Published in final edited form as:
J Phys Chem B. 2007 May 17; 111(19): 5233–5242. doi:10.1021/jp067659x.

Preferential Solvation in Urea Solutions at Different Concentrations: Properties from Simulation Studies

Hironori Kokubo* and **B. Montgomery Pettitt***
Department of Chemistry, University of Houston, Houston, Texas 77204-5003

Abstract

We performed molecular dynamics simulations of urea solutions at different concentrations with two urea models (OPLS and KBFF) to examine the structures responsible for the thermodynamic solution properties. Our simulation results showed that hydrogen-bonding properties such as the average number of hydrogen bonds and their lifetime distributions were nearly constant at all concentrations between infinite dilution and the solubility limit. This implies that the characterization of urea–water solutions in the molarity concentration scale as nearly ideal is a result of facile local hydrogen bonding rather than a global property. Thus, urea concentration does not influence the local propensity for hydrogen bonds, only how they are satisfied. By comparison, the KBFF model of urea donated fewer hydrogen bonds than OPLS. We found that the KBFF urea model in TIP3P water better reproduced the experimental density and diffusion constant data. Preferential solvation analysis showed that there were weak urea–urea and water–water associations in OPLS solution at short distances, but there were no strong associations. We divided urea molecules into large, medium, and small clusters to examine fluctuation properties and found that any particular urea molecule did not stay in the same cluster for a long time. We found neither persistent nor large clusters.

1. Introduction

Concentrations of osmolytes, electrolytes, or other small organic molecules in water mixtures can change the properties of the solution strongly. Not only macroscopic properties such as volume, free energy, viscosity, and dielectric constant but also microscopic properties such as the diffusion constant and number and type of neighbors of molecules are also changed. Indeed additives or cosolvents can have a profound consequence for macromolecules. Osmolytes can make proteins more or less stable with respect to either denaturation or precipitation. Some view this as reflecting a change in the water structure. A venerable categorization to understand these effects is to address whether the solute is a water structure maker or structure breaker.[1] More solvation shell models have considered preferential solvation or interactions.[2,3] Such theories roughly separate global solution influences from molecular associations as causes of denaturation or precipitation. Increasing evidence indicates that trying to interpret the observed effects of the concentration of third component additives or cosolvents on protein structure via only the water structure is not necessarily sufficient.[4–7] This is simply because the cosolvent may associate with or be preferentially excluded by proteins.

The stability of proteins with respect to changes in cosolvent concentration is generally a reflection of the free energy change of the solution. Water is a marginally good solvent for globular proteins of defined structure. This causes folding or phase separation sensitivity to the addition of third components, which may either improve or worsen the solvent efficacy.

*Authors to whom correspondence should be addressed. E-mail: kokubo@kitten.chem.uh.edu; pettitt@uh.edu.

Urea is a denaturant of proteins. Mechanistically, exactly how urea solutions denature proteins remains unclear and is the subject of continued discussion.[4,8,9] Before investigating the influence of urea on proteins in solution we need to understand the influence of urea on water at the molecular level. In this paper we analyze what kinds of impact urea has on water structure and correlations in detail by molecular dynamics (MD) simulations. Model dependencies could be a significant issue as we will be dealing with properties depending on small free energy differences.[10] Here, two classical urea models (OPLS and KBFF) are used to examine the dependency of force-field parameters as well. We complement a previous study of these models[11] by performing a different set of analyses. That study and the current one will be used in a subsequent free energy analysis.

In section 2 we present the details of the MD simulations and their analysis and discuss the definition of preferential solvation. In section 3 results and discussion are given for both geometric and time-dependent properties from multiple simulations spanning a range of concentrations. Section 4 is devoted to our conclusions.

2. Methods

2.1. MD Simulations

In this paper, to control for the effect of force-field variability we compared two different urea models: the OPLS model[12] and KBFF model.[8,13] Both models are fit to optimize properties in liquid solutions although different properties were targeted in each. We recount and compare the parameters of these two models in Table 1. The molecular shape and mass are the same, but the van der Waals (vdW) and electrostatic parameters are very different. The OPLS urea model has been well-studied and often used since its introduction.[12,14] However, the KBFF model was developed more recently.[8] Solvation free energy depends quite strongly on the charge distribution of the solute and its interaction with water. Other molecular properties can be less sensitive to the charge distribution.[15]

As mentioned, both force fields considered here were developed to fit a set of properties. OPLS was fit to a variety of properties including single and pair molecular properties as well as heats of solvation among others.[12] The KBFF urea model takes a more restricted and challenging approach by fitting the charge distribution to the experimental activity coefficient data.[8] The results of using a particular solute charge distribution for a two-component solution are certainly dependent on the water model used. Here we chose the TIP3P model for water in our simulations, which was used in the fitting of the OPLS parameters. This is not the model with which KBFF was parametrized, and so holding this model constant throughout our study will have some advantages and disadvantages in comparisons and controls. We note that the activity data to which KBFF was fit is well-reproduced using the TIP3P model (unpublished data).

We prepared models of urea solutions at seven different concentrations for both OPLS and KBFF models between the limits of dilute solution and the pure urea in the following way. We first estimated the approximate molecular volume of one urea from preliminary simulations to obtain the expected concentrations. We estimated that one urea molecule has approximately 2.48 times the volume of one water molecule. Through the use of that approximation the number of urea and water molecules in a volume for each concentration is determined. We prepared large pure water and urea boxes in advance. The pure urea melt was a supercooled system as pure urea would be a solid at the temperatures considered here. We randomly removed the required number of urea and water molecules from each of these pure samples. We then placed the two boxes in proximity and performed 300 ps of simulation at constant temperature while shrinking the sample to the target volume by using a constant pressure algorithm for each system. Urea and water were thus gradually mixed spontaneously. Temperature was controlled by the Nosé method to 298 K.[16] We thus obtained the initial

configurations at different concentrations. The final box sizes are close to 34 Å × 34 Å × 34 Å in every system.

We could have used the more standard method and placed urea molecules randomly in a pure water box (or vice versa) with the removal of water molecules that overlap with the solute molecules. However, we found that the systems rapidly mixed, thus providing a uniform concentration profile before the end of the equilibration. This in some ways reflects the expected solubility of urea in water. In addition, we required no iterations of the number of urea and water molecules as is common in the removal and replacement technique to obtain the expected concentrations for predecided numbers of molecules.

For the production runs the Nosé–Anderson NPT method was used for the temperature (T) and pressure (P) controls.[16,17] The integration time step was 1.0 fs. The RATTLE method was used to fix the bond lengths.[18] Because all systems have a similar box size as we expected, we could use the same cutoff length of 15 Å for vdW interactions. We used an Ewald method for electrostatic interactions with the *r*-space cutoff and Ewald convergence parameter α set to 1.5 nm and 2.09 nm^{-1}, respectively.[19] A major difference between the models is the charge distribution. So it is important to precisely estimate the electrostatics to evaluate their contributions to the thermodynamic and structural properties.

Utilizing our dissolution technique, we obtained the initial configurations of the urea solutions at seven different concentrations for the two different urea models. The number of total systems prepared is thus 14. We list all of the systems for our simulations in Table 2.

2.2. Preferential Solvation

Preferential solvation or preferential interaction is used in this paper to describe deviations from a random mixture or ideal solvation model. If a solute is surrounded by two types of solvent, then one could describe an ideal state whereby one expected a number of proximal solvent molecules to be proportional to their mole fraction or mass fraction. One could also consider a deviation from an ideal state described by deviations in the mole or mass fraction in a given solvation shell. Given a thermodynamic reference state and concentration scale we use existing methods to measure molecular association (or aggregation). One flexible definition of preferential solvation is the local mole fraction minus the total mole fraction, following the Ben-Naim definition[20,21]

$$\delta_{AB}=x_{AB}(R)-x_{A} \tag{1}$$

Here the local mole fraction x_{AB} is defined as the number of A molecules divided by the number of the total molecules in a sphere on the center of a B molecule within radius *R*. From the definition the following relations are obtained

$$\begin{aligned}\delta_{AB}&=-\delta_{BB}\\ \delta_{AA}&=-\delta_{BA}\end{aligned} \tag{2}$$

Preferential solvation can be calculated from Kirkwood–Buff *G* factors. Such calculations require a sufficiently large *R* so that

$$\delta_{AB}=\frac{x_{A}x_{B}(G_{AB}-G_{BB})}{x_{A}G_{AB}+x_{B}G_{BB}+V_{c}} \tag{3}$$

where V_c is a volume with radius *R* and the Kirkwood–Buff G_{ij} is defined using the radial distribution function $g_{AB}(r)$ as

$$G_{AB}=4\pi\int_0^\infty (g_{AB}(r)-1)r^2\mathrm{d}r \tag{4}$$

The direct calculation of high-precision Kirkwood–Buff factors demands long simulation time and large simulation samples due to problematic convergence at larger distances.[13]

We consider the density-weighted integral

$$\rho_A G_{AB}=\rho_A\int_0^\infty (g_{AB}(r)-1)4\pi r^2\mathrm{d}r \tag{5}$$

In eq 5 $\rho_A g_{AB}(r)4\pi r^2\,\mathrm{d}r$ measures the average number of A molecules in the spherical shell of width dr at radius r centered on a B molecule. However, $\rho_A 4\pi r^2\,\mathrm{d}r$ measures the average number of A molecules in a spherical shell with the origin chosen at random or equivalently for a random (ideal) distribution of A. Therefore, eq 5 can be considered as the excess (or deficit) number of A molecules around a B molecule versus random.

Since the distribution function approaches unity at large distances, eq 5 can be approximated as

$$\begin{aligned}\rho_A G_{AB} &\approx \rho_A\int_0^{R_c}(g_{AB}(r)-1)4\pi r^2\mathrm{d}r\\ &=\rho_A\int_0^{R_c} g_{AB}(r)4\pi r^2\mathrm{d}r-\rho_A\tfrac{4}{3}\pi R_c{}^3\\ &=\overline{N}_{AB}(R_c)-\rho_A\tfrac{4}{3}\pi R_c{}^3\end{aligned} \tag{6}$$

where R_c is a sufficiently large distance beyond which the correlations are small so $g_{AB}\approx 1$ and N_{AB} is the number of molecules of A molecules around a B molecule within R_c. When R_c is small, the first approximation in eq 6 breaks down, and the right-hand side corresponds to the local number of molecules minus the ideal local number, or eq 1 at small R. This truncated G_{AB} can be considered to express the local information. We note, however, that only the full G_{AB} without truncation can faithfully represent thermodynamic properties such as compressibility, activity coefficients, and partial molar volumes, which are obtainable through the inversion relations of the Kirkwood–Buff theory.[24,22]

The definition of preferential solvation in eq 1 or 6 measures the excess (or deficit) number versus a random distribution within a predefined volume. Thus an excess (or deficit) so defined does not directly reflect favorable or unfavorable effective interactions. First, the choice of the origin (a random one vs the center of a specific molecule) can cause an apparent excess or deficit without disturbing any solution properties when we use the truncated G_{AB}.[23] To quantify this effect, we performed simulations of pure urea (and pure water) and randomly considered any molecule to be the solute. Thus, choosing the center of urea (or water) as the origin caused an apparent number deficit of ~1.5 (~0.5) at R = 15 Å, although such a system, by definition, can have no excess or deficit. This value was similar for both OPLS and KBFF urea solutions. This demonstrates that the choice of origin in this simple system has an effect of approximately ±1.0.

Second, when we calculate eq 5 in the case of A = B the center molecule is not usually counted. The center A molecule is also not included in the macroscopic definition of ρ_A and usually makes little difference when the system size is large enough. (The calculation of $g(r)$ is no better than $O(1/N)$). However, when we consider eq 6 in the case of A = B, it is also possible to interpret the right-hand side of eq 6 as the number of A molecules including the center A molecule minus the ideal number of A molecules. If we adopt this latter interpretation, then it increases the value of eq 5 by 1.0. (In this case the center A molecule is included in ρ_A.) When we consider small changes of preferential solvation between systems, this difference is important.

Third, as mentioned above, the excess can be caused by different molecular sizes even if the effective interaction is very similar.[24] Finally, the excess or deficit can be caused by favorable or unfavorable interactions (or correlations) such as those measured by the classical deviations from Raoult's Law. Note also that the symmetric (or dilute) ideal solution is not necessarily the same as a random distribution.

2.3. Characterization of Clusters

Consider searching for and characterizing clusters in solution. Each molecule can be assigned to high-, medium-, or low-density clusters of its species. We count the number of molecules n within the sphere around the chosen molecule and compare it with the average number n_{ave}, following ref [25]

$$\begin{aligned} &\text{high if } n > n_{ave} + \delta \\ &\text{medium if } n_{ave} - \delta \le n \le +\delta \\ &\text{low if } n < n_{ave} - \delta \end{aligned} \tag{7}$$

We adopted this clustering method to examine urea self-aggregation (cluster). The sphere radius and δ are parameters. We set δ to be 20% of n_{ave} and the sphere radius to be 6 Å.

3. Results and Discussion

Table 2 lists the urea–water solutions of different concentrations with two different urea models in our simulations. The concentrations of the systems span from the "nearly" infinite dilute solution (mole fraction, 0.0007657) to the solubility limit (mole fraction, 0.2728) and beyond to the pure solute system (mole fraction, 1.00). The concentrations are equivalently expressed in units of molarity and molality. Note that only molarity depends on the urea force field because it includes the volume information and that molality is not well-defined for pure solute systems.

As a function of concentration, we first compared basic physical quantities such as the densities and diffusion constants with respect to the difference induced by force fields. The densities of urea solutions at different concentrations are shown in Figure 1 (see also Table 2). The volumes at different concentrations were calculated by using the average volumes from 1 ns NPT simulations, and we used these average volumes to calculate the densities. We see that the KBFF urea model produces somewhat better agreement with experimental data and the density of the OPLS urea solutions tends to be a little larger than the experimental data. This systematic difference for OPLS may be in part due to our use of lattice sums rather than the cutoffs with which it was initially parametrized. The extreme densities of our pure urea melt (formally metastable at this temperature) were approximately 1.36 (OPLS) and 1.35 (KBFF), both of which are close to the experimental density of the crystal at 1.32.

Note that these solution densities also depend on the water model. The densities of pure water for different water models are listed in ref [26]. The temperature and pressure of that literature study are the same as ours (298 K and 1 atm). The most dilute urea solution in our set is expected to have a density close to that of pure TIP3P water. The value of 1.01 g/cm^3 obtained from our simulation is a little larger than that of 0.982 g/cm^3 in ref [26]. Possible reasons for this difference include that our most dilute system still contains one urea molecule, the cutoff problem, and other differences in simulation conditions. Simulation methodology and computational power have changed considerably with time. Changes in sample size along with cutoffs versus Ewald could easily account for a 2% difference.

Among previous simulation studies of urea solutions authors have used Ewald[27] and variants[13] as well as other reaction field methods[28] on the long-range forces. Ewald and its

variants converge to similar energies and forces at each point, but other reaction field methods do not necessarily give similar values in general. In addition, since the water model used by other authors is not TIP3P, one should exercise caution when comparing our data with the previous simulation results directly.

The density of the urea solution using other water models is presented in refs [13, 27], and [28]. References 13 and [27] used the SPC/E water model, and ref [28] used the SPC water model. Table 1 in ref [29] and Table 3 in ref [13] give densities of 0.998 and 0.995 for the SPC/E pure water system, which are very close to the experimental data (0.997). However, the SPC model gives 0.971, which is a little small.[26] We see from these works that the combination of the OPLS or KBFF urea model and the SPC/E water model gives us a density that is very close to the experimental data and the combination of OPLS urea and SPC water gives a systematically smaller density than the experimental data. This is no doubt large due to the fact that the density of the SPC water model is smaller than that in the experimental data. In our simulations using the TIP3P water model, the KBFF urea model gave better agreement with experimental data than the OPLS model.

In Figure 2 the mean-square displacements of urea molecules for the OPLS model urea solutions at different mole fractions are shown. It is clear that urea displacements diminish greatly as the concentration increases, even if we ignore the metastable supercooled melt at the unit mole fraction. The urea displacement of the most dilute system was more noisy because of insufficient sampling; it had one urea molecule and 1305 water molecules.

The diffusion coefficients can be estimated from the slope of the mean-square displacements via

$$D=\lim_{t\to 0}\langle |r(t)-r(0)|^2\rangle/6t \tag{8}$$

In Figure 3 we show the diffusion coefficients of OPLS and KBFF urea molecules (top panel) and TIP3P water molecules (bottom panel). We see that urea diffusion constants are similar in both the OPLS and the KBFF urea models. Interestingly, water diffusion constants with KBFF urea are always smaller than those of OPLS urea for the same water model. Therefore, it seems that the KBFF urea model impedes or traps TIP3P water for a longer time than OPLS. The diffusion coefficients are always larger than the experimental values. This is a well-known property of the TIP3P water model, which is known to diffuse faster than experiment as does the SPC water model.[29,30] Our diffusion constants of OPLS urea in TIP3P are in good accord with those measured in SPC water.[28] The SPC/E water model is a revised version of the SPC model and shows good agreement for the diffusion constant of urea with experimental values although the water still moves a bit too fast.[13,27] Due to a greater slope, as the urea concentration increases, diffusion constants of both urea and water are in somewhat better agreement with experimental values at high concentrations.

As hydrogen bonds define much of the local geometry and therefore are often seen as the cause of many solution properties, we consider them in some detail. Table 3 shows the average number of hydrogen bonds for one urea molecule, one water molecule, and their donor–acceptor components. We use a common, arbitrary definition, and thus a hydrogen bond here is defined when the distance between heavy atoms (D and O, D–H···O) is within 3.0 Å and the angle 180°– ∠D–H···O is less than 60°.[33]

We find it noteworthy that the number of hydrogen bonds for urea–water solutions does not decrease much (is nearly constant) as the concentration increases for either model. This is consistent with the experimental infrared spectroscopy, X-ray diffraction, and NMR studies that imply that urea does not shift the water OH stretching band,[34] interrupt the hydrogen-

bond network of water,[35] or strongly change water diffusion.[36] The simulation results suggest that water can make hydrogen bonds with urea as easily as with water when the concentration increases. It also indicates that the interactions of urea and water are sufficiently similar so that the solvation free energy values of urea in different urea concentration solutions are similar. This is all consistent with urea's well-known property of forming a nearly ideal solution in the molar scale.[37] The volumes of our systems for all concentrations are similar (Table 2). Thus, we see that the number of total hydrogen bonds in the same volume is decreasing as the urea concentration is increasing. This is because the volume of urea is approximately 2.5 times larger than that of water though the number of hydrogen bonds for each urea and water molecule is almost constant in Table 3.

Table 3 shows that OPLS urea solutions make more hydrogen bonds than KBFF urea solutions at the same concentration ("total"). OPLS urea makes more hydrogen bonds than KBFF urea ("1 urea"). Hydrogen atoms of the OPLS urea model make hydrogen bonds much more often than those of the KBFF model ("4 H of urea"). However, the oxygen atom of KBFF urea makes slightly more hydrogen bonds than that of OPLS urea ("O of urea"). As a result, the OPLS urea solution makes more hydrogen bonds at the same concentration than the KBFF solution ("1 urea"). Hydrogen atoms and oxygen atoms of water in OPLS urea solution make a similar number of hydrogen bonds at different urea concentrations ("2 H of water" and "O of water"), with an insignificant decrease as the concentration increases. However, the hydrogen atoms of water in the KBFF solution make slightly more hydrogen bonds as the urea concentration increases. The hydrogen atoms of KBFF urea are less likely to make hydrogen bonds, and the hydrogen atoms of water are instead selected as hydrogen-bond partners more often.

Mountain and Thirumalai also compared the OPLS and KBFF urea models and proposed that this tendency is due to vdW excluded volume effects of KBFF urea hydrogens and thus the difference of excluded volumes rather than the charge distributions, which also would affect the tendency of urea to self-associate.[11] Oxygen atoms of the water in KBFF solutions make fewer hydrogen bonds as the urea concentration increases. This is in part because the oxygens atom of KBFF urea have a larger negative charge and so have a better ability to make hydrogen bonds than water oxygen atoms (see charges in Table 1).

These analyses suggest that OPLS urea is more similar to water in effective interactions than KBFF urea because the change in the number of hydrogen bonds of water in OPLS solutions is smaller than that in KBFF solutions ("2 H of water" and "O of water"). It also indicates that OPLS urea apparently dissolves in water better than KBFF urea. In the systems of pure urea (mole fraction, 1.0), we can confirm that the average number of hydrogen bonds of four hydrogen atoms of urea is the same as that of one oxygen atom of urea as it should be.

In Figure 4 we calculated the preference of hydrogen-bond partners for urea oxygen, urea hydrogen, water oxygen, and water hydrogen. This figure shows the deviation of the probability from the ideal case, where we defined the ideal probability of hydrogen-bond partners to be proportional to the probability of donor or acceptor atoms. The probability is defined by the number of total urea hydrogen (oxygen) atoms divided by the number of total hydrogen (oxygen) atoms in the system. Because KBFF urea hydrogens have less ability to make hydrogen bonds, urea and water oxygens thus make hydrogen bonds with water much more often than with urea (Figures 4a and 4c). Water hydrogen is chosen as a hydrogen-bond partner more often for both urea's oxygen and water's oxygen (Figures 4a and 4c). However, urea's oxygen is chosen as a hydrogen-bond partner more often for both urea's hydrogen and water's hydrogen (Figures 4b and 4d). In OPLS urea solution, we see that water hydrogens choose urea's oxygen and water's oxygen almost ideally (Figure 4d) up to the solubility limit.

In Table 4 we show the average lifetimes of the hydrogen bonds of all pairs at different concentrations. Considering the size and similarity of the standard deviations (data not shown), the average lifetime of each hydrogen bond is essentially independent of concentrations. This fact suggests that urea concentration does not influence the local lifetimes or propensity for hydrogen bonds, only how they are satisfied. We also see that the hydrogen bond of KBFF urea's hydrogen with urea's oxygen or water's oxygen is more kinetically unstable than that of OPLS. This is consistent with the previous observation in Table 3 that hydrogen atoms of KBFF urea are less ideal hydrogen-bond donors. However, the oxygen atom of KBFF urea retains water hydrogens in hydrogen bonds for a longer time than OPLS urea. The average lifetime of a water–water hydrogen bond is similar in both OPLS and KBFF urea solutions.

To confirm that the hydrogen-bond lifetime dynamics are similar at different concentrations, we calculated the complete distribution of water–water hydrogen-bond lifetimes at different urea concentrations for OPLS urea solutions, as shown in Figure 5. We see that the distributions change little for all of the systems considered. This confirms that even at high urea concentrations the lifetimes of water–water hydrogen bonds are not strongly affected. The same result was obtained in the case of the KBFF potential (figure not shown). In addition, not only the lifetime of water–water hydrogen bonds but also those of all other hydrogen-bond pairs showed a remarkable concentration independence. Even in the case of our supercooled pure urea melt, the histogram of residence times for urea–urea hydrogen bonds was almost the same.

In Figure 6 we find a remarkable difference in comparing the histograms of the lifetimes of the various hydrogen bonds of the OPLS urea solution with those of a KBFF urea solution. The system chosen is close to the solubility limit (mole fraction, 0.2728), but as we commented above the distributions show little concentration dependence. We see that in the case of the OPLS solution all hydrogen-bond pairs show little variation. This suggests that the OPLS urea model is similar to water in many hydrogen-bonding characteristics. Figure 6a shows that the histogram of the OPLS urea–urea hydrogen bond is bimodal; however, that of KBFF is monotonically decreasing. The average hydrogen-bond lifetime for KBFF is generally shorter than that of OPLS. Figure 6b shows the similar characteristics about the hydrogen bond between urea oxygen and water hydrogen. This behavior is understandable because the hydrogen atoms of KBFF urea model have less ability to make hydrogen bonds in this solution. In Figure 6c for the water hydrogen and urea oxygen we see that the histograms of both OPLS and KBFF have two similar peaks, but the valley between the two peaks for the KBFF force field is deeper, which causes the longer average water donor to urea acceptor lifetime in KBFF solutions. KBFF oxygen binds water hydrogen more strongly than OPLS oxygen because the negative charge of oxygen in KBFF urea is larger than that for OPLS urea. Figure 6d shows that the lifetimes of water–water hydrogen bonds in both OPLS and KBFF solutions are similar. This observation is consistent with our suggestion that facile local hydrogen bonding is generally responsible for the properties of urea solutions.

In Figure 7 we show the preferential solvation of urea around urea or water by using eq 1 versus the radius R of the sphere. All preferential solvation values approach zero at large R (15 Å), which means that urea and water are mixing almost uniformly beyond a radius of 15 Å. This behavior is consistent with our finding no large scale clusters or other long-range correlations (see below).

In Figure 7a we see that in OPLS urea solution there are excess urea molecules around urea at short distances (≤6.0 Å), which roughly correspond to the first solvation shell. That is, the number of waters around OPLS urea within a 6.0 Å sphere is smaller than that expected from the bulk mole fraction. There is also a small excess of urea molecules at >6.0 Å. However, Figure 7c shows that in KBFF urea solution there is a deficit of urea molecules at 7 Å. This means that the number of water molecules around KBFF urea within a 7 Å sphere is larger

than that in the bulk. The deficit at 7 Å rapidly becomes small and approaches the bulk mole fraction at distances of more than 9 Å. In summary, OPLS urea has excess urea molecules in the first solvation shell around 5 or 6 Å, and KBFF urea has an excess of water molecules at 7 Å. This is a central difference in the fundamental solution behavior between the two models.

Figures 7c and 7d show the excess or deficit of urea around water for OPLS and KBFF solutions, respectively. We see that there is always a deficit of urea molecules around water, especially in OPLS solutions within 5–7 Å, which means that water is preferentially solvated by water.

The number of excess or deficit urea molecules can be estimated roughly from preferential solvation based on eq 1. The largest values are in the case of the solubility limit concentration (red bar) in OPLS solution in Figures 7a and 7b. In this case the preferential solvation at a sphere with a radius of 6 Å is less than 0.03. There are approximately 6 urea molecules and 16 water molecules in the sphere with a radius of 6 Å for the bulk. The excess of $x = 0.03$ corresponds to an excess of approximately 0.65 urea molecules in this sphere. Therefore, the excess or deficit of urea around urea is approximately 10% in this largest case. In other cases, the degree of the excess or deficit is much smaller. Therefore, although OPLS and KBFF showed some different tendencies, we could not find any strong deviation from an ideal random distribution for either urea solutions, and there was no strong urea–urea association.

The preferential solvation based on eq 5 measures the excess or deficit number of molecules within the sphere large enough to satisfy $g(r) = 1.0$. Although our simulation time (several nanoseconds) and sample size (box length, 34 Å) were not enough for the direct convergence of the Kirkwood–Buff G, our preliminary analysis suggests that no clear evidence of a urea–urea attractive preferential interaction exists.

To understand this behavior of the preferential solvation we show the radial distribution functions of centers of mass in Figure 8. It is clear that the first peaks in the OPLS solution are higher than those in the KBFF solution except the pure urea melt system. This makes the preferential solvation of OPLS solution at short distances positive (Figure 7). In the case of the KBFF solutions, the first peaks are lower and the valleys between the first and the second peaks are more pronounced. This causes the preferential solvation of the KBFF solution between 7 and 8 Å to be negative and those at less than 6 Å to be slightly positive (Figure 7). The pure urea melts (black lines in parts a and b) have correlations beyond 14 Å that may reflect their supercooled nature. The small changes in the water–water distributions are due to simple concentration changes and not local geometric changes as illustrated above.

Up to this point we have seen no evidence for urea–urea clusters. Urea molecules in each system were divided into high, medium, or low urea density clusters by eq 7.[25] We set δ to be 20% of n_{ave} and the sphere radius to be 6 Å and focused on only urea molecules. Figure 9 shows the time series of the number of urea molecules for high-, medium-, and low-density clusters. Although preferential solvation analysis measures only the properties averaged over all urea molecules, the advantage of the clustering method is that it can detect differences in the local environment of urea molecules by dividing them into clusters. The number of medium clusters in the OPLS urea solution tends to be larger than that in the KBFF one. Correspondingly, the number of high and low clusters in the KBFF urea solution is larger than that for the OPLS solution. The fact that there is no system that has a statistically large number of high-density clusters implies that large urea aggregation is not occurring. However, some increase in medium scale clustering clearly occurs near the solubility limit.

From Figure 9, we see that the average number of molecules in each cluster is well represented on the time scale of our sampling. There are two mechanisms that maintain the number of molecules in each cluster: (1) Once the cluster is made, the cluster does not break up, and the

same molecules are included in the same cluster for a long time. (2) Molecules may exchange to or from the cluster. We show a time series of the cluster change to which one random urea molecule belongs in Figure 10. Other urea molecules behaved similarly. It is clear that each urea changes its cluster frequently (every 4 or 5 ps on the average). Therefore, there were few static clusters in urea solution for both OPLS and KBFF models. Our simulation result suggests that urea and water mix well spatially and dynamically.

Figure 11 shows the typical snapshots of urea at the solubility limit concentration ($x = 0.2728$) for (a) OPLS and (b) KBFF solutions. We see that the molecules in the high-, medium-, and low-density clusters are distributed almost uniformly and there are no very large (percolating) clusters. The molecules in the medium-density cluster exchange between high- and low-density clusters. There is a force-field dependence to the distribution however. We find that the number of molecules in medium-density clusters in OPLS solution is larger than that in KBFF solution (Figure 9, o3 and k3). These snapshots are consistent with the observation in Figure 10 that each molecule changes cluster frequently.

4. Conclusions

The character of urea solutions and the mechanism of protein denaturation by urea have long been studied. While protein denaturation could be seen as a nonideal solution effect, measures of nonideality for urea depend strongly on the reference state and concentration scale used.[4] Urea can look essentially ideal over a large concentration range in the molar scale. In this paper we performed molecular dynamics simulations of urea solutions at different concentrations comparing two urea models (OPLS and KBFF) and examined various properties of these model urea solutions.

The KBFF urea model reproduced the experimental density better than the OPLS model with the water model and simulation setup used. The diffusion constants of urea and water in the simulations were predictably larger than the experimental ones mainly because of the use of the TIP3P water model. When the concentration increases, the diffusion constants of both urea and water differentially decreased and became closer to the experimental data. The diffusion constants of water of the KBFF urea solutions were smaller than those of the OPLS solutions.

Hydrogen-bond analysis showed that not only the average number of hydrogen bonds but also the average lifetimes do not change very much at different concentrations for either model of urea. In addition, the distribution of hydrogen-bond lifetimes also did not change very much at different concentrations. This suggests that urea concentration does not influence the local propensity for hydrogen bonds, only how they are satisfied. The distributions of hydrogen-bond lifetimes for urea have two peaks, which is similar to those of TIP3P water except for the case of the hydrogen atoms of KBFF urea. Only the KBFF urea hydrogen is unimodal in hydrogen-bond lifetimes.

Preferential solvation analysis showed no strong preferential solvation at long distances. Namely, when we look at the sphere around each urea with a radius of more than ~7 Å there exists an almost ideal number of urea and water molecules in the sphere, which suggests that urea and water are uniformly mixing at every concentration up to the solubility limit. However, when we looked at the vicinity around each urea there were weak urea–urea and water–water associations in the OPLS solution at short distances of ~6 Å. However, there is a weak deficit of urea molecules around urea at ~7 Å in KBFF solutions, but there is almost no excess or deficit at other distances.

The behavior of this local preferential solvation definition was reinforced by inspection of the radial distribution functions. OPLS urea has a slightly stronger tendency to self-aggregate than KBFF urea in the first solvation shell.

Cluster analysis showed that there were neither rigid clusters nor large clusters. In general urea molecules did not stay in the same cluster for a long time and changed clusters on average every ~5 ps. Our simulation analysis implies that urea behaves almost neutrally toward water except the trivial fact that the diffusion of urea and water become slower at higher urea concentrations. We found no evidence that urea is a structure breaker of water. These general results were largely independent of the urea models used, which were parametrized using very different criteria.

Acknowledgements

We gratefully acknowledge the National Institutes of Health (Grant No. GM37657) and the R. A. Welch foundation (Grant No. E-1028) for partial financial support of this work. The computations were performed in part using the Molecular Science Computing Facility at the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the Office of Biological and Environmental Research of the U. S. Department of Energy, located at the Pacific Northwest National Laboratory, and the Pittsburgh Supercomputing Center on the National Science Foundation teragrid.

References and Notes

1. Frank HS, Franks F. J Chem Phys 1968;48:4746.
2. Timasheff SN. Annu Rev Biophys Biomol Struct 1993;22:67. [PubMed: 8347999]
3. Schellman JA. Annu Rev Biophys Biophys Chem 1987;16:115. [PubMed: 3297085]
4. Rögen J, Pettitt BM, Bolen DW. Biochemistry 2004;43:14472. [PubMed: 15533052]
5. Rösgen J, Pettitt BM, Bolen DW. Biophys J 2005;89:2988. [PubMed: 16113118]
6. Wright WW, Guffanti GT, Vanderkooi JM. Biophys J 2003;85:1980. [PubMed: 12944311]
7. Dashnau JL, Nucci NV, Sharp KA, Vanderkooi JM. J Phys Chem B 2006;110:13670. [PubMed: 16821896]
8. Weerasinghe S, Smith PE. J Chem Phys 2003;118:5901.
9. Bennion BJ, Daggett V. Proc Natl Acad Sci USA 2003;100:5142. [PubMed: 12702764]
10. Shirts MR, Pitera JW, Swope WC, Pande VS. J Chem Phys 2003;119:5740.
11. Mountain RD, Thirumalai D. J Phys Chem B 2004;108:6826–6831.
12. Duffy EM, Severance DL, Jorgensen WL. Isr J Chem 1993;33:323.
13. Weerasinghe S, Smith PE. J Phys Chem B 2003;107:3891.
14. Ikeguchi M, Nakamura S, Shimizu K. J Am Chem Soc 2001;123:677. [PubMed: 11456580]
15. Pettitt BM, Karplus M. J Am Chem Soc 1985;107:1166.
16. Nosé S. J Chem Phys 1984;81:511.
17. Andersen HC. J Chem Phys 1980;72:2384.
18. Andersen HC. J Comput Phys 1983;52:24–34.
19. Ewald P. Ann Phys 1921;64:253.
20. Ben-Naim A. J Phys Chem 1989;93:3809.
21. Ben-Naim, A. Statistical Thermodynamics for Chemists and Biochemists. Plenum Press; New York: 1992.
22. Ben-Naim A. J Chem Phys 1977;67:4884.
23. Hirata F, Rossky PJ. J Chem Phys 1981;74:5324.
24. Chitra R, Smith PE. J Phys Chem B 2002;106:1491.
25. Hernan LM, Ravi R, Susan CT. J Chem Phys 1996;104:1067.
26. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J Chem Phys 1983;79:926.
27. Sokolic F, Idrissi A, Perera A. J Chem Phys 2002;116:1636.
28. Smith LJ, Berendsen HJC, van Gunsteren WF. J Phys Chem B 2004;108:1065.
29. Berendsen HJC, Grigera JR, Straatsma TP. J Phys Chem 1987;91:6269.
30. Tirado-Rives J, Jorgensen WL. J Am Chem Soc 1990;112:2773.
31. Albright JG, Mills R. J Phys Chem 1965;69:3120.

32. Gosting LJ, Akeley DF. J Am Chem Soc 1952;74:2058.
33. Costa MTCM. J Mol Struct (THEOCHEM) 2005;729:47–52.
34. Sharp KA, Madan B, Manas E, Vanderkooi JM. J Chem Phys 2001;114:1791.
35. Finney JL, Soper AK, Turner J. Physica B 1989;156:151.
36. Mayele M, Holz M. Phys Chem Chem Phys 2000;2:2429.
37. Rösgen J, Pettitt BM, Perkyns J, Bolen DW. J Phys Chem B 2004;108:2048.

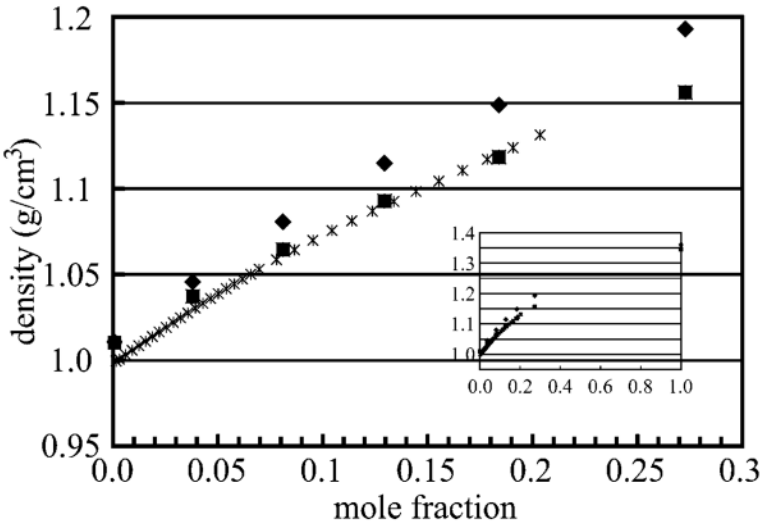

Figure 1.

Densities of urea solution as a function of mole fraction. The closed diamond represents the OPLS urea model, and the closed square represents the KBFF model, which were obtained from NPT molecular simulations at T = 298.0 K and P = 1 atm. The asterisk represents the experimental data from the CRC Handbook of Chemistry and Physics.

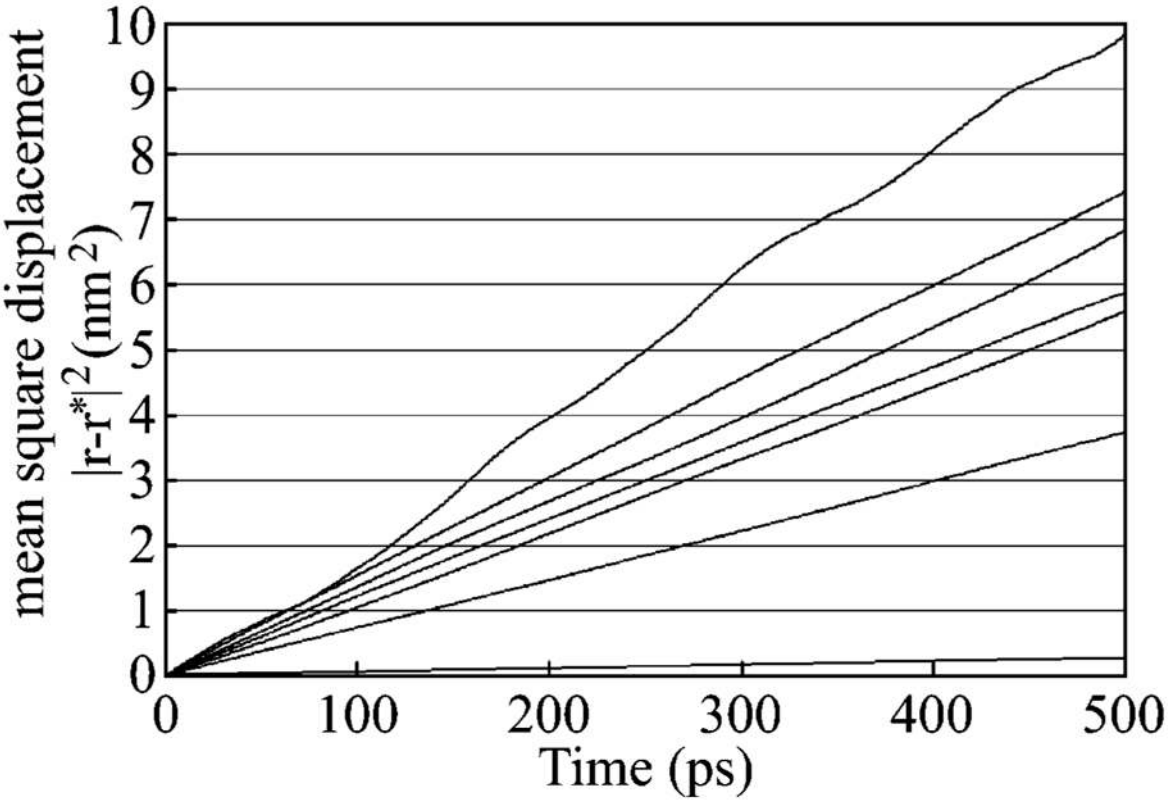

Figure 2.
Mean-square displacement of OPLS urea molecules as functions of time. The lines correspond to the urea mole fractions of 0.0007657, 0.03806, 0.08106, 0.1294, 0.1840, 0.2728, and 1.00, respectively, from top to bottom.

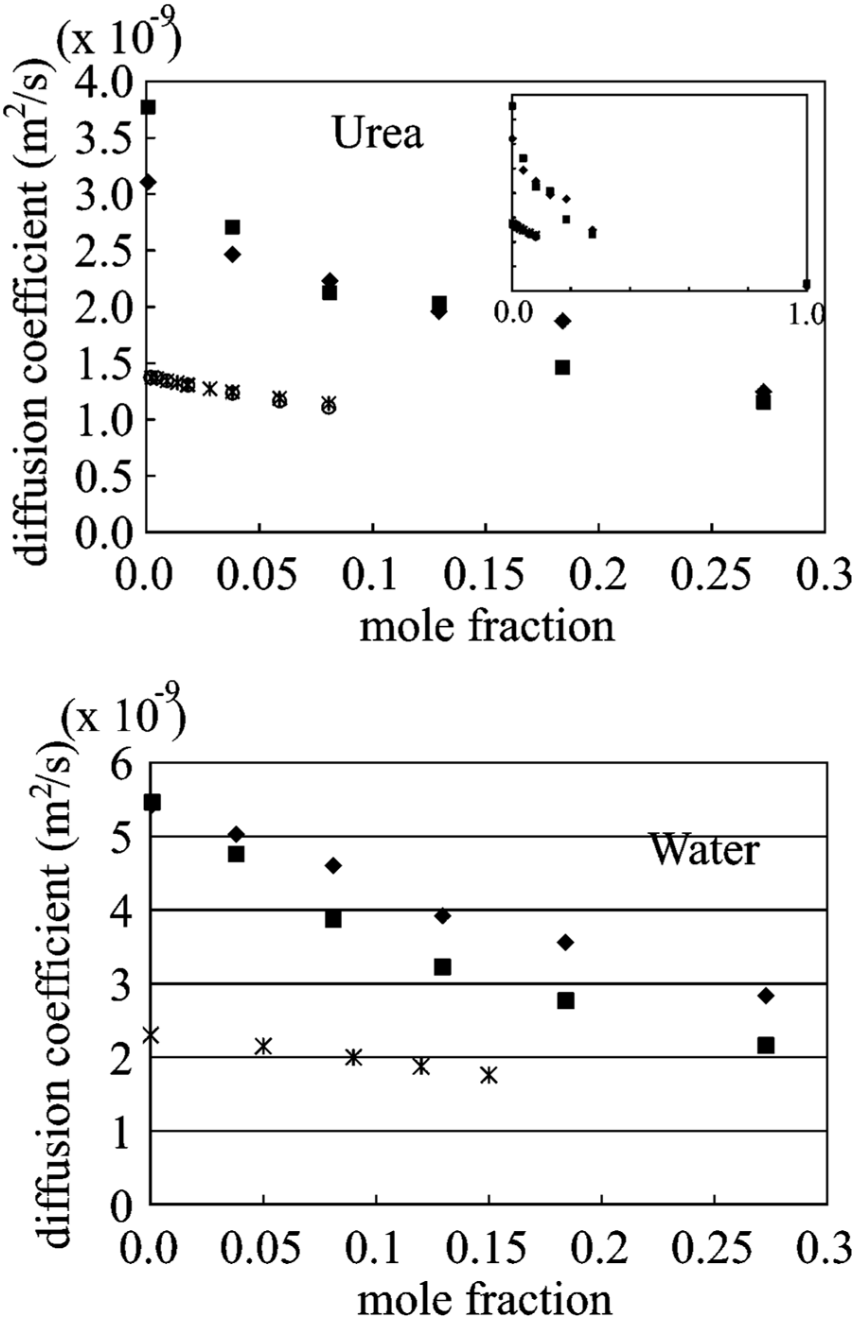

Figure 3.
Diffusion coefficients of urea and water molecules. The closed diamond represents the OPLS urea model, and the closed square represents the KBFF model, which were obtained from NPT molecular simulations at $T = 298.0$ K and $P = 1$ atm. Water is the TIP3P model in both cases. The open circle[31] and asterisk[32] in the urea figure represent the experimental data. The asterisk in the water figure represents the experimental data.[36]

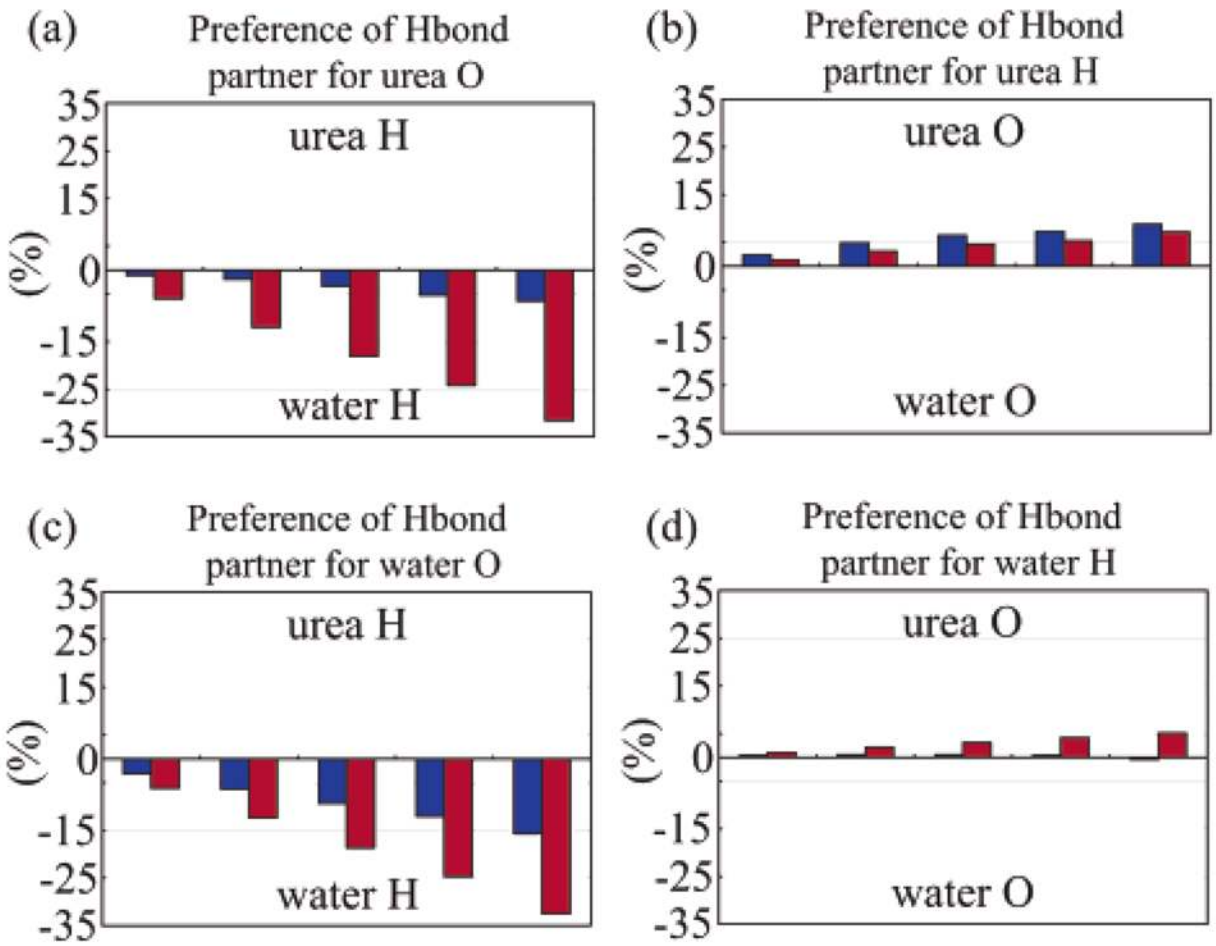

Figure 4.

Preference of hydrogen-bond partner vs concentration for (a) urea oxygen, (b) urea hydrogen, (c) water oxygen, and (d) water hydrogen. The blue bars are OPLS, and the red are KBFF. The concentrations correspond to mole fractions $x = 0.03806, 0.08106, 0.1294, 0.1840$, and 0.2728, from the left to right. A positive percentage means a preference for urea as the hydrogen-bond partner relative to ideal. The ideal probability that the oxygen (hydrogen) atom chooses urea's hydrogen as a hydrogen-bond partner is defined by the number of total urea hydrogen (oxygen) atoms divided by the number of total hydrogen (oxygen) atoms in the system.

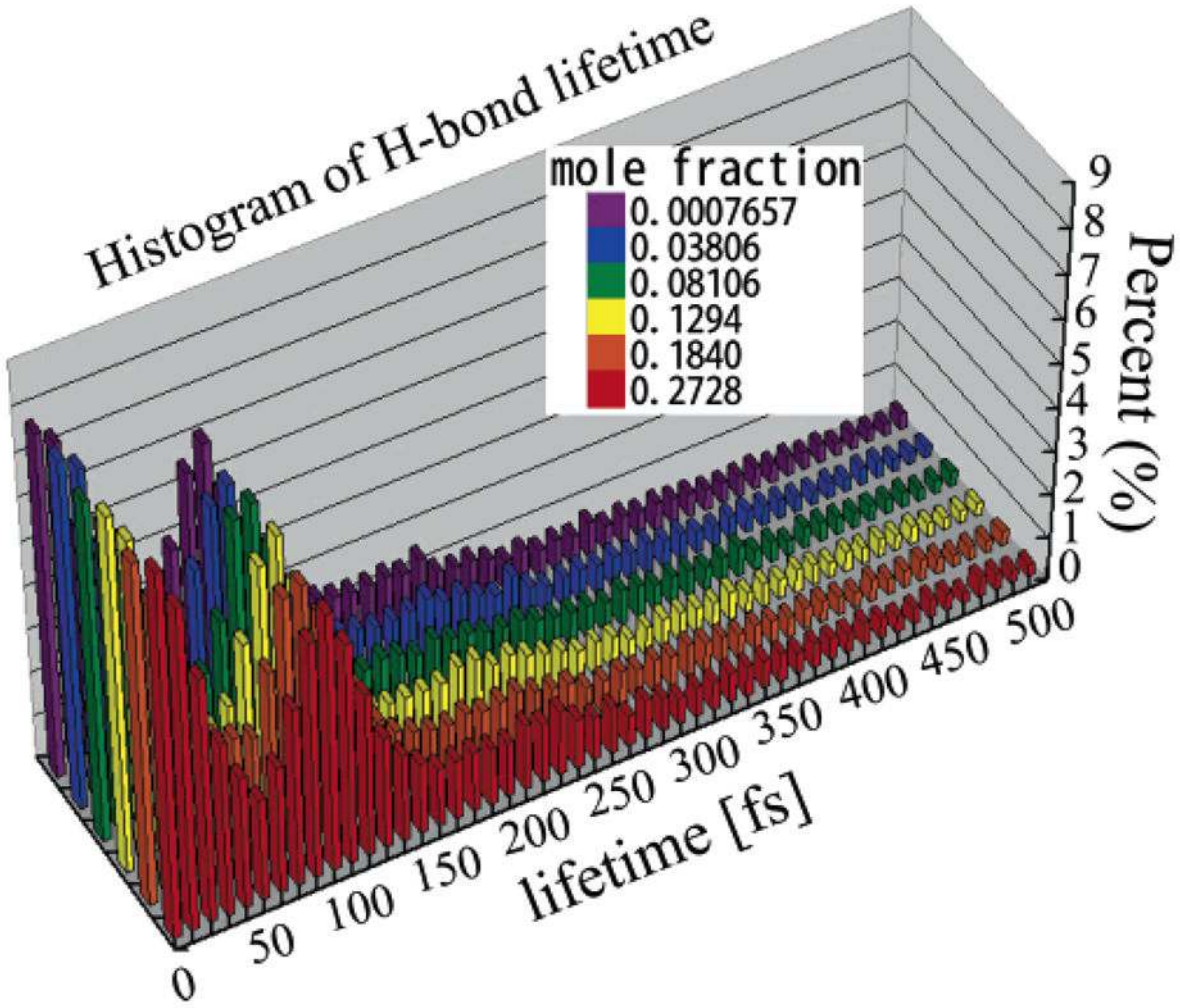

Figure 5.
Comparison of histograms of hydrogen-bond residence times between water hydrogen and water oxygen at different concentrations of OPLS urea solution. The bin size is 10 fs.

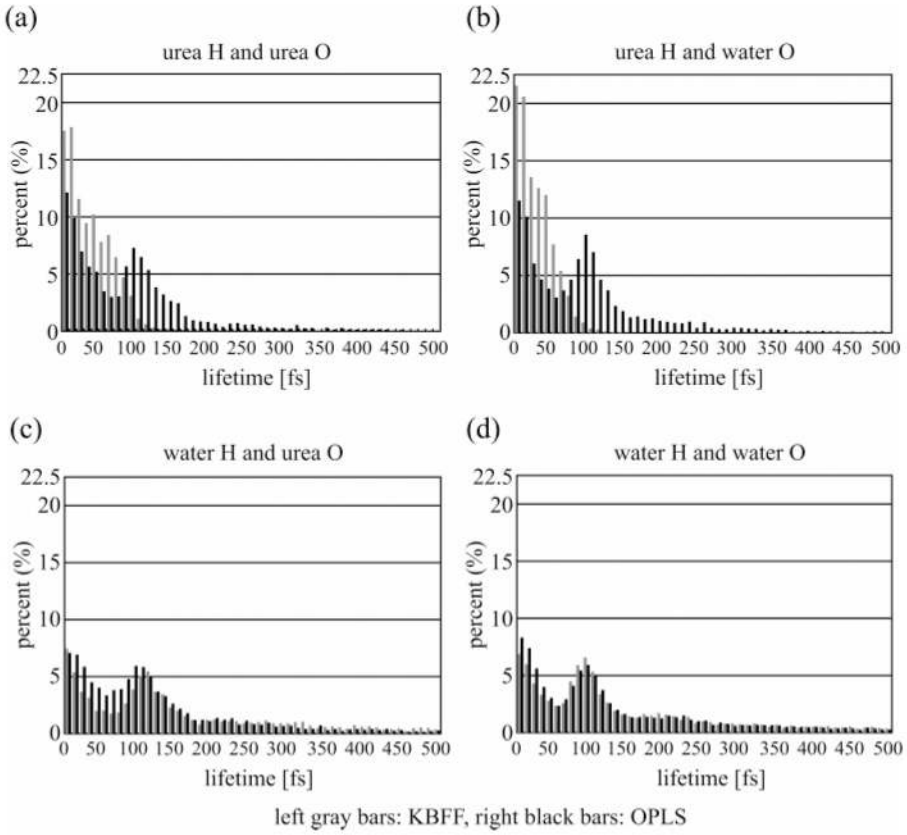

Figure 6.

Comparison of histograms of the lifetime of hydrogen bonds of OPLS and KBFF urea solutions at solubility concentrations ($x = 0.2728$): (a) hydrogen bond between urea oxygen and urea hydrogen; (b) hydrogen bond between urea oxygen and water hydrogen; (c) hydrogen bond between water oxygen and urea hydrogen; (d) hydrogen bond between water oxygen and water hydrogen. Gray bars are KBFF urea solution, and black bars are OPLS urea solution. The bin size is 10 fs.

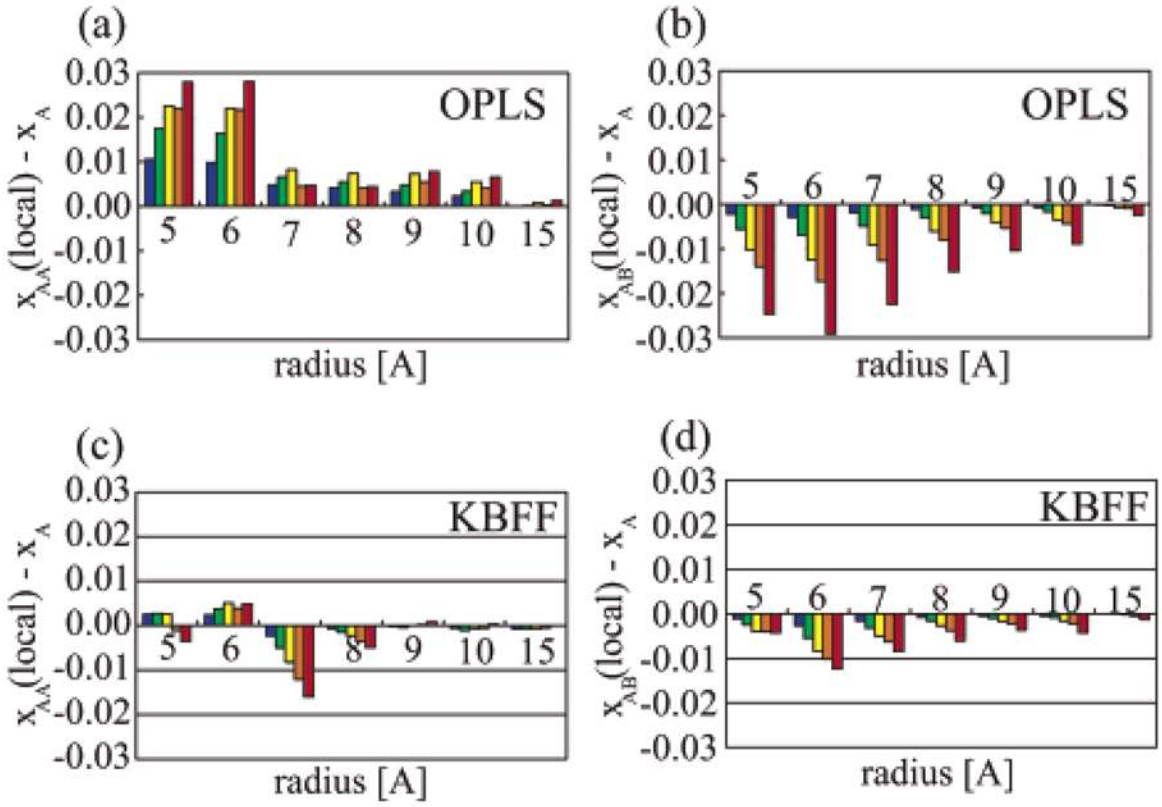

Figure 7.

Preferential solvation of (a) urea around urea and (b) urea around water for OPLS solutions at different concentrations. Parts c and d are those for KBFF solutions. Every 5 bars at the same radius express the preferential solvation at mole fractions of x = 0.03806, 0.08106, 0.1294, 0.1840, and 0.2728, from left to the right.

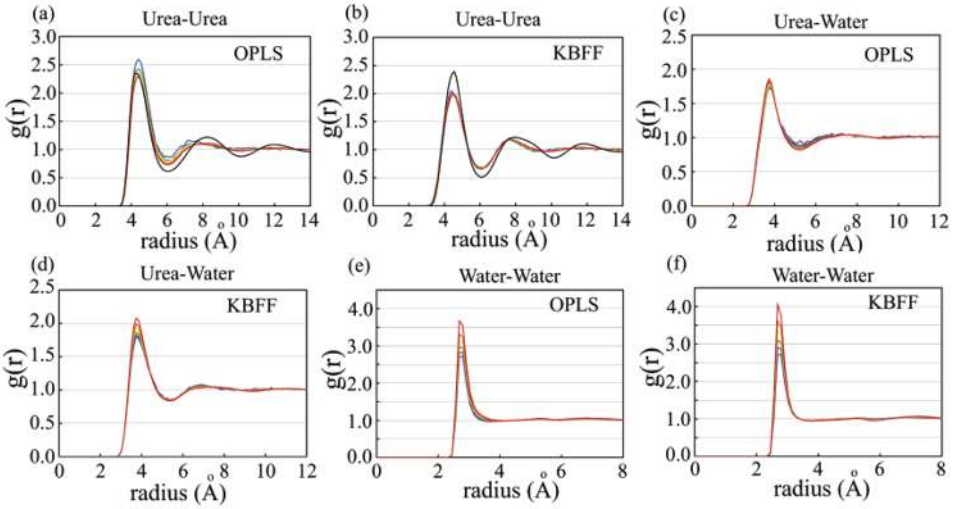

Figure 8.

Radial distribution functions of (a and b) urea–urea centers of mass, (c and d) urea–water centers of mass, and (e and f) water–water centers of mass. Parts a, c, and e are the OPLS urea solutions, and parts b, d, and f are the corresponding KBFF urea solutions. The different colors correspond to the different mole fractions, such as mole fraction 0.03806 (blue), 0.08106 (green), 0.1294 (yellow), 0.1840 (orange), 0.2728 (red), and 1.00 (black).

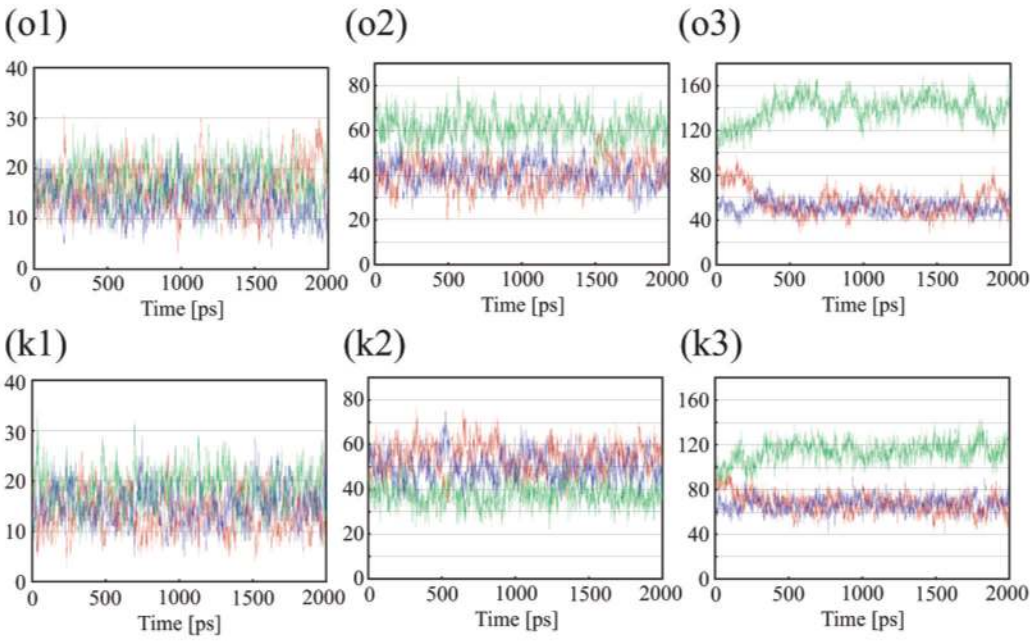

Figure 9.

Time series of the number of molecules that belong to high (red line), medium (green line), and low (blue line) urea density clusters. o1, o2, and o3 represent the OPLS solution, and k1, k2, and k3 represent the KBFF solution. o1 and k1 are the system at mole fraction $x = 0.03806$, o2 and k2 are at $x = 0.1294$, and o3 and k3 are at $x = 0.2728$, respectively. The molecule that has more than 20% excess molecules around it than in the bulk belongs to a high-density cluster, the molecule that has less than 20% deficit molecules around it belongs to a low-density cluster, and the rest belongs to a medium-density cluster.

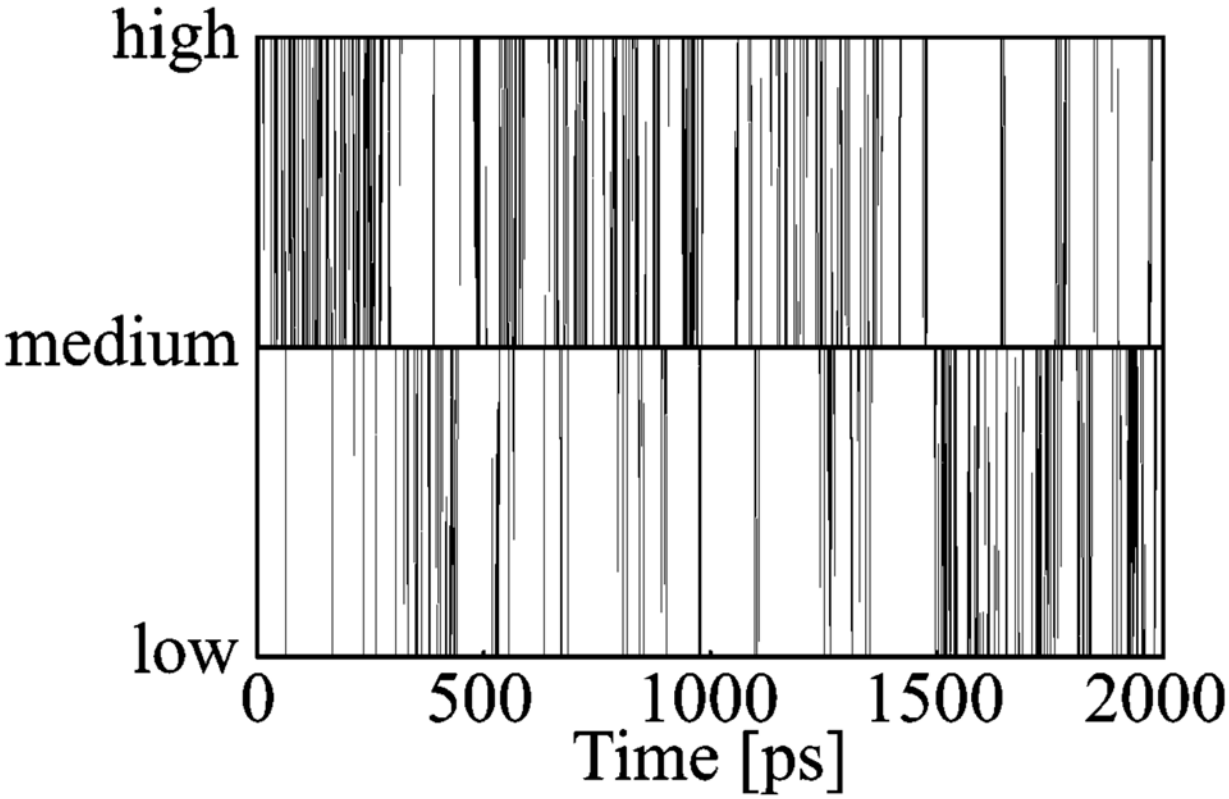

Figure 10.
Time series of the change of clusters to which one typical urea molecule belongs (from KBFF solution at $x = 0.2728$).

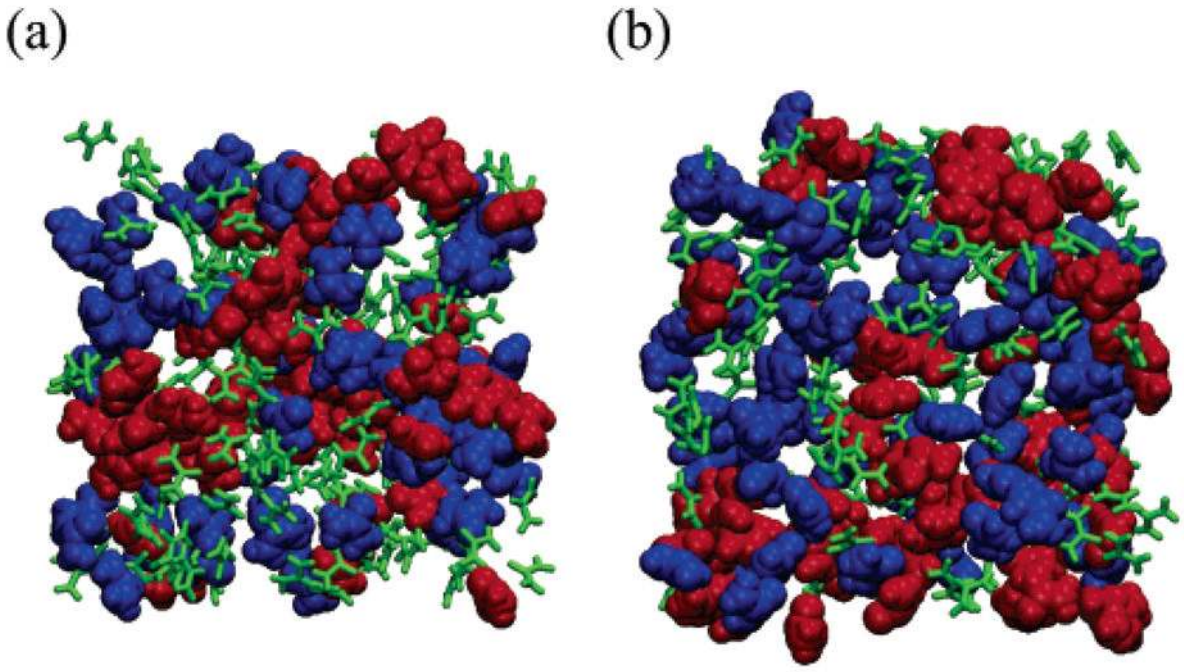

Figure 11.
Typical snapshots of (a) OPLS and (b) KBFF urea solutions at the solubility limit concentration (x = 0.2728). Red, green, and blue urea molecules belong to high-, medium-, and low-density clusters, respectively. High- (red) and low-density (blue) cluster molecules were drawn by space-filling models, and medium-density (green) cluster molecules were drawn by ball-and-stick models.

TABLE 1

Force-Field Parameters for the OPLS[12] and KBFF[13] Models

	mass	charge	ε (kJ/mol)	σ (nm)
		OPLS		
O	15.999	−0.390	0.87864	0.296
C	12.011	0.142	0.43932	0.375
N1	14.007	−0.542	0.71128	0.325
H11	1.008	0.333	0.00000	0.000
H12	1.008	0.333	0.00000	0.000
N2	14.007	−0.542	0.71128	0.325
H21	1.008	0.333	0.00000	0.000
H22	1.008	0.333	0.00000	0.000
		KBFF		
O	15.999	−0.675	0.56000	0.310
C	12.011	0.921	0.41700	0.377
N1	14.007	−0.693	0.50000	0.311
H11	1.008	0.285	0.08800	0.158
H12	1.008	0.285	0.08800	0.158
N2	14.007	−0.693	0.50000	0.311
H21	1.008	0.285	0.08800	0.158
H22	1.008	0.285	0.08800	0.158

TABLE 2

Number of Urea Molecules (N_u) and Number of Water Molecules (N_w)[a]

N_u	N_w	mole fraction	molality	molarity	volume	density
			OPLS			
1	1305	0.0007657	0.04253	0.04288	38.73	1.011
47	1188	0.03806	2.196	2.029	38.47	1.046
95	1077	0.08106	4.896	4.089	38.58	1.081
142	955	0.1294	8.254	6.152	38.32	1.115
189	838	0.1840	12.52	8.209	38.23	1.149
248	661	0.2728	20.83	11.04	37.31	1.193
530	0	1.0		22.65	38.85	1.361
			KBFF			
1	1305	0.0007657	0.04253	0.04286	38.75	1.010
47	1188	0.03806	2.196	2.013	38.78	1.037
95	1077	0.08106	4.896	4.027	39.17	1.064
142	955	0.1294	8.254	6.031	39.10	1.093
189	838	0.1840	12.52	7.993	39.27	1.118
248	661	0.2728	20.83	10.70	38.50	1.156
530	0	1.0		22.41	39.27	1.346

[a]The units of molality, molarity, volume, and density are mol/kg, mol/L, nm^3, and g/cm^3, respectively.

TABLE 3

Average Number of Hydrogen Bonds[a]

mole fraction	total	1 urea	1 water	4 H of urea	2 H of water	1 O of urea	1 O of water
			OPLS				
0.0007657	1819.3	3.13	2.79	1.55	1.39	1.58	1.39
0.03806	1716.8	3.15	2.77	1.57	1.38	1.58	1.38
0.08106	1625.1	3.11	2.74	1.56	1.37	1.55	1.37
0.1294	1515.9	3.07	2.72	1.54	1.36	1.53	1.36
0.1840	1416.7	3.04	2.69	1.53	1.35	1.52	1.35
0.2728	1250.8	3.00	2.66	1.51	1.33	1.49	1.33
1.000	710.5	2.68		1.34		1.34	
			KBFF				
0.0007657	1818.0	2.27	2.78	0.43	1.39	1.83	1.39
0.03806	1696.7	2.19	2.77	0.43	1.41	1.76	1.36
0.08106	1584.5	2.15	2.75	0.44	1.43	1.71	1.32
0.1294	1459.6	2.10	2.74	0.45	1.46	1.66	1.28
0.1840	1328.2	2.05	2.71	0.46	1.48	1.59	1.23
0.2728	1122.5	1.96	2.66	0.48	1.52	1.47	1.14
1.000	363.8	1.37		0.69		0.69	

[a]"Total" is the average number of total hydrogen bonds of each system, "1 urea" is the average number of hydrogen bonds per 1 urea molecule, similarly, "1 water" is the average number of hydrogen bonds per 1 water molecule, "4 H of urea" is the average number of hydrogen bonds per the 4 hydrogen atoms of urea (one urea contains 4 hydrogen atoms), and "2 H of water" is the average number of hydrogens bonds one per the 2 hydrogen atoms of water, "1 O of urea" is the average number of hydrogen bonds per the 1 oxygen atom of urea, and "1 O of water" is the average number of hydrogen bonds per the 1 oxygen atom of water.

TABLE 4

Average Lifetimes of Hydrogen Bonds[a]

mole fraction	urea H and urea O	urea H and water O	water H and urea O	water H and water O
		OPLS		
0.0007657		9.9×10^1	1.4×10^2	1.8×10^2
0.03806	1.0×10^2	9.4×10^1	1.5×10^2	1.8×10^2
0.08106	9.2×10^1	9.7×10^1	1.5×10^2	1.8×10^2
0.1294	9.4×10^1	9.6×10^1	1.5×10^2	1.8×10^2
0.1840	9.0×10^1	1.0×10^2	1.4×10^2	1.9×10^2
0.2728	9.5×10^1	9.9×10^1	1.5×10^2	1.9×10^2
1.00	9.7×10^1			
		KBFF		
0.0007657		3.7×10^1	2.1×10^2	1.8×10^2
0.03806	4.3×10^1	2.8×10^1	2.5×10^2	1.8×10^2
0.08106	3.2×10^1	2.9×10^1	2.6×10^2	1.9×10^2
0.1294	3.8×10^1	3.0×10^1	2.7×10^2	2.0×10^2
0.1840	3.7×10^1	2.9×10^2	2.7×10^2	2.0×10^2
0.2728	3.9×10^1	3.0×10^1	2.9×10^2	2.2×10^2
1.00	4.0×10^1			

[a]"Urea H and urea O" is for the hydrogen bond between the urea hydrogen and the urea oxygen. Others are defined similarly. The lifetimes are given in units of fs.