DOI: 10.1021/bi9020082

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Preferential Interactions between Small Solutes and the Protein Backbone: A Computational Analysis[†]

Liang Ma,[‡] Laurel Pegram,[§] M. T. Record, Jr.,[‡,§] and Qiang Cui*[,‡]

[‡]*Graduate Program in Biophysics and Department of Chemistry, University of Wisconsin, University Avenue, Madison, Wisconsin 53706, and* [§]*Department of Biochemistry, University of Wisconsin, 433 Babcock Drive, Madison, Wisconsin 53706*

Received November 23, 2009; Revised Manuscript Received January 27, 2010

ABSTRACT: To improve our understanding of the effects of small solutes on protein stability, we conducted atomistic simulations to quantitatively characterize the interactions between two broadly used small solutes, urea and glycine betaine (GB), and a triglycine peptide, which is a good model for a protein backbone. Multiple solute concentrations were analyzed, and each solute−peptide−water ternary system was studied with ∼200−300 ns of molecular dynamics simulations with the CHARMM force field. The comparison between calculated preferential interaction coefficients (Γ_{23}) and experimentally measured values suggests that semiquantitative agreement with experiments can be obtained if care is exercised to balance interactions among the solute, protein, and water. On the other hand, qualitatively incorrect (i.e., wrong sign in Γ_{23}) results can be obtained if a solute model is constructed by directly taking parameters for chemically similar groups from an existing force field. Such sensitivity suggests that small solute thermodynamic data can be valuable in the development of accurate force field models of biomolecules. Further decomposition of Γ_{23} into group contributions leads to additional insights regarding the effects of small solutes on protein stability. For example, use of the CHARMM force field predicts that urea preferentially interacts with not only amide groups in the peptide backbone but also aliphatic groups, suggesting a role for these interactions in urea-induced protein denaturation; quantitatively, however, it is likely that the CHARMM force field overestimates the interaction between urea and aliphatic groups. The results with GB support a simple thermodynamic model that assumes additivity of preferential interaction between GB and various biomolecular surfaces.

Biomolecules exist in a heterogeneous cellular environment, and it is well-known that small solutes can play important roles in regulating the stability, structure, and therefore function of biomolecules (*1*−*3*). In the context of protein stability, small solutes can be classified as denaturants or protectants according to their effect of destabilizing or stabilizing, respectively, the native state of proteins (*3*, *4*). Urea and guanidinium are two well-known protein denaturants (*5*, *6*) and widely used in protein denaturation experiments in vitro; glycine betaine, which is rich in mammalian kidney and some vascular plants (*7*−*9*), and trimethylamine *N*-oxide (TMAO), which is accumulated in cartilaginous fish and the coelacanth (*9*−*11*), act as protein protectants. Although much is known about the interaction between small solutes and biomolecules (*2*, *12*), the precise denaturing−protecting mechanisms of these small solutes at the molecular level remain controversial. For example, multiple proposals have been put forward regarding how urea denatures proteins, and they range from modifying water structure (*13*−*16*) to exhibiting strong interactions with protein backbone through either polar (*17*−*19*) or dispersive forces (*17*, *20*−*22*).

One important avenue for probing the mechanism of small solute effects is to conduct systematic thermodynamic analysis. A key quantity in this context is the "*m* value", defined as (*2*)

$$m \text{ value} = \frac{\partial \Delta G^{o}_{obs}}{\partial m_3} \quad (1)$$

in which ΔG^{o}_{obs} is the standard free energy change for the unfolding process and m_3 is the solute concentration on the molar scale; throughout the paper, we use the following conventions to label the components in a ternary system: component 1 is water, component 2 is the protein, and component 3 is the small solute. The *m* value is closely related to the difference in interactions between the solute and the unfolded and native states. For example, if the solute interacts more strongly with the unfolded state than with the native state, increasing the solute concentration will shift the equilibrium toward the unfolded state, leading to protein denaturation.

The *m* value (eq 1) is equal to $\Delta\mu_{23}$, the change in μ_{23} for the unfolding process, where

$$\mu_{23} = \frac{\partial \mu_2}{\partial m_3}\Big|_{m_2, T, P} \quad (2)$$

is the derivative of the chemical potential of component 2, μ_2, with respect to m_3, at a constant protein concentration on the molar scale (m_2), constant temperature, and constant pressure. μ_{23} refers to the change in the protein's chemical potential in response to the change in the solute concentration; i.e., it

[†]The research has been supported by the National Institutes of Health (Grant R01-GM071428 to Q.C. and Grant R01-GM047022 to M.T.R.).
*To whom the correspondence should be addressed. E-mail: cui@chem.wisc.edu. Phone: (608) 262-9801. Fax: (608) 262-9918.

 Published on Web 02/01/2010

quantifies the stabilizing or destabilizing effect of the solute on a specific (native or unfolded) state of the protein.

Another thermodynamic quantity closely related to μ_{23} is the preferential interaction coefficient (*2*), Γ_{23}, between the solute and the protein:

$$\Gamma_{23} = -\frac{\mu_{23}}{\mu_{33}} \quad (3)$$

in which μ_{33} is the counterpart of μ_{23} in the solute−water binary system. The physical meaning of Γ_{23} can be understood on the basis of the two-domain model (*23*, *24*) (see ref *25* and references cited therein for discussions of Γ_{23} in different ensembles). Because of the presence of the protein (component 2), the distributions of water (component 1) and solute (component 3) in local domain II (close to the protein) are different from those in bulk domain I (far from the protein). Γ_{23} can be cast in the following form (*23*):

$$\Gamma_{23} = \left\langle n_3^{II} - n_1^{II}\frac{n_3^{I}}{n_1^{I}} \right\rangle \quad (4)$$

which clearly characterizes the extent of accumulation or exclusion of the solute from the protein surface relative to the bulk. In eq 4, the brackets indicate ensemble average; n_3^{II} and n_1^{II} are the numbers of solute and water molecules, respectively, in local domain II, while n_3^{I} and n_1^{I} are the corresponding values in bulk domain I. Γ_{23} is positive when the solute accumulates around the protein relative to the bulk and negative when the solute is excluded from the protein. Therefore, depending on the degree of preferential association with the native versus unfolded states, a solute may stabilize or destabilize the native state of a protein. By systematically probing how Γ_{23} or μ_{23} depends on the types (e.g., polar, nonpolar, or anionic) of protein surfaces, rather quantitative relations can be established regarding the effect of small solutes on the stability of proteins (*25*−*27*).

In principle, atomistic simulations are ideally suited for in-depth analysis of small solute effects, and a significant number of simulation studies have indeed been conducted over the years (*16*, *28*−*31*). In practice, however, the reliability of atomistic simulations using existing force fields is not straightforward to evaluate considering the relatively subtle nature of solute effects. Therefore, it is important to explicitly benchmark atomistic simulations using accurate experimental data for model systems. For urea, for example, careful studies have been conducted by Smith and co-workers (*30*), who have shown that popular force fields (e.g., OPLS) give qualitatively correct but quantitatively inaccurate Kirkwood−Buff integrals (*31*, *32*), which are closely related to Γ_{23}. Using the Kirkwood−Buff integrals as a guide, Smith and co-workers have developed atomic force fields for a few small molecules (*33*−*37*) such as urea (*33*) and a peptide backbone analogue (*N*-methyl acetamide) (*37*); these parameters have been able to be transferred to protein−urea−water ternary systems.

In this work, we conduct a simulation study on the preferential interactions of urea and glycine betaine (GB)[1] with a triglycine (TGLY) peptide, which is a good model for a protein backbone. There are several motivations for this study. First, Record et al. (*25*, *27*) have recently conducted highly accurate measurements of μ_{23} and Γ_{23} between urea or GB and small peptides (e.g., diglycine and triglycine) using vapor-pressure osmometry techniques; these data have provided an unprecedented opportunity to test current force fields regarding intermolecular interactions. Second, although urea−protein interactions have been extensively studied using molecular simulations, no thorough simulation studies have been conducted so far for GB. Finally, we note that the small solute thermodynamic data directly reflect the interaction between functional groups in proteins relative to their interactions with solvent, the underlying driving force behind the assembly of amino acids in proteins. The molecular (as opposed to atomic) nature of these data makes them particularly suited for guiding the development of a protein model at the coarse-grained level, which is an active area of research. Therefore, another goal of this study is to explore the degree to which the computed Γ_{23} value and its approximate group decomposition are sensitive to parameters in the force field, which will determine the feasibility of using Γ_{23} for coarse-grained model development.

The computational details are summarized in Computational Methods; the results are presented and discussed in Results and Discussion. Finally, we draw a few conclusions.

Table 1: Simulation Details of TGLY−Urea−Water and TGLY−GB−Water Systems[a]

solute	concn (m)	n_3	n_1	simulation scheme[b]	box length[c] (Å)
urea	1.21	23	1052	100 × 3 ns	36.0
urea	2.26	39	957	133 × 2 ns	35.4
urea	3.61	56	862	160 × 2 ns	34.8
GB	1.35	23	943	100 × 3 ns	35.4
GB	2.04	32	873	100 × 3 ns	35.4

[a]Throughout the paper, component 1 is water, component 2 is protein or peptide, and component 3 is the small solute. [b]Given in the format of number of independent trajectories (e.g., 100) × the length of each trajectory (e.g., 3 ns). [c]Average box length with the rhombic dodecahedron periodic boundary condition.

COMPUTATIONAL METHODS

The basic computational methods are similar to those reported in the literature (*29*). We pay particular attention to the amount of sampling for properly converging calculated Γ_{23} values. Moreover, we find that although using CHARMM27 parameters (*38*) for urea leads to the correct sign in Γ_{23} and the dependence of Γ_{23} on urea concentration, the computed value for GB around TGLY has the wrong sign if GB is described by simply taking force field parameters from CHARMM27 for functional groups. A simple procedure is adopted to adjust the force field parameters, which we also briefly describe.

Molecular Dynamics (MD) Simulations. As summarized in Table 1, multiple solute concentrations have been studied for each small solute to achieve better comparisons with experiments. For the urea−TGLY interactions, three urea concentrations are studied: 1.21, 2.26, and 3.61 M. For the GB−TGLY interactions, two GB concentrations are studied: 1.35 and 2.04 M. To minimize the size of the simulation system, only one copy of TGLY is included with the proper number of urea or GB and water molecules. The initial structure of TGLY is constructed as the ideal geometry in the CHARMM27 force field (*38*) using the IC BUILD module in CHARMM (*39*). The peptide is first solvated with a pre-equilibrated rhombic dodecahedron (RHDO) solvent box with a box length of ~35 Å. The proper number of solutes (urea or GB) are then randomly inserted into

[1]Abbreviations: GB, glycine betaine; TGLY, triglycine; COM, center of mass; vdW, van der Waals; ASA, accessible surface area.

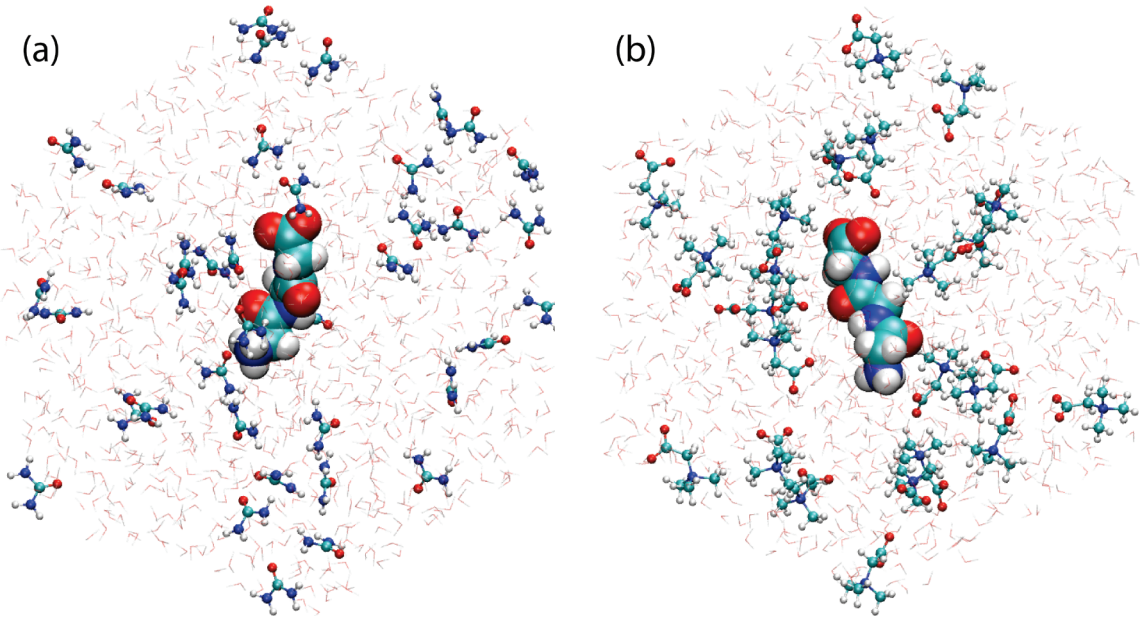

FIGURE 1: Snapshots from MD trajectories for (a) the TGLY−urea−water system at 2.26 m urea and (b) the TGLY−GB−water system at 1.35 m GB. TGLY and water molecules are shown in van der Waals form and line schemes, respectively. Urea and GB molecules are shown in the CPK scheme. Figures were created using VMD (*60*).

the simulation box according to the solute concentration; water molecules that overlap with the solutes are removed. For representative snapshots, see Figure 1.

TGLY and urea are described using the CHARMM27 force field (*38*), and the modified TIP3P model (*40*) is adopted for water molecules. There are no existing parameters for GB in the CHARMM27 force field; thus, a simple procedure is used to develop a model (see Force Field Parameters for GB). A switching scheme (*41*) for interatomic distances between 10 and 12 Å is used for van der Waals interactions. For electrostatic interactions, the particle mesh Ewald scheme (PME) (*42*) is used. All simulations were conducted under the NPT ensemble: temperature controlled using the Nose−Hoover algorithm (*43*, *44*), with a mass of 250 kcal mol^{-1} ps^2 for the thermostat and a reference temperature of 300 K; pressure controlled with the Andersen algorithm (*45*, *46*), using a mass of 500 amu for the pressure piston, a reference pressure of 1 atm, a Langevin piston collision frequency of 10 ps^{-1}, and a Langevin piston bath temperature of 300 K.

To ensure convergence in the calculated Γ_{23} values, for each ternary system, ~100 independent trajectories of 2−3 ns long were carried out. These independent trajectories start with the same equilibrated structure (~5 ns) but with different initial velocities at 300 K. Test calculations have shown that this is an effective scheme (in part because of the highly parallel nature) for converging Γ_{23}. Throughout the simulations, the SHAKE algorithm (*47*) is used to allow an integration time step of 2 fs. All simulations were conducted using CHARMM (*39*).

To supplement the TGLY−urea−water simulations and improve our understanding of the preferential interaction between urea and aliphatic (-CH_2-) groups, a set of propane−urea−water simulations were also performed at 2.26 m urea. Specifically, 40 independent trajectories with a duration of 2 ns long were conducted following the same MD simulation protocol used for the TGLY−urea−water systems; the propane is described with the standard CHARMM27 (*38*) force field.

Calculations of Solute−Water Radial Distribution Function g(r) with Respect to the Peptide Surface. To characterize the distributions of small solutes and water around the peptide, the solute−water radial distribution functions, $g(r)$, relative to the peptide surface are calculated. $g(r)$ values are computed numerically in a manner similar to that of our recent studies of ion distributions around proteins (*48*, *49*). The shortest distances r between the center of mass (COM) of each solute or water molecule and the van der Waals (vdW) surface of the peptide are collected from MD trajectories and binned to generate $\delta N(r)$ for each bin (r, dr). The volume factor $[\delta V(r)]$ corresponding to the bin (r, dr) is estimated using the volume function in CHARMM (COOR VOLU) with all peptide atoms having a radius of $r_{vdW} + r \pm dr/2$, and $\delta V(r)$ is calculated as $\delta V(r) = V(r_{vdW} + r + dr/2) - V(r_{vdW} + r - dr/2)$, where r_{vdW} is the vdW radius of each atom. Bin size dr is set to 0.1 Å for all $g(r)$ calculations. The solute−water−peptide surface radial distribution function $g(r)$ is then calculated as

$$g(r) = \frac{1}{\rho_{bulk}} \frac{\delta N(r)}{\delta V(r)} \frac{1}{N_{frame}} \quad (5)$$

where N_{frame} is the number of simulation snapshots used for calculations and ρ_{bulk} is the bulk number density of solute or water and is estimated by dividing the number of solute or water molecules by the average volume of the simulation box throughout all MD trajectories.

Calculations of Γ_{23}. Γ_{23} for each ternary system is calculated on the basis of a two-domain model (*23*, *24*) according to eq 4. MD simulations generate an ensemble of snapshots, and for each snapshot, the number of solute or water molecules in the local and bulk domains is counted on the basis of the shortest distance (r) from the COM of each solutewater molecule to the peptide vdW surface; i.e., r is the same as that in the calculations of $g(r)$. To define the local and bulk domains, a boundary r^* has to be specified, which is determined on the basis of radial distribution functions $g(r)$. As shown in Figures 2 and 5, both solute and water reach their bulk concentrations at ~7−8 Å; thus, an r^* of 8 Å is chosen as the boundary radius. Note that the calculated Γ_{23} is not sensitive to the precise value of r^*; this relative insensitivity

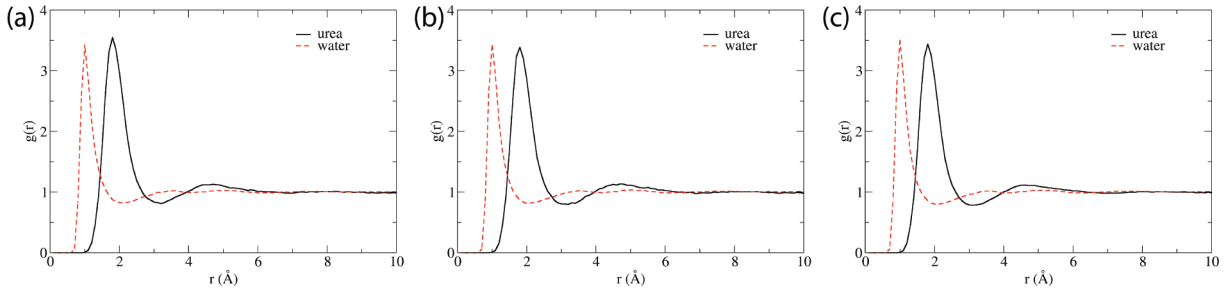

FIGURE 2: Radial distribution function, $g(r)$, of urea and water molecules around TGLY at different urea concentrations: (a) 1.21, (b) 2.26, and (c) 3.61 m urea. The distance r is the shortest distance between the center of mass of a urea or water molecule and the van der Waals surface of TGLY.

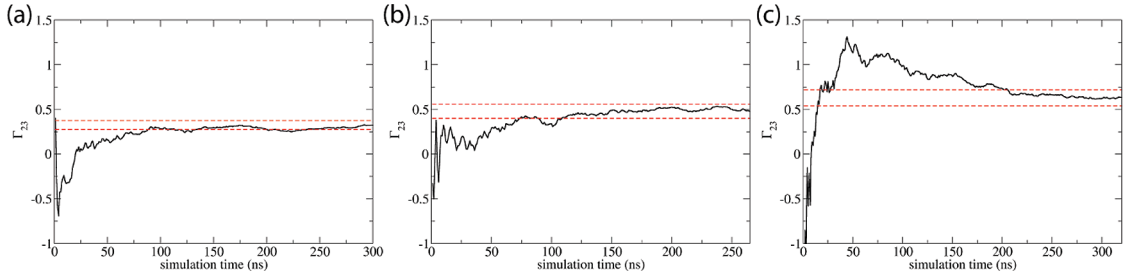

FIGURE 3: Convergence behaviors of Γ_{23} with respect to accumulated simulation time for TGLY−urea−water systems at different urea concentrations: (a) 1.21, (b) 2.26, and (c) 3.61 m urea. The red dashed lines indicate the converged Γ_{23} value ± the estimated standard deviation.

of the computed Γ_{23} to the precise value of r^* (as far as it is larger than ~4 Å) was also found in a recent computational study by Trout and co-workers (*50*).

Force Field Parameters for GB. Because there are no existing parameters for GB in the CHARMM27 force field (*38*), we started building a model for GB by selecting parameters for similar chemical groups. Specifically, the headgroup of the lipid molecule POPC (palmitoyloleoylphosphatidylcholine) is used to describe the -$N(Me)_3$ group, and the CTER patch is used to model the carboxylate group. This simple scheme is adopted with an attempt to maximize the compatibility between the GB model and other components (TGLY and water) in the system. Test calculations indicate, however, that this "raw" model leads to qualitatively incorrect Γ_{23} values. At the GB concentration of 1.35 M, for example, the calculated value is 0.40, which differs in sign from the experimental value of −0.16 (*27*); i.e., this simple model leads to GB accumulation near TGLY rather than exclusion as quantified by the experimental data.

As a simple way to improve the model (to maintain the compatibility with the CHARMM force field in general), the partial charges on GB are scaled down by a uniform factor to achieve a proper balance of interactions among GB, TGLY, and water. As shown in the Supporting Information, we consider six gas phase molecular complexes formed by GB, TGLY, and water in different relative orientations. The internal geometry of the monomers is held fixed to that optimized at the HF/6-31G(d) level (*51*) in the gas phase using Gaussian03 (*52*); this particular level of theory is chosen because it is consistent with the parametrization procedure of the CHARMM27 force field. With the relative orientation also held fixed, the distance between monomers in each complex is then varied to determine the optimal separation at the HF/6-31G(d) level in the gas phase. The interaction energy between the monomers in each complex at the optimal separation is then multiplied by a factor of 1.16 to account for the lack of electronic polarization in the CHARMM27 force field; interaction energies for the six molecular complexes are used to determine a single scaling factor for the raw GB partial charges. As shown in the Supporting Information, the scaled partial charges for functional groups are in fact rather close to the natural bonding orbital charges in GB.

RESULTS AND DISCUSSION

In the following, we first present results for urea−TGLY and then results for GB−TGLY interactions. Finally, we discuss these data in the context of relevant computational and experimental studies in the literature.

Preferential Interaction between Urea and TGLY. (*i*) *Γ_{23} between Urea and TGLY*. As shown in Figure 2, radial distribution functions $g(r)$ for urea and water around TGLY vary only slightly among three urea concentrations. The peak height of $g(r)$ is similar for urea and water, although the width of the first peak is substantially broader for urea. Qualitatively, these results highlight a higher degree of local accumulation of urea around TGLY (relative to bulk) than around water.

The convergence behaviors of computed Γ_{23} values between urea and TGLY for the three urea concentrations are shown in Figure 3. The red dashed lines represent the converged Γ_{23} ± the estimated standard deviation using block average (*53*). Evidently, it takes ~100 ns for Γ_{23} to converge, which is substantially (and surprisingly) longer than the time scale of a few nanoseconds reported in previous studies (*29*, *31*, *50*). This difference is probably due to the small protein (peptide) size, and therefore the small protein surface area, in our study, although it also highlights the importance of conducting extensive simulations for a quantitative computation of Γ_{23}.

The calculated Γ_{23} values between urea and TGLY are compared to the experimental results (*25*) in Table 2. In qualitative agreement with experiments, the calculated Γ_{23} values are positive, which indicates that urea accumulates around TGLY relative to the bulk; as the urea concentration increases, both calculated and measured Γ_{23} values increase in a linear fashion. The linear trend can be rationalized by noting that μ_{23}

and μ_{33} in eq 3 can be expanded into polynomials of m_2 and m_3 (*25*):

$$\mu_{23} = RT(a_{11} + a_{21}m_2 + a_{12}m_3) \quad (6)$$

$$\mu_{33} = RT\left(\frac{1}{m_3} + 2a_{02} + a_{12}m_2 + 3a_{03}m_3\right) \quad (7)$$

where a_{11}, a_{21}, etc., are constants specific for a given ternary system. Since m_2 is much smaller than m_3 and a_{03} is much smaller than a_{12} for the TGLY−urea−water ternary system, the major contribution of m_3 to Γ_{23} comes from $a_{12}m_3$, which gives rise to the linear dependence of Γ_{23} on m_3.

At the quantitative level, however, the computed values are too large by a factor of ~2; for example, at a urea concentration of 1.21 M, the calculated Γ_{23} is 0.33 while the measured value is 0.12. These results indicate that the standard CHARMM27 force field can be further improved to balance interactions among urea, the protein backbone (TGLY does not have a side chain), and water (see below).

(*ii*) *Group Decomposition of Urea−Protein Backbone Preferential Interaction*. To gain additional insights into the molecular details of urea−protein backbone interactions, we decompose the overall Γ_{23} into group contributions from TGLY. We define seven groups in TGLY: the N-terminus (NH_3^+), the C-terminal carboxylate, two amide groups, and three -CH_2- groups. The group-specific preferential interaction, Γ_{23}^i, is defined like the overall Γ_{23} as

$$\Gamma_{23}^i = \left\langle n_3^{\mathrm{II},i} - n_1^{\mathrm{II},i}\frac{n_3^{\mathrm{I}}}{n_1^{\mathrm{I}}}\right\rangle \quad (8)$$

where $n_3^{\mathrm{II},i}$ and $n_1^{\mathrm{II},i}$ are the number of solute or water molecules in local domain II that are assigned to group *i*. A solute or water molecule is assigned to a specific group on the basis of the shortest surface-to-surface distance between a solute or water molecule and different groups in TGLY. In this way, the sum of Γ_{23}^i over all seven groups is identical to the overall Γ_{23}.

Such calculated Γ_{23}^i values between urea and TGLY are summarized in Table 3. The same trends are observed at three urea concentrations: Γ_{23}^i is negligible at two termini (-NH_3^+ and -COO^-) and positive around both amide groups and the aliphatic -CH_2- groups. In other words, CHARMM simulations predict that urea accumulates at not only amide groups, as deduced from previous experimental (*4*, *19*, *25*, *54*) and simulation (*16*, *18*) studies, but also aliphatic -CH_2- groups, consistent with recent simulation results (*21*, *22*, *28*, *55*).

As emphasized in previous studies (*56*), a group-based decomposition of Γ_{23} is not necessarily straightforward because of potential cooperative effects and the complex nature of peptide surfaces. Therefore, to probe the preferential interaction between urea and aliphatic (-CH_2-) groups directly, a set of propane−urea−water simulations (80 ns in total) were also conducted with the CHARMM27 (*38*) force field; it is worth noting that while the nonpolar hydrogen atoms in propane and TGLY bear the same partial charge of +0.09 (a convention adopted in the CHARMM22 and -27 force fields), the net charge of the -CH_2 group in TGLY is +0.16 while the -$CH_{2,3}$ groups in propane are all charge-neutral. Γ_{23} between urea and propane is then calculated from these simulations following the same procedure that was used for urea and TGLY. As shown in Figure 4, the calculated Γ_{23} between urea and propane is essentially converged within 30 ns. This faster convergence compared to that of the urea−TGLY system is presumably due to the more homogeneous chemical nature of propane compared to TGLY. The calculated Γ_{23} between urea and propane is 0.15 ± 0.07 at 2.26 m urea, which is qualitatively consistent with the sum of contributions from three -CH_2- groups in Table 4 for TGLY (0.25 ± 0.04). Therefore, it is confirmed that, with the CHARMM27 force field, urea preferentially interacts with the aliphatic

Table 2: Calculated and Experimentally Measured (*25*) Γ_{23} Values for All Simulated Systems[a]

system	1.21 m urea	2.26 m urea	3.61 m urea	1.35 m GB	2.04 m GB
$\Gamma_{23}^{\mathrm{cal}}$	0.33 ± 0.05	0.48 ± 0.08	0.63 ± 0.09	−0.24 ± 0.04	−0.33 ± 0.05
$\Gamma_{23}^{\mathrm{exp}}$	0.12	0.18	0.20	−0.16	−0.21

[a]The error bar for experimentally determined Γ_{23} is < 0.01.

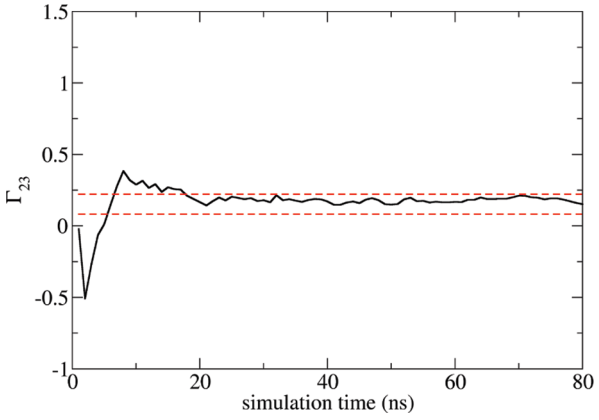

FIGURE 4: Convergence behavior of Γ_{23} with respect to accumulated simulation time for the propane−urea−water system at 2.26 m urea. The red dashed lines indicate the converged Γ_{23} value ± the estimated standard deviation.

Table 3: Group-Specific Γ_{23} Values for All Simulated Systems following the First Decomposition Scheme That Partitions TGLY into Seven Groups

system	1.21 m urea	2.26 m urea	3.61 m urea	1.35 m GB	2.04 m GB
Γ_{23} NH_3^+	0.01 (0.01)[b]	0.01 (0.02)	−0.03 (0.02)	0.09 (0.01)	0.12 (0.02)
Γ_{23} CH_2(1)[a]	0.04 (0.01)	0.05 (0.02)	0.07 (0.02)	−0.06 (0.01)	−0.08 (0.01)
Γ_{23} CH_2(2)	0.08 (0.01)	0.10 (0.01)	0.17 (0.01)	−0.04 (0.01)	−0.09 (0.01)
Γ_{23} CH_2(3)	0.05 (0.01)	0.10 (0.01)	0.13 (0.01)	−0.01 (0.01)	0 (0.01)
Γ_{23} CONH(1)	0.06 (0.01)	0.11 (0.02)	0.16 (0.01)	0.01 (0.01)	0.04 (0.01)
Γ_{23} CONH(2)	0.07 (0.01)	0.11 (0.02)	0.15 (0.01)	0.02 (0.01)	0.02 (0.01)
Γ_{23} COO^-	0.03 (0.02)	0.01 (0.03)	−0.01 (0.03)	−0.25 (0.02)	−0.34 (0.02)

[a]CH_2(1) refers to the first -CH_2- group close to the N-terminus of TGLY. [b]Numbers in parentheses are standard deviations.

-CH_2- groups in a manner independent of the peptide backbone groups (see below for additional discussions).

Preferential Interaction between GB and TGLY. As mentioned in Computational Methods, the distribution of GB near TGLY is qualitatively incorrect without scaling the partial charges of raw CHARMM27 parameters. Therefore, we focus on results obtained with the scaled partial charges (summarized in the Supporting Information).

(*i*) Γ_{23} *between GB and TGLY*. Radial distribution functions $g(r)$ of GB and water around TGLY for two GB concentrations are shown in Figure 5. As in the urea case, $g(r)$ exhibits very small variation with respect to GB concentration. The $g(r)$ of GB has a much lower peak height than water, which suggests that GB is locally more excluded from TGLY than water. The convergence behaviors of calculated Γ_{23} values between GB and TGLY for two GB concentrations are shown in Figure 6. Once again, it takes ~100 ns to obtain converged Γ_{23} values, similar to that in the case of urea.

As shown in Table 2, with the scaled partial charges, calculated Γ_{23} values agree rather nicely with experimental data for both GB concentrations. For example, the calculated Γ_{23} at 1.35 m GB is approximately -0.24 ± 0.04, compared to the experimental value of approximately -0.16. The negative value indicates that GB is excluded from the TGLY surface (protein backbone), which is consistent with the fact that GB acts as a protein protectant. Moreover, the concentration dependence of Γ_{23} (i.e., Γ_{23} decreases as the GB concentration increases) is also qualitatively reproduced by MD simulations, which is encouraging considering the simple scheme used to construct a force field model for GB.

(*ii*) *Group Decomposition of GB−Protein Backbone Preferential Interaction*. Following the same group definitions used for characterizing urea−TGLY interactions, group decomposition is performed for the GB−TGLY Γ_{23}. As shown in Table 3, the main exclusion region for GB is the carboxylate group of the peptide, which is consistent with experimental results for GB and an anionic (protein carboxylate or DNA phosphate) biopolymer surface (*7*); to a much reduced level, GB is excluded from the aliphatic -CH_2- groups. By contrast, GB weakly accumulates around the -NH_3^+ group and, to an even smaller degree, around amide groups.

In a recent set of analyses, Record et al. (*27*) measured μ_{23} values between GB and a broad series of small solutes and proteins. On the basis of global fitting of these data, they obtained the μ_{23}/ASA values between GB and different types of surfaces by assuming that μ_{23} values for different surfaces are additive and proportional to ASA; here, ASA stands for (solvent) accessible surface area. They found that this model can predict μ_{23} values for a set of small solutes and proteins in overall good agreement with experiments. Motivated by these findings, we perform a second type of group decomposition for the GB−TGLY Γ_{23}, in which TGLY is divided into five instead of seven groups on the basis of surface type: carboxylate O, amide O, cationic N, and amide N and C. The hydrogen atoms are assigned to the heavy atoms to which they are bonded, and

Table 4: Group-Specific Γ_{23} Values for the TGLY−Urea−Water System with the Second Decomposition Scheme That Partitions TGLY into Five Types of Surfaces

system	1.21 m urea	2.26 m urea	3.61 m urea
Γ_{23} carboxylate O	0.01 (0.02)	−0.01 (0.03)	−0.05 (0.03)
Γ_{23} amide O	0.08 (0.01)	0.10 (0.02)	0.18 (0.02)
Γ_{23} cationic N	0.01 (0.01)	~0 (0.02)	−0.03 (0.02)
Γ_{23} amide N	0.03 (~0)	0.06 (~0)	0.07 (~0)
Γ_{23} C	0.20 (0.02)	0.34 (0.04)	0.46 (0.03)
Γ_{23} total	0.33 (0.05)	0.48 (0.08)	0.63 (0.09)

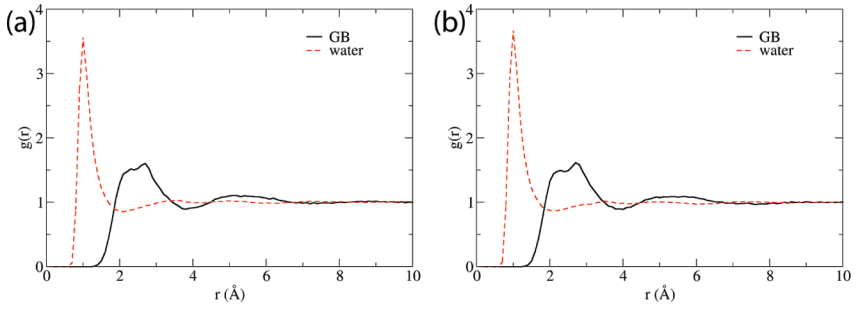

FIGURE 5: Radial distribution function $g(r)$ of GB and water molecules around TGLY at different GB concentrations: (a) 1.35 and (b) 2.04 m GB. The distance r is the shortest distance between the center of mass of a GB or water molecule and the van der Waals surface of TGLY.

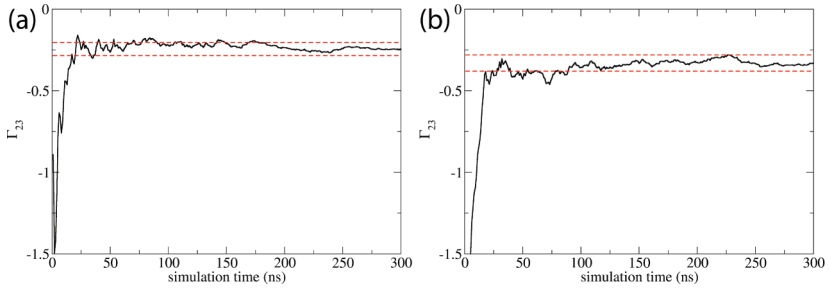

FIGURE 6: Convergence behaviors of Γ_{23} with respect to accumulated simulation time for TGLY−GB−water systems at different GB concentrations: (a) 1.35 and (b) 2.04 m GB. The red dashed lines indicate the converged Γ_{23} value ± the estimated standard deviation.

Table 5: Group-Specific Γ_{23} Values for TGLY−GB−Water System with the Second Decomposition Scheme That Partitions TGLY into Five Types of Surfaces

system	1.35 m GB (sim[a])	1.35 m GB (pred[b])	2.04 m GB (sim)	2.04 m GB (pred)
Γ_{23} carboxylate O	−0.25 (0.02)	−0.29	−0.34 (0.02)	−0.38
Γ_{23} amide O	0 (0.01)	−0.15	0 (0.01)	−0.20
Γ_{23} cationic N	0.09 (0.01)	0.04	0.12 (0.02)	0.05
Γ_{23} amide N	0.02 (0.003)	0.11	0.04 (0.004)	0.15
Γ_{23} C	−0.10 (0.02)	−0.05	−0.16 (0.02)	−0.06
Γ_{23} total	−0.24 (0.04)	−0.35	−0.33 (0.05)	−0.43

[a]Group-specific Γ_{23} values from MD simulations. Those in parentheses are standard deviations. [b]Group-specific Γ_{23} values predicted from a model based on μ_{23}/ASA fitted to experimental data (see the text for details).

surface C includes all carbon atoms. The group decomposition of Γ_{23} with this scheme is shown in Table 5 for two GB concentrations, along with the predicted ones based on μ_{23}/ASA values fitted to experimental results. Overall, the calculated Γ_{23}^i values from MD simulations agree well with predictions based on μ_{23}/ASA. The only exception is for amide O, for which the predicted Γ_{23}^i is negative while the calculated value based on MD is nearly zero. Nevertheless, the key feature that GB is significantly excluded from carboxylate O is captured well by MD simulations. Moreover, we note that the total Γ_{23} between GB and TGLY predicted on the basis of μ_{23}/ASA does not agree perfectly with experimental measurement. For example, the predicted Γ_{23} at 1.35 m GB is approximately −0.35, which is further from the experimental value of −0.16 compared to MD simulations (−0.24).

Implications. This study has important mechanistic and technical implications regarding the effects of small solutes on protein stability.

With regard to the thermodynamic basis of protein denaturation by urea, the group decomposition of CHARMM-predicted values of Γ_{23} (Table 3) indicates that urea interacts favorably relative to water with both the amide and aliphatic -CH_2- groups. The favorable urea−amide preferential interaction has been observed in simulations (*16*, *18*) and has been quantified in recent analyses of experimental data (*19*, *25*, *54*). For example, from an analysis of experimental osmometric (isopiestic distillation) values of Γ_{23} for interactions of urea with small peptides (e.g., GlyGly, TGLY, and GlyAla), Cannon et al. (*25*) concluded that the dominant preferential interaction of urea with the protein backbone is a favorable interaction with amide groups and were able to quantify this interaction. In a recent NMR study, Lim et al. (*19*) concluded that urea forms hydrogen bonds with the amide group on the basis of the dependence of the hydrogen exchange rate on urea concentration. The urea−aliphatic group preferential interaction observed here is qualitatively consistent with limited data from solubility experiments, which show that aliphatic hydrocarbons (longer than ethane) are more soluble in 7 M urea than in water (*57*, *58*). Additionally, the authors of an osmotic stress/X-ray scattering study of the effect of urea on forces between hydroxypropylcellulose polymers concluded that urea interacts favorably (but very weakly) with the mostly aliphatic surface of these polymers. Moreover, several recent simulation studies (*21*, *22*, *28*, *55*) also proposed that the dispersion interaction between urea and aliphatic groups could be a driving force for denaturation. The most direct computational evidence is the observation that urea weakens the hydrophobic interactions between pure hydrophobic particles and unfolds pure hydrophobic polymers (*22*). Only our study, however, has explicitly and quantitatively compared simulation results with experimental thermodynamic data, which is essential for evaluating the findings from simulations. For example, with the CHARMM27 force field, the calculated preferential coefficient of urea near propane is 0.15 ± 0.07 at 2.26 m urea, which is rather large considering the experimental observation that the butane solubility as a function of urea concentration (*57*) indicates a very small positive Γ_{23}. Overestimation of preferential interaction between urea and aliphatic groups with the CHARMM27 force field is consistent with the overestimated Γ_{23} for urea near TGLY (e.g., at 1.21 m urea, the calculated Γ_{23} is 0.33 while the measured value is 0.12); as noted above, the -$CH_{2/3}$ groups in propane and TGLY have fairly different overall charges, although polar hydrogens have the same partial charges. To further validate this conjecture, comparing computed and measured group decomposition of Γ_{23} for urea and peptide (E. Guinn and M. T. Record, Jr., work in progress) as we have done for GB and peptide is valuable, and the results can provide guidance to the improvement of the CHARMM27 force field. In this context, we note that even a small positive preferential interaction between urea and aliphatic groups still may indicate a significant contribution of such effects in protein unfolding since the amount of exposed aliphatic hydrocarbon surface upon unfolding is very large.

For the stabilizing mechanism of GB, this study reproduces the overall and group-specific Γ_{23} values between GB and TGLY reasonably well compared to experimental analyses. The major exclusion region for GB (see Table 3 and 5) is the carboxylate group, which is consistent with the experimental evidence (*7*) which shows that GB is excluded from anionic biopolymer surfaces. GB is also slightly excluded from aliphatic groups (-CH_2- groups in Table 3 and aliphatic C surfaces in Table 5), which is in line with findings from recent X-ray scattering experiments, which showed that GB is excluded from the nonpolar surface of hydroxypropylcellulose (*59*). The semiquantitative agreement between the surface-specific Γ_{23} and the prediction of Record et al. based on μ_{23}/ASA (Table 5) supports the simple thermodynamic model that assumes additivity of preferential interactions between GB and various biomolecular surfaces.

On the technical side, this study serves as a benchmark for CHARMM (*38*), a popular classical, nonpolarizable force field, in terms of describing interactions among small solutes (urea and GB), a protein backbone (TGLY peptide), and water. It is encouraging that the CHARMM27 force field is shown to reproduce Γ_{23} between urea and TGLY on a semiquantitative level, although there is clearly room for further improvements. For GB, however, the computed Γ_{23} has even the wrong sign if parameters are not carefully chosen to balance the interactions among GB, TGLY, and water. This result clearly highlights that

care needs to be exercised when constructing force field models based on existing parameters for chemically similar groups, which is commonly done in the literature on the basis of the consideration of force field compatibility.

The sensitivity of Γ_{23} to force field parameters suggests that small solute thermodynamic data can be valuable in guiding the development of accurate molecular models, which has already been recognized by several authors (*33*, *37*) but yet to be broadly appreciated. In force field development, emphasis is often placed on solvation free energies, which measure the interaction between functional groups and water. By contrast, small solute data measure interactions between functional groups relative to their interactions with water, which are precisely the driving force behind many biological processes such as ligand binding and protein folding and association. Therefore, small solute thermodynamic data nicely complement solvation free energies in force field development. Considering the molecular (as opposed to atomic) nature of small solute data (e.g., as illustrated by the group decompositions), this is particularly true for model development at the coarse-grained level, which we plan to pursue in the near future with joint experimental and computational efforts. Along this line, we note that to reliably describe the thermodynamics of processes that involve a significant change in the local environment of amino acids (e.g., protein unfolding), it might be important to include electronic polarization so that the effective interaction between groups can adequately adjust to the local environment.

CONCLUSIONS

Motivated by recent experimental efforts to understand the effects of small solutes on protein stability, we have conducted extensive MD simulations to characterize the interaction between two widely used small solutes, urea and glycine betaine (GB), and a triglycine peptide that represents a protein backbone. The key quantity of interest is the preferential interaction coefficient, Γ_{23}, which characterizes the degree of solute accumulation near or exclusion from the protein surface relative to the bulk. To ensure convergence and a thorough comparison with experiments, Γ_{23} is calculated for multiple solute concentrations and each solute−peptide−water ternary system is studied with 200−300 ns of atomistic simulations using the CHARMM force field. The results show that good agreement with experiments can be obtained for both solutes if care is exercised to balance the interactions among water, solute, and protein. On the other hand, a model based on directly taking parameters for chemically similar groups from an existing force field leads to qualitatively incorrect results (i.e., wrong sign in Γ_{23}), which highlights the value of small solute thermodynamic data in guiding the development of accurate force fields for biomolecules.

Despite potential caveats noted in a previous study (*56*), we find that decomposing the calculated Γ_{23} values into group contributions allows us to gain new insights regarding the mechanism of small solute effects. Use of the CHARMM force field predicts that urea preferentially interacts with both amide and aliphatic hydrocarbon groups (CH_2 groups in the triglycine peptide); this supports the conclusion from several recent simulation studies (*21*, *22*, *55*) that interaction between urea and aliphatic groups also plays an important role in urea-induced denaturation. Quantitatively, however, comparison to experimental solubility (*57*, *58*) and X-ray scattering (*59*) data indicates that the CHARMM force field and perhaps others likely overestimate the interaction between urea and aliphatic groups. With regard to GB, this study finds that it is excluded mainly from the carboxylate groups and weakly from nonpolar groups, results that are consistent with recent experimental findings (*7*, *59*). The group decomposition of Γ_{23} also supports an additive thermodynamic model for preferential interactions between GB and various biomolecular surfaces (*27*).

ACKNOWLEDGMENT

We thank Dr. I. A. Shkel for stimulating discussions. Computational resources from the National Center for Supercomputing Applications at the University of Illinois (Urbana, IL) are greatly appreciated.

SUPPORTING INFORMATION AVAILABLE

Detailed results (charges and interaction energies) for the development of GB parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wyman, J. (1964) Linked functions and reciprocal effects in hemoglobin: A second look. *Adv. Protein Chem. 19*, 223–286.
2. Record, M. T., Jr., Zhang, W. T., and Anderson, C. F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv. Protein Chem. 51*, 281–353.
3. Timasheff, S. N. (1998) Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated. *Adv. Protein Chem. 51*, 355–432.
4. Courtenay, E. S., Capp, M. W., Anderson, C. F., and Record, M. T., Jr. (2000) Vapor pressure osmometry studies of osmolyte-protein interactions: Implications for the action of osmoprotectants in vivo and for the interpretation of "osmotic stress" experiments in vitro. *Biochemistry 39*, 4455–4471.
5. Hopkins, F. G. (1930) Denaturation of proteins by urea and related substances. *Nature 126*, 328–330, 383−384.
6. Tanford, C. (1968) Protein denaturation. *Adv. Protein Chem. 23*, 121–282.
7. Felitsky, D. J., Cannon, J. G., Capp, M. W., Hong, J., Van Wynsberghe, A. W., Anderson, C. F., and Record, M. T., Jr. (2004) The exclusion of glycine betaine from anionic biopolymer surface: Why glycine betaine is an effective osmoprotectant but also a compatible solute. *Biochemistry 43*, 14732–14743.
8. Storey, R., and Jones, R. G. W. (1975) Betaine and choline levels in plants and their relationship to nacl stress. *Plant Sci. Lett. 4*, 161–168.
9. Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982) Living with water-stress: Evolution of osmolyte systems. *Science 217*, 1214–1222.
10. Lin, T. Y., and Timasheff, S. N. (1994) Why do some organisms use a urea-methylamine mixture as osmolyte: Thermodynamic compensation of urea and trimethylamine n-oxide interactions with protein. *Biochemistry 33*, 12695–12701.
11. Yancey, P. H., and Somero, G. N. (1980) Methylamine osmoregulatory solutes of elasmobranch fishes counteract urea inhibition of enzymes. *J. Exp. Zool. 212*, 205–213.
12. Volker, J., and Breslauer, K. J. (2005) Communication between noncontacting macromolecules. *Annu. Rev. Biophys. Biomol. Struct. 34*, 21–42.
13. Frank, H. S., and Franks, F. (1968) Structural approach to solvent power of water for hydrocarbons: Urea as a structure breaker. *J. Chem. Phys. 48*, 4746.
14. Vanzi, F., Madan, B., and Sharp, K. (1998) Effect of the protein denaturants urea and guanidinium on water structure: A structural and thermodynamic study. *J. Am. Chem. Soc. 120*, 10748–10753.
15. Idrissi, A., Sokolic, F., and Perera, A. (2000) A molecular dynamics study of the urea/water mixture. *J. Chem. Phys. 112*, 9479–9488.
16. Bennion, B. J., and Daggett, V. (2003) The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. U.S.A. 100*, 5142–5147.

17. Zou, Q., Habermann-Rottinghaus, S. M., and Murphy, K. P. (1998) Urea effects on protein stability: Hydrogen bonding and the hydrophobic effect. *Proteins: Struct., Funct., Genet. 31*, 107–115.
18. Tobi, D., Elber, R., and Thirumalai, D. (2003) The dominant interaction between peptide and urea is electrostatic in nature: A molecular dynamics simulation study. *Biopolymers 68*, 359–369.
19. Lim, W. K., Rosgen, J., and Englander, S. W. (2009) Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proc. Natl. Acad. Sci. U.S.A. 106*, 2595–2600.
20. Paul, S., and Patey, G. N. (2007) The influence of urea and trimethylamine-n-oxide on hydrophobic interactions. *J. Phys. Chem. B 111*, 7932–7933.
21. Hua, L., Zhou, R. H., Thirumalai, D., and Berne, B. J. (2008) Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc. Natl. Acad. Sci. U.S.A. 105*, 16928–16933.
22. Zangi, R., Zhou, R. H., and Berne, B. J. (2009) Urea's action on hydrophobic interactions. *J. Am. Chem. Soc. 131*, 1535–1541.
23. Shearwin, K. E., and Timasheff, S. N. (1992) Linkage between ligand-binding and control of tubulin conformation. *Biochemistry 31*, 8080–8089.
24. Record, M. T., Jr., and Anderson, C. F. (1995) Interpretation of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a 2-domain model. *Biophys. J. 68*, 786–794.
25. Cannon, J. G., Anderson, C. F., and Record, M. T., Jr. (2007) Urea-amide preferential interactions in water: Quantitative comparison of model compound data with biopolymer results using water accessible surface areas. *J. Phys. Chem. B 111*, 9675–9685.
26. Hong, J., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (2005) Use of urea and glycine betaine to quantify coupled folding and probe the burial of DNA phosphates in lac repressor−lac operator bindings. *Biochemistry 44*, 16896–16911.
27. Capp, M. W., Pegram, L. M., Saecker, R. M., Kratz, M., Riccardi, D., Wendorff, T., Cannon, J. G., and Record, M. T., Jr. (2009) Interactions of glycine betaine with molecular surfaces in water: Thermodynamics, structural interpretation and prediction of *m*-values. *Biochemistry 48*, 10372–10379.
28. Caflisch, A., and Karplus, M. (1999) Structural details of urea binding to barnase: A molecular dynamics analysis. *Struct. Folding Des. 7*, 477–488.
29. Baynes, B. M., and Trout, B. L. (2003) Proteins in mixed solvents: A molecular-level perspective. *J. Phys. Chem. B 107*, 14058–14067.
30. Smith, P. E. (2004) Cosolvent interactions with biomolecules: Relating computer simulation data to experimental thermodynamic data. *J. Phys. Chem. B 108*, 18716–18724.
31. Pierce, V., Kang, M., Aburi, M., Weerasinghe, S., and Smith, P. E. (2008) Recent applications of Kirkwood−Buff theory to biological systems. *Cell Biochem. Biophys. 50*, 1–22.
32. Kang, M., and Smith, P. E. (2007) Preferential interaction parameters in biological systems by Kirkwood−Buff theory and computer simulation. *Fluid Phase Equilib. 256*, 14–19.
33. Weerasinghe, S., and Smith, P. E. (2003) A Kirkwood−Buff derived force field for sodium chloride in water. *J. Chem. Phys. 119*, 11342–11349.
34. Weerasinghe, S., and Smith, P. E. (2003) Kirkwood−Buff derived force field for mixtures of acetone and water. *J. Chem. Phys. 118*, 10663–10670.
35. Weerasinghe, S., and Smith, P. E. (2004) A Kirkwood−Buff derived force field for the simulation of aqueous guanidinium chloride solutions. *J. Chem. Phys. 121*, 2180–2186.
36. Weerasinghe, S., and Smith, P. E. (2005) A Kirkwood−Buff derived force field for methanol and aqueous methanol solutions. *J. Phys. Chem. B 109*, 15080–15086.
37. Kang, M., and Smith, P. E. (2006) A Kirkwood−Buff derived force field for amides. *J. Comput. Chem. 27*, 1477–1485.
38. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B 102*, 3586–3616.
39. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem. 4*, 187–217.
40. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926–935.
41. Steinbach, P. J., and Brooks, B. R. (1994) New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem. 15*, 667–683.
42. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577–8593.
43. Nose, S. (1984) A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys. 81*, 511–519.
44. Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A 31*, 1695–1697.
45. Andersen, H. C. (1980) Molecular-dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys. 72*, 2384–2393.
46. Feller, S. E., Zhang, Y. H., Pastor, R. W., and Brooks, B. R. (1995) Constant-pressure molecular-dynamics simulation: The Langevin piston method. *J. Chem. Phys. 103*, 4613–4621.
47. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of Cartesian equations of motion of a system with constraints: Molecular-dynamics of n-alkanes. *J. Comput. Phys. 23*, 327–341.
48. Formaneck, M. S., Ma, L., and Cui, Q. (2006) Effects of temperature and salt concentration on the structural stability of human lymphotactin: Insights from molecular simulations. *J. Am. Chem. Soc. 128*, 9506–9517.
49. Ma, L., and Cui, Q. (2006) The temperature dependence of salt-protein association is sequence specific. *Biochemistry 45*, 14466–14472.
50. Shukla, D., Shinde, C., and Trout, B. L. (2009) Molecular computations of preferential interaction coefficients of proteins. *J. Phys. Chem. B 113*, 12546–12554.
51. Harihara, P. C., and Pople, J. A. (1973) Influence of polarization functions on molecular-orbital hydrogenation energies. *Theor. Chim. Acta 28*, 213–222.
52. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Jr., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) Gaussian 03, revision C.02, Gaussian, Inc., Wallingford, CT.
53. Frenkel, D., and Smit, B. (2002) Understanding Molecular Simulation: From Algorithms to Applications, Academic Press, San Diego.
54. Courtenay, E. S., Capp, M. W., and Record, M. T., Jr. (2001) Thermodynamics of interactions of urea and guanidinium salts with protein surface: Relationship between solute effects on protein processes and changes in water-accessible surface area. *Protein Sci. 10*, 2485–2497.
55. Stumpe, M. C., and Grubmuller, H. (2007) Interaction of urea with amino acids: Implications for urea-induced protein denaturation. *J. Am. Chem. Soc. 129*, 16126–16131.
56. Aburi, M., and Smith, P. E. (2004) A combined simulation and Kirkwood−Buff approach to quantify cosolvent effects on the conformational preferences of peptides in solution. *J. Phys. Chem. B 108*, 7382–7388.
57. Wetlaufer, D. B., Malik, S. K., Stoller, L., and Coffin, R. L. (1964) Nonpolar group participation in the denaturation of proteins by urea and guanidinium salts. Model compound studies. *J. Am. Chem. Soc. 86*, 508–514.
58. Graziano, G. (2001) On the solubility of aliphatic hydrocarbons in 7 M aqueous urea. *J. Phys. Chem. B 105*, 2632–2637.
59. Stanley, C., and Rau, D. C. (2008) Assessing the interaction of urea and protein-stabilizing osmolytes with the nonpolar surface of hydroxypropylcellulose. *Biochemistry 47*, 6711–6718.
60. Humphrey, W., Dalke, A., and Schulten, K. (1996) Vmd: Visual molecular dynamics. *J. Mol. Graphics 14*, 33–38.