special case, unequivocally determined by changing the value of $\mathbf{q}^r$ to the only possible value in the adjacent end state while keeping the configuration of the environment subsystem $\mathbf{q}^e$

$$\mathcal{M}: \mathbf{q}^e \mathbf{q}_0^r \to \mathbf{q}^e \mathbf{q}_1^r |_{\mathbf{q}_0^r, \mathbf{q}_1^r \text{const}} \tag{15}$$

This mapping is bijective and volume preserving; i.e., there is a one-to-one correspondence of each differential volume element $d\mathbf{q}_0$ to exactly one differential volume element $d\mathbf{q}_1$. The size of the differential volume elements $d\mathbf{q}_0$ and $d\mathbf{q}_1$ and the total volume of the configuration subspaces mapped onto each other is equal. The derived one-to-one mapping can be perceived as defining a unique FEP move which is defined by changing the configuration of the reactive subsystem $\mathbf{q}^r$ from the only set of values allowed in the initial ensemble to the respective only set of values allowed in the final ensemble. This FEP move can be described as a switching of the energy function,[11] where the microstate energies E_i specific to the end states $i = \{0,1\}$ would be conceived as two different energy functions whose range of validity is defined by the configuration of the reactive

subsystem (see setup A of Bennett's model system in section 3.2 for an example). Consequently, the familiar FEP equation as given in ref 10 follows as a special case of eq 13 in the case of just one possible configuration of the reactive subsystem

$$\exp(-\beta \Delta F_{0\to1}) = \int \rho_0 \exp[-\beta(E_1^r(\mathbf{q}^e, \mathbf{q}_1^r) - E_0^r(\mathbf{q}^e, \mathbf{q}_0^r))] \, d\mathbf{q}^e|_{\mathbf{q}_0^r, \mathbf{q}_1^r \mathrm{const}} \quad (16)$$

where the configuration $\mathbf{q}^e$ is often written as $\mathbf{q}$ or $\mathbf{q}_0$ as in eq 2, while $\mathbf{q}^r$ is perceived as a parameter that defines the initial or final state and the corresponding energy function. A volume preserving, bijective (one-to-one) mapping $\mathscr{M}: \mathbf{q}_0 \to \mathbf{q}_1$ of each volume element in the initial configuration subspace to a volume element in the final configuration subspace can also be defined if the total configuration space volume (see Figure 1) of the reactive subsystem is equal for the initial and the final state. However, the Zwanzig equation makes no statement on how to define this mapping—it does not need to correspond to a mapping in real physical space.[12-14] Such a defined mapping leads, for example, from setup B to setup A of Bennett's model transformation presented in section 3.2. These considerations make it clear that the Zwanzig equation[10] implicitly assumes a system in which an unequivocal, volume preserving, bijective mapping of initial and final configuration subspaces exists. This prerequisite limits the direct applicability of the Zwanzig equation to transformations between end states of equal total configuration space volume.

2.3. Simulations Conducted with Single FEP Moves. Equation 13 has the disadvantage that at each configuration sampled from the equilibrium distribution of the initial state one would have to evaluate the integral $\int \exp[-\beta(E_1^r - E_0^r)] \, d\mathbf{q}_1^r$, which might be difficult or impossible especially in the case of a continuous configuration space. In the case of a discrete configuration space, the evaluation of this integral simplifies to collecting as many state energy samples $E_1 - E_0$ as there are configurations of the reactive subsystem in the final state. To circumvent this difficulty, we can simplify eq 13 to use only one perturbation energy sample per data collection step and correct afterward for the deviation from the true free energy difference. Such a form of eq 13 is more practical to use, but requires a decision on how to choose the FEP move to be executed and an appropriate correction. Two possibilities to choose among the FEP moves are explored here. The choice can be made randomly as pursued in the next section or based on the equilibrium distribution of the final state as explored in the subsequent section.

2.3.1. Perform Random Single FEP Moves. To obtain a simpler expression as eq 13, we randomly select one configuration of the reactive subsystem $\mathbf{q}_1^r$ at the final state instead of realizing all possible switching moves

$$\Delta \tilde{F}_{0\to1}^{\mathrm{rand}} = -\beta^{-1} \ln\left[\frac{1}{\Psi_1^r} \int \rho_0 \exp[-\beta(E_1^r - E_0^r)] \, d\mathbf{q}^e \, d\mathbf{q}_0^r \, d\mathbf{q}_1^r\right] \quad (17)$$

This equation corresponds to weighting all switching moves by the same uniform probability $1/\Psi_1^r$. In doing so, we do not grant the system time to relax toward equilibrium. Hence, this process can be described as an infinitely fast switching from initial to final state. Note that eq 17 reduces to the Zwanzig equation (eq 16) if the total configuration space volume of the reactive

subsystem is unity in the initial and final states. The difference between the true free energy difference eq 5 and eq 17 is

$$\beta \Delta \Delta F_{0\to1}^{\mathrm{rand}} = \beta(\Delta F_{0\to1} - \Delta \tilde{F}_{0\to1}^{\mathrm{rand}})$$
$$= -\ln \frac{Q_1}{Q_0} + \ln \int \frac{\rho_0}{\Psi_1^r} \exp[-\beta(E_1^r - E_0^r)] \, d\mathbf{q}^e \, d\mathbf{q}_1^r \, d\mathbf{q}_0^r \quad (18)$$

using eqs 8 and 12 we obtain

$$\beta \Delta \Delta F_{0\to1}^{\mathrm{rand}} = -\ln \frac{Q_1}{Q_0} + \ln \int \frac{\exp[-\beta E_0]}{Q_0} \frac{1}{\Psi_1^r} \times \exp[-\beta(E_1 - E_0)] \, d\mathbf{q}^e \, d\mathbf{q}_0^r \, d\mathbf{q}_1^r \quad (19)$$

$$\beta \Delta \Delta F_{0\to1}^{\mathrm{rand}} = -\ln Q_1 + \ln \int \frac{1}{\Psi_1^r} \exp[-\beta E_1] \, d\mathbf{q}^e \, d\mathbf{q}_0^r \, d\mathbf{q}_1^r \quad (20)$$

Since there are no terms depending on $\mathbf{q}_0^r$ remaining, integrating over all $d\mathbf{q}_0^r$ gives the total configuration space volume of the reactive subsystem in the initial state Ψ_0^r

$$\beta \Delta \Delta F_{0\to1}^{\mathrm{rand}} = -\ln Q_1 + \ln \int \exp[-\beta E_1] \, d\mathbf{q}^e \, d\mathbf{q}_1^r + \ln \frac{\Psi_0^r}{\Psi_1^r} \quad (21)$$

Finally, we obtain the deviation of our free energy estimate from the true free energy difference as

$$\Delta \Delta F_{0\to1}^{\mathrm{rand}} = -\beta^{-1} \ln \frac{\Psi_1^r}{\Psi_0^r} \quad (22)$$

This correction term can often easily be calculated analytically especially in the case of a discrete configuration space. Thus, we can use eqs 17 and 22 to obtain the free energy difference from

$$\Delta F_{0\to1} = \Delta \tilde{F}_{0\to1}^{\mathrm{rand}} + \Delta \Delta F_{0\to1}^{\mathrm{rand}} \quad (23)$$

2.3.2. Perform Single FEP Moves to an Equilibrated System. In this section, the single FEP move is performed from the configuration sampled from the equilibrium distribution of the initial state to a configuration sampled from the equilibrium distribution of the final state. Thus, each energy difference in the exponential average of this simulation scheme corresponds to an infinitely slow transition between the corresponding initial and final configurations. That is, the system has enough time to relax to equilibrium during the transition. The free energy estimate obtained from this simulation scheme is described by

$$\Delta \tilde{F}_{0\to1}^{\mathrm{equi}} = -\beta^{-1} \ln \int \rho_0 \rho_1 \exp[-\beta(E_1 - E_0)] \, d\mathbf{q}_0 \, d\mathbf{q}_1 \quad (24)$$

The integral is taken over all possible pairs of microstates from the initial and final ensemble, where in contrast to the random single move defined by eq 17, the configurations from both

ensembles are weighted by their equilibrium occupation probability. In principle, one could simulate actual trajectories for very slow transitions between initial and final state, but this is not required. In contrast to the Jarzynski equation,[19] eq 24 averages over state energy differences between the initial and final microstate of each transition instead of the amounts of work done along actual trajectories. Since the energy differences in eq 24 make no reference to actual trajectories connecting the microstates $\mathbf{q}_0$ and $\mathbf{q}_1$, there are different equivalent ways to implement the simulation scheme in practice. One could randomly select a microstate from the final ensemble as described in the previous section and equilibrate the system afterward. Alternatively, one could run independent equilibrium simulations of both end states in parallel and take microstate energy differences in certain intervals chosen such as to ensure uncorrelated samples.

The difference between the true free energy difference eq 5 and eq 24 is

$$
\begin{aligned}
\beta\Delta\Delta F_{0\to1}^{\text{equi}} &= \beta(\Delta F_{0\to1} - \Delta\tilde{F}_{0\to1}^{\text{equi}}) \\
&= -\ln\frac{Q_1}{Q_0} + \ln\int\rho_0\rho_1\exp[-\beta(E_1 - E_0)]\,d\mathbf{q}_0\,d\mathbf{q}_1
\end{aligned}
\tag{25}
$$

and using eqs 8 and 12 we obtain

$$
\begin{aligned}
\beta\Delta\Delta F_{0\to1}^{\text{equi}} = -\ln\frac{Q_1}{Q_0} + \ln\int\frac{\exp[-\beta E_0]}{Q_0}\rho_1\times \\
\exp[-\beta(E_1 - E_0)]\,d\mathbf{q}_0\,d\mathbf{q}_1
\end{aligned}
\tag{26}
$$

which can be further simplified to

$$
\beta\Delta\Delta F_{0\to1}^{\text{equi}} = -\ln Q_1 + \ln\int\rho_1\exp[-\beta E_1]\,d\mathbf{q}_0\,d\mathbf{q}_1
\tag{27}
$$

Again, there is no term depending on the initial configuration $\mathbf{q}_0$ left, so that integrating over all $d\mathbf{q}_0$ gives the total configuration space volume of the whole system in the initial state Ψ_0. Writing $\int\rho_1\exp[-\beta E_1]\,d\mathbf{q}_1$ as ensemble average $\langle\exp[-\beta E_1]\rangle_1$, we obtain

$$
\beta\Delta\Delta F_{0\to1}^{\text{equi}} = -\ln Q_1 + \ln\langle\exp[-\beta E_1]\rangle_1 + \ln\Psi_0
\tag{28}
$$

At first sight, this expression looks very discouraging, since we want to use the FEP simulation exactly because the partition function is difficult, if not impossible, to calculate for a large system. But when looking at the equation in more detail, we see something familiar. The first two terms on the right-hand side are very similar to the definition of the total configurational entropy of the final ensemble $-\beta TS_1 = -\ln Q_1 - (\beta E_1)_1$, with the average state energy exchanged by the exponential average of the state energy. Thus, we can define a modified configurational entropy of the final ensemble as

$$
-\beta_1 T\tilde{S}_1 \overset{\text{def}}{=} -\ln Q_1 - (-\ln\langle\exp[-\beta E_1]\rangle_1)
$$

A possible interpretation of this quantity is examined in more detail in the following two sections. The last term in eq 28 is

the free energy contribution of the ideal, maximum reachable configurational entropy of the system in the initial state. This maximum configurational entropy would be reached if all microstates had equal energy (see section 2.4). Using the definition of the modified entropy $\tilde{S}$, eq 28 can be written as

$$
\beta\Delta\Delta F_{0\to1}^{\text{equi}} = -\beta T\tilde{S}_1 + \ln\Psi_0
\tag{29}
$$

We define the effective configuration space volume

$$
\tilde{\Omega}^{\text{def}} = \exp[\beta T\tilde{S}_1]
$$

and rewrite eq 29 as

$$
\beta\Delta\Delta F_{0\to1}^{\text{equi}} = -\ln\frac{\tilde{\Omega}_1}{\Psi_0}
\tag{30}
$$

Using eq 30, the free energy difference $\Delta\tilde{F}_{0\to1}$ is given by

$$
\Delta F_{0\to1} = \Delta\tilde{F}_{0\to1}^{\text{equi}} + \Delta\Delta F_{0\to1}^{\text{equi}}
\tag{31}
$$

The correction term eq 30 involves the effective configuration space volume of the final state $\tilde{\Omega}_1$. The interpretation of this quantity as volume of the thermally accessible portions of the total configuration space volume of the final state is addressed in detail in the following sections. In the remainder of this section, we will be concerned with the derivation of an expression for the calculation of this quantity.

A possibility for the calculation of $\tilde{\Omega}$ follows from the expressions for the free energy difference $\Delta F_{0\to1}$ derived for our two single-move simulation schemes. Combining the right-hand sides of eqs 31 and 23 and solving for $\tilde{\Omega}$ gives

$$
\frac{\tilde{\Omega}_1}{\Psi_1} = \exp[-\beta(\Delta\tilde{F}_{0\to1}^{\text{rand}} - \Delta\tilde{F}_{0\to1}^{\text{equi}})]
\tag{32}
$$

Direct use of this equation results in a large statistical uncertainty of the obtained estimate of $\tilde{\Omega}_1$. A better expression can be found by noticing that the left-hand side is independent of the initial state 0 used in computing the averages on the right-hand side. The statistical uncertainty of free energy estimates computed with free energy perturbation theory is known to be minimized by maximizing the similarity or overlap of the probability distributions of the initial and final state.[11,15,20] Thus, one would expect to maximize the efficiency of the calculation by using the same state as initial and final state in the calculation of the averages $\tilde{F}_{0\to1}^{\text{rand}}$ and $\tilde{F}_{0\to1}^{\text{equi}}$

$$
\frac{\tilde{\Omega}_1}{\Psi_1} = \exp[-\beta(\Delta\tilde{F}_{1\to1}^{\text{rand}} - \Delta\tilde{F}_{1\to1}^{\text{equi}})]
\tag{33}
$$

where $\Delta\tilde{F}_{1\to1}^{\text{rand}}$ is equal to 0. Expanding $\Delta\tilde{F}_{1\to1}^{\text{equi}}$ according to eq 24 and using $\exp[\ln x] = x$ gives

$$
\frac{\tilde{\Omega}_1}{\Psi_1} = \frac{1}{\int\rho'_1\rho_1\exp[-\beta(E'_1 - E_1)]\,d\mathbf{q}'_1\,d\mathbf{q}_1}
\tag{34}
$$

Generalized Free Energy Perturbation Theory

J. Phys. Chem. B, Vol. 115, No. 3, 2011 **513**

The integration on the right-hand side can be separated into two independent integrals

$$\frac{\tilde{\Omega}_1}{\Psi_1} = \frac{1}{\left(\int \rho'_1 \exp[-\beta E'_1]\, d\mathbf{q}'_1 \right)\left(\int \rho_1 \exp[\beta E_1]\, d\mathbf{q}_1 \right)} \quad (35)$$

This equation results in a reduced statistical uncertainty of the calculated $\tilde{\Omega}_1$ as expected. Equation 35 formally allows one to calculate the effective configuration space volume from a set of N uncorrelated samples drawn from an equilibrium simulation as

$$\tilde{\Omega}_1 = \frac{\Psi_1 N^2}{\sum_i^N \exp[-\beta E_i] \sum_j^N \exp[\beta E_j]} \quad (36)$$

The estimate of $\tilde{\Omega}_1$ obtained from such a simulation still converges slowly to the correct absolute value but does so significantly faster than an estimate based on eq 32. Besides the more favorable intrinsic convergence properties, also the computation of $\Delta \tilde{F}_{0 \to 1}^{rand}$ is not required when using eq 36, resulting in an extra saving of computation time.

2.4. Configurational Entropy and the Effective Configuration Space Volume. In this section, we want to highlight one of the numerous possible interpretations of the quantity entropy, which will enable us to derive an interpretation of the correction terms derived in sections 2.3.1 and 2.3.2. The meaning of the correction terms is examined in detail in section 2.5.

The configurational part of the Boltzmann entropy is calculated from

$$-TS = \beta^{-1} \int \rho \ln \rho \, d\mathbf{q} \quad (37)$$

The meaning of the quantity entropy can be interpreted in various different ways. Two of them could be useful here. One follows from comparison with the ideal, maximum achievable configurational entropy of the system, which is also encountered in the microcanonical ensemble

$$-TS = -\beta^{-1} \ln \Omega \quad (38)$$

where Ω is the partition function of a microcanonical ensemble. Ω is often interpreted as the volume of the thermally accessible configuration space, hence also termed effective configuration space volume.

The second possible interpretation of the entropy follows from rearranging eq 37

$$\begin{aligned} -TS &= \beta^{-1} \int \rho \ln \rho \, d\mathbf{q} \\ &= \beta^{-1} \int \rho \ln \frac{\exp[-\beta E]}{Q} d\mathbf{q} \\ &= \beta^{-1} \left(-\ln Q + \int \rho \ln \exp[-\beta E]\, d\mathbf{q} \right) \\ &= \beta^{-1} \left(-\ln Q - \int \rho[\beta E]\, d\mathbf{q} \right) \end{aligned} \quad (39)$$

Writing this as

$$-TS = -\beta^{-1} \ln Q - \langle E \rangle \quad (40)$$

we can interpret the entropy as the mean deviation of the state energy from the absolute free energy. In section 2.3.2, a quantity similar to the configurational entropy emerged from our derivation, in which the arithmetic ensemble average of the state energy is substituted with the exponential ensemble average

$$-T\tilde{S} = -\beta^{-1} \ln Q + \beta^{-1} \ln\langle \exp[-\beta E] \rangle \quad (41)$$

This modified entropy can be rearranged to be written in a form similar to the Boltzmann entropy in eq 37

$$\begin{aligned} -T\tilde{S} &= -\beta^{-1} \ln Q + \beta^{-1} \ln\langle \exp[-\beta E] \rangle \\ &= -\beta^{-1} \ln Q + \beta^{-1} \ln \int \rho \exp[-\beta E]\, d\mathbf{q} \\ &= -\beta^{-1} \ln \int \rho \frac{\exp[-\beta E]}{Q} d\mathbf{q} \end{aligned} \quad (42)$$

from which we finally obtain

$$-T\tilde{S} = \beta^{-1} \ln \int \rho^2 \, d\mathbf{q} \quad (43)$$

Interestingly, the form of this result shows analogy to the corresponding equation for the Boltzmann entropy eq 37. A practical advantage of this entropy formulation over eq 37 is that it does not suffer numerical problems in the case of zero probabilities. By analogy to eq 38, we can write now a measure of the effective configuration space volume

$$\tilde{\Omega} = \exp[\beta T\tilde{S}] = \frac{1}{\int \rho^2 \, d\mathbf{q}} \quad (44)$$

This result has been obtained before by Sussmann in a quantum mechanical context as a measure of uncertainty about the position of a particle in phase space.[21] This interpretation is absolutely analogous to that as effective configuration or phase space volume in the present context. The difference is that a quantum mechanical system can often even be described by a blend of different microstates at a single point in time, while the uncertainty about the position of the system in phase space in the classical case arises from occupation of different microstates at different points in time. A similar interpretation of entropy, as measure of the incompleteness of knowledge about the current microstate of the system, was already provided by the seminal works of Edwin Jaynes, which established a formulation of statistical mechanics in terms of information theory.[22] See ref 28 for a review on the relation of entropy and information in physics. We note that the functional form of the Sussmann entropy eq 43 is formally identical to that of the information theoretic Rényi entropy of order two.[24]

Regarding the relation of Boltzmann entropy and Sussmann entropy, we can see from comparison of eqs 40 and 41 that the two entropy measures are equal if the exponential average of the state energy equals its arithmetic average. This condition is met in the microcanonical ensemble or in a degenerate system in which all configurations have the same state energy. From the considerations in this section, it is not yet entirely clear to us if there is more physical meaning hidden within the difference of Boltzmann and Sussmann entropy than the obvious math-
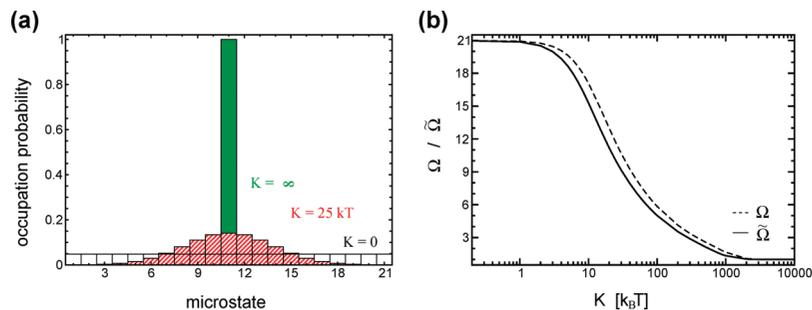
Figure 2. Occupation probability and effective (thermally accessible) configuration space volume available to a classical particle in a one-dimensional box with a discretized harmonic potential. See text for a detailed description of the system. (a) The occupation probability distribution of the model system is plotted for its 21 microstates using three different harmonic potentials of different strength determined by the parameter K. (b) The effective configuration space volume $\tilde{\Omega} = \exp[\beta T \tilde{S}]$ is plotted as a function of K (solid curve). For comparison, the plot contains also $\Omega = \exp[\beta T S]$ (dashed curve).

ematical differences tell. Since this question is outside the immediate scope of this paper, we will not pursue it further here. The interested reader is referred to section C of the Supporting Information for a more elaborate analysis of this question. Nevertheless, we can deduce a possible interpretation of the Sussmann entropy as effective configuration space volume occupied by a system in a specific macrostate, i.e., an ensemble of microstates using eq 44. This interpretation is fortified in the next section by an analogy argument regarding the meaning of the derived correction terms eqs 22 and 30. The meaning of the effective configuration space volume and its dependence on the shape of the state energy landscape is illustrated by a simple example system at the beginning of the section Examples and by a slightly more complicated example in the Supporting Information.

2.5. Meaning of the Correction Terms for Simulations Conducted with Single FEP Moves. For the random single-move scheme outlined in section 2.3.1, the correction term eq 22 describes the relative total volume of the initial and final configuration subspace. The interpretation of the correction term for the single FEP move scheme with selection of the target state from the equilibrium distribution eq 30 is analogous. We restate eqs 22 and 30 here to emphasize the analogy. The correction term for the equilibrated single-move scheme is

$$\Delta\Delta F_{0\rightarrow 1}^{\text{equi}} = -\beta^{-1} \ln \frac{\tilde{\Omega}_1}{\Psi_0} \qquad (45)$$

The configurational entropy can be interpreted as measure for the effective volume of the configuration space (see section 2.4). In the case of a discrete configuration space, $\tilde{\Omega}$ is the effective number of configurations thermally accessible to the system. The correction term for the random single-move scheme reads

$$\Delta\Delta F_{0\rightarrow 1}^{\text{rand}} = -\beta^{-1} \ln \frac{\Psi_1}{\Psi_0} \qquad (46)$$

where we multiplied the numerator and the denominator of the fraction in eq 22 with Ψ^e and substituted the products of the total configuration space volumes of the two subsystem with the total configuration space volume of the whole system (see eq 7). Comparing eqs 45 and 46, one can easily spot the analogy between $\tilde{\Omega}_1$ and Ψ_1. Since we collected only one state energy difference per sampling step corresponding to one transition instead of integrating over all possible FEP moves as prescribed by eq 13, we have to correct for the actual number of possible

transitions. The ratio of the numbers of possible forward and backward transitions between initial and final state is given by the relative configuration space volumes appearing in the correction terms. The numbers of possible transitions is determined by the numbers of configurations in each configuration subspace which are accessible as initial or final configuration of the perturbation. The accessibility of a configuration as final point of the perturbation is determined by the probability of selecting a configuration in the respective single-move FEP scheme. For the random single-move scheme (eq 23), there is a uniform probability for all configurations of the final configuration subspace that is given by Ψ_1^{-1} (or by $(\Psi_1^r)^{-1}$ if the variables are subdivided into environment and reactive subsystem). For the equilibrated single-move FEP scheme (eq 31) the probability to select a configuration as the final point of the perturbation is given by its equilibrium probability.

3. Examples

This section has the purpose to illustrate the theory derived in the previous section and its application with the aid of some model systems. These model systems are simple enough to be analytically explored on a piece of paper but sufficiently complex to underline the importance of our theoretical findings.

3.1. Effective Configuration Space Volume Occupied by a Classical Particle in a Box with a Discretized Harmonic Potential. Consider a model system consisting of 21 discrete microstates in a one-dimensional box evenly distributed along the coordinate $0 \leq \mathbf{q} \leq 1$. The energy of the microstate i is given by the harmonic potential

$$E_i = K(0.5 - \mathbf{q}_i)^2 \qquad (47)$$

The potential is symmetric around the center of the length interval at $\mathbf{q} = 0.5$ and its strength is defined by the constant K. The energy of the microstates is evaluated at the discrete points $\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_{21}\} = \{0, 1/20, ..., 1\}$.

The population probability of the microstates and the effective configuration space volume $\tilde{\Omega}$ are plotted in Figure 2. For comparison, the plot also contains the analogous quantity Ω derived from the Boltzmann entropy for this model system. Figure 2b shows the effect of K on the effective configuration space volume $\tilde{\Omega}$. If the energy landscape is very flat ($K \rightarrow 0$), all microstates are thermally accessible and so $\tilde{\Omega} = 21$. If on the other hand the harmonic potential is very steep, the particle is trapped at the center of the box and will essentially never populate a state other than the central one, which is reflected by the convergence of $\tilde{\Omega}$ to 1 for $K \rightarrow \infty$. In the extreme limits,
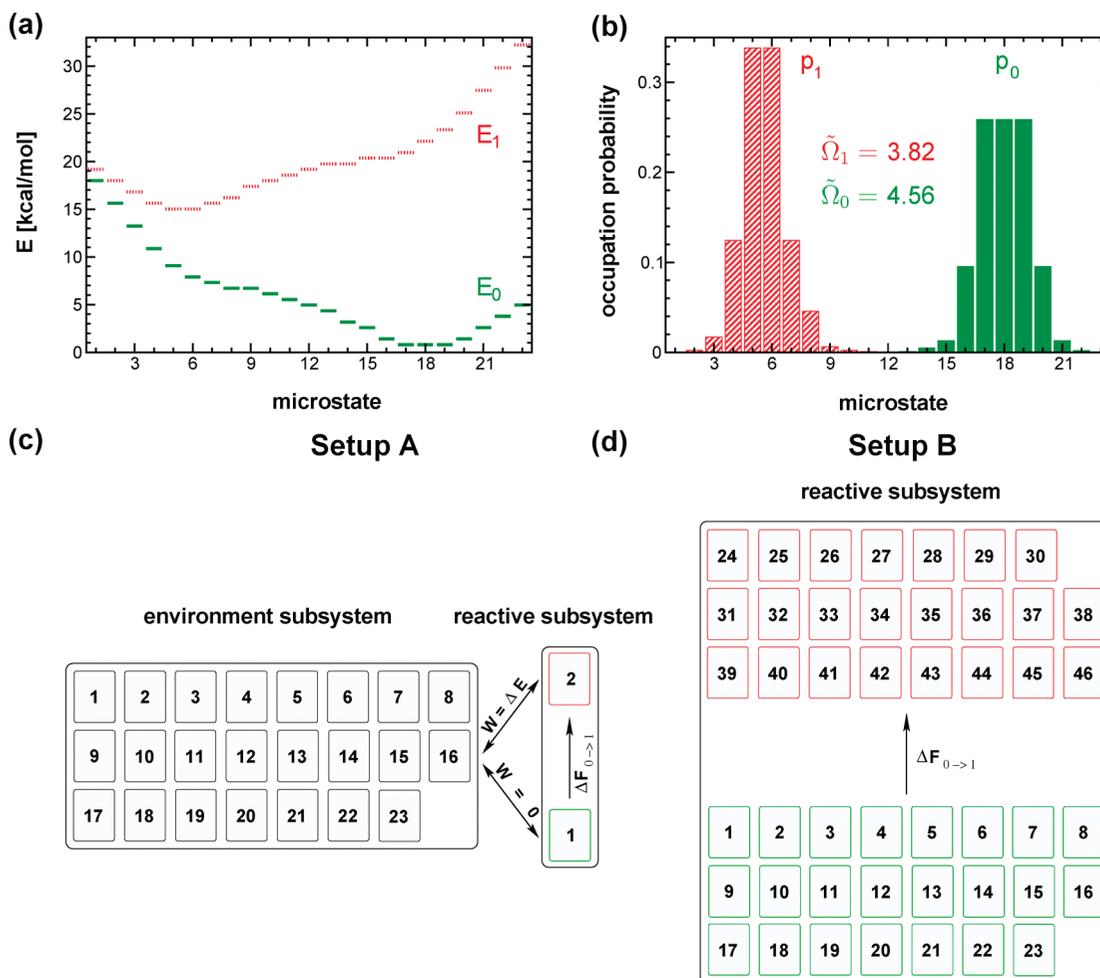
Figure 3. Bennett's model system.[11] The system can be set up in two different, yet fully equivalent, ways. The setups exemplify the consequences of a reactive subsystem with just one or multiple possible configurations in the initial and final state for the outcome of FEP simulations performed in different ways. See the text for a detailed description of the system. (a) The green solid and the red dashed profiles show the energies of the 23 microstates of the initial state and the final state, respectively. (b) The green filled columns show the occupation probabilities of the 23 microstates of the initial state while the red hatched columns show those of the final state. (c) Setup A: An unequivocal one-to-one mapping of the initial and final configurations exists, making the application of the Zwanzig equation possible. (d) Setup B: There are multiple possibilities to define a one-to-one mapping of the initial and final configurations. Mapping the configurations in the order in which they are numbered in the figure leads to setup A. In (c) and (d) the configurations are symbolized by the rectangles with the numbering inscribed.

the meaning of $\tilde{\Omega}$ is thus easily understandable. In the intermediate regime, however, fractional values of $\tilde{\Omega}$ can occur. These fractional values of $\tilde{\Omega}$ in case of discrete microstates can be understood considering eq 44. One realizes that $\tilde{\Omega} = 1/(\int p^2 \, d\mathbf{q})$ is the inverse integral of the squared occupation probability carried out over the whole configuration space. The integration can be visualized as summation of the squares of the microstate occupation probabilities p depicted for our example system by the bars in Figure 2a. The interpretation of the integral $\int p^2 \, d\mathbf{q}$ is straightforward—it is the mean probability to find the system in the same microstate when comparing two uncorrelated samples picked from the ensemble.

3.2. Bennett's Model Transformation. The second example that we want to present here is a model transformation taken from an article of Charles Bennett.[11] In this article, Bennett introduced his acceptance ratio method which allows to extract the most probable value of the free energy difference from a given set of FEP samples. This method can be used to minimize the statistical error of the free energy estimate obtained from a FEP simulation.[11,25] An outline of the method in the context of our formalism is given in section B2 of the Supporting Information. We will not make use of the method in this paper, but we wish to use the model transformation proposed by

Bennett as an example for the calculation of the transformation free energy from the different simulation schemes derived in the theory section.

The system consists of an ensemble of 23 microstates in the initial state and an ensemble of 23 microstates in the final state. The microstate energy surfaces of the initial and final state are plotted in Figure 3 along with the resulting occupation probability distributions. This example system has the nice property that it can be formulated in two equivalent ways: one in which the reactive subsystem has just one possible configuration in the initial and final states (setup A) and one in which the reactive subsystem has multiple possible configurations in both states (setup B). In setup A shown in Figure 3c, the system consists of an environment subsystem with 23 possible configurations and a reactive subsystem with 2 possible configurations where occupation of configuration 1 defines the initial state and occupation of configuration 2 defines the final state. In setup A, the energy functions of the initial and final state are given by

$$E_0 = E_i^{\text{intr}}; \quad E_1 = E_i^{\text{intr}} + W_i \qquad (48)$$

Each configuration i of the environment subsystem contributes a specific energy E_i^{intr} to the total energy of the system, which is equal in the initial and final state. This energy contribution is equal to the energy of the corresponding microstate from the initial ensemble as depicted in Figure 3a. The switching of the energy function from state 0 to state 1 is realized in setup A by an interaction of configuration 2 of the reactive subsystem with the environment subsystem. The reactive subsystem does not interact with the environment subsystem if it adopts configuration 1. The magnitude of the interaction energy W_i is specific to the configuration i of the environment subsystem and equal to the respective microstate energy difference between the initial and final states given for each of the 23 microstates in the appendix of ref 11 (see Figure 3a). In setup B, the energy of the system is given by the energy of each of the 23 configurations in initial or final ensemble given in the appendix of ref 11 (see Figure 3a).

The two possible setups of the system are ideally suited to illustrate the mapping between the configuration subspaces of the end states. In setup A, there is an unequivocal one-to-one mapping of each configuration from the initial state to exactly one configuration of the final state and vice versa. The mapping $\mathcal{M}$: $\mathbf{q}^e\mathbf{q}_0^r \to \mathbf{q}^e\mathbf{q}_1^r$ is unequivocal because there is only one configuration of the reactive subsystem $\mathbf{q}^r$ for each of the end states. Changing the configuration of the reactive subsystem constitutes the unique switching move between the initial and final configurations derived in section 2.2 as prerequisite for the applicability of the Zwanzig equation, eq 1. Picking up the original description of FEP by Zwanzig,[10] the configuration of the reactive subsystem would be interpreted as control parameter which perturbs the energy function, while the configuration of the environment subsystem would be perceived as the configuration $\mathbf{q}$ appearing in eq 1. The switching move can then be interpreted as changing the energy function, while keeping the configuration of the system as described by Bennett.[11] In contrast to the situation in setup A, there is no unequivocal one-to-one mapping of the initial and final configurations in setup B, because there are multiple configurations of the reactive subsystem at each end state, leading to multiple possible FEP moves between the initial and final state. However, since the number of configurations (the total configuration space volume) is the same in the initial and final state, it is possible to define a one-to-one mapping of the initial and final configurations by forming pairs of configurations from the initial and final state. One of the possible one-to-one mappings can be chosen to make the application of the Zwanzig equation possible.[13] In our example, the configurations of setup B are simply paired in the order in which they are listed (1 is mapped to 24, 2 to 25, ...), leading back to setup A, which is the original arrangement given by Bennett.[11,26]

Table 1 shows analytical results of FEP calculations performed according to the different schemes developed in section 2 and compares them to the exact free energy difference obtained from eq 5. The multimove FEP formula given by eq 13 demands collecting state energy difference samples for all possible FEP moves and leads to the correct results with both setups. Simulations performed with single FEP moves without the appropriate correction term do not always lead to correct results. A simulation performed with random single FEP moves leads to the correct results without the correction term for both, setup A and B. Why this is the case becomes clear upon examination of the correction term corresponding to this simulation scheme given by eq 22. The correction term adopts a value of zero for the two setups, because the number of theoretically possible

TABLE 1: Results from Perfect FEP Simulations at Setup A and Setup B of Bennett's Model System[a]

|		free energy difference (kcal/mol)			
		forward	−reverse	mean	difference
		Setup A			
$F_{0\to1}$	exact/multimove	14.37791	14.37791	14.37791	0
$\tilde{F}^{rand}_{0\to1}$	uncorrected	14.37791	14.37791	14.37791	0
	deviation	0	0	0	
	corrected	14.37791	14.37791	14.37791	0
	deviation	0	0	0	
$\tilde{F}^{equi}_{0\to1}$	uncorrected	15.17150	13.47887	14.32519	1.69262
	deviation	0.79359	−0.89904	−0.05272	
	corrected	14.37791	14.37791	14.37791	0
	deviation	0	0	0	
		Setup B			
$\tilde{F}^{rand}_{0\to1}$	uncorrected	14.37791	14.37791	14.37791	0
	deviation	0	0	0	
	corrected	14.37791	14.37791	14.37791	0
	deviation	0	0	0	
$\tilde{F}^{equi}_{0\to1}$	uncorrected	14.92590	13.72447	14.32519	1.201425
	deviation	0.65343	−0.54799	−0.05272	
	corrected	14.37791	14.37791	14.37791	0
	deviation	0	0	0	

[a] The data is calculated analytically using eqs 17 and 23 for the uncorrected and corrected $\tilde{F}^{rand}_{0\to1}$ and eqs 24 and 31 for the uncorrected and corrected $\tilde{F}^{equi}_{0\to1}$. The results are checked to be exactly reproducible in actual simulations within statistical error. The calculated values are compared to the exact statistical mechanics result $F_{0\to1}$ calculated from eq 5. The exact values are identical to those calculated from the multimove FEP formula eq 13.

configurations of the reactive subsystem Ψ^r is equal in the initial and final states for both setups. In case of a simulation with single FEP moves selected according to the equilibrium probability distribution, the situation is different. Deviations arise for both setups without the correction term given by eq 30. It can be observed that the correction term for simulations conducted in forward direction $0 \to 1$ differs from that for simulations conducted in the reverse direction $1 \to 0$. The reason for this asymmetry is the different effective configuration space volume populated by the system in the initial and final state as can be seen from eq 30 and Figure 3b. One can see that the deviations from the exact result obtained with this simulation scheme can be fully corrected for by the appropriate correction term eq 30.

3.3. Relative Binding Free Energy from Alchemical FEP. In this model example, we go one step further and examine a system with a differing number of reactive subsystem configurations at the two end states. The example is adopted from an application of FEP simulations frequently encountered in computational drug design pioneered by Tembe and McCammon.[27]

The aim is to calculate the relative binding free energy of two ligands to a receptor. For simplicity, we assume that the chemical potential of both ligands in solution is equal, so that the task reduces to the calculation of the free energy change upon exchanging ligand A with ligand B in the binding site. The model system setup for the relative ligand binding free energy example, as depicted in Figure 4, consists of a receptor with the ligand binding site and a pH-titratable group. The titratable group can adopt two proton binding forms (bound or unbound, one of which is charged) and two rotamers each oriented toward one of the two niches of the ligand binding site. The ligands A and B bind exclusively to the two distinct but overlapping niches of the ligand binding site (i.e., only one

Generalized Free Energy Perturbation Theory

J. Phys. Chem. B, Vol. 115, No. 3, 2011 **517**
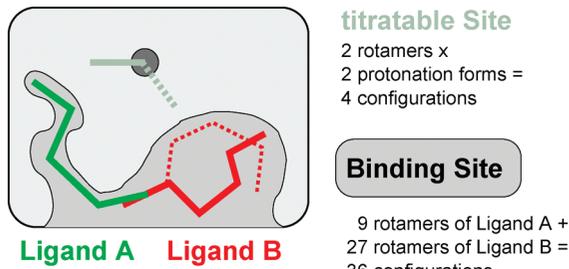
Figure 4. Model system setup for the relative ligand binding free energy example consists of a receptor with the ligand binding site and a pH-titratable group. The transformation is defined by the exchange of ligand A with ligand B in the binding site. See text for detailed explanation.

of them at a time) and interact favorably with the charged form of the titratable site. The flexibility of the ligands is treated by considering three rotamers of each dihedral angle. Since ligand A possesses two rotatable bonds and ligand B possesses three rotatable bonds, this discretization results in nine rotameric forms of ligand A and 27 rotameric forms of ligand B. Ligand B can be accommodated by the binding site in all of its rotameric forms, while ligand A can only be adopted in one of its rotameric forms by its binding niche, but is bound more tightly and interacts also more favorably with the charged form of the titratable site. The state energy of this model is given by

$$E(i,j) = E_i^{\text{intr}} + E_j^{\text{intr}} + W_{i,j} \tag{49}$$

with the energy contributions

$$E_i^{\text{intr}} = \begin{cases} -3.0 \text{ kcal/mol} & \text{if } i = 1 \\ 0.0 \text{ kcal/mol} & \text{if } 2 \le i \le 9 \\ -2.0 \text{ kcal/mol} & \text{if } 10 \le i \le 36 \end{cases}$$

$$E_j^{\text{intr}} = \begin{cases} 0.0 \text{ kcal/mol} & \text{if } j = \{1, 2\} \\ 2.0 \text{ kcal/mol} & \text{if } j = \{3, 4\} \end{cases}$$

$$W_{i,j} = \begin{cases} -4.3 \text{ kcal/mol} & \text{if } j = 3; i = 1 \\ -4.0 \text{ kcal/mol} & \text{if } j = 4; 10 \le i \le 36 \\ 0.0 \text{ kcal/mol} & \text{otherwise} \end{cases} \tag{50}$$

where i is the configuration of ligand A or B in the ligand binding site and j is the configuration occupied by the titratable site. The first $i = \{1, .., 9\}$ configurations of the binding site are the rotameric forms of ligand A, while the remaining $i = \{10, .., 36\}$ configurations of the binding site correspond to the rotameric forms of ligand B. The titratable site is oriented toward the binding niche of ligand A in the configurations $j = \{1,3\}$ and oriented toward the binding niche of ligand B when adopting the configurations $j = \{2,4\}$.

The results of FEP simulations conducted according to our three different simulation schemes are given in Table 2. It can be seen that all simulation schemes yield the correct result. FEP simulations employing a random single-move scheme (eq 17) or a single-move scheme according to the target ensemble equilibrium distribution (eq 24) without the appropriate correction term lead to severe deviations from the true relative binding free energy. Clearly, the observed deviations in the uncorrected FEP simulations conducted with single moves originate from the failure to account properly for the relative statistical weight of the ensembles. The appropriate correction terms derived in section 2 fully correct for the deviations. The

TABLE 2: Results from Perfect FEP Simulations at the Relative Ligand Binding Free Energy Example Calculated Analytically Using Eqs 17 and 23 for the Uncorrected and Corrected $\tilde{F}_{0 \to 1}^{\text{rand}}$ and Eqs 24 and 31 for the Uncorrected and Corrected $\tilde{F}_{0 \to 1}^{\text{equi}}$ [a]

		free energy difference (kcal/mol)			
		forward	−reverse	mean	difference
$F_{0 \to 1}$	exact/multimove	−0.67	−0.67	−0.67	0
$\tilde{F}_{0 \to 1}^{\text{rand}}$	uncorrected	−0.02	−0.02	−0.02	0
	deviation	0.65	0.65	0.65	
	corrected	−0.67	−0.67	−0.67	0
	deviation	0	0	0	
$\tilde{F}_{0 \to 1}^{\text{equi}}$	uncorrected	0.06	1.24	0.65	−1.17
	deviation	0.73	1.90	1.32	
	corrected	−0.67	−0.67	−0.67	0
	deviation	0	0	0	

[a] The results are checked to be exactly reproducible in actual simulations within statistical error. The calculated values are compared to the exact statistical mechanics result $F_{0 \to 1}$ calculated from eq 5. The exact values are identical to those calculated from the multimove FEP formula eq 13.

importance of such contributions to the free energy of binding has already been well recognized in the literature[8,28,29] and is further underlined by our example. We show that application of FEP with a single move per data collection point chosen randomly or according to target state equilibrium distribution without proper correction terms can lead to quantitatively and even qualitatively wrong results.

3.4. Relative Free Energy of Two Classical Particle in the Box Models with a Single-Well or a Double-Well Harmonic Potential in Continuous Coordinates. This model example demonstrates the application of our formalism to a system formulated in nondimensionless, continuous coordinates. The model system consists of a particle of mass m in a one-dimensional box of length L. The spatial coordinate $\mathbf{q}$ is defined on the interval $-L/2 \le \mathbf{q} \le L/2$ and constitutes the reactive subsystem. The momentum $\mathbf{p}$ must not be confused with the occupation probability p. The range of values available to the momentum variable is not restricted at either end state and does consequently not differ among them. Thus, the momentum constitutes the environment subsystem. The potential energy is given by a single-well or double-well harmonic potential denoted by sw or dw, respectively. The Hamiltonian of the system is given by

$$H^{\text{sw}} = \frac{\mathbf{p}^2}{2m} + \frac{m\omega^2 \mathbf{q}^2}{2}$$

$$H^{\text{dw}} = \frac{\mathbf{p}^2}{2m} + \begin{cases} \dfrac{m\omega^2}{2}\left(\mathbf{q} + \dfrac{L}{4}\right)^2 & \text{if } \mathbf{q} \le 0 \\ \dfrac{m\omega^2}{2}\left(\mathbf{q} - \dfrac{L}{4}\right)^2 & \text{if } \mathbf{q} > 0 \end{cases} \tag{51}$$

where the kinetic energy is given by the first terms on the right-hand sides, whereas the potential energy is given by the second terms on the right-hand sides. The strength of the harmonic potential is determined by the angular frequency ω of the oscillator. The single-well model is equivalent to a (confined) classical harmonic oscillator.

In a classical formulation, as adopted here, the Hamiltonian can be separated into the momentum-dependent and configuration-dependent terms which are mutually independent. This property, together with the unrestricted range of values available to the two end states, has the consequence that all momentum-
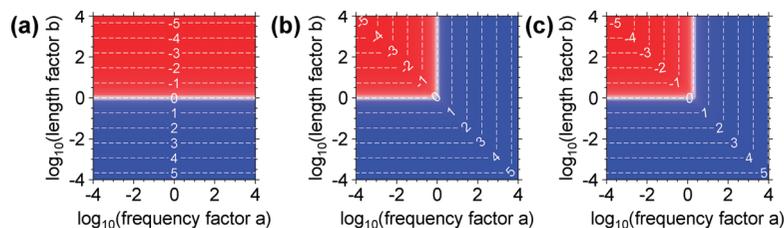
Figure 5. Correction terms for the single-move FEP schemes as function of the frequency scaling factor a and the box length scaling factor b. Positive and negative values are indicated by blue and red color, respectively. The contour values are given in kcal/mol. The particle mass is set to $m = 2.0 \times 10^{-26}$ kg. The box length of end states 0 and 1 are $L_0 = 1.6 \times 10^{-9}$ m and $L_1 = bL_0$, where $b > 0$ is a constant factor. The corresponding angular frequencies are $\omega_0 = 10^{-12}$ s^{-1} and $\omega_1 = a\omega_0$ where $a > 0$ is a constant factor. (a) $\Delta\Delta F^{\mathrm{rand}}$: given otherwise equal parameters, the correction term is identical for transformations leading from single-well to single-well or from single-well to double-well harmonic potential. (b) $\Delta\Delta F^{\mathrm{equi}}$ for transformations of a single-well to a single-well harmonic potential. (c) $\Delta\Delta F^{\mathrm{equi}}$ for transformations involving a transition from a single-well to a double-well harmonic potential.

dependent terms cancel from free energy differences. Thus, the kinetic energy terms can be omitted from the calculation.

A detailed derivation of the theory underlying this example, plots of the potential energy and the probability distributions in configuration or phase space are given in the Supporting Information. Additionally, the Supporting Information contains example calculations and plots of the effective phase and configuration space volume as function of the system parameters.

Figure 5 shows the dependency of correction terms for the single-move FEP simulation schemes defined by eqs 45 and 46 for transformations which can involve a change in box length, angular frequency, and potential type (single well or double well). It can be seen that the correction term for the random single-move scheme depends only on differences in the box length between the end states, but not on changes in the angular frequency or in the harmonic potential type of the final state. This is because $\Delta\Delta F^{\mathrm{rand}}$ depends only on the ratio of the total configuration space volumes of the end states, which is given by the box length. In contrast, the correction terms for the equilibrated single-move scheme depend also on the harmonic potential type and on the angular frequency of the final state, because these correction terms depend on the ratio of the effective configuration space volume of the final state and the total configuration space volume of the initial state. While the total configuration space volume is given by the box length, the effective configuration space volume is a function of all system parameters (see Table 1 of the Supporting Information).

Comparison of Figure 5, b and c, shows that the correction terms for the transformations of a system with a single-well to another system with a single-well potential are very similar to those of the corresponding transformations which substitute the single-well with a double-well potential. This can be explained by the logarithmic dependence of $\Delta\Delta F^{\mathrm{equi}}$ on the effective configuration space volume of the final state, considering that the effective configuration space volume of the double-well model differs at most by a factor 2 from that of the corresponding system with a single-well potential (see Figure 2 of the Supporting Information).

Furthermore, it can be seen that the correction terms can reach considerable values, especially if the box length or the frequency differs widely between the end states. Especially for the equilibrated single-move scheme, the correction terms are often comparable in magnitude to the total transformation free energies themselves (plots of the transformation free energies can be found in the Supporting Information). The Supporting Information also contains further example calculations and plots showing the free energy changes and correction terms for two more example transformations involving a change in box length and particle number or in angular frequency and particle number.

4. Discussion

In this section we will discuss applications for which the use of our FEP theory might be advantageous, advantages and disadvantages of the different simulation schemes, and possible perspectives for further developments of the FEP theory.

4.1. Applications. Figure 6 shows a schematic depiction of applications for which our formalism could be advantageous. The formalism is especially well suited for studying transformations with end states of differing configuration space volume. The use of our simulation schemes might, however, also be favorable for transformations with end states of equal configuration space volume that are treatable with traditional FEP according to the Zwanzig equation.

Figure 6a shows the protonation of a carboxylic acid residue with multiple tautomers of the protonated form. This problem was already used as motivating example in the introduction of this article. The problem is a simple representative of a whole class of free energy calculations on systems that are defined on discrete configuration spaces. The number of configurations for either of the end states can be much larger, as for example if rotameric forms of protein side chains or multiple redox and protonation forms of protein cofactors are considered. Our novel formalism has already been fruitfully applied in such free energy calculations on different systems. One application was the calculation of free energy changes in coupled protonation and reduction processes in the small electron carrier protein azurin from *Pseudomonas aeruginosa*. The results of these structure-based calculations are in very good aggreement with experimental results[30,31] and provide a microscopic picture of the underlying thermodynamics. Another application was the calculation of free energy differences for small molecule binding processes to study the permeation of of ammonium and ammonia through the membrane transporter protein Amt-1 from *Archaeoglobus fulgidus*. The calculations were based on the experimental crystal structure of the transporter determined by Andrade and co-workers.[32] The results of our calculations were compared to experimental results of the same group and aid in their interpretation. A particularly interesting part of both applications is the calculation of free energy measures of cooperativity.[33] These works are in preparation for publication and will be presented elsewhere. The FEP simulation schemes of the present article are implemented in our Monte Carlo simulation suite GMCT that will be made publicly available as open-source software.[9]

Figure 6b shows the binding of a flexible ligand to a flexible receptor as representative for the calculation of absolute binding free energies. Such calculations are often done with molecular mechanics models to account for the flexibility of the binding
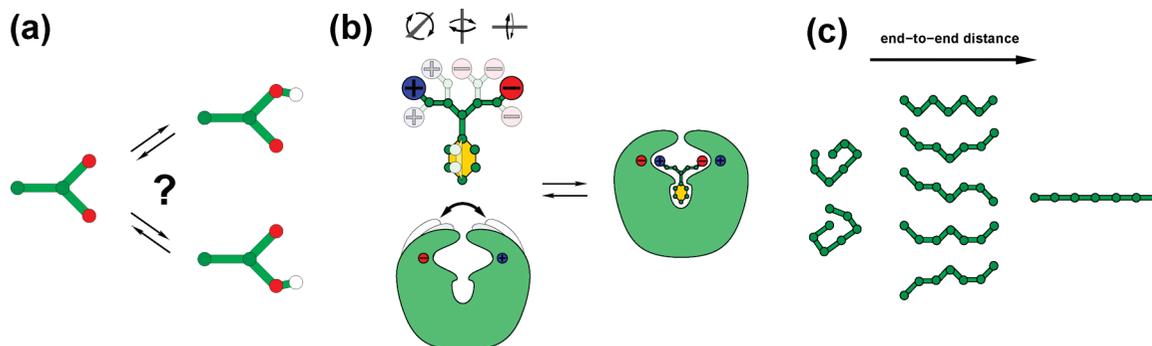
Figure 6. Three example applications of free energy simulations for which it could be crucial to consider the principles presented in the present article. (a) Calculation of protonation free energies with multiple tautomers. One can recognize the similarity to the motivating example from section 1. Similar situations arise in the treatment of other amino acids (e.g., histidine) and many important biological cofactors like flavines and quinones.[9] (b) General binding free energy calculations with flexible ligands and/or flexible receptors. (c) Calculation of (un)folding free energies for flexible polymers like proteins or nucleic acids.

partners.[8,34] The necessity to consider the change in the total configuration space volume upon binding was recognized earlier and investigated in depth by Luo and Sharp[29] and Hermans and Wang.[35] In their formalism, Ψ occurs as a spatial volume, relating the volume of the simulated system in the bound state to some reference volume in the unbound state. The reference volume corresponds to some reference concentration of the unbound ligand. Our formalism automatically includes this concept and generalizes it to include all degrees of freedom of the system. The change in the total configuration space volume can be conceptually described by three components. The first component is caused by the confinement of the ligand to the vicinity of the binding site and restricts the translational degrees of freedom. The second component can arise from a restriction of the rotational degrees of freedom of the ligand relative to the receptor that is compatible with the binding mode. The third component arises from the restriction of internal degrees of freedom in ligand and receptor due to geometric constraints that are imposed by the necessity to form specific contacts of receptor moieties with ligand moieties.

Figure 6c shows the stretching or force induced unfolding of a macromolecule. The end states of transformations in which the polymer changes its length will often have a different total configuration space volume.[36] There will be a smaller number of possible configurations in the folded state which is characterized by a compact shape (small length or end-to-end distance) and many defined intramolecular contacts, which further restricts the available configuration space. The complementary extreme is constituted by the fully elongated polymer, which has a very small total configuration space volume. In contrast to the folded and fully elongated states, there is a much larger number of configurations compatible with an intermediate elongation. Additionally, there can be restrictions of the configuration space imposed by the need to form the native intramolecular contacts which stabilize the folded state. A simplified model of the latter example is constituted by the freely jointed chain model of a macromolecular polymer examined in detail in ref 37.

4.2. Efficiency Comparison of the Simulation Schemes and Possible Perspectives for Further Developments. All the problems presented in the example section are solved analytically, that is, with perfect sampling giving exact results. The efficiency of a given simulation scheme is, however, a very important topic in actual free energy calculations from simulation data and requires careful consideration. The choice of the simulation scheme has practical consequences regarding the statistical uncertainty of a free energy estimate obtained from a given number of energy difference samples. The statistical

uncertainty of a single estimate determines the number of individual free energy estimates or the length of a simulation needed to reach a given target value of the standard deviation for the overall free energy estimate.

The equilibrated single-move scheme (eq 31) leads to much larger statistical uncertainties than the other simulation schemes. This is mainly due to the high statistical uncertainty of the correction term given by eq 36 which converges slowly and often in a very unsteady fashion. Because of these properties, the equilibrated single-move scheme is in the present form of little value for practical simulations. The true value of this simulation scheme might be less immediate as we will discuss in the end of this section. The choice between the two remaining simulation schemes has two aspects. On the one hand, the statistical uncertainty is smaller when collecting energy difference samples for all possible FEP moves in a simulation according to eq 13, because always the whole available information is used including the FEP moves with lowest energy difference that dominate the exponential average. On the other hand, the computational effort per sampling step is higher with the multimove scheme than with the random single-move scheme (eq 23) due to the higher number of energy difference samples that need to be calculated in each collection step. It is difficult to decide a priori which simulation scheme to choose, because the efficiency of the simulations schemes depends on the actual probability distributions of initial and final states; i.e., the relative efficiency varies with system, energy function, and transformation. For systems defined on a discrete configuration space, one is in principle free to choose one of the simulation schemes. However, for systems with a continuous configuration space it might be difficult to evaluate the integral over all possible FEP moves as demanded by the multimove simulation scheme eq 13.

Staging[16] and the Bennett acceptance ratio method[11] can be used to increase the efficiency of both simulation schemes as with traditional FEP.[7,8,25] The use of these two methods with our FEP simulation schemes is outlined in section B of the Supporting Information.

Although the equilibrated single-move scheme is of limited value for practical simulations because of its low efficiency, it is interesting in its own right from a theoretical point of view. In addition, comparison to the random single-move scheme might hint at possible perspectives for further developments of the FEP theory. The two single-move schemes use stochastic mappings that employ different probability distributions. It is tempting to ask whether other probability distributions could

be used to obtain FEP simulation schemes with optimized efficiency. It seems worthwhile to pursue this idea in future work.

5. Summary

We have presented a generalized FEP theory that extends the traditional FEP theory by some novel concepts and possibilities. Our formalism is inspired by Monte Carlo methods and previous interpretations of FEP in their terminology.[11,12] The novel formalism is applicable to a general class of transformations. Transformations in which the end states have a differing configuration space volume can be considered for systems defined on continuous or discrete configuration spaces.

The main results of this paper are three simulation schemes for free energy calculations with FEP. All simulation schemes partition the variables that define the position of the system in configuration space into two types to make the energy calculations more efficient. The first variable type has differing value ranges at the two end states (reactive subsystem) whereas the second variable type does not have these differing value ranges (environment subsystem). Each simulation scheme starts from microstates sampled from the equilibrium distribution of the initial state and collects energy differences for transitions or FEP moves to microstates of the final state. A move between the end states amounts to altering one or more variables of the reactive subsystem from values permitted in the initial state to values permitted in the final state. The domain of possible moves is defined by the ranges of values permissible for the variables forming the reactive subsystem of the final state. The way to select from the possible moves differs among the simulation schemes. The multimove FEP simulation scheme uses energy differences for all possible transitions and requires no correction. The two remaining simulation schemes use single FEP moves like traditional FEP based on the Zwanzig equation. The novelty in comparison with previous FEP methods[10,14] is that these simulation schemes are based on a stochastic instead of a deterministic and bijective mapping of the two end states. Traditional FEP maps each differential configuration space element of the initial state to exactly one configuration space element of the final state and vice versa. Our new approach selects the configuration space element of the final state to be mapped to a given configuration space element of the initial state based on a probability distribution. The first single-move scheme randomly selects the microstates from the final state for the FEP move. The free energy estimate obtained from this simulation scheme must be corrected for the relative total configuration space volume of the end states. The second single-move scheme selects the microstates according to the equilibrium distribution of the final state. The free energy estimate obtained from this simulation scheme must be corrected for the ratio of the effective configuration space volume of the final state and the total configuration space volume of the initial state.

We have discussed the relevance of our method to applications of general interest in the field of free energy calculations on biomolecular systems. Our theoretical findings are in line with established knowledge and provide a general, formal integration of this knowledge into FEP theory. Furthermore, we pointed out advantages of our simulation schemes in practical free energy calculations. Our discussion specifically addressed the calculation of binding free energies and free energies of polymer folding.

The novel formalism provides a generally applicable theoretical framework for FEP that promises to have potential for further developments and applications.

Supporting Information Available: Detailed derivations and additional plots for the classical particle in the box models, a brief outline of how to use staging and the Bennett acceptance ratio method in the framework of our formalism, and a discussion of the relation between Boltzmann and Sussmann entropy. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.
(2) Kollmann, P. *Chem. Rev.* **1993**, *93*, 2395–2417.
(3) Rickman, J. M.; LeSar, R. *Annu. Rev. Mater. Res.* **2002**, *32*, 195–117.
(4) Kofke, D. A. *Fluid Phase Equilib.* **2005**, *228−229*, 41–48.
(5) Oostenbrink, C.; van Gunsteren, W. F. *Chem. Phys.* **2006**, *323*, 102–108.
(6) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. In *Alchemical free energy calculations: Ready for prime time?*; Elsevier: Amsterdam, 2007; Vol. 3, Chapter 4, pp 41−59.
(7) Chipot, C.; Pohorille, A. *Free Energy Calculations*; Springer: Berlin, 2007; Vol. 86.
(8) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235–10253.
(9) Ullmann, R. T.; Essigke, T.; Ullmann, G. M. "GMCT Monte Carlo simulation software for macromolecular receptors with multiple ligands, membrane potential and conformational flexibility", to be submitted for publication.
(10) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
(11) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.
(12) Voter, A. F. *J. Chem. Phys.* **1985**, *82*, 1890–1899.
(13) Severance, D. L.; Essex, J. W.; Jorgensen, W. *J. Comput. Chem.* **1995**, *16*, 311–327.
(14) Jarzynski, C. *Phys. Rev. E* **2002**, *65*, 046122.
(15) (a) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103. (b) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 084109.
(16) Valleau, J. P.; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457–5462.
(17) Hill, T. L. *Statistical Mechanics Principles and Selected Applications*; Dover Publ., Inc.: New York, 1987.
(18) Zhou, H.-X.; Gilson, M. K. *Chem. Rev.* **2009**, *109*, 4092–4107.
(19) (a) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693. (b) Jarzynski, C. *Phys. Rev. E* **1997**, *56*, 5018–5035.
(20) Ferguson, D. M. *J. Comput. Phys.* **1993**, *99*, 10086–10087.
(21) Süssmann, G. *Z. Naturforsch. A* **1997**, *52a*, 49–52.
(22) (a) Jaynes, E. T. *Phys. Rev.* **1957**, *106*, 620–630. (b) Jaynes, E. T. *Phys. Rev.* **1957**, *108*, 171–190. (c) Jaynes, E. T. In *Information Theory and Statistical Mechanics*; W. A. Benjamin: New York, 1963; Vol. 3, pp 181−218.
(23) Maruyama, K.; Nori, F.; Vedral, V. *Rev. Mod. Phys.* **2009**, *81*, 1–23.
(24) Rényi, A. On measures of information and entropy. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960*, Vol. 1: Contributions to the Theory of Statistics, 1961; pp 547−561.
(25) (a) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601. (b) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740–5761. (c) Shirts, M. R.; Pande, V. *J. Chem. Phys.* **2005**, *122*, 134508.
(26) From Figure 3b it can be seen that the probability distributions of the two ensembles have very little overlap. In terms of efficiency of a FEP simulation with the Zwanzig equation, however, it would have been much better to reverse the ordering (1 is mapped to 46, 2 to 45,. .). This mapping would result in a much better overlap of the probability distributions, and consequently in a much smaller variance of the free energy estimate obtained from such a simulation.[11,14,15,20]
(27) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281–283.
(28) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047–1069.
(29) Luo, H.; Sharp, K. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10399–10404.

(30) Nar, H.; Messerschmidt, A.; Huber, R.; van de Kamp, M.; Canters, G. W. *J. Mol. Biol.* **1991**, *221*, 765–772.

(31) Jeuken, L. J. C.; Wisson, L.-J.; Armstrong, F. A. *Inorg. Chim. Acta* **2002**, *331*, 216–223.

(32) Andrade, S. L. A.; Dickmanns, A.; Ficner, R.; Einsle, O. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14994–14999.

(33) Ben-Naim, A. *Cooperativity and Regulation in Biochemical Processes*; Kluwer Academic/Plenum Publishers: New York, 2001.

(34) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234–2246.

(35) Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, *119*, 2707–2714.

(36) Chan, H. S.; Dill, K. A. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 447–90.

(37) Hill, T. L. *An Introduction to Statistical Thermodynamics*; Dover Publ. Inc.: New York, 1960.