

# Aqueous Urea Solutions: Structure, Energetics, and Urea Aggregation

Martin C. Stumpe and Helmut Grubmüller\*

Theoretical and Computational Biophysics Department, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

Received: October 2, 2006; In Final Form: March 19, 2007

Urea is ubiquitously used as a protein denaturant. To study the structure and energetics of aqueous urea solutions, we have carried out molecular dynamics simulations for a wide range of urea concentrations and temperatures. The hydrogen bonds between urea and water were found to be significantly weaker than those between water molecules, which drives urea self-aggregation due to the hydrophobic effect. From the reduction of the water exposed urea surface area, urea was found to exhibit an aggregation degree of ca. 20% at concentrations commonly used for protein denaturation. Structurally, three distinct urea pair conformations were identified and their populations were analyzed by translational and orientational pair distribution functions. Furthermore, urea was found to strengthen water structure in terms of hydrogen bond energies and population of solvation shells. Our findings are consistent with a direct interaction between urea and the protein as the main driving force for protein denaturation. As an additional, more indirect effect, urea was found to enhance water structure, which would suggest a weakening of the hydrophobic effect.

## 1. Introduction

Aqueous urea solutions are widely used for protein denaturation. They therefore have received a lot of interest in the past and are still subject of ongoing research. Despite its ubiquitous use, only little is known about the molecular mechanism underlying urea-induced protein denaturation. However, in order to explain how denaturants work, the delicate balance between intra-protein, protein–solvent, and intra-solvent interactions has to be understood on a qualitative and structural level. Of these three, we here will focus on the intra-solvent interactions.

Urea water solutions show a number of remarkable properties, in particular, enhanced solubility of hydrocarbons,<sup>1</sup> decreased micelle formation,<sup>2</sup> and, most importantly, the ability to denature proteins.<sup>3</sup> In an attempt to explain these properties, mainly two models were proposed in the 1960s on the basis of thermodynamic arguments and still set the framework for ongoing discussions. The SKSS model,<sup>4–6</sup> proposed by Schellman, Kresheck, Sheraga, and Stokes, attributes the properties of aqueous urea solutions to dimerized or oligomerized urea. In contrast, the FF model,<sup>7</sup> suggested by Frank and Franks, focuses on the changes in the water network induced by urea and regards it as a “structure breaker” for water. Since the 1960s, a wealth of new information has been provided by calorimetry,<sup>8,9</sup> circular dichroism spectroscopy,<sup>10,11</sup> neutron scattering,<sup>12</sup> NMR spectroscopy,<sup>13</sup> fluorescence measurements,<sup>14</sup> the time resolved optical Kerr effect,<sup>15</sup> IR and Raman spectroscopy,<sup>16–18</sup> and a number of molecular dynamics simulation studies.<sup>19–48</sup> For a more comprehensive bibliographical and historical overview, see refs 49 and 50.

Today, it is widely accepted that urea exhibits a certain tendency to self-aggregate in aqueous solution.<sup>23,34</sup> We will address this point on the structural level in terms of urea pair conformations and quantify the aggregation. In many other cases, however, experimental data are still somewhat contradic-

tory. One of the most controversial issues is the effect of urea on water structure. In some works urea is suggested to disrupt the natural water structure and is termed a “structure breaker” or “chaotrope”.<sup>51,52</sup> Others find urea to enhance the water structure and coined the terms “structure maker” or “kosmotrope”.<sup>28,34</sup> Both of the two views attribute the peculiarities of urea to changes in the water structure and are in agreement with the FF model. However, others find no or only negligible changes in the water structure and suggest this finding as evidence for the SKSS model.<sup>16,31,32,37,53</sup> A number of recent studies also suggest a combination of direct and indirect effects.<sup>9,21,28,33</sup>

So far, most studies have focused on radial distribution functions as a means to analyze both the influence of urea on water structure and urea self-aggregation. These two effects, apparently, are closely related to the question at hand but, in light of the molecular complexity involved, probably fall short of capturing the whole picture. In particular, we here argue therefore that not only the distribution functions but also the energetics of urea–water interactions need to be investigated in more detail.

We will address these issues along three lines by molecular dynamics simulations. First, we will characterize the short-range structural properties of urea–water mixtures and urea aggregation by calculating generalized density distribution functions including three translational and three rotational molecular degrees of freedom. Second, the energetics will be addressed in terms of hydrogen bond interactions. Third, structural perturbations imposed by urea molecules on water structure will also be characterized by density distributions. Going beyond previous studies, the temperature and urea concentration dependence will be covered here for the full experimentally accessible range.

To evaluate the accuracy of the urea force field used, atom–atom distribution functions, calculated from our simulations, will be compared to neutron scattering data. Additionally, we will compare mass densities to experimental values. In summary,

\* Corresponding author. Phone: +49-551-201-2301. Fax: +49-551-201-2302. E-mail: hgrubmu@gwdg.de.

TABLE 1: Setup of the Simulation System<sup>a</sup>

number of water molecules	number of urea molecules	mole fraction	urea concentration (mol/L)					
			280 K	300 K	320 K	340 K	360 K	380 K
895	0	0.00	0.0	0.0	0.0	0.0	0.0	0.0
844	16	0.02	1.0	1.0	1.0	1.0	0.9	0.9
787	33	0.04	2.1	2.1	2.1	2.0	2.0	1.9
730	49	0.06	3.2	3.2	3.1	3.1	3.0	2.9
690	65	0.09	4.3	4.2	4.2	4.1	4.0	3.9
612	81	0.12	5.6	5.5	5.4	5.3	5.2	5.1
567	98	0.15	6.8	6.7	6.6	6.5	6.4	6.3
518	114	0.18	8.0	7.9	7.8	7.7	7.6	7.4
493	130	0.21	9.0	8.9	8.8	8.6	8.5	8.3
437	146	0.25	10.4	10.2	10.1	9.9	9.8	9.6
401	163	0.29	11.5	11.3	11.2	11.0	10.8	10.7

<sup>a</sup> Listed are the number of water and urea molecules in the system, urea mole fractions, and urea concentrations. Due to changes in the box volume, the effective concentrations slightly differ with temperature. The saturation limit at 300 K is at a mole fraction of 0.27.<sup>75</sup>

a comprehensive picture of urea–water mixtures at the molecular level is obtained, which has implications for putative mechanisms of urea-induced protein denaturation.

## 2. Methods

**2.1. Simulation Setup.** All simulations were performed using the Gromacs<sup>54,55</sup> program suite, versions 3.1.4 and 3.2.1, with the OPLS-all-atom force field.<sup>56</sup> The TIP4P water model<sup>57</sup> was used, and the urea force field was adopted from Smith et al.,<sup>22</sup> which is a refined version of a previous one by Duffy et al.<sup>58</sup> A cutoff of 1.0 nm was used for short-range Coulomb as well as Lennard-Jones interactions. Particle mesh Ewald summation<sup>59,60</sup> was used to calculate the long-range electrostatic interactions with a grid spacing of 0.12 nm and an interpolation order of 4. All simulations were performed in the  $NpT$  ensemble using Berendsen-type temperature coupling<sup>61</sup> with a coupling coefficient of  $\tau_T = 0.1$  ps and Berendsen-type pressure coupling<sup>61</sup> at 1 bar with a coupling coefficient of  $\tau_p = 1$  ps. An integration time step of 1 fs was used. The initial size of the periodic rectangular box was set to  $(3 \text{ nm})^3$  in all simulations. To set up the simulation systems for various urea concentrations ranging from 0 to 11.3 mol/L, 0 to 163 (non-overlapping) urea molecules were placed at random positions within the simulation box, which was subsequently filled up with TIP4P water molecules using the genbox routine of the Gromacs package. All 11 systems shown in Table 1 were simulated for 10 ns each at six different temperatures ranging from 280 to 380 K, totaling about 1  $\mu$ s of simulation time.

**2.2. Density Distributions.** To characterize the short-range order of urea and water, six-dimensional (three translational and three rotational degrees of freedom) density distributions,  $\rho(\Delta x, \Delta y, \Delta z, \alpha_x, \alpha_y, \alpha_z)$ , were calculated for water–water, water–urea, and urea–urea. To this end, the relative position and orientation of molecules with a center-of-mass distance smaller than 0.5 nm was computed pairwise. The orientation of one molecule with respect to the other was expressed in axis-angle notation; that is, a three-dimensional vector  $(\alpha_x, \alpha_y, \alpha_z)$  denotes the direction of the rotation axis, and the length of this vector defines the rotation angle. For example,  $(\pi, 0, 0)$  describes a  $180^\circ$  rotation around the  $x$ -axis. The coordinate system was defined as shown in Figure 1.

All position histograms were built using 100 bins in each of the three dimensions and smoothed with a three-dimensional Gaussian function of 0.01 nm width,<sup>62</sup> which was chosen to trade off resolution and statistical noise. Orientation histograms for subsets of positions were calculated similarly. All histograms were normalized such that the sum over all bins was 1.

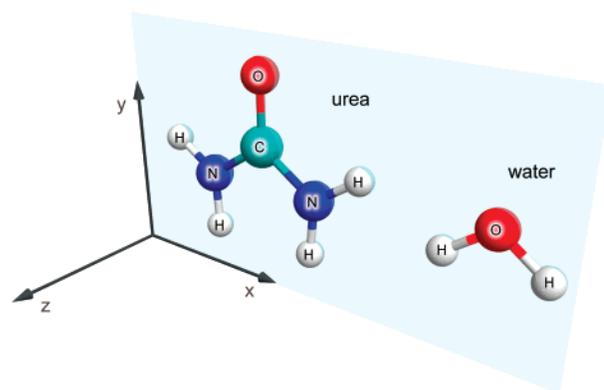


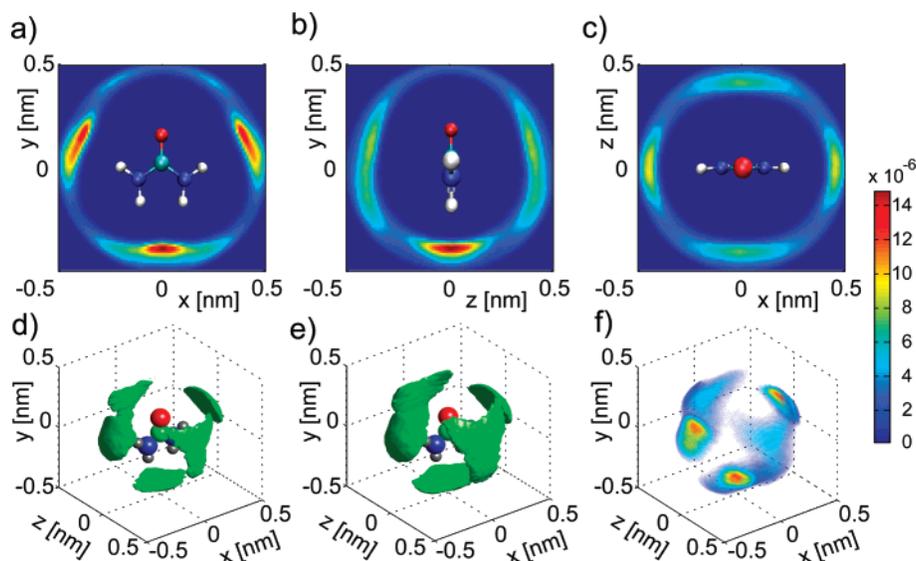
Figure 1. Definition of the coordinate system used for calculations of positional and orientational histograms.

**2.3. Hydrogen Bonds.** The number of hydrogen bonds per water or urea molecule was calculated using the standard Gromacs tools with a cutoff radius of 0.35 nm between donor and acceptor and a cutoff angle of  $30^\circ$ .<sup>63</sup> Energies of hydrogen bonds were estimated using the empirical function,<sup>64</sup> where  $d$  denotes the distance between hydrogen atom and acceptor atom.

We note that we here rely only on the widely accepted<sup>65–69</sup> fact that this formula provides at least a semiquantitative measure for the H-bond strength. In particular, our conclusions are based on the monotonic dependence of H-bond energy on distance (in the considered distance range), rather than on accurate numbers.

**2.4. Urea Aggregation.** The tendency of urea molecules to self-aggregate was analyzed and quantified in terms of the reduction of total water accessible surface exposed by urea (“interface surface”) with respect to nonaggregated urea. To this aim, a sphere of 0.14 nm radius was used to probe the surface with the double cubic lattice method.<sup>70</sup> The interface surface area was used as a measure because its minimization is assumed to be the main driving force for the aggregation. Furthermore, this quantity was found to be more sensitive to the size of the contact area than other measures such as Kirkwood–Buff integrals<sup>71</sup> or cluster analysis (data not shown).

The accurate assessment of aggregation is complicated by the fact that, also in the absence of any interaction and, therefore, also aggregation, random contacts between urea molecules would already reduce this solvent exposed surface area. To distinguish this “geometric aggregation” from real aggregation, we performed two types of additional simulations. A first set of simulations was carried out with completely uncharged (i.e., super-hydrophobic) urea molecules to maximize the hydrophobic



**Figure 2.** Density distributions of urea molecules with respect to a reference urea molecule: (a) in the  $xy$  plane; (b) in the  $yz$  plane; (c) in the  $xz$  plane; (d) isosurfaces enclosing 15% of all urea molecules within a 0.5 nm distance; (e) same with 25%; (f) color- and transparency-coded density distribution enclosing 50%. Low densities, blue and most transparent; high densities, red and least transparent (see color bar for parts a–c and f). Five preferential coordination sites are seen, three of which actually differ from each other.

effect and enforce maximal aggregation. The obtained surface area served to define the 100% aggregation level. To preclude artifacts caused by periodic boundary conditions (see the Appendix), a box size of  $(5 \text{ nm})^3$  was used for the higher concentrations.

For a second set of simulations, both water and urea molecules were completely uncharged to remove any hydrophobic effect from the simulation, such that the opposite extreme of purely stochastic clustering was achieved, which defined the 0% aggregation level. To prevent evaporation of the resulting van der Waals liquid in this case, these simulations were performed under constant volume conditions.

To assess the statistical accuracy of the estimated interface surface area, its autocorrelation function was calculated, and an autocorrelation time of about 100 ps was obtained. Hence, for each of the 10 ns trajectories, an effective number of 100 independent measurements can be assumed, implying a statistical error of  $1/N^{1/2} = 1/10$  of the data obtained from the simulations.

**2.5. Comparison to Experiments, Controls.** Mass densities were compared to the experimental fitting function presented by Sokolic et al.<sup>27</sup> Radial distribution functions (rdf's) for selected atom pairs were compared to experimental based data obtained from neutron scattering experiments<sup>12</sup> with empirical potential structure refinement.<sup>72</sup> For this comparison, we performed a simulation with a mole fraction of 0.2 (500 water molecules, 125 urea molecules) which corresponds to the experimental concentration. The influence of the water model on the coordination geometries was assessed with a pure water simulation at 300 K using the TIP5P model.<sup>73</sup> To detect possible artifacts caused by periodic boundary conditions or finite box size, a larger system of  $(6 \text{ nm})^3$  box size with 508 urea molecules and 5520 water molecules (4.2 M) was simulated at 300 K for 20 ns, and the results obtained from this simulation were compared to those from the  $(3 \text{ nm})^3$  system of the same concentration.

### 3. Results and Discussion

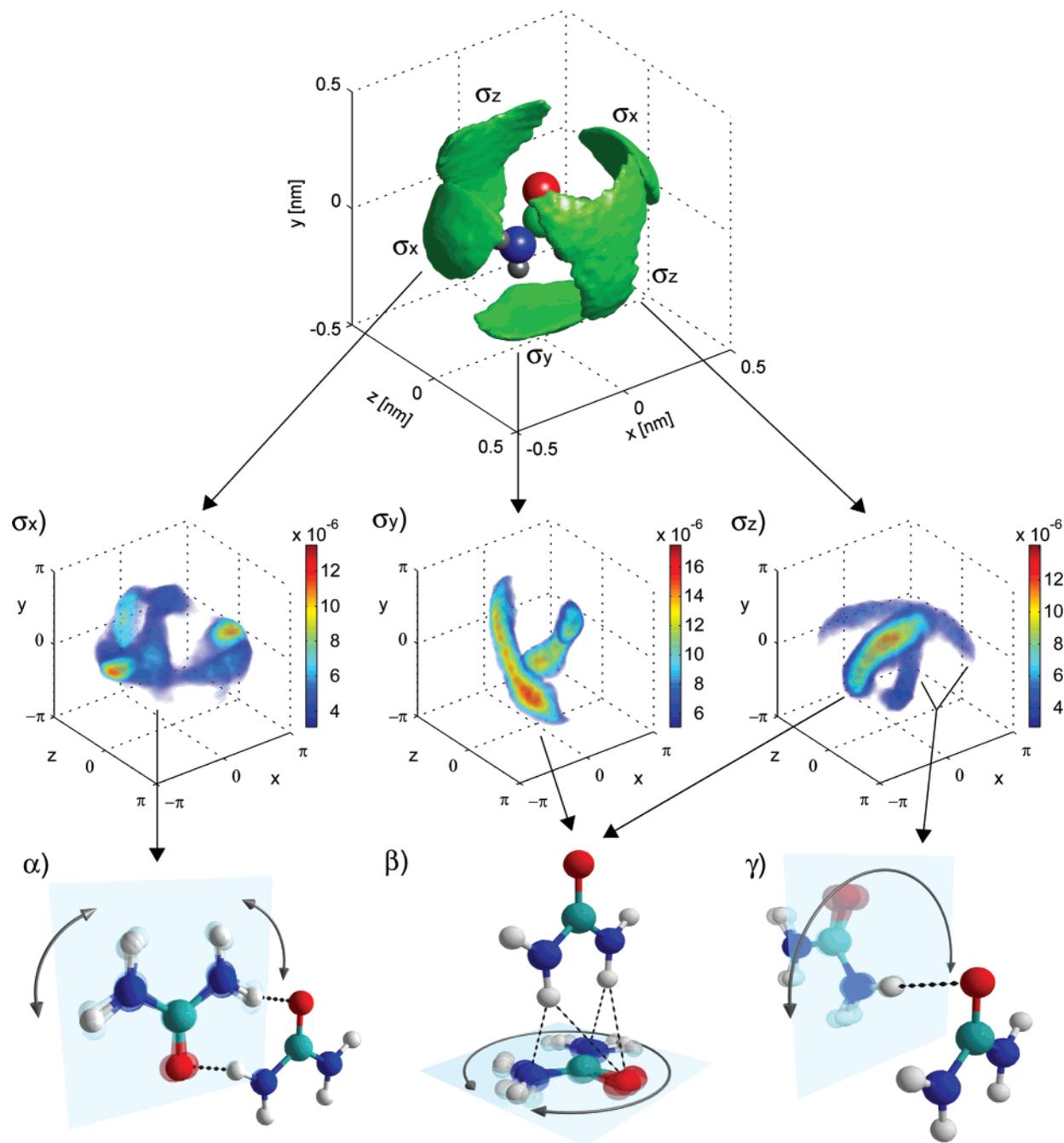
**3.1. Spatial Density Distributions.** To characterize the geometry and short-range order of urea–water solutions, spatial

density distributions were extracted from 10 ns molecular dynamics simulations. In the following, the results for 300 K and a mole fraction of 0.21 (8.9 M), which is in the common range for denaturation, are presented.

We first focus on the relative position of the urea molecules with respect to each other. Positions and orientations were determined to be pairwise, although larger connected networks or clusters of urea were seen, without evidence for dimerization to be preferred over multimerization. Figure 2 shows the spatial density distributions within a distance of 0.5 nm from a reference molecule (center of mass). Panels a–c show the densities in planes through the center of mass of the reference molecule. High densities appear in red, and low densities, in blue color. The full three-dimensional structure of the preferred coordination positions is shown in panels d and e via isodensity surfaces. Panel f shows the same data color- and transparency-coded. Five distinct high-occupancy regions could be identified. Due to the  $C_{2v}$  symmetry of the urea molecule, only three of these are actually different from each other. According to their positions, we labeled these regions “ $\sigma_x$ ”, “ $\sigma_y$ ”, and “ $\sigma_z$ ”, as indicated in Figure 3.

As a next step, the orientation relative to the reference urea molecule was examined. The orientational distributions (Figure 3,  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$ ) were relatively localized, implying strong orientational preferences. The highest density corresponds to those orientations that are prevalent at the preferred positions  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  (Figure 3,  $\alpha$ ,  $\beta$ ,  $\gamma$ ). The size of the high-density regions can be used to compare the accessible phase space volume and, hence, entropies of the three corresponding conformations.

The orientation histogram of molecules in the  $\sigma_x$  position (Figure 3,  $\sigma_x$ ) exhibits two regions of high density at  $(0, 0, \pm\pi)$  as well as two maxima at  $(\pm\pi, 0, 0)$ . Due to the symmetry of the urea molecule, these four rotations are equivalent and correspond to the same conformation where the urea molecules are rotated by about  $180^\circ$  with respect to the reference molecule around the  $z$ -axis. This conformation (“ $\alpha$ ”) has been termed cyclic urea<sup>37,44</sup> and is the enthalpically most favorable one, as it is stabilized by two hydrogen bonds (depicted as dotted lines) between O as acceptor and N as donor. Entropically, in contrast,



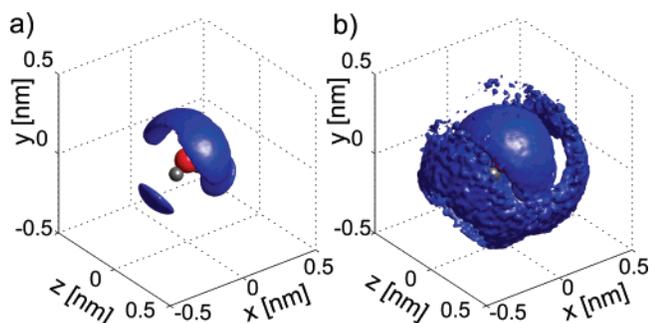
**Figure 3.** Three dominant urea pair conformations are seen in the simulations. Panels  $x$ ,  $y$ , and  $z$  show the orientational density distribution at the respective positions,  $x$ ,  $y$ ,  $z$ , in axis-angle representation. Panels  $\alpha$ ,  $\beta$ , and  $\gamma$  illustrate the three different urea pair conformations obtained from the respective orientational density distributions. The length of the arrows in panels  $\alpha$ ,  $\beta$ , and  $\gamma$  depicts the rotability of the molecule in the cyan plane. Hydrogen bonds are marked by black dotted lines. In conformation  $\beta$ , the smaller dots depict weak electrostatic interactions.

it is the least favorable one of the three conformations, because the two hydrogen bonds restrict the relative mobility of the molecules (as indicated by the short arrows in Figure 3,  $\alpha$ ). Due to the entropic penalty, it is expected that this conformation becomes less populated at higher temperature, which was in fact found in the simulations (data not shown).

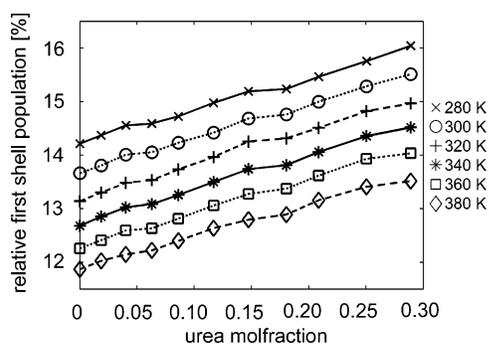
The two curved volumes in the orientational density distribution of the  $\sigma_y$  position (Figure 3,  $\sigma_y$ ) correspond to a conformation in which the urea molecule is tilted by  $90^\circ$  around the  $x$ -axis with respect to the reference molecule and can freely rotate around its  $z$ -axis (as indicated by the long arrow in Figure 3,  $\beta$ ). This conformation (" $\beta$ ") has weak hydrogen bonds between donor N and acceptors O and N. Due to the free rotability, it is

the entropically most favorable one, and its population does not decline at higher temperatures.

Finally, position  $\sigma_z$  (Figure 3,  $\sigma_z$ ) shows three curved density clouds for the orientational degrees of freedom. The most populated one also corresponds to conformation  $\beta$ , but here with the other molecule as reference molecule. The other two density clouds represent a third conformation (" $\gamma$ ") in which the urea molecule is rotated by  $90^\circ$  around the  $y$ -axis and then inclined by a few degrees around its new  $z$ -axis so that a hydrogen bond between N as donor and O as acceptor can form (Figure 3,  $\gamma$ ). Urea molecules in this conformation are seen to rotate in a range from  $-90^\circ$  to  $90^\circ$  around the hydrogen bond. In terms of enthalpy and entropy, it falls between conformations  $\alpha$  and  $\beta$ .



**Figure 4.** Spatial density distribution of water with respect to the reference water molecule: (a) first hydration shell (isurface enclosing 15% of all molecules in a 0.5 nm distance); (b) second hydration shell (isurface enclosing 40% of all molecules in a 0.5 nm distance).



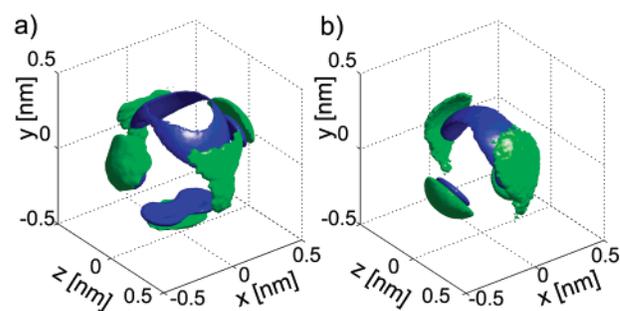
**Figure 5.** Relative population of the first hydration shell around water molecules (normalized to the total number of water molecules within 0.5 nm distance).

There is one relatively strong hydrogen bond with O as acceptor, and the rotational entropy is larger than that for  $\alpha$  but smaller than that for  $\beta$ . As expected, the population of this conformation also decreased at higher temperatures, but not as strongly as that of conformation  $\alpha$ .

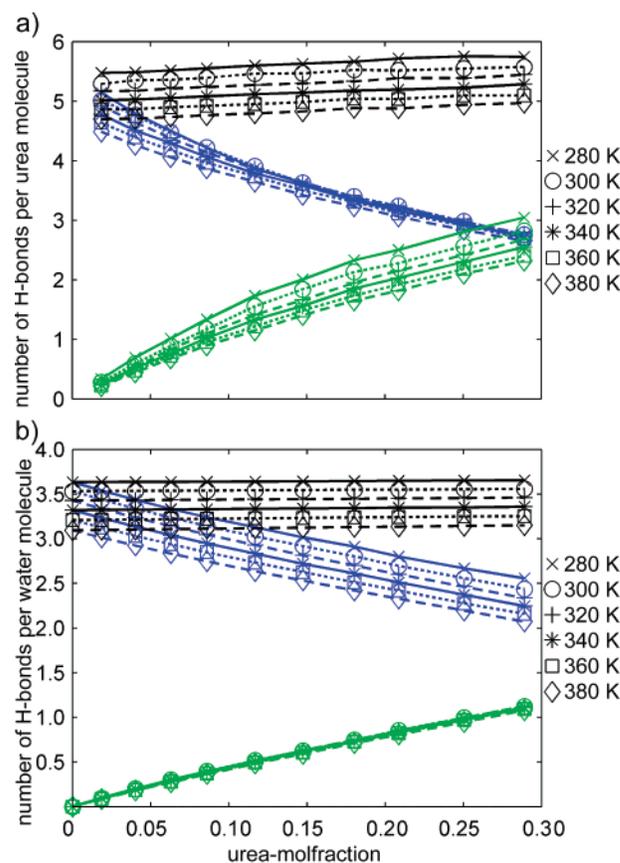
We now focus on the position of water molecules with respect to other water molecules. Two hydration shells can be identified in the respective density distribution (Figure 4). The first, inner hydration shell contains the four tetrahedrally arranged sites for hydrogen bonding. As a result of the single point charge representing the two free electron pairs on the oxygen atom in the TIP4P model, the region between the two acceptor positions is also populated. A perfectly tetrahedral density distribution was observed in a test simulation using TIP5P water,<sup>73</sup> where each of the two free electron clouds of the oxygen is represented by a dummy atom (results not shown).

As a measure for short-range order in the water structure, we used the relative population of the first solvation shell. This property was defined as the fraction of water molecules whose center of mass is enclosed in the volume shown in Figure 4, out of all water molecules which are within 0.5 nm center-of-mass distance. For illustration, this property is expected to be maximal in ice due to the perfect tetrahedral arrangement of the molecules. We find this property to increase with decreasing temperature as well as increasing urea concentration (Figure 5). In this regard, we find urea to strengthen water structure and make it more icelike. The hydrogen bond analysis further below will corroborate this finding.

Finally, we turn to the density distributions of urea and water with respect to each other, shown in Figure 6. As can be seen, the preferred coordination positions of urea and water widely overlap. Because of the larger size of the urea molecule, the urea density maxima are located a bit farther away from the reference center.



**Figure 6.** Spatial density distributions of urea (green) and water (blue) with respect to (a) urea and (b) water, displayed as isosurfaces enclosing 15% of all molecules within a 0.5 nm distance.

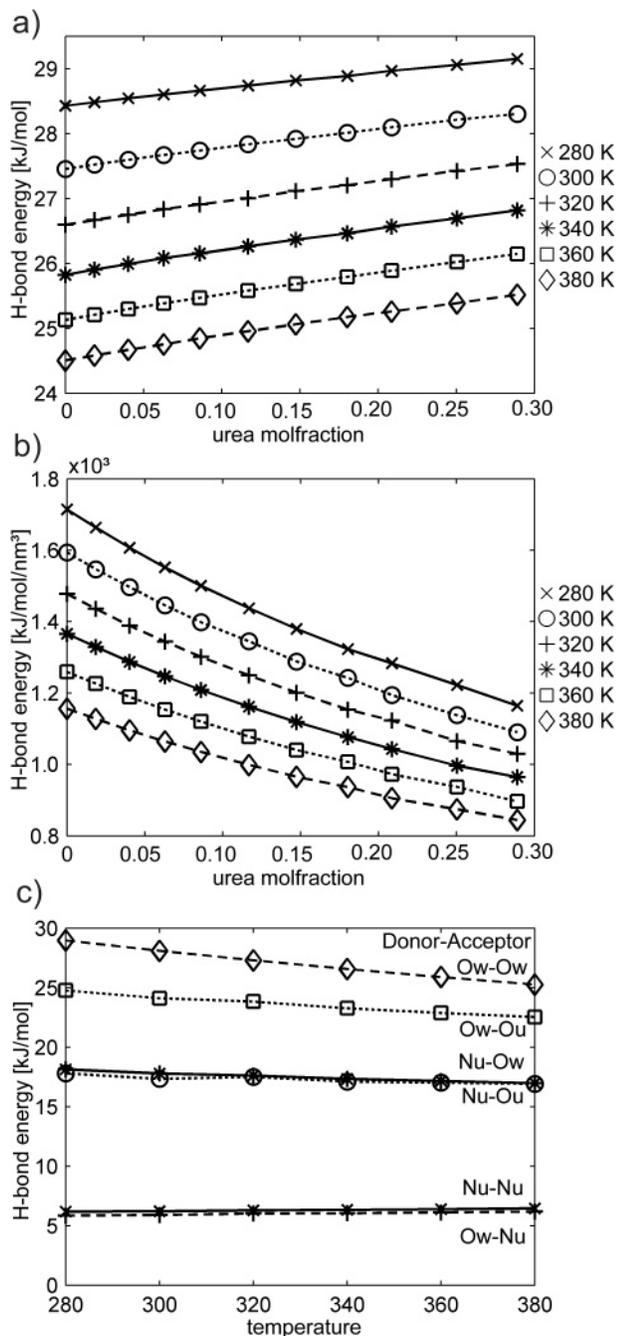


**Figure 7.** Average number of hydrogen bonds, normalized (a) per urea molecule and (b) per water molecule. Blue, hydrogen bonds to water; green, hydrogen bonds to urea; black, sum.

The large overlap between the coordination geometries of urea and water together with the aforementioned only minor perturbations in the water structure even at high urea concentrations likely contributes to the high solubility of urea in water.

**3.2. Hydrogen Bonds.** Figure 7 shows the number of hydrogen bonds per molecule for different urea concentrations. As expected, for all concentrations, a decrease of the number of hydrogen bonds with increasing temperature is observed due to the enthalpic nature of the hydrogen bonds.

The number of hydrogen bonds per urea molecule increases with urea concentration (Figure 7a), since more urea–urea hydrogen bonds are formed than urea–water hydrogen bonds are lost. For water, however, the average number of hydrogen bonds per molecule increases only marginally with increasing urea concentration from 3.53 in pure water to 3.56 in 11.3 M

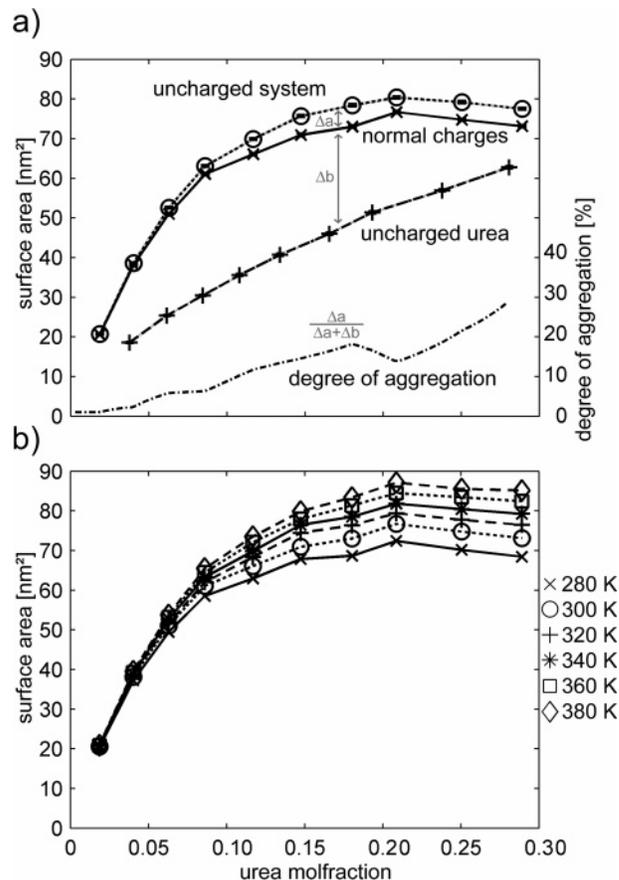


**Figure 8.** (a) Water–water hydrogen bond energies. (b) Total hydrogen bond energy per volume. (c) Mean energy per hydrogen bond for all donor–acceptor combinations (at 0.21 mole fraction).

urea solution at 300 K (Figure 7b). In this respect, urea substitutes well for water and perturbs the hydrogen bond network only slightly.

Considering not only the numbers but also the energetics of the hydrogen bonds, a significant effect of urea concentration becomes evident. Figure 8a shows that the water–water hydrogen bonds become stronger with increasing urea concentration. This is in agreement with the aforementioned strengthening of the water structure (see Figure 5) and supports the view of urea being a structure maker rather than a structure breaker for water.

However, despite the strengthening of single water–water hydrogen bonds with increasing urea concentration, the total hydrogen bond energy per volume of the solution decreases significantly (Figure 8b).



**Figure 9.** Interface surface area (a) with regular charges (solid line) within the range of minimal aggregation (uncharged system, dotted line) and maximal aggregation (uncharged urea, dashed line) at 300 K. The lower, dash–dotted line displays the degree of aggregation in percent (right axis). (b) For 8.9 M at different temperatures. The statistical error for the surface areas is about 3 nm<sup>2</sup>.

An analysis of the individual hydrogen bond energies between all donor–acceptor pairs (Figure 8c) provides further insight and also explains this effect. The strongest hydrogen bonds were formed between water molecules with an average energy of 27.6 kJ/mol (300 K, 8.9 M). Between water and urea, the strongest hydrogen bonds were those formed between the water oxygen atom ( $O_W$ ) as donor and the urea oxygen atom ( $O_U$ ) as acceptor with a mean energy of 24.1 kJ/mol. The weakest hydrogen bonds existed between  $O_W$  as donor and the urea nitrogen atom ( $N_U$ ) as acceptor (5.9 kJ/mol) and between  $N_U$  as donor and acceptor (6.2 kJ/mol).

This progression of water–water, urea–water, and then urea–urea from strongest to weakest hydrogen bonds has two consequences. First, due to the decreasing number of water–water hydrogen bonds in favor of the less energetic water–urea and urea–urea hydrogen bonds with increasing urea concentration (compare also Figure 7), the total hydrogen bond energy per volume decreases, as shown in Figure 8b. Second, urea self-aggregation due to the hydrophobic effect is expected. The latter will be analyzed in more detail below.

**3.3. Urea Aggregation.** We quantified urea self-aggregation by measuring the interface surface area between water and urea, as described in the Methods section. Figure 9a shows the mean interface surface area between urea and water for different hypothetical partial charges.

The lower limit for aggregation (0%) was calculated with both urea and water completely uncharged and exhibits the largest interface surface area (dotted line in Figure 9a). In the

absence of any contacts between urea molecules, the interface surface area is expected to increase linearly, by 1.41 nm<sup>2</sup> per urea molecule. However, due to the limited volume, random contacts (“stochastic clustering”) between urea molecules occur, and thus, the interface surface area increases less than linearly with the concentration. Beyond a mole fraction of 0.2, the interface surface area decreases. This effect becomes obvious by considering that a system with pure urea would not have any interface surface at all. The upper limit for aggregation (100%) was defined using a set of simulations with completely uncharged urea but water with regular charges (dashed line in Figure 9a).

The two extremes of purely stochastic clustering (dotted line in Figure 9a) and maximal aggregation for uncharged urea (dashed line) are subsequently used to define a scale to quantify the degree of urea aggregation. Maximal aggregation was taken as 100% urea aggregation and purely stochastic clustering as 0%.

The degree of urea aggregation within this scale is depicted by the lower (dash-dotted) line in Figure 9a. At typical concentrations used for protein denaturation (about 8–9 M, mole fraction of 0.2), an aggregation of ca. 20% is seen. At these concentrations, the difference in surface area was 5.6 nm<sup>2</sup>, corresponding to the surface of four urea molecules.

Figure 9b shows that urea aggregation decreases with temperature, which corresponds to higher solubility. We note that part of the surface area increase also resulted from the increase of volume with temperature, as seen from test simulations with constant volume conditions. This effect, however, was found to be small (data not shown).

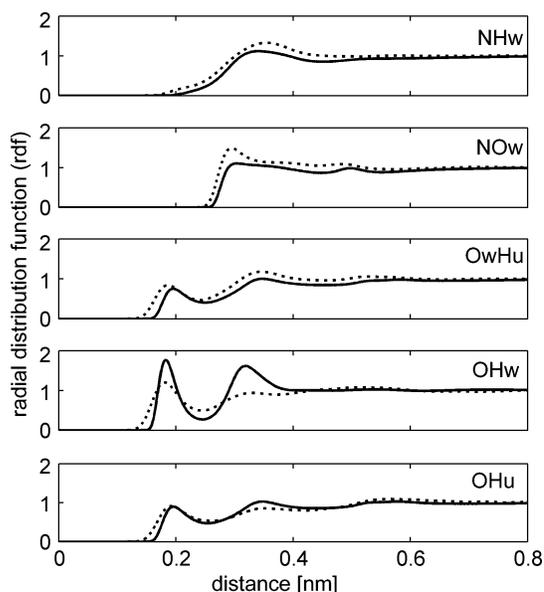
In light of the denaturing effect of urea on proteins, this tendency to self-aggregate might point toward and explain preferential binding to the peptide backbone and less polar parts of the protein which are more exposed in the denatured state.

**3.4. Comparisons to Experiments, Controls.** To assess the accuracy of the force field, we compared our simulation data to measured mass densities and radial distribution functions.

The mass densities,  $\rho$ , of the simulations at 300 K deviated by less than 1% from the experimental data using the fit function<sup>27</sup> for all urea concentrations,  $c_{\text{urea}}$ , given in mol/L (data not shown). Radial distribution functions for various atom pairs involved in hydrogen bonding were calculated and compared to neutron scattering data (Figure 10) from Soper and co-workers obtained with empirical potential structure refinement.<sup>12</sup> As can be seen, the positions of the peaks are well reproduced. In some cases, the heights of the peaks deviate. However, the peak heights are of minor relevance for our conclusions, since they are based on the energies of the hydrogen bonds, which correlate with the respective peak positions.

To test if our results are affected by possible periodic boundary or finite size artifacts, a larger system size of (6 nm)<sup>3</sup> with 4.2 M urea at 300 K was simulated and yielded virtually identical structural and energetic properties as compared to the corresponding simulation with the smaller system size. We therefore can safely exclude such artifacts for the present study.

Despite the good agreement with experimental data, we note that the choice of force field is critical. In particular, the extent of urea aggregation has been found to be sensitive to force field details.<sup>46</sup> The OPLS urea model in combination with the TIP3P and SPC/E<sup>74</sup> water models has been found to overestimate urea aggregation.<sup>46</sup> In the present study, the combination of the TIP4P water model with the refined OPLS urea<sup>22</sup> was used, which was not investigated in the study of Weerasinghe et al.<sup>46</sup> However,



**Figure 10.** Radial distribution functions between atoms involved in hydrogen bonds. Solid lines, simulation data; dotted lines, refined data from neutron scattering experiments (data taken from ref 12).

since the urea model used in the present work is based on the OPLS model, the numbers given here might be somewhat too large.

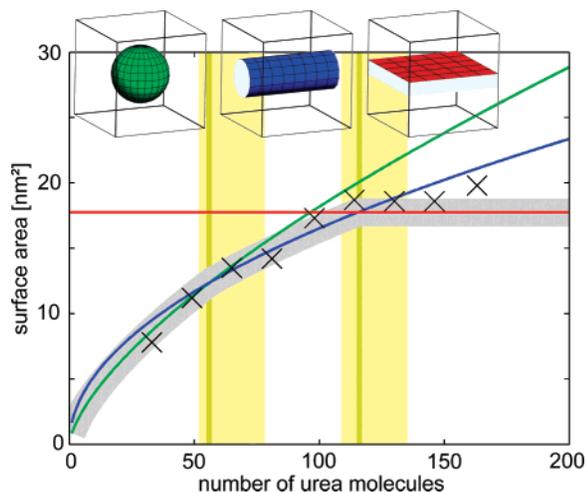
#### 4. Summary and Conclusion

This study of the structure and energetics of urea–water systems aimed at a deeper understanding of the special properties of these systems with a special focus on the ability of urea to denature proteins. We quantified the aggregation tendency of urea to ca. 20% on a scale ranging from purely stochastic clustering to full aggregation. Three different pair conformations for urea have been identified. As a result of the respective entropic and enthalpic contributions, the relative population of each conformation depends on temperature to a different extent.

Even at high urea concentrations, the mutual density distribution of water molecules was found to be changed surprisingly little even at high urea concentrations. Furthermore, the coordination position distributions between urea surrounding water and water surrounding urea, respectively, were found to be very similar. This finding, together with an almost concentration-independent number of hydrogen bonds per water molecule, shows that urea is able to substitute for water in the hydrogen bond network very well in geometric and sterical terms. In contrast, the energetics is different. In particular, urea was found to be a kosmotrope and to strengthen water structure due to stronger water–water hydrogen bonds and a more rigid occupation of the tetrahedral coordination positions.

Whereas these features did not yet provide clues for the mechanism of protein denaturation, an analysis of the hydrogen bond energies between urea and water revealed that the water–water hydrogen bonds are considerably stronger than those between water and urea or urea and urea.

While urea is able to incorporate into the water network, differences in the hydrogen bond energies lead to urea self-aggregation. Accordingly, the mechanism of urea-induced protein denaturation might be entropically dominated via hydrophobic interactions, for example, in the form that urea interfaces between water and the urea-like backbone or less polar residues of the protein. Preferential binding to the peptide backbone and urea-like residues would be another instance of



**Figure 11.** Interface area for strong aggregation with periodic boundaries as predicted by theoretical model (green, sphere; blue, cylinder; red, layer; gray, minimum surface; dark yellow, transition concentrations; black crosses, simulation data for 50% regular charges with offset correction). The yellow bars depict the transition concentrations between the shapes observed in the simulation.

the well-known rule *similia similibus solvuntur* (like solvates like). Alternatively, the strengthening of the water structure could also facilitate protein denaturation by a mechanism resembling cold denaturation, that is, the entropic penalty to solvate hydrophobic groups, and thus, the hydrophobic effect is reduced for higher ordered water. Therefore, we conclude that this indirect effect might add to the direct effect to drive protein denaturation efficiently. Although the changes in the water structure seem to be rather small compared to the more direct effect, a quantification of the balance of these thermodynamic driving forces poses an attractive challenge for future work.

**Acknowledgment.** We thank Alan K. Soper for kindly providing the neutron scattering data and Ira Tremmel for carefully reading the manuscript and for scientific discussions. M.C.S. gratefully acknowledges support from the Deutsche Volkswagen Stiftung, grant I/78 839.

## 5. Appendix

When studying urea aggregation, the use of a simulation box with periodic boundary conditions, together with a high degree of aggregation, led to an artifact which has to be accounted for, and which is described in this Appendix. A sphere has the smallest surface area ( $4\pi r^2$ , where  $r$  denotes the radius) for a given volume and hence is the preferential shape for a strongly aggregating substance. Above a certain concentration, however, a cylinder spanning the box width (see illustration in Figure 11) exhibits an even smaller surface ( $2\pi r b$ , where  $r$  is the radius and  $b$  is the box size, hence the height of the cylinder), since its caps touch each other at the periodic image and thus effectively vanish. Along similar lines, a planar layer offers the smallest surface area ( $2b^2$ , where  $b$  is the box size) at even higher concentrations.

This behavior was observed in the simulations with reduced urea partial charges, which were performed to quantify the degree of urea self-aggregation. Figure 11 shows the calculated surface areas for the three shapes sphere (green line), cylinder (blue line), and layer (red line). The points where the lines intersect mark the transition points from one shape to another (dark yellow lines) and define the minimal surface area (gray line).

A simulation set was performed with only 50% regular urea charges to obtain increased aggregation. These simulations exhibited the described geometric transitions from sphere to cylinder to layer with increasing urea concentration. The observed transition regions are marked as light yellow bars in Figure 11 and are in good agreement with the calculated transition points. The interface surface area (black crosses in Figure 11) follows the minimal surface after an offset correction to account for the roughness of the surface, deviations from ideal shapes, partly solvated urea, and box volume fluctuations.

For the situation at hand, this change in “ideal” surface area due to periodic boundary conditions would distort our aggregation scale. Therefore, those urea clusters that had formed cylinders or layers, which occurred at mole fractions larger than 0.1 in the  $(3 \text{ nm})^3$  box, were simulated in a larger box of size  $(5 \text{ nm})^3$  with additional water, such that a spherical shape was retained in the simulation. Solvated urea molecules remote from the sphere were excluded from the calculation of interface surface area, and the concentration was corrected accordingly. This procedure allowed accurate determination of the urea–water surface also above the critical concentration of about 0.1 mole fraction.

## References and Notes

- (1) Schellman, J. A. *Biophys. Chem.* **1990**, *37* (1–3), 121–140.
- (2) Bruning, W.; Holtzer, A. *J. Am. Chem. Soc.* **1961**, *83* (23), 4865–4866.
- (3) Tanford, C. *J. Am. Chem. Soc.* **1962**, *84* (22), 4240–4247.
- (4) Schellman, J. C. *R. Trav. Lab. Carlsberg Ser. Chim.* **1955**, *29*, 230–259.
- (5) Kresheck, G. C.; Scheraga, H. A. *J. Phys. Chem.* **1965**, *69* (5), 1704–1706.
- (6) Stokes, R. H. *Aust. J. Chem.* **1967**, *20*, 2087–2100.
- (7) Frank, H. S.; Franks, F. *J. Chem. Phys.* **1968**, *48* (10), 4746–4757.
- (8) Makhatadze, G. L.; Privalov, P. L. *J. Mol. Biol.* **1992**, *226* (2), 491–505.
- (9) Zou, Q.; Habermann-Rottinghaus, S. M.; Murphy, K. P. *Proteins: Struct., Funct., Bioinf.* **1998**, *31* (2), 107–115.
- (10) Caballero-Herrera, A.; Nordstrand, K.; Berndt, K. D.; Nilsson, L. *Biophys. J.* **2005**, *89* (8), 842–857.
- (11) Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. *Biochemistry* **2005**, *44* (16), 6269–6275.
- (12) Soper, A.; Castner, E.; Luzar, A. *Biophys. Chem.* **2003**, *105* (2–3), 649–666.
- (13) Yoshida, K.; Ibuki, K.; Ueno, M. *J. Chem. Phys.* **1998**, *108* (4), 1360–1367.
- (14) Nuallain, B. O.; Mayhew, S. G. *Eur. J. Biochem.* **2002**, *269*, 212–223.
- (15) Idrissi, A.; Bartolini, P.; Ricci, M.; Righini, R. *J. Chem. Phys.* **2001**, *114* (15), 6774–6780.
- (16) Grdadolnik, J.; Maréchal, Y. *J. Mol. Struct.* **2002**, *615* (1–3), 177–189.
- (17) Keuleers, R.; Desseyn, H. O.; Rousseau, B.; Alsenoy, C. V. *J. Phys. Chem. A* **1999**, *103* (24), 4621–4630.
- (18) Sharp, K. A.; Madan, B.; Manas, E.; Vanderkooi, J. M. *J. Chem. Phys.* **2001**, *114* (4), 1791–1796.
- (19) Tanaka, H.; Touhara, H.; Nakanishi, K.; Watanabe, N. *J. Chem. Phys.* **1984**, *80* (10), 5170–5186.
- (20) Oostenbrink, C.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2005**, *7* (1), 53–58.
- (21) Bennion, B. J.; Daggett, V. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (9), 5142–5147.
- (22) Smith, L. J.; Berendsen, H. J. C.; van Gunsteren, W. F. *J. Phys. Chem. B* **2004**, *108* (3), 1065–1071.
- (23) Idrissi, A.; Cinar, E.; Longelin, S.; Damay, P. *J. Mol. Liq.* **2004**, *110* (1–3), 201–208.
- (24) Tsai, J.; Gerstein, M.; Levitt, M. *J. Chem. Phys.* **1996**, *104* (23), 9417–9430.
- (25) Tobi, D.; Elber, R.; Thirumalai, D. *Biopolymers* **2003**, *68* (3), 359–369.
- (26) Ikeguchi, M.; Nakamura, S.; Shimizu, K. *J. Am. Chem. Soc.* **2001**, *123* (4), 677–682.
- (27) Sokolic, F.; Idrissi, A.; Perera, A. *J. Chem. Phys.* **2002**, *116* (4), 1636–1646.
- (28) Vanzi, F.; Madan, B.; Sharp, K. *J. Am. Chem. Soc.* **1998**, *120*, 10748–10753.

- (29) Smith, L. J.; Jones, R. M.; van Gunsteren, W. F. *Proteins: Struct. Funct. Bioinf.* **2005**, *58* (2), 439–449.
- (30) Duffy, E. M.; Kowalczyk, P. J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1993**, *115* (20), 9271–9275.
- (31) Mountain, R. D.; Thirumalai, D. *J. Am. Chem. Soc.* **2003**, *125* (7), 1950–1957.
- (32) Klimov, D.; Straub, J. E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (41), 14760–14765.
- (33) Caffisch, A.; Karplus, M. *Structure* **1999**, *7* (5), 477–488.
- (34) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2000**, *104* (24), 5854–5864.
- (35) Astrand, P.-O.; Wallqvist, A.; Karlstrom, G. *J. Chim. Phys.* **1991**, *88*, 2457–2464.
- (36) Astrand, P.-O.; Wallqvist, A.; Karlstrom, G. *J. Chem. Phys.* **1994**, *100* (2), 1262–1273.
- (37) Astrand, P.-O.; Wallqvist, A.; Karlstrom, G. *J. Phys. Chem.* **1994**, *98*, 8224–8233.
- (38) Zhang, Z.; Zhu, Y.; Shi, Y. *Biophys. Chem.* **2001**, *89* (2–3), 145–162.
- (39) Zou, Q.; Bennion, B. J.; Daggett, V.; Murphy, K. P. *J. Am. Chem. Soc.* **2002**, *124* (7), 1192–1202.
- (40) Trzesniak, D.; van der Vegt, N. F. A.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2004**, *6* (4), 697–702.
- (41) Sokolic, F.; Idrissi, A.; Perera, A. *J. Mol. Liq.* **2002**, *101* (1–3), 81–87.
- (42) Idrissi, A.; Sokolic, F. *J. Chem. Phys.* **2000**, *112* (21), 9479–9488.
- (43) Boek, E. S.; Briels, W. J.; van Eerden, J.; Feil, D. *J. Chem. Phys.* **1992**, *96* (9), 7010–7018.
- (44) Boek, E. S.; Briels, W. J. *J. Chem. Phys.* **1993**, *98* (2), 1422–1427.
- (45) Wallqvist, A.; Covell, D. G. *J. Am. Chem. Soc.* **1998**, *120*, 427–428.
- (46) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2003**, *107*, 3891–3898.
- (47) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *118* (13), 5901–5910.
- (48) Lee, M.-E.; van der Vegt, N. F. *J. Am. Chem. Soc.* **2006**, *128* (15), 4948–4949.
- (49) Idrissi, A. *Spectrochim. Acta, Part A* **2005**, *64* (1–2), 1–17.
- (50) Schellman, J. A. *Biophys. Chem.* **2002**, *96*, 91–101.
- (51) Finer, E.; Franks, F.; Tait, M. *J. Am. Chem. Soc.* **1972**, *94* (13), 4424–4429.
- (52) Hoccoart, X.; Turrell, G. *J. Chem. Phys.* **1993**, *99* (11), 8498–8503.
- (53) Tirado-Rives, J.; Orozco, M.; Jorgensen, W. L. *Biochemistry* **1997**, *36* (24), 7313–7329.
- (54) Berendsen, H.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
- (55) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
- (56) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (57) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79* (2).
- (58) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *Isr. J. Chem.* **1993**, *33*, 323–330.
- (59) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.
- (60) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.
- (61) Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. V.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.
- (62) Härdle, W.; Simar, L. *Applied Multivariate Statistical Analysis*; Springer: Berlin, 2003.
- (63) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A.; Apol, E.; Meulenhoff, P.; Tieleman, D.; Sijbers, A.; Feenstra, K.; van Drunen, R.; Berendsen, H. *Gromacs user manual*, version 3.2; www.gromacs.org, 2004.
- (64) Espinosa, E.; Molins, E.; Lecomte, C. *Chem. Phys. Lett.* **1998**, *285* (3–4), 170–173.
- (65) Arnold, W. D.; Oldfield, E. *J. Am. Chem. Soc.* **2000**, *122*, 12835–12841.
- (66) Jenkins, S.; Morrison, I. *Chem. Phys. Lett.* **2000**, *317*, 97–102.
- (67) Rozenberg, M.; Loewenschuss, A.; Marcus, Y. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2699–2702.
- (68) Alkorta, I.; Elguero, J. *J. Phys. Chem. A* **1999**, *103*, 272–279.
- (69) Gálvez, O.; Gómez, P.; Pacios, L. *J. Chem. Phys.* **2001**, *115* (24), 11166–11184.
- (70) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. *J. Comput. Chem.* **1995**, *16* (3), 273–284.
- (71) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19* (6), 774–777.
- (72) Soper, A. K. *Mol. Phys.* **2001**, *99* (17), 1503–1516.
- (73) Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112* (20), 8910–8922.
- (74) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (75) D’Ans, J.; Koch, K. J.; Zimmer, H. *Landolt-Börnstein, New Series, Group IV*; Springer-Verlag: Berlin, 1962; Vol. II/2b, p 3/442.