Correlation as a Determinant of Configurational Entropy in Supramolecular and Protein Systems

Andrew T. Fenley,[†,∥] Benjamin J. Killian,[‡,∥] Vladimir Hnizdo,[*,§] Adam Fedorowicz,[§] Dan S. Sharp,[§] and Michael K. Gilson[*,†]

[†]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093, United States

[‡]Department of Chemistry, University of Florida, 214 Leigh Hall, Gainesville, Florida 32611, United States

[§]National Institute for Occupational Safety and Health, 1095 Willowdale Road, Morgantown, West Virginia 26505, United States

ABSTRACT: For biomolecules in solution, changes in configurational entropy are thought to contribute substantially to the free energies of processes like binding and conformational change. In principle, the configurational entropy can be strongly affected by pairwise and higher-order correlations among conformational degrees of freedom. However, the literature offers mixed perspectives regarding the contributions that changes in correlations make to changes in configurational entropy for such processes. Here we take advantage of powerful techniques for simulation and entropy analysis to carry out rigorous in silico studies of correlation in binding and conformational changes. In particular, we apply information-theoretic expansions of the configurational entropy to well-sampled molecular dynamics simulations of a model host−guest system and the protein bovine pancreatic trypsin inhibitor. The results bear on the interpretation of NMR data, as they indicate that changes in correlation are important determinants of entropy changes for biologically relevant processes and that changes in correlation may either balance or reinforce changes in first-order entropy. The results also highlight the importance of main-chain torsions as contributors to changes in protein configurational entropy. As simulation techniques grow in power, the mathematical techniques used here will offer new opportunities to answer challenging questions about complex molecular systems.
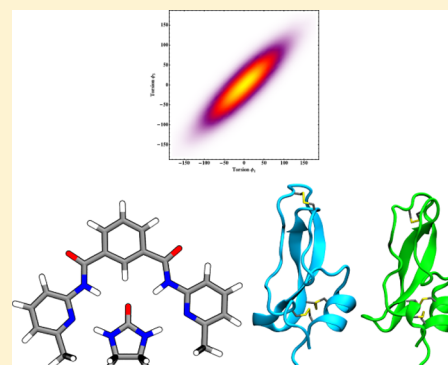
INTRODUCTION

When molecules bind in solution via noncovalent forces or proteins shift between conformational states, the associated changes in entropy can be substantial and are typically commensurate with the corresponding changes in enthalpy and free energy. As a consequence, the changes in entropy are important determinants of the equilibrium constants for such changes in state. The change in overall entropy can itself be expressed as the sum of two parts:[1,2] the change in the solvation entropy, averaged over the conformational distribution of the solute(s), and the change in the configurational entropy, which is associated with the conformational fluctuations of the solute(s). The configurational entropy depends on the overall shape and width of the joint probability density function (PDF) over the solute's internal coordinates. Prior studies indicate that changes in not only the solvent entropy but also the configurational entropy can be substantial. For example, both NMR and computational studies point to large changes in configurational entropy when proteins bind other molecules.[1,3−10]

The configurational entropy depends on the PDF of individual conformational variables, such as bond torsions, that is, on their first-order marginal PDFs; but it also depends on the correlations among these variables. In particular, one

may write the full entropy as the sum of the first-order entropy and an additional term due to correlation: $S_{\text{full}} = S_1 + S_{\text{full}}^{\text{corr}}$. Intuitively, greater correlation implies lower entropy because correlation implies less freedom to explore configurational space for a given set of marginal distributions. However, it is not yet clear whether there are any physical patterns or rules as to what types of molecular processes lead to net increases versus decreases in correlation entropy or how large the contributions from correlation are likely to be. Correlation entropy is of particular interest in relation to changes in NMR order parameters, such as when proteins change conformations or bind other molecules. These changes are often interpreted, at least semiquantitatively, in terms of changes in configurational entropy.[3,5,9−16] However, such NMR data are not directly informative about changes in correlation. As a consequence, it is of interest to consider the nature of correlation contributions to the entropy changes of interest. The identification and characterization of correlated motion is

also of broader interest, as correlations may be indicative of mechanistic couplings important in phenomena such as binding and allostery,[17−19] and knowledge of the nature and range of correlations in proteins would contribute insight into the basic mechanisms of protein motions and function.

The role of correlations as determinants of changes in configurational entropy has been the topic of a number of insightful contributions. It has been proposed, based on empirical relationships among measured NMR order parameters and binding entropies, that changes in correlation entropy on binding vary linearly with changes in measures of the entropy that neglect correlation.[9,10,15] Encouraging results have been obtained for relative binding entropies, $\Delta\Delta S$, among variants (e.g., mutants) of a given protein−protein system, and it will be interesting to learn how well such approaches work for the "absolute" entropy changes associated with binding or major conformational changes. Another notable contribution provides evidence, based on extensive molecular dynamics (MD) simulations, that inter-residue side-chain correlations lower the absolute configurational entropy by about 10−20%[16] of the first-order rotameric entropy. However, it is still of interest to probe the potential contributions from correlations involving main-chain torsions, and, again, to examine large-scale changes like binding and conformational shifts. Another contribution has argued, based on molecular simulation data, that changes in correlations do not contribute significantly to the entropy changes associated with biomolecular functions[20] and thus proposed that entropic interpretations of NMR order parameters may safely neglect any contributions from changes in correlation; and a prior study of calmodulin,[21] using quasi-harmonic analysis, reached a similar conclusion. The role of torsional correlations as determinants of binding entropy has also been addressed through application of the mutual information expansion (MIE)[22,23] to multiple long MD simulations of a protein−peptide system.[24] Interestingly, this analysis reported large entropic contributions from changes in pairwise torsion−torsion correlations, including correlations involving main-chain torsions and the six degrees of freedom (DOF) specifying the location of the peptide relative to the protein. These results would suggest that changes in torsional correlations, including those of the main chain, cannot be safely neglected. However, although this study provided a reasonably clear accounting of the strongest pairwise correlations in a large protein system, there were numerical problems for the many weak pairwise correlations. In addition, the simulations lacked the power to estimate third-order and higher correlation terms, which are also of interest. In summary, despite significant prior work, there are still important, unanswered questions regarding the role of torsional correlations as determinants of changes in configurational entropy.

Today, increasingly powerful techniques for both simulation and entropy analysis allow comparatively rigorous in silico studies of the configurational entropy changes associated with binding and conformational changes. The present study takes advantage of such techniques to revisit the important topic of correlation entropy for two molecular cases. The first is the experimentally studied association, in chloroform, of ethyleneurea with a designed host molecule (Figure 1),[25] whose small size allows correlation terms of order greater than two to be converged; the second is the 58-residue protein BPTI, where an available millisecond-duration MD simulation[26] allows calculation of convincingly converged pairwise correlation terms. This simulation of BPTI has been extensively studied
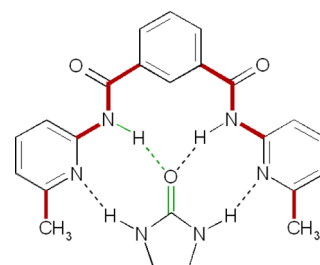

Figure 1. Host−guest complex. Bold red bonds indicate the host torsion angles for which the entropy was computed. Thin green lines define one distance variable (dashed green, H to O), two angle variables (N−H−O, H−O−C), and three torsional variables (central bonds N−H, H−O, O−C), which together define the relative position and orientation of the two molecules in their bound state. Dashed lines indicate hydrogen bonds.

in a variety of post-analysis works that include investigation of allostery,[27] isomerization rates of disulfide bonds,[28] NMR order parameters,[29] and entropy−enthalpy transduction.[30] The results presented here bear on the sign, magnitude, and determinants of correlation contributions to entropy changes and hence on the entropic interpretation of NMR order parameter data.

■ METHODS

We studied the role of correlations as determinants of entropy changes on binding and conformational change through analysis of MD simulations. The contributions of pairwise and higher-order correlations to the entropy were estimated by applying the MIE and maximum information spanning tree (MIST)[31,32] methods to the simulation trajectories. These approaches are suitable for the present purpose because they express the total entropy as a sum of terms accounting for successively higher-order correlations among the conformational variables. The following subsections detail the computational methods applied to both the host−guest and BPTI systems.

Configurational Entropy of Host−Guest Binding. *Theoretical Framework.* The configurational entropy S of a molecule or supramolecular complex of N_a atoms, at standard concentration, is given by[24,33]

$$S/k_B \equiv -\int d\vec{x}\, \rho(\vec{x})\ln[\rho(\vec{x})]$$

$$= \ln\frac{8\pi^2}{C^\circ} - \int d\vec{q}\, \tilde{\rho}(\vec{q})\ln[\tilde{\rho}(\vec{q})] + \langle \ln J(\vec{q}) \rangle \tag{1}$$

Here k_B is Boltzmann's constant, $\rho(\vec{x})$ is the PDF of the Cartesian coordinates $\vec{x} = (x_1,...,x_{3N_a})$ of the system, C° is a standard concentration (usually taken as 1 mol/L = 6.022 × 10^{-4} molecules/Å^3), $\tilde{\rho}(\vec{q}) = \rho[\vec{x}(\vec{q})]J(\vec{q})$ is the PDF of internal coordinates $\vec{q} = (q_1,...,q_{3N_a-6})$, $J(\vec{q})$ is the Jacobian of the transformation $\vec{x} \rightarrow \vec{q}$, and $\langle \cdot \rangle$ denotes an expectation value (mean). The first term in the second line of eq 1 arises from the overall translation and rotation of the system.

The change, ΔS, of the standard configurational entropy on the binding of host A and guest B to form the complex AB is then

$$\Delta S \equiv S_{AB} - (S_A + S_B)$$

$$= -k_B \ln \frac{8\pi^2}{C^\circ} + \tilde{S}_{AB} - (\tilde{S}_A + \tilde{S}_B) \qquad (2)$$

where $\tilde{S}_X$, the internal entropy of molecule X = A, B, AB, is defined by

$$\tilde{S}_X / k_B = -\int d\vec{q}\, \tilde{\rho}_X(\vec{q}) \ln[\tilde{\rho}_X(\vec{q})] + \langle \ln J_X(\vec{q}) \rangle \qquad (3)$$

which is the entropy of the PDF $\tilde{\rho}_X(\vec{q})$ of the internal coordinates of system X plus the term that arises from the Jacobian $J_X(\vec{q})$ of the transformation $\vec{x} \to \vec{q}$ from Cartesian to internal coordinates. Of the three translational−rotational terms $-k_B \ln(8\pi^2/C^\circ)$ associated with systems A, B, and AB, only one survives in the change given in eq 2. Note that an isothermal change in the configurational entropy S of a system, defined as in eq 1, equals the change in the full (i.e., spatial plus momentum) thermodynamic entropy of the system.[33] (The present notation differs from that of ref 33, where a tilde is used to indicate the internal-coordinate entropy alone.)

The first term on the right-hand side of eq 3 may be estimated by the MIE or MIST[31,32] approaches based on the Boltzmann sample of internal coordinates, $\vec{q}_i$, $i = 1,...,N_s$, afforded by a molecular simulation. The second term in eq 3 is estimated by a simple arithmetic mean,

$$\langle \ln J(\vec{q}) \rangle \equiv \int d\vec{q}\, \tilde{\rho}(\vec{q}) \ln[J(\vec{q})] \approx 1/N_s \sum_{i=1}^{N_s} \ln[J(\vec{q}_i)] \qquad (4)$$

where the sum runs over the points $\vec{q}_i$, $i = 1,...,N_s$ of the same sample.

Host−Guest Simulations and Entropy Calculations. Stiff DOF do not contribute much to entropy changes, as their PDFs change little on binding. Therefore, the present entropy analysis focused on the soft variables; that is, the eight soft torsions of the host and, for the bound complex, six additional DOF specifying the relative position and orientation of the two molecules that form the complex (Figure 1). The free guest has no soft DOF and thus was treated as an "inert" participant in the binding process so that, to a good approximation, the change in internal entropy is given by

$$\tilde{S}_{AB} - (\tilde{S}_A + \tilde{S}_B) \approx \tilde{S}'_{AB} - \tilde{S}'_A \qquad (5)$$

Here $\tilde{S}'_{AB}$ is the entropy, eq 3, of the complex computed based on only the soft torsions of the bound host and the relative position/orientation variables. S'_A is the entropy, eq 3, of the unbound host computed using soft torsions only. Thus, the configurational entropy of a total of 14 internal coordinates had to be estimated from the simulation of the bound complex and compared with the corresponding estimate of the configurational entropies of eight internal coordinates from the simulation of the unbound host. We used the generalized AMBER force field (GAFF)[34] and AM1-BCC[35,36] charges to parametrize our systems. The software package AMBER 12[37] with PMEMD GPU support[38,39] was used to run the MD simulations at constant temperature and pressure (NPT). Production simulations ran for a duration of 5 μs for both the unbound host and the bound complex, immersed in the solvent chloroform, which was treated explicitly.[40,41] The temperature was maintained at 300 K by the Langevin thermostat with the collision frequency scaling set to 1.0 ps^{-1}, and the pressure was set to 1 bar using the Berendsen barostat with a relaxation time set to 2 ps. Snapshots were saved every 1 ps, resulting in five

million sample points per simulation. The time series of torsion angles of interest were computed from the simulation data with the program CPPTRAJ.[42]

Entropies were extracted from the torsional time series by several methods. First, we applied the MIE, where instead of computing the required marginal entropies by the histogram method,[23,24] we instead used the more powerful *k*th nearest neighbor (NN) estimation of entropy;[43−45] this combined approach is termed the MIE-NN method.[46] We also directly applied the NN method to estimate the entropies of the full 8- and 14-dimensional PDFs of the host and complex, respectively, using extrapolations in time for multiple values of *k*. The extrapolation for each value of *k* used the phenomenological function form[44] $-TS_t \approx -TS_{t\to\infty} + a_k/t^p$, where a_k is fitted separately for each value of *k* but all values of *k* shared the exponent *p* and asymptote $-TS_{t\to\infty}$. Note, however, that different values of *p* and $-TS_{t\to\infty}$ were fitted for the free host and bound complex. We use the difference of both fit curves to estimate $-T\Delta S_{t\to\infty}$. Finally, we also applied the MIST approximation[31,32] in combination with the NN method. Like MIE, MIST is a systematic, dimension-reduction approximation based on an information-theoretic expansion of entropy. Unlike MIE, however, MIST approximations of increasing order are guaranteed to furnish decreasing upper bounds of the exact entropy.

Configurational Entropy of a Protein Conformational Change. For the protein analysis, we limit attention to the contributions of pairwise correlations, due to the computational expense of converging higher-order correlation terms. Even the pairwise contribution can be challenging to compute for a large system, because the second-order mutual information contributions, which appear in both the MIE and MIST, are greater than or equal to zero. Thus, even if two torsions are entirely uncorrelated, their mutual information will approach zero asymptotically from above with increased sampling. For a protein, the sum of many spurious positive correlation contributions from the many torsion−torsion pairs can yield a misleadingly large estimate of the overall pairwise contribution to the entropy. Here we address this problem by a combination of approaches.

First, to maximize convergence, we study a single, 1 ms trajectory of BPTI in explicit solvent,[26] which was generated on the Anton supercomputer.[47] This is much longer than the prior study of torsion−torsion correlations in protein−peptide binding,[24] which processed 2 μs of simulation trajectories generated by 200 independent 10 ns runs starting from the same conformation. The BPTI simulation, which provided over four million snapshots at 250 ps intervals, used the TIP4P-Ew water model,[48] and, for the protein, the AMBER ff99SB[49] force field with additional corrections to side-chain torsions of isoleucines was used. The original study of this simulation decomposed the trajectory into conformational clusters, or states, based on a kinetic clustering approach.[26] Subsequent thermodynamic analysis[30] indicated that Clusters 1 and 2 have free energies that are equal to within ∼0.5 kcal/mol, but their total entropies, including both the protein and the solvent molecules, differ by ∼3.2 kcal/mol. Their configurational entropies were estimated to differ by over 18 kcal/mol[30] based on MIST analysis. This very large difference in configurational entropy is accompanied by visibly larger conformational fluctuations for Cluster 2 versus Cluster 1. Here we expand on the prior thermodynamic analysis[30] by examining the role of torsional correlations as determinants of

Table 1. Changes in Configurational Entropy Due to Host-Guest Binding, $-T\Delta S_m$, Computed at Various Orders m = 1, 2, and 3 of the MIST and MIE Expansions and at Full Order, $-T\Delta S_{full}$, by Extrapolation of NN Results to Infinite Simulation Time (See Figure 2)[a]

method	$-T\Delta S_1$	$-T\Delta S_2$	$-T\Delta S_2^{corr}$	$-T\Delta S_3$	$-T\Delta S_3^{corr}$	$-T\Delta S_{full}$	$-T\Delta S_{full}^{corr}$
MIE-NN	6.03	8.39	2.37	8.57	2.55		
MIST-NN	6.03	7.39	1.37	7.90	1.88		
k-NN						8.43	2.40

[a]Correlation contributions at each reported order are also reported: $-T\Delta S_m^{corr} = -T\Delta S_m + T\Delta S_1$, and $-T\Delta S_{full}^{corr} = -T\Delta S_{full} + T\Delta S_1$. All values are in kilocalories per mole.

the large difference in configurational entropy between Clusters 1 and 2.

In addition to using a much longer simulation, we employ MIST instead of MIE, as the former requires postprocessing many fewer torsion−torsion pairs. This not only saves computer time but also, relative to the MIE, reduces the error from summing many small, potentially spurious pairwise contributions from large numbers of torsion−torsion pairs, as previously discussed. We also address the problem of residual pairwise contributions by repeating the calculations with a cyclic permutation technique that removes spurious entropic contributions due to inadequate sampling, as previously detailed.[30] Although the NN method allows well-converged entropy estimates to be obtained with less simulation data, we used the histogram method for this study because it allows the nearly 400 000 mutual information pairs to be processed considerably more quickly. The subtraction of spurious correlation from each MIST pair removes numeric bias[50] of the entropy estimate from the histogram method for a given cluster. Because the spurious correlation estimates are computed using the same number of frames as the original mutual information estimate for a given pair, any sample bias for a given cluster is removed before computing the difference in entropies between clusters that have different frame counts.

The full bond-angle-torsion coordinate system of BPTI comprises 889 torsion angles. Treating each of these torsions separately can lead to trivial high-order correlations;[23] we eliminated these by redefining the torsion angles (Φ_j) of all of the torsions that share the same rotatable bond as phase angles (ϕ_j) of a single, representative torsion angle (Φ_i):[51]

$$\phi_j = \Phi_j - \Phi_i \qquad (6)$$

For BPTI, 500 of the torsion angles are treated as phase angles in this manner. It should be noted that in our prior analysis of this simulation[30] we accounted only for those phase angles that had three of the four torsion atoms in common with a master torsion angle. We now also include any torsion that shares the two atoms that form the central rotatable bond. We also found very rare rotations of the NH_2 moieties within the guanidinium group of the arginine residues. The rotations are now corrected to account for their symmetry. These procedural changes caused the MIST-889c entropy estimate to slightly change from $-T\Delta S = -18.9$ to -18.1 kcal/mol (Table 2).

We also determined whether restricting the set of torsions included in the entropy calculations to only those most perturbed by the conformational change affects the entropy results. We quantified the degree to which a torsion angle's PDF changes between Clusters 1 and 2 by computing the Jensen−Shannon divergence metric (JSDM)[52,53] between the two PDFs. If P is the PDF of a torsion for Cluster 1 and Q is its PDF for Cluster 2, the JSD metric is given by

$$\text{JSDM}(P; Q) = \left[S\left(\frac{P+Q}{2}\right) - \frac{1}{2}[S(P) + S(Q)] \right]^{1/2} \qquad (7)$$

where $S()$ is the Shannon entropy of the PDF in the argument. The JSDM is zero when the two PDFs are identical and attains a maximum value of 0.83 when the two PDFs are completely nonoverlapping; we considered a torsional PDF to be significantly perturbed if its JSDM between Clusters 1 and 2 was greater than 0.083. Note that, unlike the widely known Kullback−Leibler divergence,[54] the JSDM has the merit of being unaffected by the ordering of the two PDFs. Computing the JSDM between two PDFs of the same torsion angle involves only 1-D PDFs, so we obtained it from a straightforward histogram method.

■ RESULTS

Configurational Entropy of Host−Guest Binding. When the host and guest are both free in solution, they both rotate freely relative to each other, and, for a standard concentration of 1 M, each effectively occupies its own volume of ∼1660 $Å^3$.[2,55,56] After they have bound to form a noncovalent complex, their relative rotation and translation is markedly constrained, and this reduction in rotational and translational freedom in itself contributes a loss in configurational entropy. Whether the net change in configurational entropy is unfavorable then depends on how the rest of the system responds to the binding event. For this small host−guest system, the overall change in configurational entropy is found to remain unfavorable, as $-T\Delta S_{full}$ = 8.43 kcal/mol (Table 1). (Note that this entropy change, as well as the others in this section, includes the $-k_B \ln(8\pi^2/C°)$ term in eq 2.) Most of this overall entropy change is manifested in the first-order entropy contribution, $-T\Delta S_1$ = 6.03 kcal/mol, which by definition neglects all correlation contributions. However, increased correlation clearly plays a role, given the 2.4 kcal/mol difference between the full entropy, which accounts for correlations, and the first-order entropy, which does not. Thus, we find that binding induces increased correlations, which further oppose binding. This pattern is opposite to a prior computational result, indicating that proteins with lower first-order side-chain entropies tend to have partly balancing decreases in side-chain correlations.[16]

Much of the change in entropy can be attributed to changes in the relative rotational and translational DOF of the host and guest: application of the MIE-NN method to these six key variables in the bound state yields estimates for $-T\Delta S_{RT}$ of 5.0, 5.9, and 6.1 kcal/mol, relative to the unbound state, at the first, second and third orders of the MIE, respectively. It should be noted, however, that alternative systems of internal coordinates could give different results for these quantities.[2]

The roles of pairwise and third-order correlations may be examined within both the MIE and MIST expansions (Table 1). Interestingly, the MIE at both second and third order agrees well with the estimate of the full-order entropy, indicating that, for this method, the pairwise correlations account for the bulk of the total correlations. However, it should be noted that there is no guarantee that higher order terms, if computable, would continue smoothly toward the full-order result. The MIST approach proceeds more gradually toward the full-order estimate and, at third order, already captures most of the correlation present in the full-order estimate.

The correlation contribution identified here is considerably larger than that previously reported for thermal perturbations of a series of dipeptides (<0.1 kcal/mol) and the much larger villin headpiece protein (<0.4 kcal/mol).[20] The present result corresponds to 0.17 kcal/mol change in correlation entropy per soft DOF of the complex (14 variables) or 0.30 kcal/mol per soft dihedral angle (8 dihedrals). These values may be compared with prior results for the binding of a 9-residue peptide to the 145-residue protein TSG101,[24] where changes in second-order correlation on binding led to an entropy penalty of ∼0.3 kcal/mol per residue of the protein−peptide system. Assuming an average of four soft dihedrals per residue, the correlation entropic penalty is ∼0.075 kcal/mol per soft dihedral, which is less than the present host−guest result. This smaller value probably reflects, at least in part, the fact that the protein has many DOF that are not closely coupled with the binding site.

The present results are well-converged. Thus, the MIE results for orders one, two, and three remained constant to within 0.1 kcal/mol over the simulation time interval 4 to 5 μs, and the MIST results at the third order were numerically even more stable than the corresponding third-order MIE results. It is worth noting that we also attempted to compute $-T\Delta S_m$ at orders $m \geq 4$, but convergence was not favorable, and extrapolation was not attempted, as the estimates obtained at different values of k differed from each other to a much greater extent than those in the full-dimensional k-NN estimates. It appears that these numerical problems stem in part from the much greater number of mathematical clusters of variables that must be analyzed for larger values of the order m ($m > 3$).

Configurational Entropy of a Conformational Change in BPTI. A prior 1 ms simulation of the 58-residue protein BPTI, using explicit water, yielded several different conformational clusters.[26] Subsequent thermodynamic analysis of the simulation results indicated that the total entropy of the protein−water system changes by about $-T\Delta S = -3.0$ kcal/mol on transitioning from the more crystal-structure like Cluster 1 to the more flexible Cluster 2.[30] Interestingly, the change in configurational entropy for this conformational shift is much larger: the second-order MIST estimates range from −15.1 to −18.1 kcal/mol (Table 2), depending on methodological details discussed later. Because the total entropy can be expressed as the configurational entropy plus an appropriately defined solvation entropy,[1] one may conclude that the strongly favorable increase in configurational entropy on going from Cluster 1 to Cluster 2 is largely canceled by an opposing decrease in solvent entropy. The estimated change in configurational entropy on going from Cluster 1 to Cluster 2 greatly exceeds that computed for the host−guest system (previously described). This is perhaps not surprising given the much larger size of the protein system; on the other hand, a

Table 2. Changes in Configurational Entropy when BPTI Switches from Conformational Cluster 1 to Cluster 2, Reported As $-T\Delta S$, in Kilocalories Per Mole[a]

method	$-T\Delta S_1$	$-T\Delta S_2$	$-T\Delta S_2^{corr}$
MIST-889	−21.1	−15.1	6.0
MIST-889c	−21.1	−18.1	3.0
MIST-157	−20.7	−15.4	5.4
MIST-157c	−20.7	−16.2	4.5

[a]The first set of results is computed by applying the MIST approach to all 889 protein torsions, without (MIST-889) and with (MIST-889c) a correction for possible incomplete sampling, based on cyclic permutations of the trajectory. The second set results from applying MIST to only the 157 torsions whose PDFs change most between the two clusters, based on their JSDM values. Again, results are presented without (MIST-157) and with (MIST-157c) the permutation correction. Column headers and units are as in Table 1.

change in conformational state might be expected to produce a smaller entropy change than a binding event.

It is of interest to break down the estimated change in configurational entropy between Clusters 1 and 2 into its first order and pairwise correlation contributions. At first order, the increase in torsional entropy on going from Cluster 1 to Cluster 2 is found to be about −21.1 kcal/mol (Table 2), while, as previously noted, accounting for pairwise correlations reduces this change to −15.1 to −18.1 kcal/mol. Thus, although the first-order entropy becomes much more favorable, a concurrent increase in pairwise torsional correlations effectively cancels out 3−6 kcal/mol of the first-order entropy difference. This cancelation of 15−30% of the first-order entropy by changes in correlation is in striking agreement with a prior simulation study of protein side-chain entropy, previously mentioned.[16] However, it is opposite in sense to what we observe for the host−guest system, above: there, changes in correlation reinforce, rather than balance, the first-order change in configurational entropy of binding.

We tested the robustness of the present entropy estimates by two substantial variations in the method results. First, we corrected the estimate of each pairwise mutual information used in the MIST calculations for possible spurious correlation due to inadequate sampling by subtracting out pairwise entropies computed with a permuted, and hence entirely decorrelated, trajectory.[30] As shown in Table 2, results with (MIST-889) and without (MIST-889c) this correction for possible spurious correlations due to inadequate convergence differ by only a few kilocalories per mole We also recalculated both the first- and second-order entropies using only the 157 torsions with greatest JSDMs between Clusters 1 and 2. As shown in Figure 3, the torsions with the largest JSDM values, that is, the ones whose PDFs differ most between Clusters 1 and 2, reside in the two loops toward the top of the protein in this representation. This is not surprising because increased motion of these loops is a hallmark of Cluster 2, as illustrated in Figure 1 of a prior study of this simulation.[30] The use of only 157 torsions dramatically reduces the total number of mutual information pairs (12 246 versus 394 716) that need to be calculated prior to computing the MIST estimate of the configurational entropy. Even with such a large reduction in the total number of torsions, the results are all within 2 kcal/mol of those based on all 889 torsions, for both the uncorrected (MIST-157) and permutation-corrected (MIST-157c) estimates. The robustness of the overall results to these methodological variations supports their validity.
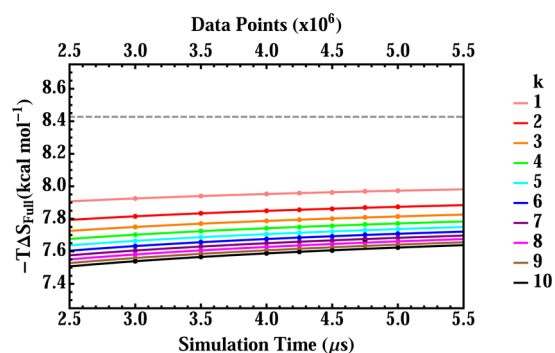
6451

dx.doi.org/10.1021/jp411588b | J. Phys. Chem. B 2014, 118, 6447−6455

Figure 2. Convergence of the full-dimensional entropy change, $-T\Delta S_{full}$, for host–guest binding, estimated by the k nearest-neighbor method, as a function of simulation time t. Each data point is the delta entropy estimate between the host and complex for a given value of t and k, and each line is the difference in the host and complex phenomenological fit functions for a given value of k, where the fits for the host and were allowed to approach their own asymptotic values as $t \to \infty$. The fitted values of the exponent p are 0.365 and 0.221, for the host and complex, respectively. The gray horizontal dashed line at $-T\Delta S_{full} = 8.43$ kcal/mol indicates delta in asymptotic values for the host and complex.
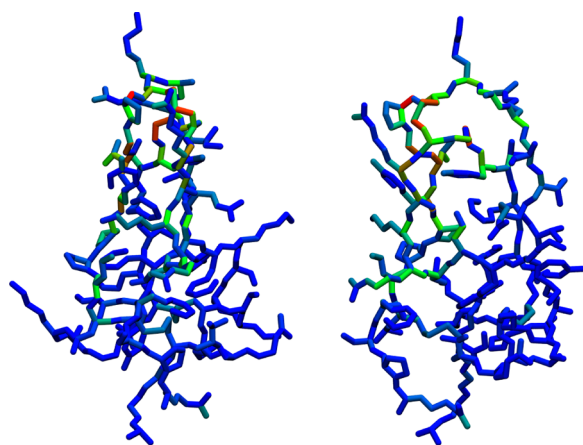

Figure 3. BPTI, from two viewpoints, with torsions colored based on their Jensen–Shannon divergence metric (JSDM) between Clusters 1 and 2. Dark blue represents the minimum possible divergence of 0.00, and red represents the maximum observed value of 0.75.

It is also of interest to determine which parts of the protein contribute most to the MIST entropy estimates. As a first level of analysis, we decomposed the change in first-order configurational entropy into main-chain and side-chain contributions and the change in correlation entropy into main-chain/main-chain, main-chain/side-chain, and side-chain/side-chain contributions. As shown in Table 3, the main-chain contribution to the first-

Table 3. Decomposition of the Configurational Entropy Change, $-T\Delta S$ (kcal/mol), between BPTI's Conformational Clusters 1 and 2 into First Order (Main-Chain (M), Side-Chain (S)) and Correlation (Main-Chain Main-Chain (MM), Main-Chain Side-Chain (MS), and Side-Chain Side-Chain (SS)) Contributions

method	M	S	MM	MS	SS
MIST-889	14.0	7.1	3.6	0.9	1.5
MIST-889c	14.0	7.1	2.4	0.6	−0.2

order change is twice the side-chain contribution, and main-chain/main-chain torsion pairs similarly provide the largest contribution to the change in correlation entropy, especially after the permutation corrections are applied. These results suggest that changes in main-chain torsions play a predominant role in determining changes in configurational entropy, at least in the present system, so that a focus on side-chain contributions[16,21] may be unduly limiting.

Figure 4 furthermore visualizes the changes in correlation throughout the protein structure. The middle panel, which displays the changes in pairwise mutual information for all torsional pairs, shows clear diagonal patterning, which indicates correlation contributions from torsions that are near each other in protein sequence and hence in space. It also shows large patches corresponding to the two loops highlighted in the left-hand panel. As shown in the graph along the top of the heat map, these loops are also subject to large changes in first-order entropy. The strong off-diagonal patches associated with loop–loop correlations mean that correlation contributions are strong for torsions that are near each other in space yet not in sequence. Torsion 154 is χ_1 of Tyr10 and has an intense stripe in the side-chain and side-chain/main-chain sections of the heat map, indicating strong correlations with many other torsions. Torsions 170 and 264, which are χ_2 and χ_3 of the disulfide bridge between the two loops (Cys14-Cys38), also have particularly intense stripes in the heat map. The right-hand panel of Figure 4 is the same as the middle one, except that it only marks those torsion pairs included in the MIST estimate of the pairwise entropy. Thus, MIST selects only highly correlated pairs within each conformational cluster, and it is interesting to see that this procedure focuses attention even more on the diagonals. Thus, torsions that are near neighbors in sequence contribute the most to the computed entropy difference.

The estimated change in pairwise correlation entropy of about 3–6 kcal/mol corresponds to about 0.05 to 0.10 kcal/mol/residue for this protein of 58 residues. This may be compared with the value of ~0.3 kcal/mol/residue computed for the TSG101-peptide binding system (above).[24] It seems reasonable that the binding reaction of the TSG101 system might lead to a larger perturbation per residue than the conformational change studied here. Inadequate sampling can cause the correlation entropy to be overestimated, and the net simulation time of 2 μs for TSG101 is much less than the 1 ms of time available for BPTI, so convergence could also play a role in the observed difference. Finally, it is worth noting that BPTI's three disulfide bridges may dampen its total configurational flexibility relative to TSG-101, which lacks such structural constraints.

■ DISCUSSION

The results presented here bear on several aspects of the contributions of torsional correlations to changes in configurational entropy. A central result is that well-converged simulations clearly show nontrivial contributions to configurational entropy changes from changes in correlations, for both binding and conformational change events. For the host–guest system, the correlation contributions amount to about 40% of the first-order entropy, and they act in the same sense as the first-order contribution; that is, they further reduce the configurational entropy of binding. For the conformational change of BPTI, the correlation contribution of 3–6 kcal/mol represents 15–30% of the first-order entropy but now acting in
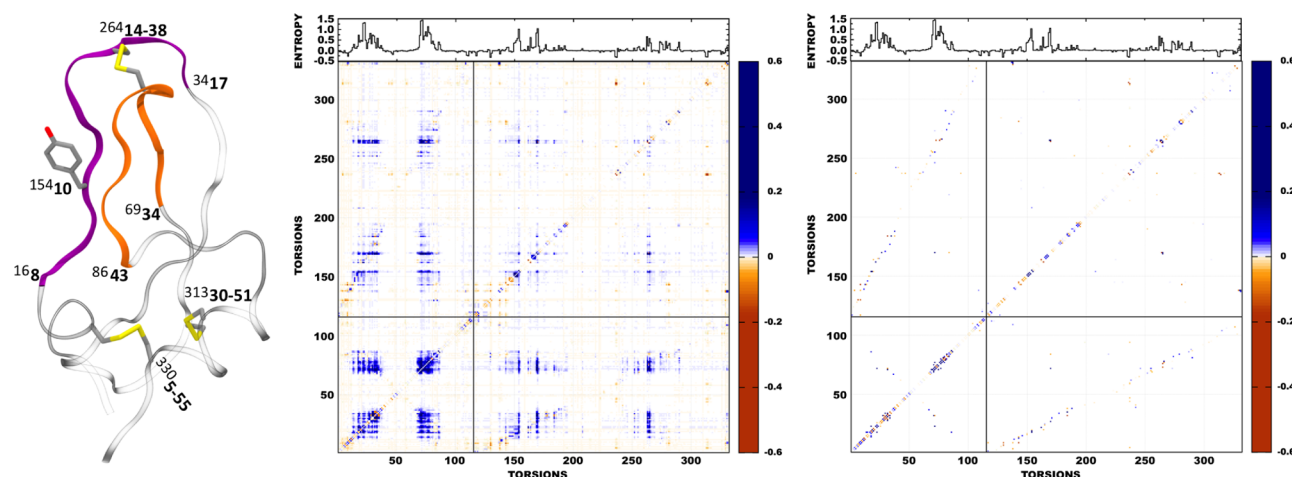
Figure 4. Structural analysis of the changes in pairwise correlation between Clusters 1 and 2 for torsions throughout BPTI. Left: cartoon representation of BPTI with structural features having large correlation annotated by the torsion number (superscripted) and residue number (bold). Loop 1 (purple) and loop 2 (orange) are connected by the central disulfide bridge (Cys14-Cys38) and show strong intra- and interloop correlation. Center: Mutual information heat map for all torsion pairs pertaining to the 332 nonphase side-chain and main-chain ϕ/ψ torsions. The first 116 torsions correspond to main-chain torsions starting at residue 1, and the following 216 are side-chain torsions, again starting at residue 1. Main-chain torsions of loop 1 start at torsion 16 (residue 8) and end at torsion 34 (residue 17). Main-chain torsions of loop 2 start at torsion 69, residue 34 and end at torsion 86, residue 43. The disulfide bridges are side-chain torsions 264, 313, and 330. Side-chain torsion 154 (Tyr10) shows a noticeable correlation fingerprint. (Right) The corresponding MIST solution for the same set of torsions. The mutual information and entropy values represent Cluster 2 minus Cluster 1 deltas and are reported in units of S/k_B. Furthermore, each delta in pairwise mutual information had the average delta in pairwise spurious mutual information (0.0057), as determined by the previous permutation analysis subtracted from its value to minimize numerical bias.

the opposite sense; that is, correlation acts opposite to the first-order term.

The conclusion that correlation contributes significantly to configurational entropy differences is consistent with a prior quasiharmonic study using Cartesian coordinates[7] and with MIE[23,24] studies using bond-angle-torsion coordinates. It is also consistent with a recent study that applied MIST to side-chain rotameric states across a series of different proteins.[16] However, it appears less consistent with a prior simulation study of peptides and a small protein, which indicated relatively small contributions from changes in correlation.[20] We conjecture that the apparent inconsistency stems largely from the different nature of the cases studied. In the prior study, changes in configurational entropy were computed for five dipeptides in solution, when T was changed from 270 to 380 K, but perhaps these small molecules were largely unstructured and uncorrelated at both temperatures. For villin headpiece, the prior study considered only a 20 K temperature change, which may have been too small to produce much change in overall motion. Moreover, the villin headpiece simulations at 300 K were terminated at 70 ns because the protein structure was beginning to change significantly after 75 ns in two of the runs. Perhaps continuing the run, so as to include the impending conformational change, would have altered the findings. Finally, for the villin study, the small magnitude of the correlation contributions observed may result in part from the study's neglect of inter-residue correlations.[20] Interestingly, another prior study, of calmodulin,[21] which also suggested that changes in correlation entropy are small, likewise studied a temperature change (295 to 346 K), rather than focusing explicitly on a conformational change or binding event. In summary, the current body of evidence seems consistent with a view that clear-cut conformational changes and binding events tend to be associated with quantitatively important entropic contributions from changes in correlation. This conclusion suggests that one cannot confidently overlook the potential importance of correlation in the entropic interpretation of NMR order parameter studies, at least for proteins undergoing well-defined binding events or conformational changes.

Recently, however, it has been suggested, based on empirical and simulation data, that the entropy contribution of correlation is roughly proportional to the first-order entropy, $\Delta S_{full}^{corr} \approx -0.17\Delta S_1$, so that a total entropy change may be estimated as $\Delta S_{full} \approx 0.83\Delta S_1$.[16] Such a regularity would clearly be useful, given that computing or measuring correlations is difficult. It also can make intuitive sense, as the entropy of a set of correlated DOF may require a larger correlation correction if they become more flexible, at least up to a point, and, indeed, the present results for BPTI fit the pattern rather well. However, the host−guest binding results do not fit the pattern. Here the change in correlation entropy has the same sign as the change in the first-order entropy, such that $\Delta S_{full}^{corr} \approx -1.4\Delta S_1$. In respect of the sign of the relationship, this result agrees with a prior simulation study of peptide binding by the protein TSG101.[24] It is perhaps worth noting that although greater correlation necessarily decreases entropy, relative to the first-order entropy, the *change* in correlation in the course of some process may work in either direction. One may speculate, based on the available data, that binding, in particular, tends to generate decreases in both first-order and correlation entropy, but only further study could establish this point. What one may conclude at this point is that the change in correlation entropy can, in general, either reinforce or compensate the change in first-order entropy. As a consequence, entropy changes derived from changes in NMR order parameters cannot be reliably assumed to represent either upper or lower limits of the total entropy change.

The present methodology also affords a detailed look at the specific torsions and torsional correlations that determine the computed entropy changes in BPTI. We find that the estimated

6453

dx.doi.org/10.1021/jp411588b | *J. Phys. Chem. B* 2014, 118, 6447−6455

entropy difference between Clusters 1 and 2 of BPTI is associated largely, though not exclusively, with changes in the conformational distributions of main-chain torsions. This holds not only at first but also at second order, where changes in pairwise torsional correlations contribute. Thus, entropy estimates that neglect main-chain contributions, including main-chain correlations, risk incurring substantial errors. It is also of interest that the most significant pairwise correlations involve torsions that are close to each other in space, either because they are sequence neighbors or because they are brought together by the 3-D fold of the protein. It will be interesting to learn whether longer-ranged correlations are important for more complex and flexible proteins than BPTI, which is small and is cross-linked by three disulfide bridges.

An innovative methodological aspect of this study is that it has presented the first application of the MIE-NN and MIST-NN methods to a binding problem, albeit a simple one, and the first MD-based estimation of the change in configurational entropy on binding at full dimensionality. Such methods can have broader applicability to larger systems as well, especially when coupled to the use of information theoretic methods, like the Jensen−Shannon divergence approach used here, to select the most relevant DOF and subsequently drastically reduce the combinatorics associated with MIE and MIST calculations. Given advancing computer power, exemplified by the milli-second protein simulation of BPTI, quantitative investigation of correlations at higher than second order in a system as large as BPTI may soon become computationally feasible.

■ AUTHOR INFORMATION

Corresponding Authors
*V.H.: E-mail: hnizdo2044@gmail.com.
*M.K.G.: E-mail: mgilson@ucsd.edu.

Author Contributions
‖A.T.F. and B.J.K. contributed equally.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Chang, C.-E. A.; Chen, W.; Gilson, M. K. Ligand Configurational Entropy and Protein Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 1534−1539.
(2) Zhou, H.-X.; Gilson, M. K. Theory of Free Energy and Entropy in Noncovalent Binding. *Chem. Rev.* **2009**, *109*, 4092−4107.
(3) Zidek, L.; Novotny, M. V.; Stone, M. J. Increased Protein Backbone Conformational Entropy Upon Hydrophobic Ligand Binding. *Nat. Struct. Biol.* **1999**, *6*, 1118−1121.
(4) Stone, M. J. NMR Relaxation Studies of the Role of Conformational Entropy in Protein Stability and Ligand Binding. *Acc. Chem. Res.* **2001**, *34*, 379−388.
(5) Homans, S. W. Probing the Binding Entropy of Ligand−Protein Interactions by NMR. *ChemBioChem.* **2005**, *6*, 1585−1591.
(6) Bingham, R. J.; Findlay, J. B. C.; Hsieh, S.-Y.; Kalverda, A. P.; Kjellberg, A.; Perazzolo, C.; Phillips, S. E. V.; Seshadri, K.; Trinh, C. H.; Turnbull, W. B.; et al. Thermodynamics of Binding of 2-Methoxy-3-isopropylpyrazine and 2-Methoxy-3-isobutylpyrazine to the Major Urinary Protein. *J. Am. Chem. Soc.* **2004**, *126*, 1675−1681.
(7) Baron, R.; McCammon, J. A. (Thermo)dynamic Role of Receptor Flexibility, Entropy, and Motional Correlation in Protein−Ligand Binding. *ChemPhysChem* **2008**, *9*, 983−988.
(8) Harpole, K. W.; Sharp, K. A. Calculation of Configurational Entropy with a Boltzmann−Quasiharmonic Model: The Origin of High-Affinity Protein−Ligand Binding. *J. Phys. Chem. B* **2011**, *115*, 9461−9472.
(9) Tzeng, S.-R.; Kalodimos, C. G. Protein Activity Regulation by Conformational Entropy. *Nature* **2012**, *488*, 236−240.
(10) 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*, 75−81.
(11) Akke, M.; Brueschweiler, R.; Palmer, A. G., III. NMR Order Parameters and Free Energy: An Analytical Approach and Its Application to Cooperative Calcium (2+) Binding by Calbindin D9k. *J. Am. Chem. Soc.* **1993**, *115*, 9832−9833.
(12) Yang, D.; Kay, L. E. Contributions to Conformational Entropy Arising from Bond Vector Fluctuations Measured from NMR-Derived Order Parameters: Application to Protein Folding. *J. Mol. Biol.* **1996**, *263*, 369−382.
(13) Li, Z.; Raychaudhuri, S.; Wand, A. J. Insights Into the Local Residual Entropy of Proteins Provided by NMR Relaxation. *Protein Sci.* **1996**, *5*, 2647−2650.
(14) Prompers, J.; Brüschweiler, R. Thermodynamic Interpretation of NMR Relaxation Parameters in Proteins in the Presence of Motional Correlations. *J. Phys. Chem. B* **2000**, *104*, 11416−11424.
(15) Marlow, M. S.; Dogan, J.; Frederick, K. K.; Valentine, K. G.; Wand, A. J. The Role of Conformational Entropy in Molecular Recognition by Calmodulin. *Nat. Chem. Biol.* **2010**, *6*, 352−358.
(16) Kasinath, V.; Sharp, K. A.; Wand, A. J. Microscopic Insights into the NMR Relaxation-Based Protein Conformational Entropy Meter. *J. Am. Chem. Soc.* **2013**, *135*, 15092−15100.
(17) Gunasekaran, K.; Ma, B.; Nussinov, R. Is Allostery an Intrinsic Property of All Dynamic Proteins? *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 433−443.
(18) Goodey, N. M.; Benkovic, S. J. Allosteric Regulation and Catalysis Emerge via a Common Route. *Nat. Chem. Biol.* **2008**, *4*, 474−482.
(19) McClendon, C. L.; Friedland, G.; Mobley, D. L.; Amirkhani, H.; Jacobson, M. P. Quantifying Correlations Between Allosteric Sites in Thermodynamic Ensembles. *J. Chem. Theory Comput.* **2009**, *5*, 2486−2502.
(20) Li, D.-W.; Brüschweiler, R. In silico Relationship between Configurational Entropy and Soft Degrees of Freedom in Proteins and Peptides. *Phys. Rev. Lett.* **2009**, *102*, 118108.
(21) Prabhu, N. V.; Lee, A. L.; Wand, A. J.; Sharp, K. A. Dynamics and Entropy of a Calmodulin-Peptide Complex Studied by NMR and Molecular Dynamics. *Biochemistry* **2003**, *42*, 562−570.
(22) Matsuda, H. Physical Nature of Higher-Order Mutual Information: Intrinsic Correlations and Frustration. *Phys. Rev. E* **2000**, *62*, 3096−3102.
(23) Killian, B. J.; Kravitz, J. Y.; Gilson, M. K. Extraction of Configurational Entropy from Molecular Simulations via an Expansion Approximation. *J. Chem. Phys.* **2007**, *127*, 024107.
(24) Killian, B. J.; Kravitz, J. Y.; Somani, S.; Dasgupta, P.; Pang, Y.-P.; Gilson, M. K. Configurational Entropy in Protein−Peptide Binding: Computational Study of TSG101 Ubiquitin E2 Variant Domain with an HIV-Derived PTAP Nonapeptide. *J. Mol. Biol.* **2009**, *389*, 315−335.
(25) Goswami, S.; Mukherjee, R. Molecular Recognition: a Simple Dinaphthyridine Receptor for Urea. *Tetrahedron Lett.* **1997**, *38*, 1619−1622.
(26) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330*, 341−346.
(27) Long, D.; Brüschweiler, R. Atomistic Kinetic Model for Population Shift and Allostery in Biomolecules. *J. Am. Chem. Soc.* **2011**, *133*, 18999−19005.

6454

dx.doi.org/10.1021/jp411588b | *J. Phys. Chem. B* 2014, 118, 6447−6455

(28) Xue, Y.; Ward, J. M.; Yuwen, T.; Podkorytov, I. S.; Skrynnikov, N. R. Microsecond Time-Scale Conformational Exchange in Proteins: Using Long Molecular Dynamics Trajectory to Simulate NMR Relaxation Dispersion Data. *J. Am. Chem. Soc.* **2012**, *134*, 2555−2562.

(29) Genheden, S.; Akke, M.; Ryde, U. Conformational Entropies and Order Parameters: Convergence, Reproducibility, and Transferability. *J. Chem. Theory Comput.* **2013**, *10*, 432−438.

(30) 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*, 20006−20011.

(31) King, B. M.; Tidor, B. MIST: Maximum Information Spanning Trees for Dimension Reduction of Biological Data Sets. *Bioinformatics* **2009**, *25*, 1165−1172.

(32) King, B. M.; Silver, N. W.; Tidor, B. Efficient Calculation of Molecular Configurational Entropies Using an Information Theoretic Approximation. *J. Phys. Chem. B* **2012**, *116*, 2891−2904.

(33) Hnizdo, V.; Gilson, M. K. Thermodynamic and Differential Entropy under a Change of Variables. *Entropy* **2010**, *12*, 578−590.

(34) 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*, 1157−1174.

(35) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132−146.

(36) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(37) Case, D.; Darden, T.; Cheatham, T., III; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K.; et al. *AMBER 12*; University of California: San Francisco, 2012.

(38) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(39) Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed Without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374−380.

(40) Fox, T.; Kollman, P. A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B* **1998**, *102*, 8070−8079.

(41) Cieplak, P.; Caldwell, J.; Kollman, P. Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases. *J. Comput. Chem.* **2001**, *22*, 1048−1057.

(42) Roe, D. R.; Cheatham, T. E., III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.

(43) Singh, H.; Misra, N.; Hnizdo, V.; Fedorowicz, A.; Demchuk, E. Nearest Neighbor Estimates of Entropy. *Am. J. Math. Manage. Sci.* **2003**, *23*, 301−321.

(44) Hnizdo, V.; Darian, E.; Fedorowicz, A.; Demchuk, E.; Li, S.; Singh, H. Nearest-Neighbor Nonparametric Method for Estimating the Configurational Entropy of Complex Molecules. *J. Comput. Chem.* **2007**, *28*, 655−668.

(45) Misra, N.; Singh, H.; Hnizdo, V. Nearest Neighbor Estimates of Entropy for Multivariate Circular Distributions. *Entropy* **2010**, *12*, 1125−1144.

(46) Hnizdo, V.; Tan, J.; Killian, B. J.; Gilson, M. K. Efficient Calculation of Configurational Entropy from Molecular Simulations by Combining the Mutual-Information Expansion and Nearest-Neighbor Methods. *J. Comput. Chem.* **2008**, *29*, 1605−1614.

(47) Shaw, D. E. et al. Anton, A Special-Purpose Machine for Molecular Dynamics Simulation. In *Proceedings of the 34th Annual International Symposium on Computer Architecture*; D. E. Shaw Research: New York, 2007; pp 1−12.

(48) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665.

(49) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.

(50) Numata, J.; Knapp, E.-W. Balanced and Bias-Corrected Computation of Conformational Entropy Differences for Molecular Trajectories. *J. Chem. Theory Comput.* **2012**, *8*, 1235−1245.

(51) Abagyan, R.; Totrov, M.; Kuznetsov, D. ICMA New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation. *J. Comput. Chem.* **1994**, *15*, 488−506.

(52) Lin, J. Divergence Measures Based on the Shannon Entropy. *IEEE Trans. Inf. Theory* **1991**, *37*, 145−151.

(53) Endres, D.; Schindelin, J. A New Metric for Probability Distributions. *IEEE Trans. Inf. Theory* **2003**, *49*, 1858−1860.

(54) Kullback, S.; Leibler, R. A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79−86.

(55) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: a Critical Review. *Biophys. J.* **1997**, *72*, 1047−1069.

(56) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J. Phys. Chem. B* **2003**, *107*, 9535−9551.