How Molecular Conformational Changes Affect Changes in Free Energy

Mazen Ahmad,[†] Volkhard Helms,[‡] Thomas Lengauer,[†] and Olga V. Kalinina*,[†]

[†]Department for Computational Biology and Applied Algorithms, Max Planck Institute for Informatics, Campus E1 4, 66123 Saarbrücken, Germany

[‡]Center for Bioinformatics, Saarland University, 66123 Saarbrücken, Germany

S *Supporting Information*

ABSTRACT: A simple quantitative relationship between the molecular conformational changes and the corresponding changes in the free energy is presented. The change in free energy is the sum of that part of the enthalpic change that is due to the externally applied work (perturbation) and of that part of the entropic change, termed dissipative entropy, that is related to the conformational changes. The dissipative entropy is equivalent to the relative entropy, a concept from information theory, between the distributions of the conformations in the initial and the final states. The remaining change in entropy (nondissipative) cancels exactly with the remaining enthalpic change. The calculation of the dissipative entropy is demonstrated to pose the main difficulty in free energy computation. The straightforward decomposition of the dissipative entropy into contributions from different parts of the system promises to improve the understanding of the role of conformational changes in biochemical reactions.

$$\Delta F^0 = \langle U_p \rangle_B + kT \sum P_B \ln\left[\frac{P_B}{P_A}\right]$$
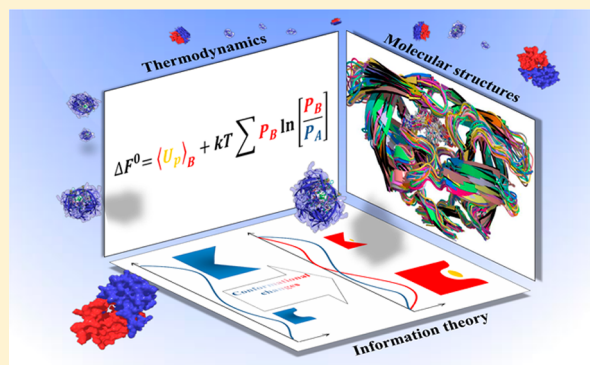
1. INTRODUCTION

The change of free energy in biomolecular interactions plays a central role in many biological and chemical fields[1,2] such as solvation[3] and ligand binding[4] processes as well as when considering changes of protein stability or ligand binding affinity due to mutations in a protein.[2] In most cases, conformational changes are assumed to play an essential role in the function of the biomolecules and to influence the changes of free energy of the biochemical reactions.[5,6] Understanding the role of conformational changes on the basis of the static structures of the system at the end points of the reaction is impaired by the loss of important data on the dynamics of the molecular changes, information that can be in part regained using NMR spectroscopy.[5,7]

The connection between thermodynamic measurements and the conformational changes of biomolecules received increasing attention in recent years.[6,8,9] The conformational changes are usually connected to the entropic changes of biochemical reactions.[6,10,11] Although the observed relationship between the dynamics of the conformational changes and their interpretation in terms of entropic changes is very insightful, our understanding of this relationship, especially in quantitative terms, is still incomplete. The complexity of the problem is increased by the fact that elucidating the relationship between enthalpic and entropic changes and their partial compensation in the free energy change is hindered by the limited understanding of the phenomenon of enthalpy–entropy compensation (EEC).[12] EEC is evident from frequent experimental observations of considerable changes in enthalpy and entropy canceling each other, resulting in a comparatively small change in free energy. This leads to substantial difficulties in interpreting the respective thermodynamic changes.[12,13] Such compensation has been observed in numerous studies, such as on the optimization of HIV-1 protease inhibitors,[14,15] peptides binding to the Grb2 SH2 domain,[16] and the binding of DNA to the catabolite activator protein.[6] The highly variable magnitude of EEC in different molecular systems[17] raises the important issues of how to separate the part of the thermodynamic change that contributes to the free energy from that which is compensated and to identify the molecular mechanisms that control the relationship between the enthalpy and the entropy.

We have recently studied the relationship between the conformational changes in biomolecular systems and the compensation between the corresponding thermodynamic changes.[18] In that work, we extended a concept for the compensation of enthalpy and entropy (EEC) due to solvent reorganization that was introduced by Ben-Naim[19] and Yu and Karplus[20] to include all occurring conformational changes in the solutes and in the solvent.[18] Here, we will use the term general EEC (GEEC) to refer to this extended interpretation of EEC to include the conformational changes.[18] GEEC uses the formalism of perturbation theory to show that the molecular

conformational changes dictate the compensation between enthalpy and entropy. The internal enthalpic and entropic changes with respect to the original (unperturbed) system cancel exactly and, thus, do not contribute to the free energy. Examples include the reorganization of the water molecules in the case of solvating a solute, conformational changes of a molecule upon solvation, and conformational changes in a ligand and a receptor upon binding of the ligand. The free energy change consists of the enthalpic contribution from the external work (perturbation) and the entropic contribution from the perturbed potential. Dissipation of the external work (perturbation) was previously related to the conformational changes.[18]

In this study, we now investigate the quantitative relationship between the conformational changes and the entropic changes, which we term here the dissipative entropy, for they are responsible for the dissipation of the external work exerted by the perturbation. The main finding in this work is to explain the role of the conformational changes on free energy changes by using a direct relationship between the dissipative entropy and relative entropy[21] between the distributions of the conformations in the initial and the final states of the molecular system. The relative entropy between two distributions is a concept from information theory. Also, we present a straightforward decomposition of the dissipative entropy into contributions from different parts of the system, and we will discuss that the molecular basis of the dissipative entropy has a fundamental impact on methods for calculating the free energy and on identifying the source of error in free energy estimation methods.

2. BACKGROUND

We consider a system with a fixed number of molecules at a given volume V and temperature T. The probability density $P_A(X)$ of observing the system in configuration X in state A is given by[22]

$$P_A(X) = \frac{\exp[-\beta U_A(X)]}{\int dX \exp[-\beta U_A(X)]} \quad (2.1)$$

Here $\beta = (kT)^{-1}$ where k is the Boltzmann constant. The total potential energy of the system $U_A(X)$ in state A depends on conformation X, which includes all degrees of freedom of the system. For convenience, we will only consider a perturbation in the potential energy here. In the general case, when the masses of the atoms are also perturbed, this should be substituted by the Hamiltonian. Next, we consider a reaction that transforms the system into another state B. This reaction is driven by a perturbation in the potential energy of the system:

$$U_p = U_p(X) = U_B(X) - U_A(X) \quad (2.2)$$

The perturbation of the system can be represented by a switching variable λ that allows switching from the canonical ensemble of state A to that of state B by changing $U_p(X; \lambda) = \lambda U_p(X)$ from ($\lambda = 0$) for state A to ($\lambda = 1$) in state B:

$$U_\lambda(X) = U_A(X) + \lambda U_p(X) = U_{up}(X) + U_p(X; \lambda) \quad (2.3)$$

Here, we use the notation $U_{up}(X) = U_A(X)$ for the potential of the initial state A (unperturbed system). The free energy perturbation (FEP) equation[23,24] can be used to compute the change in the free energy $\Delta F^0(A \to B) = F^0(B) - F^0(A)$ between states A and B. The FEP method can be used either in the forward or in the backward direction by sampling the distributions of either state, initial or final, to calculate the difference between the free energies in the two states:[22]

$$\exp[-\beta \Delta F^0(A \to B)] = \langle \exp[-\beta U_p] \rangle_A \quad \text{(Forward FEP)} \quad (2.4)$$

$$\exp[-\beta \Delta F^0(B \to A)] = \exp[\beta \Delta F^0(A \to B)]$$
$$= \langle \exp[\beta U_p] \rangle_B \quad \text{(Backward FEP)} \quad (2.5)$$

where $\langle \cdots \rangle_\alpha$ represents the average over the canonical ensemble of state α. Another closely related method for computing free energy changes is thermodynamic integration[25,26] (TI):

$$\Delta F^0(A \to B) = \int_0^1 \left\langle \frac{\partial U_\lambda(X)}{\partial \lambda} \right\rangle_\lambda d\lambda = \int_0^1 \langle U_p \rangle_\lambda d\lambda \quad \text{(TI)} \quad (2.6)$$

According to our previous work,[18] the changes in enthalpy and entropy can be decomposed into contributions due to the perturbed potential and contributions originating from the original unperturbed potential (of the initial state). The total enthalpic changes ΔE^0 are given by[18]

$$\Delta E^0(A \to B) = \langle U_B(X) \rangle_B - \langle U_A(X) \rangle_A = \Delta E_{up}(A \to B)$$
$$+ \Delta E_p(A \to B) \quad (2.7)$$

$$\Delta E_{up}(A \to B) = \langle U_{up}(X) \rangle_B - \langle U_{up}(X) \rangle_A \quad (2.8)$$

$$\Delta E_p(A \to B) = \langle U_p(X) \rangle_B \quad (2.9)$$

Here, ΔE_{up} and ΔE_p are the contributions to the enthalpy due to the unperturbed and the perturbed potentials, respectively. An analogous decomposition is applied to the entropy:[18]

$$T\Delta S^0(A \to B) = T\Delta S_{up}(A \to B) + T\Delta S_p(A \to B) \quad (2.10)$$

$$T\Delta S_{up}(A \to B) = \langle U_{up}(X) \rangle_B - \langle U_{up}(X) \rangle_A \quad (2.11)$$

$$T\Delta S_p(A \to B) = -\beta^{-1}\ln\langle \exp[\beta\{U_p(X) - \langle U_p(X) \rangle_B\}] \rangle_B \quad (2.12)$$

We will term the entropic change that is related to the perturbation ΔS_p the *dissipative entropy*, due to its analogy to the dissipated work in nonequilibrium free energy calculations.[27] GEEC shows that the change in enthalpy $\Delta E_{up}(A \to B)$ due to the unperturbed part of the potential always equals the corresponding entropic change $T\Delta S_{up}(A \to B)$, i.e., these two terms cancel in the free energy change:[18]

$$\Delta E_{up}(A \to B) = T\Delta S_{up}(A \to B)$$
$$= \langle U_{up}(X) \rangle_B - \langle U_{up}(X) \rangle_A$$
$$= -\beta^{-1}\int dX[P_B(X) - P_A(X)]\ln[P_A(X)] \quad (2.13)$$

These compensating energetic changes are related to the conformational changes in the system which are described by the changes of the probability densities $P(X)$ upon perturbation. Thus, the change of the free energy comprises the noncanceling terms, namely, the changes of enthalpy and entropy due to perturbation in the potential:[18]

$$\Delta F^0(A \to B) = \Delta E_p(A \to B) - T\Delta S_p(A \to B) \quad (2.14)$$

The corresponding thermodynamic changes for the backward process $(B \rightarrow A)$ can be found in analogy to 2.14 to be

$$\Delta F^0(A \rightarrow B) = -\Delta F^0(B \rightarrow A)$$
$$= \langle U_p(\boldsymbol{X}) \rangle_A + T \Delta S_p(B \rightarrow A)$$

$$\Delta E_p(B \rightarrow A) = -\langle U_p(\boldsymbol{X}) \rangle_A \qquad (2.15)$$

Equations 2.14 and 2.15 emerging from the GEEC theory are comparable to equations introduced by Wu and Kofke when they discussed the impact of phase space overlaps on the accuracy of free energy calculations[21] (see eq 6 in their work[21]). The entropic change due to the perturbed potential ΔS_p is the same as the relative entropy[28] between the initial and the final distributions of the potential perturbation in their previous work.[21] The relative entropy was suggested to quantify the overlap between the phase spaces of the initial and the final states in free energy calculations. However, the usage of the relative entropy to quantify the overlap between the phase spaces was found to be inapplicable due to the need of prior knowledge of the free energy, which in turn is the quantity of interest.[21] The relationship between the relative entropy and the number of samples needed for the convergence of free energy calculations was briefly discussed by Jarzynski.[29] The role of the relative entropy was previously discussed in the context of conformational fluctuations of a single molecule[30] and of the irreversibility of Markov random processes.[31] The physical significance of the relative entropy and its relationship to the irreversibility of physical processes also received attention in works by Crooks who studied nonequilibrium free energy changes.[32−35] Relative entropy is a widely used concept for measuring the changes in the distributions in many fields. Interestingly, relative entropy was used recently to compare conformational ensembles in torsion space in order to quantify the "population shifts" in the torsion angles in protein.[36] As we will show here, the relative entropy between the one-dimensional distributions of the perturbed potential U_p of the initial and the final states in the work of Wu and Kofke[21] equals the relative entropy of the distributions of the multidimensional system conformations $\boldsymbol{X}$ of the initial and the final states. Thus, expressing the dissipative entropy using the relative entropy formalism provides a simple interpretation of the relationship between the dissipative entropy and the conformational changes in the biomolecular system.

3. DISSIPATIVE ENTROPY AND THE MOLECULAR CONFORMATIONAL CHANGES

We showed in our previous work[18] that in response to a perturbation a thermodynamic system undergoes conformational changes that dissipate the effect of the perturbation (ΔE_p) by the entropic amount $T\Delta S_p$ (see eq 2.14), while the enthalpic and entropic changes originating from the unperturbed (original) part of the potential cancel out (eq 2.13). The relationship between the conformational changes and the compensation of external work $\Delta E_p = \langle U_p \rangle$ by the dissipative entropy $T\Delta S_p$ was briefly discussed in our previous study.[18] Here, we quantitatively investigate the relationship between ΔS_p and the conformational changes in the system. First, we note that eq 2.12 does not provide a clear interpretation of the molecular process bringing about the entropy change ΔS_p. Therefore, it is more useful to formulate an equation for the dissipative entropy that demonstrates its relationship to the initial and the final distributions. We start

with the relationship between the probabilities of the initial and final distributions:[24]

$$\frac{P_B(\boldsymbol{X})}{P_A(\boldsymbol{X})} = \exp[\beta \Delta F^0(A \rightarrow B) - \beta U_p(\boldsymbol{X})] \qquad (3.1)$$

Using eqs 3.1 and 2.14 in 2.9 we obtain (see Supporting Information):

$$T\Delta S_p(A \rightarrow B) = -\frac{1}{\beta} \int d\boldsymbol{X} P_B(\boldsymbol{X}) \ln\left[\frac{P_B(\boldsymbol{X})}{P_A(\boldsymbol{X})}\right]$$
$$= -\frac{1}{\beta} D(P_B(\boldsymbol{X}) \| P_A(\boldsymbol{X})) \qquad (3.2)$$

$$D(B\|A) = D(P_B(\boldsymbol{X})\|P_A(\boldsymbol{X})) = \int d\boldsymbol{X} P_B(\boldsymbol{X}) \ln\left[\frac{P_B(\boldsymbol{X})}{P_A(\boldsymbol{X})}\right] \qquad (3.3)$$

Here, $D(B \| A)$ denotes the relative entropy,[28] which is also known as the Kullback−Leibler divergence in information theory. The relative entropy is a measure of the "distance" between the two distributions and also reflects the possibility to sufficiently sample distribution A through sampling distribution B. Equation 3.3 for the relative entropy between the multidimensional distributions of the conformations $\boldsymbol{X}$ of the initial and the final states equals the relative entropy between the corresponding one-dimensional distributions of the perturbed potential U_p. Using the relationship[37] $P_B(U_p) = \exp[\beta \Delta F^0(A \rightarrow B) - \beta U_p]P_A(U_p)$ and eq 3.1, we obtain

$$\frac{P_B(U_p = U_p(\boldsymbol{X}))}{P_A(U_p = U_p(\boldsymbol{X}))} = \frac{P_B(\boldsymbol{X})}{P_A(\boldsymbol{X})} = \exp[\beta \Delta F^0(A \rightarrow B) - \beta U_p]$$

$$P_B(U_p) = \int d\boldsymbol{X} P_B(U_p, \boldsymbol{X}) = \int d\boldsymbol{X} P_B(\boldsymbol{X}) P_B(U_p|\boldsymbol{X})$$

$$D(P_B(U_p)\|P_A(U_p)) = \int dU_p P_B(U_p) \ln\left[\frac{P_B(U_p)}{P_A(U_p)}\right]$$

$$= \int dU_p \int d\boldsymbol{X} P_B(\boldsymbol{X}) P_B(U_p|\boldsymbol{X}) \ln\left[\frac{P_B(\boldsymbol{X})}{P_A(\boldsymbol{X})}\right]$$

$$= \int d\boldsymbol{X} P_B(\boldsymbol{X}) \ln\left[\frac{P_B(\boldsymbol{X})}{P_A(\boldsymbol{X})}\right] \qquad (3.4)$$

Here $P_B(U_p|\boldsymbol{X}) = P_B(U_p,\boldsymbol{X})/P_B(\boldsymbol{X})$ is the conditional probability of U_p given $\boldsymbol{X}$, and the probability density $P_B(U_p)$ of observing a value U_p of the perturbation is written as the marginal distribution of the perturbation over the joint distribution $P_B(U_p,\boldsymbol{X})$ of the perturbation and the conformation.[38]

The relative entropy between the initial and the final distributions of the perturbed potential was suggested as a measurement of the phase-space overlap in FEP calculations[21] and for estimating the number of samples needed for the convergence of free energy calculations.[29] It is well known that a sufficient overlap of the initial and final distributions is crucial for the accuracy of free energy computation.[21,29,39] The interpretation of the relationship of relative entropy and dissipative entropy in eq 3.2 is of high importance as it provides a clear and a simple way to explain the relationship between the conformational changes and the change in free energy. As pointed out above, the dissipative entropy is the only entropic
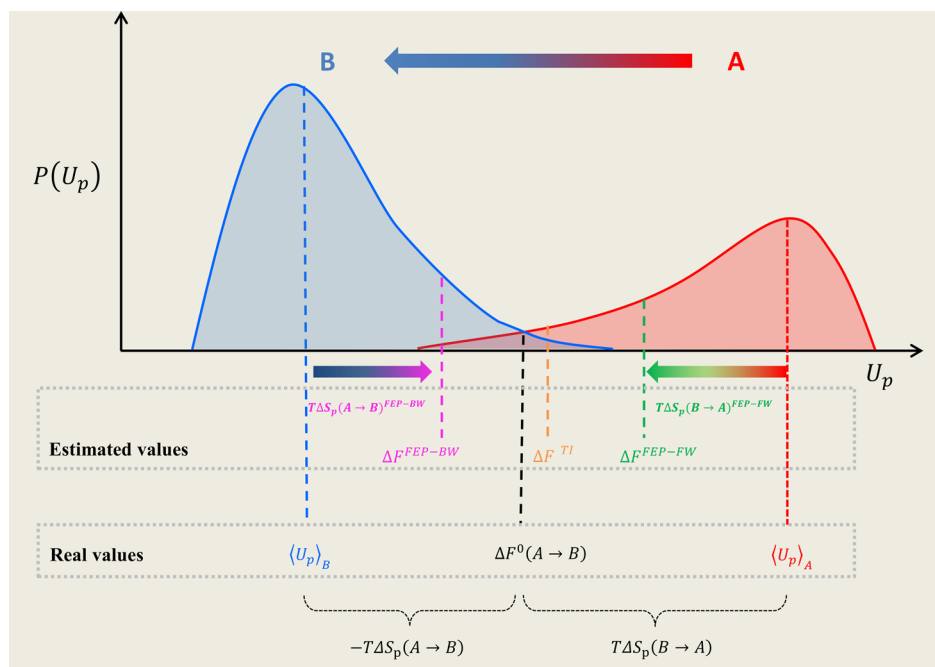
Figure 1. Schematic representation of the free energy calculation between two states A and B. The values for forward FEP (ΔF^{FEP-FW}) and backward FEP (ΔF^{FEP-BW}) are colored green and magenta, respectively. The value at the midpoint (the thermodynamic integration ΔF^{TI}) is colored orange.

term that contributes to the free energy because the rest of the entropy is known to be exactly compensated due to GEEC. Equation 3.2 shows that the conformational changes represented by the logarithmic distance $\ln[P_B(X)/P_A(X)]$ between the probabilities of the initial and final state, respectively, amount exactly to the dissipation of the external work ΔE_p; see eq 2.14. The dominating contributions to the relative entropy come from the conformations that have significantly different probabilities in the two states (conformational changes).

The relationship to the relative entropy provides several unique features of the dissipative entropy. The dissipative entropy is always nonpositive due to the positive value of the relative entropy.[28] It equals zero only if no conformational changes take place, and the initial and the final distributions are identical[28] $P_A(X) = P_B(X)$. Furthermore, it is important to notice that the dissipative entropy is not symmetric and rather depends on the direction of the reaction. When computing the free energy change from the backward direction, the dissipative entropy can be written in analogy to eqs 3.2 and 3.3 (see the derivation in the Supporting Information);

$$T\Delta S_p(B \rightarrow A) = -\frac{1}{\beta} \int dX P_A(X) \ln\left[\frac{P_A(X)}{P_B(X)}\right]$$

$$= -\frac{1}{\beta} D(A\|B) \tag{3.5}$$

Interestingly, the sum of the dissipative entropies of the forward and the backward directions is the distance between the means of the perturbed potential at the initial and the final states (see Figure 1). Using eqs 2.14 and 2.15 we get

$$T\Delta S_p(A \rightarrow B) + T\Delta S_p(B \rightarrow A) = \langle U_p \rangle_B - \langle U_p \rangle_A \tag{3.6}$$

Decomposing the Dissipative Entropy. The dissipative entropy in eqs 3.2 and 3.5 is related to the conformational changes in the system (X) through the relative entropy

computed between the distributions of the configurations of the total system X in the initial and the final states. However, in biochemical applications, one is often more interested in the role of the conformational changes of parts of the system rather than those of the full system. For example, it is of high importance to quantify the role of the conformational changes in the biomolecules independent from the solvent. Here, we show using simple properties of the relative entropy that it is straightforward to decompose the dissipative entropy into contributions from subsets of the system, such as the ones originating from the conformational changes in the biomolecules and from the solvent, while keeping the simple interpretability.

Let us consider a system of a solute R and solvent W. The configuration of the system X is defined by the joint distribution of its components $X = (R,W)$. The probability density $P_B(X)$ of observing the system in configuration X in the state B is the probability density of observing the solute configuration R and the solvent configuration W. It is given by

$$P_B(X) = P_B(R, W) = P_B(R)P_B(W|R) \tag{3.7}$$

Here, $P_B(R)$ is the probability density of observing the solute in a configuration R and $P_B(W|R)$ is the conditional probability density of observing the solvent in configuration W, given the solute in configuration R. The relative entropy $D(P_B(X) \| P_A(X))$ between the distributions of the system when the system undergoes a change from state A to B can be decomposed into the contributions from the components (R,W) using the chain rule for the relative entropy[28] and eq 3.7:

$$D(A\|B) = D(P_B(X)\|P_A(X)) = D(P_B(R)\|P_A(R))$$
$$+ D(P_B(W|R)\|P_A(W|R)) \tag{3.8}$$

Here, $D(P_B(W|R) \| P_A(W|R))$ is the conditional relative entropy of the distributions of W conditional to the distribution of R. Now, we will present examples to illustrate the straightforward interpretability of eq 3.8. Let us consider the

D

solvation of a biomolecule R in a solvent W. The solvation process is defined as the transfer of a molecule from a fixed position in an ideal gas phase (G) into a fixed position in the solvated phase (S).[40] Therefore, the translational and the rotational degrees of freedom of the solute do not contribute to the changes in free energy. The solvation free energy ΔF^* can be written using eqs 2.14 and 3.2:

$$\Delta F^* = \langle U_p(X) \rangle_S + \beta^{-1} \int dX P_S(X) \ln \left[\frac{P_S(X)}{P_G(X)} \right]$$

$$= \langle U_p(X) \rangle_S + \beta^{-1} D(P_S(X) \| P_G(X)) \qquad (3.9)$$

Here $\langle \cdots \rangle_s$ denotes the average over the canonical distribution in the solvated state. The perturbation potential $U_p(X) = U_S(X) - U_G(X)$ represents the solute–solvent interactions. $P_S(X)$ describes the probability density of configuration X in the solvated state when the perturbation $U_p(X)$ is turned on. $P_G(X)$ is the probability density in the initial state when the interactions between the solute and the solvent are turned off so that the solute is in the gas phase while the water is in its pure state. Using eq 3.8, we can write eq 3.9 as a sum of all terms contributing to the free energy:

$$\Delta F^* = \langle U_p(X) \rangle_S + \beta^{-1} D(P_S(R) \| P_G(R))$$
$$+ \beta^{-1} D(P_S(W|R) \| P_G(W|R)) \qquad (3.10)$$

$$\Delta F^* = \langle U_p(X) \rangle_S + \beta^{-1} \int P_S(R) \ln \left[\frac{P_S(R)}{P_G(R)} \right] dR$$
$$+ \beta^{-1} \int P_S(R, W) \ln \left[\frac{P_S(W|R)}{P_G(W|R)} \right] dW dR$$

The same equation is derived in the Supporting Information using the traditional statistical thermodynamic formalism. The comparison between the straightforward derivations of eqs 3.10 and the lengthy derivation using the statistical thermodynamic formalism shows the power of using the properties of the relative entropy. In the following, we will discuss the individual terms in eq 3.10 and their simple interpretations.

The dissipative entropy due to the conformational changes in the solute is also related to the relative entropy:

$$D(P_S(R) \| P_G(R)) = \int P_S(R) \ln \left[\frac{P_S(R)}{P_G(R)} \right] dR \qquad (3.11)$$

$$\ln \left[\frac{P_S(R)}{P_G(R)} \right] = \beta \Delta F^* - \beta \Delta F_R^*$$

$$\Delta F_R^* - \Delta F^* = kT \ln \left[\frac{P_G(R)}{P_S(R)} \right] \qquad (3.12)$$

Equation 3.12, which is derived in the Supporting Information, provides a clear interpretation of the contribution to the dissipative entropy from a given fixed conformation of the solute R, which is related to the difference in the free energy of solvating a solute in a fixed conformation ΔF_R^* (fixed R) and the free energy of solvating the flexible solute ΔF^* with the conformational changes allowed. The contributions from the different conformations are averaged using the probability distribution of the final state S.

The contributions to the dissipation due to the solvent are given as the conditional relative entropy where $\int P_S(R,W)$ $\ln[P_S(W|R)/P_G(W|R)] dW$ is the dissipation due to solvent reorganization given a conformation R of the solute. The total contribution to the dissipation from solvent reorganization is the average over all conformations of the solute:

$$D(P_S(W|R) \| P_G(W|R)) = \int P_S(R, W) \ln \left[\frac{P_S(W|R)}{P_G(W|R)} \right] dW dR \qquad (3.13)$$

$$\ln \left[\frac{P_S(W|R)}{P_G(W|R)} \right] = \beta \Delta F_R^* - \beta U_p(R, W) \qquad (3.14)$$

Here $U_p(R,W) = U_p(X)$ is the perturbation of the total configuration $X = (R,W)$ upon the solvation process. Equation 3.14 is also derived in the Supporting Information and provides a clear interpretation of the dissipation due to water reorganization upon solvating a fixed conformation of the solute where the dissipation of a given fixed configuration of the solvent given the fixed conformation of the solute $W|R$ is related to the difference in the solute–solvent interactions $U_p(R,W)$ of configuration $X = (R,W)$ and the solvation free energy when the water is allowed to reorganize while the solute is still fixed ΔF_R^*.

It is important to notice that the decomposition of the total configuration of the system X into two components R and W is arbitrary. Hence, R can be chosen to be a part of the solute or its degrees of freedom (for example, one or a group of dihedral angles), and W will comprise the remaining degrees of freedom. It is also interesting to show that if we only consider a part of the system (R for example), then the contribution to the dissipative entropy from the rest of the system (W) is buried implicitly in the free energy changes of the individual conformations of the considered subset (ΔF_R^*). More details are provided in the Supporting Information:

$$\Delta F^* = \int P_S(R) \Delta F_R^* dR + \beta^{-1} \int P_S(R) \ln \left[\frac{P_S(R)}{P_G(R)} \right] dR$$

$$\Delta F_R^* = \int P_S(W|R) U_p(R, W) dW$$
$$+ \beta^{-1} \int P_S(W|R) \ln \left[\frac{P_S(W|R)}{P_G(W|R)} \right] dW \qquad (3.15)$$

Interestingly, the conditional relative entropy in eq 3.8 can be written in the form

$$D(P_S(W|R) \| P_G(W|R)) = \int P_S(R, W) \ln \left[\frac{P_S(W|R)}{P_G(W|R)} \right] dW dR$$

$$= \int P_S(R, W) \ln \left[\frac{P_S(W|R)}{P_G(W)} \right] dW dR$$

$$= \int P_S(R, W) \ln \left[\frac{P_S(R, W)}{P_S(R) P_S(W)} \frac{P_S(W)}{P_G(W)} \right] dW dR$$

$$= D(P_S(W) \| P_G(W)) + I_S(R, W) \qquad (3.16)$$

Here, $I_S(R,W) = \int P_S(R,W) \ln[P_S(R,W)/P_S(R)P_S(W)] dW dR$ is the mutual information between R and W in the solvated state.

$P_G(W|R) = P_G(W)$ because the conformations of W and R are independent in the initial (gas) state.

Similarly to eqs 3.8 and 3.16, we can also decompose the contributions of the conformational changes in a protein–ligand complex to the dissipation of their association (bound/free) into contributions from the protein R and from the ligand L:

$$D(P_{bound}(R, L) \| P_{free}(R, L)) = D(P_{bound}(R) \| P_{free}(R))$$
$$+ D(P_{bound}(L|R) \| P_{free}(L|R))$$

$$= D(P_{bound}(R) \| P_{free}(R))$$
$$+ D(P_{bound}(L) \| P_{free}(L))$$
$$+ I_{bound}(R, L)$$

$$(3.17)$$

Here $P_{free}(L|R) = P_{free}(L)$ because the conformations of the ligand and the protein are independent in the unbound state.

4. DISSIPATIVE ENTROPY IS AT THE ROOT OF THE DIFFICULTIES IN FREE ENERGY CALCULATIONS

In the following, we will discuss the relationship between the dissipative entropy and the two fundamental equations for free energy calculations: the free energy perturbation[3] (FEP) and thermodynamic integration[41] (TI). Although both the FEP[3] and TI[41] methods have been well established for a long time, the application of these methods is still hampered by the high computational cost involved. The importance of the overlap between the phase spaces of the initial and final states was recognized early on,[39,42] and the introduction of intermediate states was suggested to improve the overlap between the two distributions.[39] The analysis of the error in free energy calculations progressed considerably in the last years[29,43−45] in particular with respect to understanding the systematic bias of FEP methods due to insufficient sampling of the points providing large contributions to the free energy change.[29,43−45] The applications of the free energy calculation methods require the introduction of a pathway of nonphysical intermediates with overlapping phase-spaces.[37] As pointed out above, the relative entropy was suggested as a measure for the phase space overlap from the theoretical point of view.[21] In the following, we will use the exponential average equation of the relative entropy (eq 2.12) to provide a simple pictorial illustration of the impact of the dissipative entropy on the performance of FEP and the source of error from a thermodynamic point of view. First, we decompose the thermodynamic components in the FEP by recalling the forward FEP eqs 2.4 and using eq 2.15:

$$\Delta F^0(A \to B) = -\beta^{-1} \ln \langle \exp[-\beta U_p] \rangle_A$$

$$= -\beta^{-1} \ln \langle \exp[\beta \langle U_p \rangle_A - \beta \langle U_p \rangle_A - \beta U_p] \rangle_A$$

$$= \langle U_p \rangle_A - \beta^{-1} \ln \langle \exp[\beta \langle U_p \rangle_A - \beta U_p] \rangle_A$$

$$= -\Delta E_p(B \to A) + T \Delta S_p(B \to A) \quad (4.1)$$

Equation 4.1 is equivalent to the expansion of the FEP[24] where the first linear term $\langle U_p \rangle_A = -\Delta E_p(B \to A)$ is the enthalpic part and the higher terms in the exponential term are the entropic parts:

$$\exp\left[\frac{-\Delta S_p(B \to A)}{k}\right] = \langle \exp[\beta \langle U_p \rangle_A - \beta U_p] \rangle_A \quad (4.2)$$

Similarly, the backward FEP 2.5 reads:

$$-\Delta F^0(B \to A) = \Delta F^0(A \to B) = \beta^{-1} \ln \langle \exp[\beta U_p] \rangle_B$$

$$= \beta^{-1} \ln \langle \exp[\beta U_p + \beta \langle U_p \rangle_B - \beta \langle U_p \rangle_B] \rangle_B$$

$$= \langle U_p \rangle_B + \beta^{-1} \ln \langle \exp[\beta U_p - \beta \langle U_p \rangle_B] \rangle_B$$

$$= \Delta E_p(A \to B) - T \Delta S_p(A \to B)$$

$$(4.3)$$

$$\exp\left[\frac{-\Delta S_p(A \to B)}{k}\right] = \langle \exp[\beta U_p - \beta \langle U_p \rangle_B] \rangle_B \quad (4.4)$$

Equations 4.1 and 4.3 present the considered thermodynamic changes in FEP calculations. Interestingly, the FEP calculation is equivalent to the calculation of the thermodynamic changes in the reverse direction where the forward calculation $(A \to B)$ in eq 4.1 accounts for the thermodynamic changes of the reverse direction $(B \to A)$ and vice versa. Clearly, the FEP computes the thermodynamic changes due to the perturbation and avoids the unperturbed parts which are canceled out in the free energy changes 2.11.

The dissection of the FEP equations into thermodynamic components can be used to show that the dissipative entropy is behind the difficulties in FEP free energy calculations. Figure 1 illustrates a free energy calculation between two states A and B that are represented by the probability distributions of the perturbation U_p. First, we consider the forward FEP 4.1 which provides an estimation of the free energy changes through the exponential averaging $\langle \exp[-\beta U_p] \rangle_A$ by sampling from the initial distribution A. If the full conformational space was sampled, the FEP calculation would, in principle, return the exact free energy difference. The systematic bias in the FEP estimator $\langle \exp[-\beta U_p] \rangle_A$ due to the insufficient sampling is well recognized:[29,43−45] the important points for the exponential average when sampling from one distribution (A) have low probability in this state (A) and are located close to the bulk of the second distribution (B).[29] A substantial overlap between distributions A and B is required to overcome the systematic error due to insufficient sampling.[29] The expanded terms in 4.1 provide a simpler explanation of this situation. The mathematical average of the first "enthalpic" linear term $\langle U_p \rangle_A$ is the mean value of the perturbed potential over the distribution and is easily estimated compared to the exponential "entropic" terms $-kT \ln \langle \exp[\beta \langle U_p \rangle_A - \beta U_p] \rangle_A$. The contribution to the exponential averages from the sampled points grows exponentially with their distances from the mean of the sampled distribution $\langle U_p \rangle_A - U_p$. Unfortunately, the important points are rarely sampled introducing an underestimation of the absolute value of the entropic term. A similar behavior is observed when considering the backward FEP 4.3 when sampling from distribution B where the importance of the sampled points increases exponentially with their distances from the mean of the distribution $U_p - \langle U_p \rangle_B$ toward the bulk of the initial distribution A. Therefore, the systematically biased estimation of the dissipative entropy is behind the systematic bias in FEP estimation.[21,29,43−45]

The TI method can be shown to approximate the free energy change via the midpoint between $\langle U_p \rangle_B$ and $\langle U_p \rangle_A$. Consider
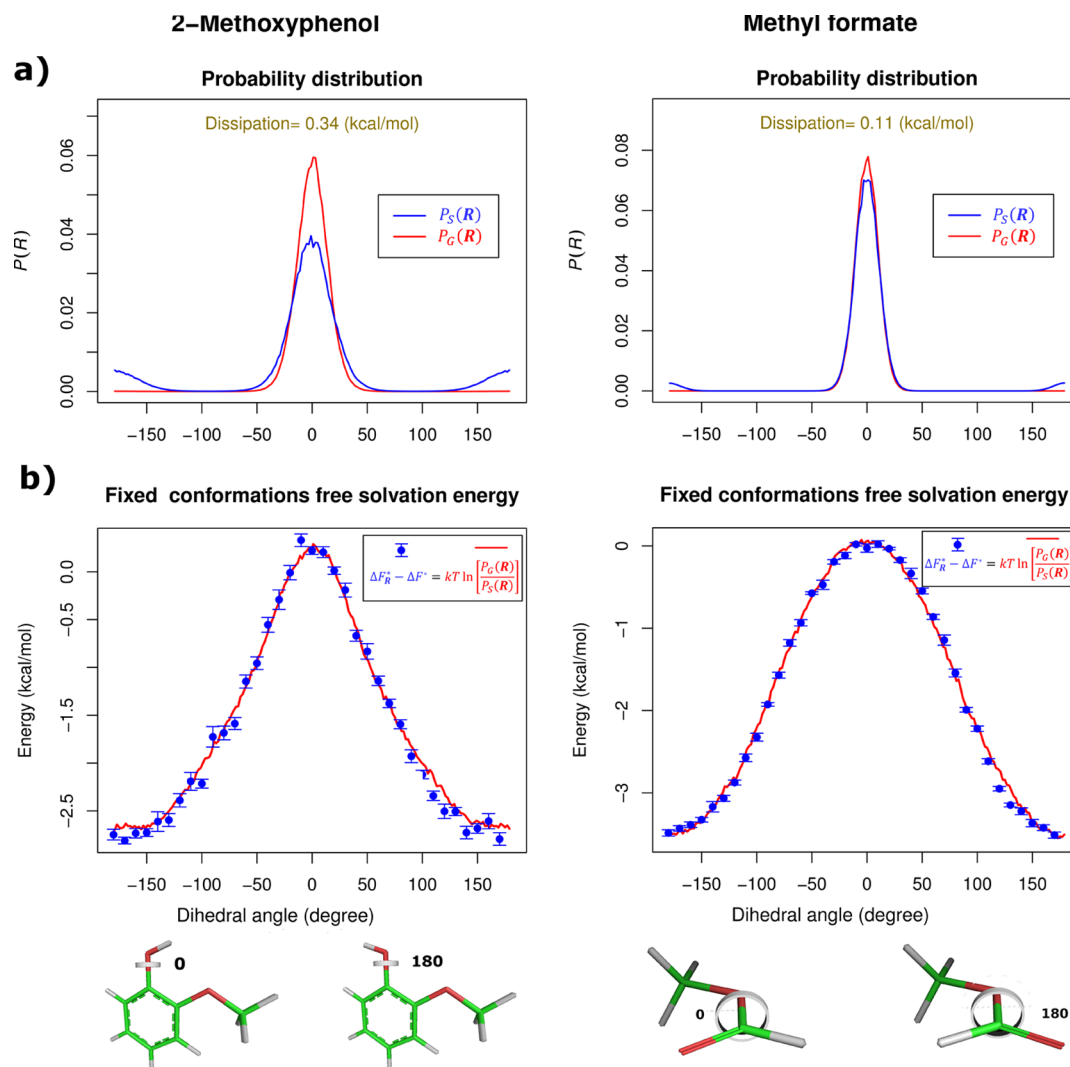
Figure 2. Dissipation due to the conformational changes and conformational dependency of solvation free energies. (a) Probability distributions of the dihedral angle in the gas phase (the red line) and in solution (the blue line). The rotatable angles are indicated in the stick representations of both 2-methoxyphenol and methyl formate. The distributions were obtained via umbrella sampling MD simulations and WHAM. The dissipative entropy $kT \times D(P_S(\mathbf{R}) \parallel P_G(\mathbf{R}))$ due to the conformational changes in the distributions of the dihedral angle was computed using eq 3.11. (b) The solvation free energies were computed after fixing (via a dihedral restraint) the conformation ΔF_R^*. The differences $\Delta F_R^* - \Delta F^*$ relative to the computed free solvation energy of the flexible molecule (ΔF^*) are plotted (the blue circles) in comparison with the values (the red line) predicted from the ratios between the probabilities using eq 3.12 $\Delta F_R^* - \Delta F^* = kT \ln[P_G(\mathbf{R})/P_S(\mathbf{R})]$.

the contribution to the free energy from a given step $\lambda_n \rightarrow \lambda_{n+1}$ along the pathway between the initial state A and the final state B (see eq 2.6):

$$
\begin{aligned}
\Delta F^0(\lambda_n &\rightarrow \lambda_{n+1}) \\
&= \int_{\lambda_n}^{\lambda_{n+1}} \left\langle \frac{\partial U_\lambda(\mathbf{X})}{\partial \lambda} \right\rangle_\lambda d\lambda \\
&= \int_{\lambda_n}^{\lambda_{n+1}} \langle U_p \rangle_\lambda d\lambda \\
&\approx \frac{\Delta \lambda}{2} [\langle U_p \rangle_{\lambda_{n+1}} + \langle U_p \rangle_{\lambda_n}] \\
&= \frac{1}{2} [\langle \Delta \lambda U_p \rangle_{\lambda_{n+1}} + \langle \Delta \lambda U_p \rangle_{\lambda_n}]
\end{aligned}
\tag{4.5}
$$

The perturbation to the potential at this step $\lambda_n \rightarrow \lambda_{n+1}$ is related to the total perturbation $U_p(\mathbf{X})$ between states A and B:

$$
U_{\lambda_{n+1}}(\mathbf{X}) - U_{\lambda_n}(\mathbf{X}) = (\lambda_n \rightarrow \lambda_{n+1}) U_p = \Delta \lambda U_p
\tag{4.6}
$$

Therefore, $\langle \Delta \lambda \, U_p \rangle_{\lambda_{n+1}}$ and $\langle \Delta \lambda \, U_p \rangle_{\lambda_n}$ are the means of the perturbed potential at the considered step $\lambda_n \rightarrow \lambda_{n+1}$, and the free energy change is estimated in eq 4.5 in terms of the midpoint between the two distributions. This implies that the TI method calculates the same approximate values of the dissipative entropy in the forward and backward directions.

A general illustration of the performance of the FEP and TI methods is shown in Figure 1. The simplest rough estimation for the free energy change is given by the Gibbs−Bogoliubov inequality[46,42] which reads:

$$
\langle U_p \rangle_B \leq \Delta F^0(A \rightarrow B) \leq \langle U_p \rangle_A
\tag{4.7}
$$

This inequality states that the value of the free energy change ΔF^0 is located between the means of the two histograms of the perturbed potential in the initial and the final states. The distance, $\langle U_p \rangle_B - \langle U_p \rangle_A$, provides an estimate of the worst-case scenario in the free energy calculation and is related to the dissipative entropies of the forward and the backward

directions, $T\Delta S_p(A \rightarrow B) + T\Delta S_p(B \rightarrow A) = \langle U_p \rangle_B - \langle U_p \rangle_A$; see eq 3.6. The FEP estimation of the free energy change easily defines the mean value of the sampled histogram $\langle U_p \rangle_A$ or $\langle U_p \rangle_B$, which is the first linear "enthalpic" term in eqs 4.1 and 4.3. The rest are the nonlinear "entropic" terms which always carry a negative sign (see above) and contribute to displacing the energy estimate toward the actual value of the free energy change (see the arrows between the centers of the histograms and estimated values of ΔF^{FEP} in Figure 1). The systematic error in estimating the nonlinear term, the dissipative entropy, due to the insufficient sampling leads to an overestimation of the free energy changes in the forward calculation and to an underestimation in the backward calculation (see the deviation between the estimated and the real values of free energy in Figure 1).

5. EXAMPLES OF THE RELATIONSHIP BETWEEN THE CONFORMATIONAL CHANGES AND THE DISSIPATIVE ENTROPY

Conformational Changes and Solvation Free Energy. We have studied the relationship between the solvation free energy and the conformational changes in the solute for the two small molecules 2-methoxyphenol and methyl formate; see Figure 2. We selected these two molecules because previous authors found that due to the rotatable dihedral angles around their C−O bonds the computed solvation free energies for both molecules varied significantly for different conformations of these molecules.[47] Hence, although these two molecules will likely only undergo small conformational changes upon solvation, the possibility of computing their free solvation energy for many different conformations provides an interesting case study for obtaining insight into the relationship between the free energy changes and the conformation of the solute.

The simulations were started using the force field parameters (the general AMBER force field[48]) and initial structures of the molecules from the FreeSolv database.[47] The TIP3P water model[49] was used for the solvent. The solutes were solvated in a rectangular parallelepiped box such that water extended at least 1.2 nm beyond the solute surface, and periodic boundary conditions were used. The production simulations were performed using a leapfrog stochastic dynamics integrator[50] and a time step of 2.0 fs. Simulations were performed in the isothermal−isobaric ensemble (1 atm and 300 K) by keeping the pressure constant through a Berendsen pressure bath.[51] The particle-mesh Ewald method[52] was used to compute the long-range electrostatic interactions. A cutoff of 1.2 nm was used for computing the van der Waals interactions and short-range electrostatic interactions. The probability distributions in the gas phase and in the solution were obtained using umbrella sampling[53] along the central dihedral angle. For this, we used 36 windows where the centers of the umbrella potentials were spaced by 10° (see Figure 1 in the Supporting Information), and the dihedral restraints were applied through a harmonic potential using a force constant of 200 kJ/mol/rad². The weighted histogram analysis method (WHAM) was used to obtain the unbiased distributions using a bin size of 2°.[54,55] Similar results were obtained when using a bin size of 5°. Solvation free energies were calculated from alchemical free energy simulations of decoupling the interactions between the solvent and solute using the Gromacs 4.6.5 simulation package.[56] The decoupling was performed in 10 steps for the Coulomb interactions followed by 10 steps for the decoupling of the Lennard-Jones interactions. Each simulation was

performed for 2 ns. The solvation free energy was computed for both the flexible molecule (no restraint is applied) and for the fixed conformations by applying dihedral restraints through a harmonic potential using a force constant of 600 kJ/mol/rad². The data were collected every 50 steps. The free energy changes were computed using the Bennett Acceptance Ratio method using the Gromacs g_bar tool.[56]

Figure 2a shows the probability distributions of the conformations of the studied dihedral angle in gas phase and in solution. The different contributions to the solvation free energy according to eq 3.10 are listed in Table 1. The total

Table 1. Interactions−Dissipation upon Solvation (kcal/mol)[a]

	ΔF^*	$\langle U_p \rangle_S$	$kT \times D(P_S(\boldsymbol{R}) \| P_G(\boldsymbol{R}))$	$kT \times D(S \| G)$	$kT \times D(P_S(\boldsymbol{W}| \boldsymbol{R}) \| P_G(\boldsymbol{W}| \boldsymbol{R}))$
methyl-formate	−3.12	−15.78	0.11	12.66	12.55
2-methoxyphenol	−4.51	−24.83	0.34	20.32	19.98

[a]ΔF^* is the free solvation energy, and $\langle U_p \rangle_S$ is the average of the solute−solvent interactions. $kT \times D(S \| G) = kT \times D(P_S(\boldsymbol{X}) \| P_G(\boldsymbol{X}))$ is the total dissipative entropy. $kT \times D(P_S(\boldsymbol{R}) \| P_G(\boldsymbol{R}))$ is the dissipation due to the solute dihedral angle conformational changes. $kT \times D(P_S(\boldsymbol{W}|\boldsymbol{R}) \| P_G(\boldsymbol{W}|\boldsymbol{R}))$ is the conditional dissipation due to the remaining conformational changes.

dissipative entropy was computed using eq 2.14. The dissipative entropy due to the conformational changes in the studied dihedral angle ($\boldsymbol{R}$) and the conditional dissipation due to the remaining conformational changes in the system which are mostly due to solvent reorganization were calculated using eq 3.10. The contributions from the conformational changes in the solute (the dihedral angle) to the dissipative entropy turned out to be quite small because the contributions of each conformation to the dissipative entropy are related to its probability in the final phase (the solution). The favorable conformation of methyl formate (0° in Figure 2) results from the potentials of the dihedral angles around the C−O bond. The favorable conformation of 2-methoxyphenol is due to the formation of an internal hydrogen bond which is expected to contribute around 5.4 kcal/mol.[57] Such large contributions are the reason for the relatively small changes in the probability distributions in spite of the large contributions from the solvation energies (up to 3 kcal/mol; see Figure 2b).

As mentioned above, the solvation free energies of 36 fixed conformations were computed for these molecules by applying a dihedral restraint to fix the conformation of the central dihedral angle at 36 different values. Figure 2b shows the differences between the computed free solvation energies of the fixed conformations ΔF_R^* and the computed solvation energy of the flexible molecules ΔF^* (when no restraint is applied) which are shown to vary up to 3 kcal/mol. Interestingly, these differences are accurately computed from the ratio of the probability distributions $kT \ln[P_G(\boldsymbol{R})/P_S(\boldsymbol{R})]$ using eq 3.12. The higher solvation free energies of the conformations (180°) indicate cooperative solvation of the hydrophilic group, which is expected to be due to the water bridging between them.[58] Discussing the role of water mediation is beyond the scope of this work. These examples show how to consider a part of the degrees of freedom ($\boldsymbol{R}$ is here the dihedral angle) when studying biomolecular reactions and how the remaining degrees
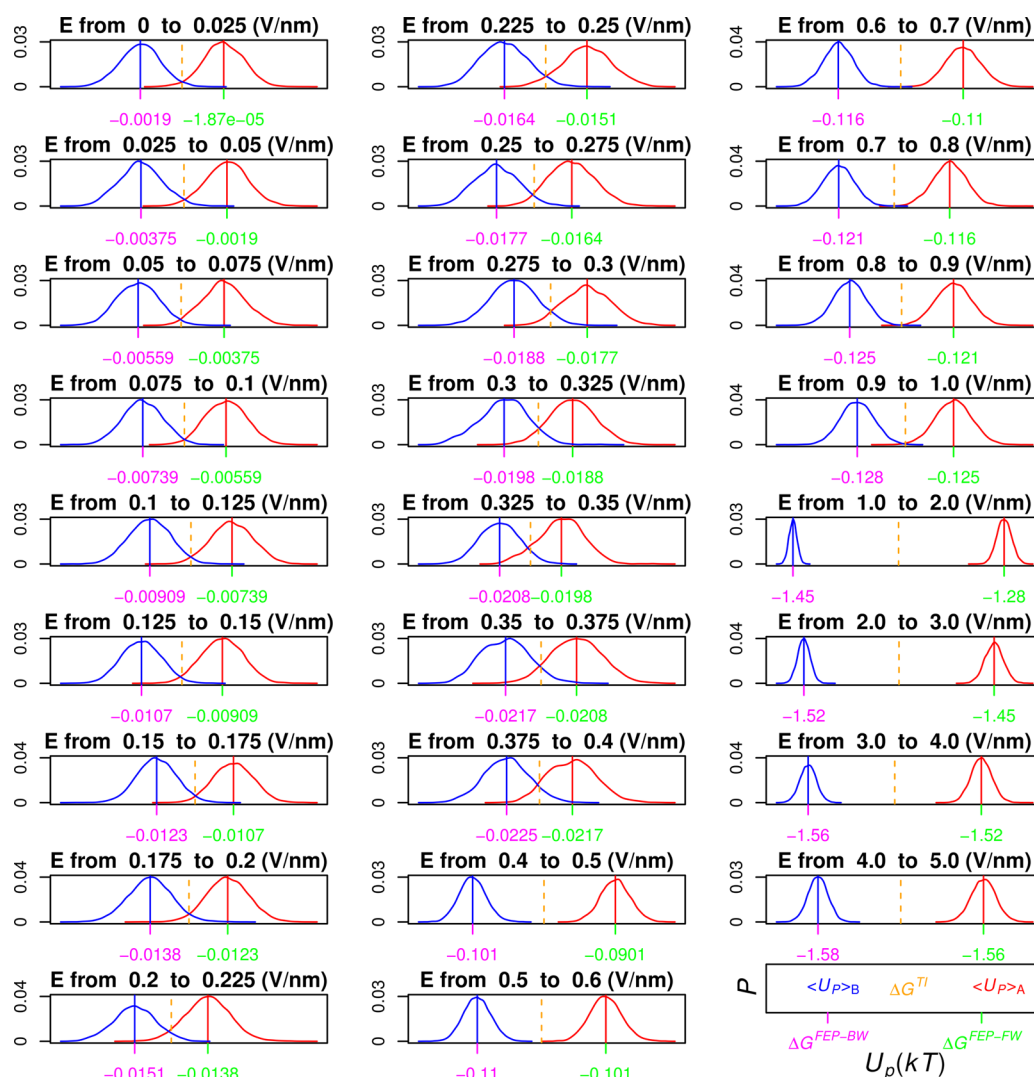
Figure 3. Calculation of the free energy changes in water due to an external electric field. The forward FEP (ΔF^{FEP-FW}) and the backward FEP (ΔF^{FEP-BW}) values (per step) are colored green and magenta, respectively. The value at the midpoint (the thermodynamic integration ΔF^{TI}) is colored orange. The red and blue straight lines represent the averages of the perturbed potential over the distributions A and B, respectively. The distributions were plotted using the kernel density plot function in R.[59] The represented values of free energy changes correspond to the step of changing the intensity of the electric field. The analysis was performed using snapshots collected every 0.5 ps over 4 ns of simulation times after skipping the first 0.25 ns.

of freedom (W) can be considered implicitly in the fixed conformation free energy changes.

Water in an External Electric Field. In the following, we will illustrate the relationship between the dissipative entropy and the fundamental equations of free energy calculation (FEP and TI). Examples from molecular dynamics simulations of a water box in an external electric field are presented. We selected this system due to the possibility of describing the conformational changes of the system in terms of the changes in the total dipole moment as a response to the perturbation (the external field). The simulations of the bulk water in an external electric field were performed using a rectangular box of bulk water of dimensions ($10.3 \times 5.5 \times 4.8$ nm^3) which contains 8998 water molecules. Twenty-seven MD simulations were performed using different intensities of the external field (0 to 5 V/nm). The external electric field was applied homogeneously in the direction of the y-axis. The TIP4P water model[49] was used for the explicit solvent. All simulations were run with the Gromacs 4.6.5 simulation package.[56] At first, the systems were energeti-

cally minimized using 500 steps of steepest-descent energy minimization. Then, the NPT ensemble [1 atm and 310 K] was applied through the Berendsen external bath[51] using coupling times of 0.1 ps^{-1} and 1 ps^{-1} for the temperature and the pressure, respectively. The particle-mesh Ewald method[52] was used for the long-range electrostatic interactions. van der Waals interactions were computed within a 1.2 nm cutoff. A time step of 2.0 fs was used. The analysis was performed using snapshots collected every 0.5 ps over 4 ns of simulation times after skipping the first 0.25 ns.

The perturbation in the potential for a given conformation X is related to the intensity of the external field E and to the total dipole moment M_y along the direction of the field: $U_p(E,X) = -EM_y(X)$. Therefore, the conformational changes in the system defined by M_y can be represented using the corresponding value of the perturbation in the potential U_p. The perturbation of the potential due to increasing the electric field from $E_1 \rightarrow E_2$ is given by

$$U_p(E_1 \rightarrow E_2) = -M_y E_2 - (-M_y E_1) = -M_y(E_2 - E_1)$$

$$(5.1)$$

Here M_y is the total dipole moment of the system along the y-axis.

Figure 3 represents different steps of free energy changes in the bulk water as the intensity of the external electric field is increased (0 to 5 V/nm). The free energy change may be roughly estimated visually from the overlap between the distributions of the perturbed potential since $\Delta F^0 = U_p$ when the two distributions are equal;[37] see eq 3.1. The means of the perturbed potentials $\langle U_p \rangle_A$ and $\langle U_p \rangle_B$ for the initial and final distributions of each step are represented by red and blue vertical lines. The free energy changes at each step of changing the field intensity are computed using different methods. The estimated value of the dissipative entropy equals the deviation of the estimated value of ΔF^0 from the corresponding mean $\langle U_p \rangle$ toward the actual values of ΔF^0 (see Figure 1). The estimated values of ΔF^0 calculated by the forward FEP, backward FEP, and the TI methods are shown in green, magenta, and orange, respectively. In all steps, the estimated values of ΔF^0 using the FEP calculations are very close to the corresponding mean $\langle U_p \rangle$ indicating difficulties in the calculation of the dissipative entropy. Interestingly, the FEP calculations fail to estimate the dissipative entropy in spite of the visually considerable overlap between the distributions in many steps (see Figure 3). However, the shapes of the distributions, which are close to Gaussian, and the small standard deviations (the first nonlinear term) can explain the failure of the exponential averaging for computing the dissipative entropy; see eqs 4.2 and 4.4. The values resulting from TI which approximate the dissipative entropy using the midpoint values appear to be more accurate than FEP in these examples.

The dissipative ability of a thermodynamic system in response to the work performed by an external perturbation can be represented by the portion of the external interactions $\Delta E_p(A \rightarrow B) = \langle U_p \rangle_B$ that is dissipated and not converted to free energy. Thus, we calculated the ratio $\Delta F^0/\Delta E_p$, which quantifies the conversion of the perturbation to free energy (the productivity of the perturbation) for all the steps of changing the electric field intensity. Figure 4 shows that the conformational changes, quantified by the change in the total dipole moment, decrease when the intensity of the electric field is increased, and they saturate at high intensity of the electric field. The ratio $\Delta F^0/\Delta E_p$ is low when the conformational changes are taking place at the low values of electric field intensity. The ratio approaches 1 when the conformational changes are small, the dissipative resistance of the system is lost, and almost all external work by the perturbation is converted to free energy (high productivity). This point is made clearer by Figure 5 which shows the loss of the dissipative ability of the system as we increase the perturbation in five equal-size steps by increasing the intensity of the electric field. The same horizontal scale was used in all steps. The dissipative ability can be estimated from the distance between the means $\langle U_p \rangle_B - \langle U_p \rangle_A$ of the initial and final distributions at each step; see eq 3.6. The dissipative ability of the system is very high in the first step in comparison to the following steps. The loss of the dissipative ability of the system at high values of the perturbation is of apparent advantage for the estimation of the free energy changes (Figure 5). The dissipative entropic changes are small at these steps as indicated by the distance
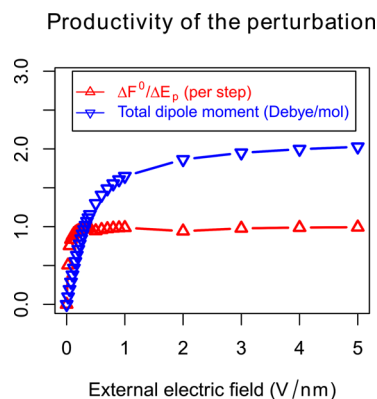

Figure 4. Productivity of the external perturbation. The red line represents the ratio $\Delta F^0/\Delta E_p$ using the values of the corresponding step of changing the intensity of the perturbation (the electric field). The blue line represents the change of the total dipole moment as a measurement of the conformational changes due to the perturbation. ΔF^0 values are estimated using the thermodynamic integration method.
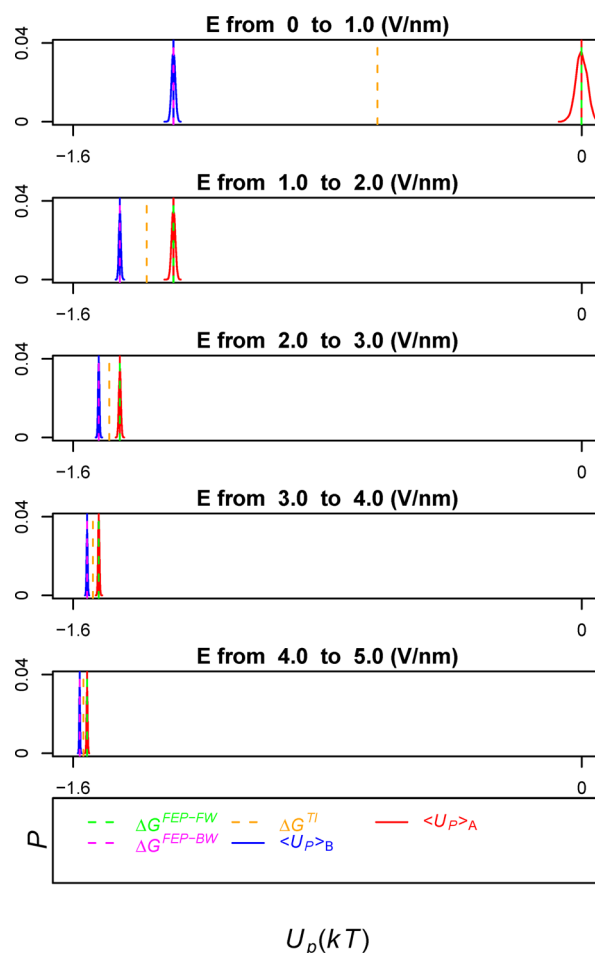

Figure 5. Loss of the dissipative ability of water for increased perturbations (external electric field) through five equal levels. The free energy changes are shown for the corresponding step.

between the histogram, and the error estimated for the values of ΔF^0 is small with both FEP and TI methods in spite of the poor overlap between the distributions (see Figure 3).
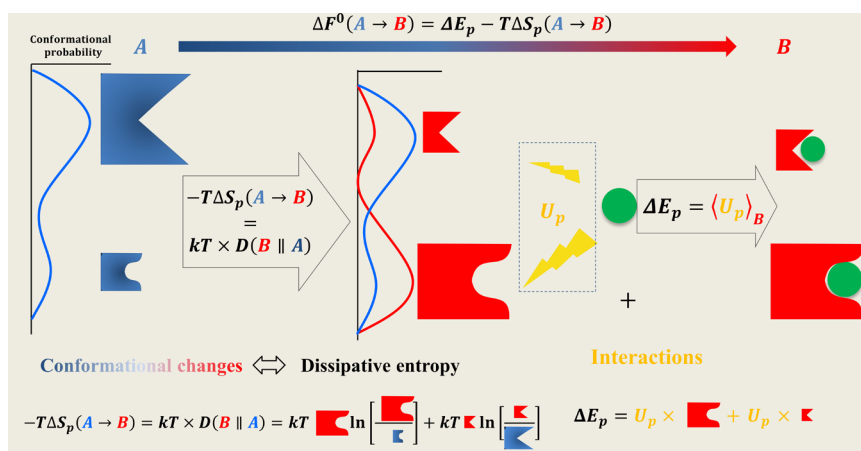
Figure 6. Interpretation of the dissipative entropy. The free energy change between two states A (blue colored) and B (red colored) is decomposed into two terms. The dissipative entropy $-T\Delta S_p$ is the free energy of changing the distributions of the conformations from state A to state B. The dissipative entropy is the negative of the relative entropy between the distributions of the conformations in the initial (the blue line) and the final states (the red line). The conformational changes are induced by the external interactions (yellow colored) and are represented by the size of the shape which accounts for the probabilities of corresponding conformations. The external interactions $\Delta E_p = \langle U_p \rangle_B$ are averaged over the new distribution B. The probability of a conformation is related to the size of its symbol.

6. CONCLUSIONS

In this study, we presented a quantitative relationship between the molecular conformational changes and the dissipative entropy. The analogy among the dissipative entropy from the thermodynamic side, the relative entropy from the information theory side, and the conformational changes from the molecular side sheds light on the physical significance of the relative entropy as a powerful concept to improve our understanding of the role of conformational changes in biochemical reactions. The analysis presented in this work shows that the relative entropy is at the basis of understanding the quantitative relationship between the conformational changes and the free energy changes. The dissipative entropy is found to be the part of the total entropy that contributes to the free energy changes. GEEC states that the rest of the total entropy, the entropy of the unperturbed part, is exactly compensated by the corresponding changes in enthalpy. This study shows that the problems of free energy calculations are related to the quantification of the relative entropy, and this is the part that is difficult to compute. The simple and clear interpretation of the relative entropy enables us to summarize our finding about the relationship between the conformational changes and the free energy change in Figure 6, which illustrates the free energy change in a system undergoing conformational changes between two states A and B. The free energy changes can be decomposed into two different terms based on eqs 2.14 and 3.2. The first term is related to the free energy needed to induce the conformational changes between states A and B and equals the dissipative entropy $-T\Delta S_p(A \to B)$. The second term $\Delta E_p(A \to B)$ is related to the average of the interactions of the new distribution B with the external perturbation.

The introduced concept of the dissipative ability of a system in response to external interactions and its relationship to the conformational changes have the potential to improve our understanding of many biochemical processes. For example, the conformational changes in both solvent and biomolecules upon binding of a ligand to a receptor are related to the dissipative entropy that compensates the contribution of the ligand–protein interactions to the free energy changes. The dissipation

of the protein–ligand interactions is caused by the contributions from the dissipative entropy due to the conformational changes of the system and can be inferred from the difference between the measured free binding energy and the computed protein–ligand interactions. The conformational changes in the biomolecule may not be observable due to the lack of the experimentally resolved structures of the free protein or the ligand. For example, many crystal structures of the HIV-1 protease are available in complexes with inhibitors, while the structures of the free enzyme are usually not available.[60] Large conformational changes are expected in these systems upon the association of an inhibitor due to the closure of the flaps around the pocket.[61] Interestingly, the thermodynamic measurements for these systems showed considerable compensation between the enthalpy and the entropy.[15] The interaction–dissipation model is a promising tool for relating the conformational changes of the receptor and the ligand to the dissipation of the interaction energy, which controls the efficiency of utilizing the ligand–protein interactions in improving the ligand affinity to the protein. The straightforward decomposition of the contributions of the parts of the system to the dissipative entropy shows the power of using the well-established properties of relative entropy in information theory to provide a simple interpretation of the conformational changes.

◼ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00235.

◼ AUTHOR INFORMATION

Corresponding Author
*E-mail: kalinina@mpi-inf.mpg.de.

Notes
The authors declare no competing financial interest.

◼ REFERENCES

(1) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303* (5665), 1813−1818.

(2) Chipot, C.; Mark, A. E.; Pande, V. S.; Simonson, T. Applications of Free Energy Calculations to Chemistry and Biology. In *Free Energy Calculations*; Chipot, C., Pohorille, A., Eds.; Springer Series in Chemical Physics; Springer: Berlin, Germany, 2007; pp 463−501.

(3) Jorgensen, W. L. Free Energy Changes in Solution. In *Encyclopedia of Computational Chemistry*; John Wiley & Sons, Ltd.: New York, 2002.

(4) Gilson, M. K.; Zhou, H.-X. Calculation of Protein-Ligand Binding Affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21−42.

(5) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The Ensemble Nature of Allostery. *Nature* **2014**, *508* (7496), 331−339.

(6) Tzeng, S.-R.; Kalodimos, C. G. Protein Activity Regulation by Conformational Entropy. *Nature* **2012**, *488* (7410), 236−240.

(7) Cembran, A.; Kim, J.; Gao, J.; Veglia, G. NMR Mapping of Protein Conformational Landscapes Using Coordinated Behavior of Chemical Shifts upon Ligand Binding. *Phys. Chem. Chem. Phys.* **2014**, *16* (14), 6508−6518.

(8) Veglia, G.; Cembran, A. Role of Conformational Entropy in the Activity and Regulation of the Catalytic Subunit of Protein Kinase A. *FEBS J.* **2013**, *280* (22), 5608−5615.

(9) Wand, A. J. The Dark Energy of Proteins Comes to Light: Conformational Entropy and Its Role in Protein Function Revealed by NMR Relaxation. *Curr. Opin. Struct. Biol.* **2013**, *23* (1), 75−81.

(10) Forman-Kay, J. D. The "Dynamics" in the Thermodynamics of Binding. *Nat. Struct. Mol. Biol.* **1999**, *6* (12), 1086−1087.

(11) Zídek, L.; Novotny, M. V.; Stone, M. J. Increased Protein Backbone Conformational Entropy upon Hydrophobic Ligand Binding. *Nat. Struct. Mol. Biol.* **1999**, *6* (12), 1118−1121.

(12) Chodera, J. D.; Mobley, D. L. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annu. Rev. Biophys.* **2013**, *42* (1), 121−142.

(13) Fenley, A. T.; Muddana, H. S.; Gilson, M. K. Entropy−Enthalpy Transduction Caused by Conformational Shifts Can Obscure the Forces Driving Protein−Ligand Binding. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (49), 20006−20011.

(14) Lafont, V.; Armstrong, A. A.; Ohtaka, H.; Kiso, Y.; Mario Amzel, L.; Freire, E. Compensating Enthalpic and Entropic Changes Hinder Binding Affinity Optimization. *Chem. Biol. Drug Des.* **2007**, *69* (6), 413−422.

(15) King, N. M.; Prabu-Jeyabalan, M.; Bandaranayake, R. M.; Nalam, M. N. L.; Nalivaika, E. A.; Özen, A.; Haliloğlu, T.; Yılmaz, N. K.; Schiffer, C. A. Extreme Entropy−Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease. *ACS Chem. Biol.* **2012**, *7* (9), 1536−1546.

(16) DeLorbe, J. E.; Clements, J. H.; Whiddon, B. B.; Martin, S. F. Thermodynamic and Structural Effects of Macrocyclic Constraints in Protein−Ligand Interactions. *ACS Med. Chem. Lett.* **2010**, *1* (8), 448−452.

(17) Olsson, T. S. G.; Ladbury, J. E.; Pitt, W. R.; Williams, M. A. Extent of Enthalpy−entropy Compensation in Protein−ligand Interactions. *Protein Sci.* **2011**, *20* (9), 1607−1618.

(18) Ahmad, M.; Helms, V.; Lengauer, T.; Kalinina, O. V. Enthalpy−Entropy Compensation upon Molecular Conformational Changes. *J. Chem. Theory Comput.* **2015**, *11* (4), 1410−1418.

(19) Ben-Naim, A. Hydrophobic Interaction and Structural Changes in the Solvent. *Biopolymers* **1975**, *14* (7), 1337−1355.

(20) Yu, H.-A.; Karplus, M. A Thermodynamic Analysis of Solvation. *J. Chem. Phys.* **1988**, *89* (4), 2366−2379.

(21) Wu, D.; Kofke, D. A. Phase-Space Overlap Measures. I. Fail-Safe Bias Detection in Free Energies Calculated by Molecular Simulation. *J. Chem. Phys.* **2005**, *123* (5), 054103.

(22) Pohorille, A.; Jarzynski, C.; Chipot, C. Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* **2010**, *114* (32), 10235−10253.

(23) Jorgensen, W. L.; Thomas, L. L. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.* **2008**, *4* (6), 869−876.

(24) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22* (8), 1420−1426.

(25) Gray, C. G.; Gubbins, K. E. *Theory of Molecular Fluids*; Oxford Univ. Press: Oxford, U.K., 2011; Vol. 2.

(26) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3* (5), 300−313.

(27) Crooks, G. E. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.* **1998**, *90* (5−6), 1481−1487.

(28) Cover, T. M.; Thomas, J. A. *Elements of Information Theory*, 2nd ed..; Wiley-Interscience: Hoboken, NJ, 2006.

(29) Jarzynski, C. Rare Events and the Convergence of Exponentially Averaged Work Values. *Phys. Rev. E* **2006**, *73* (4), 046105.

(30) Qian, H. Relative Entropy: Free Energy Associated with Equilibrium Fluctuations and Nonequilibrium Deviations. *Phys. Rev. E* **2001**, *63* (4), 042103.

(31) Gaspard, P. Time-Reversed Dynamical Entropy and Irreversibility in Markovian Random Processes. *J. Stat. Phys.* **2004**, *117* (3−4), 599−615.

(32) Crooks, G. Beyond Boltzmann-Gibbs Statistics: Maximum Entropy Hyperensembles out of Equilibrium. *Phys. Rev. E* **2007**, *75* (4).

(33) Feng, E.; Crooks, G. Length of Time's Arrow. *Phys. Rev. Lett.* **2008**, *101* (9).

(34) Still, S.; Sivak, D. A.; Bell, A. J.; Crooks, G. E. Thermodynamics of Prediction. *Phys. Rev. Lett.* **2012**, *109* (12).

(35) Sivak, D. A.; Crooks, G. E. Near-Equilibrium Measurements of Nonequilibrium Free Energy. *Phys. Rev. Lett.* **2012**, *108* (15).

(36) McClendon, C. L.; Hua, L.; Barreiro, A.; Jacobson, M. P. Comparing Conformational Ensembles Using the Kullback-Leibler Divergence Expansion. *J. Chem. Theory Comput.* **2012**, *8* (6), 2115−2126.

(37) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: New York, 2001.

(38) Wasserman, L. *All of Statistics: A Concise Course in Statistical Inference*; Springer: New York, 2004.

(39) Valleau, J. P. Monte Carlo Estimation of the Free Energy by Multistage Sampling. *J. Chem. Phys.* **1972**, *57* (12), 5457.

(40) Ben-Naim, A.; Marcus, Y. Solvation Thermodynamics of Nonionic Solutes. *J. Chem. Phys.* **1984**, *81* (4), 2016−2027.

(41) Kollman, P. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93* (7), 2395−2417.

(42) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22* (2), 245−268.

(43) Zuckerman, D. M.; Woolf, T. B. Theory of a Systematic Computational Error in Free Energy Differences. *Phys. Rev. Lett.* **2002**, *89* (18), 180602.

(44) Lu, N.; Kofke, D. A. Accuracy of Free-Energy Perturbation Calculations in Molecular Simulation. I. Modeling. *J. Chem. Phys.* **2001**, *114* (17), 7303−7311.

(45) Gore, J.; Ritort, F.; Bustamante, C. Bias and Error in Estimates of Equilibrium Free-Energy Differences from Nonequilibrium Measurements. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (22), 12564−12569.

(46) Singer, K.; Smith, W. R. Perturbation Theory in Classical Statistical Mechanics of Fluids. In *Statistical Mechanics*; Royal Society of Chemistry: Cambridge, U.K., **1973**; Vol. *1*, pp 71−133.

(47) Mobley, D. L.; Guthrie, J. P. FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. *J. Comput.-Aided Mol. Des.* **2014**, *28* (7), 711−720.

(48) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157−1174.

(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(50) Goga, N.; Rzepiela, A. J.; de Vries, A. H.; Marrink, S. J.; Berendsen, H. J. C. Efficient Algorithms for Langevin and DPD Dynamics. *J. Chem. Theory Comput.* **2012**, *8* (10), 3637−3649.

(51) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.

(52) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.

(53) Torrie, G. M.; Valleau, J. P. Monte Carlo Study of a Phase-Separating Liquid Mixture by Umbrella Sampling. *J. Chem. Phys.* **1977**, *66* (4), 1402−1408.

(54) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011−1021.

(55) Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135* (1), 40−57.

(56) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701−1718.

(57) Abraham, R. J.; Mobli, M. An NMR, IR and Theoretical Investigation of 1H Chemical Shifts and Hydrogen Bonding in Phenols. *Magn. Reson. Chem.* **2007**, *45* (10), 865−877.

(58) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem.* **1994**, *98* (8), 2198−2202.

(59) R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.

(60) Hornak, V.; Simmerling, C. Targeting Structural Flexibility in HIV-1 Protease Inhibitor Binding. *Drug Discovery Today* **2007**, *12* (3−4), 132−138.

(61) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. HIV-1 Protease Flaps Spontaneously Open and Reclose in Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (4), 915−920.