

# Computer Simulation of Urea–Water Mixtures: A Test of Force Field Parameters for Use in Biomolecular Simulation

Lorna J. Smith,<sup>†</sup> Herman J. C. Berendsen,<sup>‡</sup> and Wilfred F. van Gunsteren<sup>\*,§</sup>

*Oxford Centre for Molecular Sciences, Central Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QH, United Kingdom, Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands, and Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zurich, ETH Hönggerberg, CH-8093 Zürich, Switzerland*

*Received: April 28, 2003; In Final Form: October 15, 2003*

A molecular model for urea to be used in conjunction with the simple point charge (SPC) model for liquid water in protein denaturation studies is validated by comparison of molecular dynamics (MD) simulation results to experimental data at 298 K as a function of urea mole fraction. The density, enthalpy of mixing, free enthalpy of urea hydration, and urea diffusion show very good agreement with the experimental values. The experimental error in the free enthalpy of hydration, which is dominated by the inaccuracy of the vapor pressure of solid urea, is larger than the simulation error. This limited accuracy does not allow a check of the nonideality of the solution. Free enthalpies have been obtained by thermodynamic integration. The importance of a correct use of the combinatorial factor in the partition function when interpreting simulation results obtained by thermodynamic integration is discussed. The tested, GROMOS96 compatible, force field parameters form a good basis for biomolecular simulations using urea–water mixtures.

## I. Introduction

Understanding the driving forces for protein folding and unfolding in aqueous solution is one of the major challenges in computational biochemistry. Protein unfolding or denaturation can be induced by changing the temperature or pressure, or by changing the composition of the solvent. Admixture of so-called denaturants as cosolvents to water induces proteins to unfold. Denaturants such as guanidinium chloride or urea are much used to study protein (un)folding experimentally. When simulating the onset of protein unfolding by molecular dynamics (MD) computer simulation, neutral urea molecules are the denaturants of choice, because the appropriate treatment of long-range electrostatic interactions is simpler for polar than for highly charged systems. Urea-induced unfolding of the protein barnase has been studied using MD simulation of this protein in an 8 M urea solution starting from the native, folded structure.<sup>1,2</sup> A similar study has been performed for helix unfolding.<sup>3</sup> The free energy of association of methane with methane,<sup>4,5</sup> and of urea with aromatic hydrocarbons<sup>6</sup> in urea–water mixtures has been investigated by Monte Carlo and MD simulation with an eye to understanding the basic interactions responsible for protein folding and denaturation.

However, to obtain thermodynamically correct results, the molecular models used for the protein molecule, for the urea molecules and for the water molecules should be consistent. They should contain similar force field terms and their force field parameters should have been calibrated similarly. The thermodynamic properties of the water models that are generally used to simulate proteins in aqueous solution, such as SPC,<sup>7</sup> TIP3P,<sup>8</sup> SPC/E,<sup>9</sup> and TIP4P,<sup>8</sup> are well-known.<sup>10</sup> For the urea

models used this is not the case. The barnase study of Tirado-Rives et al.<sup>1</sup> uses the AMBER/OPLS force field for the protein, the TIP3P water model, and the OPLS urea model proposed by Duffy et al.<sup>11</sup> The barnase study of Caflisch and Karplus<sup>2</sup> uses the CHARMM22 force field for the protein, a modified TIP3P model for water, and for urea a model with parameters derived in analogy to those for Asn and Glu side chains in proteins. The methane–methane association study<sup>4</sup> is based on the SPC water model and the OPLS urea model, and the urea–aromatic hydrocarbons association study<sup>6</sup> on the TIP4P water model and the OPLS model for urea and the solutes. For none of the combinations of water and urea models has the thermodynamics of mixing, i.e., the density, enthalpy of mixing, and free enthalpy of hydration for urea, been reported and compared to experimental values.

During the past decade a number of urea models have been proposed for use in simulations or urea–water mixtures. The parameters of the urea models used by Boek et al.<sup>12,13</sup> were taken from general purpose biomolecular force fields and used in conjunction with SPC water<sup>12</sup> and SPC/E water.<sup>13</sup> The simulations were analyzed in structural terms: the total nitrogen scattering function  $G_N(r)$  was compared to neutron scattering data, and it was concluded that reproduction of  $G_N(r)$  is not providing a very severe test of the force field used. Hernandez-Cobos and Ortega-Blake<sup>14</sup> presented MC simulations of urea–water mixtures as a function of urea concentration. They used their own ab initio data based models for water and urea and analyzed their simulations in terms of radial distribution functions and energies of the components of the systems. Duffy et al.<sup>11</sup> proposed an OPLS compatible urea model and tested it in conjunction with TIP4P water by comparing the difference in chloroform/water partition coefficients for urea and acetamide with the experimental number. They did not give a comparison of the calculated free energy of hydration of their urea model with an experimental value. Astrand et al.<sup>15</sup> presented a

\* Corresponding author. E-mail: wfvgn@igc.phys.chem.ethz.ch. Phone: +41-1-632-5501. Fax: +41-1-632-1039.

<sup>†</sup> University of Oxford.

<sup>‡</sup> University of Groningen.

<sup>§</sup> Swiss Federal Institute of Technology Zurich.

polarizable urea model and analyzed a simulation of a urea–water mixture mainly in structural terms.<sup>16</sup> Tsai et al.<sup>17</sup> proposed two urea models that are compatible with the ENCAD biomolecular force field<sup>18</sup> and the F3C water model.<sup>19</sup> MD simulations of urea–water mixtures were analyzed in terms of hydrogen-bond occurrence and lifetimes, radial distribution functions, and diffusion constants. The latter two data were compared to experimental values. Vanzi et al.<sup>20</sup> presented MC simulations of one urea molecule in aqueous solution using the OPLS urea model<sup>11</sup> and the TIP4P water model<sup>10</sup> and analyzed the hydrogen bonding behavior. Idrissi and Sokolic<sup>21</sup> studied the short time dynamics of urea–water mixtures using the simple urea model of ref 17 and the SPC water model. Translational and rotational dynamic properties of urea and water molecules were reported. Chitra and Smith<sup>22</sup> reported the density, radial distribution functions, diffusion coefficients, and dielectric properties of urea–water mixtures and compared these to experimental values. They used the OPLS urea model<sup>11</sup> in conjunction with SPC/E water.<sup>9</sup> Other recent simulation studies of urea consider mainly structural properties of the solution.<sup>23–26</sup>

To simulate urea-induced protein denaturation using the GROMOS96 general-purpose biomolecular force field,<sup>27,28</sup> which is compatible with the SPC water model, we set out to select from the literature or compose a urea model consistent with the GROMOS96 force field and SPC water. We chose the OPLS urea model<sup>11</sup> and tested its properties in conjunction with the SPC water model. Simulated values for the density, enthalpy of mixing, free enthalpy of urea hydration, and urea diffusivity are compared to experimental values as a function of urea mole fraction. The comparison of simulated with experimental values for the enthalpy of mixing and the free enthalpy of urea hydration is given in detail, because such a comparison has not been given before<sup>11</sup> and is not trivial. Because of the limited lengths of our simulations, dielectric properties were not calculated.

In section II the molecular models for urea and water are specified together with the simulation setup and protocols. The results are reported and compared to experimental data in section III. There, the theory of thermodynamic integration for mixtures, which leads to a contribution to the free enthalpy of solvation of urea due to the combinatorial factor in the partition function, is given. Corrections resulting from nonideality behavior, which are small, are derived in an appendix. Section IV contains a discussion of the results.

## II. Methods

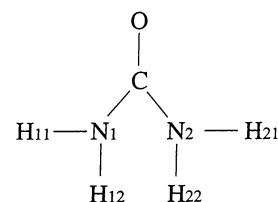
**A. Models for Urea and Water.** The rigid urea model of Duffy et al.<sup>11</sup> was converted to a flexible one by introducing bonded, intramolecular interaction terms. Duffy et al.<sup>11</sup> based their ideal urea geometry on the X-ray structure. The values for the force field parameters as defined in terms of the GROMOS96 force field<sup>27,28</sup> are specified in Table 1 and Figure 1. In the simulations the bond lengths of the urea molecules were kept rigid. For water, the completely rigid simple point charge (SPC) model<sup>7</sup> was used. The van der Waals interactions between urea and water molecules were modeled using the combination rules as used in the GROMOS96 force field.<sup>27,28</sup>

**B. Simulation Setup and Protocols.** The MD simulations were performed with the GROMOS96 biomolecular simulation package.<sup>27,29</sup> One pure water system and four different urea–water mixtures, with the urea concentration ranging from 0.08 molar to 10.04 M, were considered. The latter corresponds to a saturated urea solution. The five systems termed A–E are

**TABLE 1: Force Field Parameters for Urea<sup>a</sup>**

atom	description	$C_6^{1/2} 10^{-3} \text{ (kJ mol}^{-1} \text{ nm}^6)^{1/2}$	$C_{12}^{1/2} 10^{-3} \text{ (kJ mol}^{-1} \text{ nm}^{12})^{1/2}$	$q, e$
C	carbon	69.906	3.6864	0.142
O	oxygen	48.620	1.2609	−0.390
N	nitrogen	57.903	1.9877	−0.542
H	hydrogen	0.0	0.0	0.333
bond	$b_o, \text{ nm}$	$K_b, 10^6 \text{ kJ mol}^{-1} \text{ nm}^{-4}$		
C–O	0.1265	13.1		
C–N	0.135	10.3		
N–H	0.100	18.7		
bond angle	$\theta_o \text{ (deg)}$	$K_\theta \text{ (kJ mol}^{-1})$		
O–C–N	121.4	690		
N–C–N	117.2	636		
C–N–H	120.0	390		
H–N–H	120.0	445		
improper dihedral angle	$\xi_o \text{ (deg)}$	$K_\xi \text{ (kJ mol}^{-1} \text{ deg}^{-2})$		
C –N1 –N2 –O	0.0	0.0510		
N1–H11–H12–C	0.0	0.0510		
N2–H21–H22–C	0.0	0.0510		
torsional dihedral angle	$\cos(\delta)$	$m$	$K_\varphi \text{ (kJ mol}^{-1})$	
O–C–N1–H11	−1.0	2	41.8	
O–C–N2–H21	−1.0	2	41.8	

<sup>a</sup> The functional form of the potential energy function and the definition of the parameters are specified elsewhere.<sup>27,29</sup> The atom names used to specify the dihedral angles are given in Figure 1.



**Figure 1.** Chemical formula of urea and atom names used in the definition of the force field parameters in Table 1.

**TABLE 2: Simulated Urea–Water Mixtures<sup>a</sup>**

mixture name	no. of molecules (mol)			urea mole fraction, $x$	$T$ (K)	vol $\langle V \rangle$ (nm <sup>3</sup> )	urea concn $\langle c \rangle$ (M)
	$n_u$	$n_w$	$n_{\text{system}}$				
A	0	658	658	0.0	297.8	20.322	0.0
B	1	658	659	0.001517	297.8	20.371	0.08
C	36	1908	1944	0.018519	299.9	61.299	1.0
D	200	1100	1300	0.153846	296.5	48.216	6.9
E	432	1296	1728	0.250000	295.1	70.760	10.04

<sup>a</sup>  $n_u$ : number of urea molecules.  $n_w$ : number of water molecules.  $n_{\text{system}}$ : total number of molecules in the system. The urea mole fraction is denoted by  $x$ . The average temperature is denoted by  $T$ . The average volume of the computational box in the constant pressure simulations is denoted by  $\langle V \rangle$ . The average urea concentration  $\langle c \rangle$  is calculated using the formula  $\langle c \rangle = \langle \rho \rangle x (xM_u + (1-x)M_w)$  in which the average density  $\langle \rho \rangle$  of the system is obtained using the formula  $\langle \rho \rangle = (n_u M_u + n_w M_w) / (N_A \langle V \rangle)$ , in which  $N_A$  is Avogadro's number and  $M_u = 60.0558 \text{ g mol}^{-1}$  and  $M_w = 18.0154 \text{ g mol}^{-1}$ . System B corresponds to a pseudo-infinite dilution.

specified in Table 2. In all cases periodic boundary conditions were used. Systems A and B were contained in a truncated octahedron, whereas systems C–E were rectangular. The initial configurations were generated as follows. The initial configuration for system A was taken from an equilibrated SPC water simulation at room temperature and pressure. For system B one urea molecule was superimposed onto an equilibrated box with water molecules and subsequently the water molecules overlapping with the urea molecule were removed and the box size

was chosen to keep  $n_w = 658$  water molecules. For system C 1 urea molecule was immersed in a small box with 53 water molecules. Thirty-six periodic copies of this system were assembled to obtain an initial configuration. For system D a similar procedure was carried out starting from a small system containing 2 urea molecules and 11 water molecules. System E was obtained similarly starting from 1 urea molecule surrounded by 3 water molecules.

The OH bond lengths, the HOH bond angle (or intramolecular HH distance) of the water molecules, and the bond lengths of the urea molecules were kept rigid by applying the SHAKE algorithm<sup>30</sup> with a relative geometric tolerance of  $10^{-4}$ . The temperature of the systems was weakly coupled to a bath at 298 K with a relaxation time of 0.1 ps.<sup>31</sup> The pressure was calculated with a molecular virial and held constant at 1 atm using the weak coupling method with a relaxation time of 0.5 ps.<sup>31</sup> The isothermal compressibility of the system was set to a value of  $4.575 \times 10^{-4} \text{ (kJ mol}^{-1} \text{ nm}^{-3})^{-1}$ . The equations of motion were integrated using the leapfrog algorithm and a time step of 2 fs. The nonbonded interaction between atoms of pairs of molecules was calculated according to a spherical triple-range cutoff scheme. Short-range van der Waals and electrostatic interactions were evaluated at every time step by using a molecular-pair list that was generated with a short-range cutoff radius of 0.8 nm between water oxygen atoms and/or centers of geometry of the urea molecules. Longer range van der Waals and electrostatic interactions, between molecules with the water oxygen and/or urea center of geometry distances larger than 0.8 nm and shorter than 1.4 nm, were evaluated every fifth time step, at which point the pair list was also updated, and were kept unchanged between these updates. Electrostatic interactions beyond the long-range cutoff radius of 1.4 nm were approximated using a reaction-field term<sup>32</sup> with a dielectric permittivity  $\epsilon_{\text{RF}} = 54$  for all mixtures.<sup>33</sup> At the beginning of the simulations the velocities of the atoms were assigned from a Maxwell distribution at 298 K. For each system, a period of 100 ps of equilibration was followed by 100 ps for the calculation of the various properties, such as the density, energy, and diffusivity. During these simulations configurations were saved every 0.1 ps.

The free enthalpies of solvation of a urea molecule in the various mixtures were obtained through thermodynamic integration starting from equilibrated system configurations. The interaction of one designated urea molecule with all other molecules was made a function of a coupling parameter  $\lambda$  according to the standard GROMOS96 procedure.<sup>27,29</sup> Soft-core interactions were used with soft-core parameters  $\alpha_L = 0.5$  and  $\alpha_C = 0.5 \text{ nm}^2$ , and the power of  $\lambda^n$  in the  $\lambda$ -dependent prefactor was taken as  $n = 1$ . Ensemble averages of  $\partial H / \partial \lambda$  were obtained from 100 ps sampling periods (after 5 ps of equilibration at each  $\lambda$  value), for 23  $\lambda$  values in the range [0,1]. For each of the systems C–E three free enthalpy calculations were carried out, in each case designating another urea molecule of the 36, 200, and 432 urea molecules, respectively, as the one for which the interactions with the other molecules are switched off in function of  $\lambda$ . Trapezoid summation was used for the integration over  $\lambda$ .

### III. Results and Comparison with Experiment

**A. Density.** In Table 2 the average volumes of system A–E are reported, and in Table 3 average densities, which are compared with experimental densities. The latter values have been interpolated from the *CRC Handbook of Chemistry and*

**TABLE 3: Thermodynamic Properties, Density  $\rho(x)^a$**

mixture name	mole fraction $x$	density $\rho(x)$ (g cm <sup>-3</sup> )		
		simulation	experiment <sup>34</sup>	experiment <sup>35</sup>
A	0.0	0.9686	0.9982	0.9982
B	0.001517	0.9712	0.9995	0.9995
C	0.018519	0.9897	1.0136	1.0140
D	0.153846	1.0961	1.1036	1.1030
E	0.250000	1.1567	1.1577	1.1470

<sup>a</sup> For the definition of the symbols see the footnote for Table 2.

**TABLE 4: Thermodynamic Properties, Energy of Solution and Heat of Solvation<sup>a</sup>**

mixture name	mole fraction $x$	$E_{\text{pot}}(x)$ (kJ mol <sup>-1</sup> )	$E_{\text{pot}}^{\text{ex}}(x)$ (kJ mol <sup>-1</sup> )		$\Delta H_{\text{solv}}(\text{u,s to sol},x)$ (kJ mol <sup>-1</sup> )
		sim	sim	exp	exp
A	0.0	-41.60	0	0	15.1
B	0.001517	-41.62	-0.02	-0.05	15.1
C	0.018519	-42.02	-0.42	-0.64	15.0
D	0.153846	-47.27	-5.67	-5.45	14.0
E	0.250000	-51.57	-9.97	-8.96	13.6

<sup>a</sup>  $E_{\text{pot}}(x)$ : nonbonded potential energy of the system divided by the number of molecules in the system.  $E_{\text{pot}}^{\text{ex}}(x) = E_{\text{pot}}(x) - E_{\text{pot}}(x=0)$ . When the excess potential energy  $E_{\text{pot}}^{\text{ex}}(x)$  is calculated, the simulated SPC value  $E_{\text{pot}}(0) = -41.60 \text{ kJ mol}^{-1}$  is used to obtain the simulated and experimental excess potential energies, because we are interested in the performance of the urea model with respect to SPC water.  $H_{\text{solv}}$ : heat of solvation. For the definition of the other symbols see the footnote of Table 2.

Physics tables<sup>34</sup> (pp 8–83). For comparison, the values according to the equation of Kawahari and Tanford,<sup>35</sup>

$$\rho(w_u) = \rho(0)[1 + 0.2658w_u + 0.0330w_u^2] \quad (1)$$

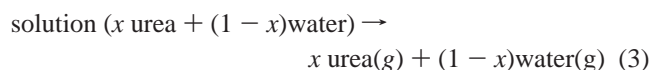
are given as well. Here  $w_u$  represents the weight fraction of urea in the mixture, and  $\rho(0)$  is the density of water. The values computed with eq 1 agree to the third decimal place with those of the *CRC Handbook* tables, except for the saturated solution (system E), where the value from eq 1 is 1% lower. The largest discrepancy between simulation and experiment is due to the density of SPC water, which is somewhat lower than that of real water. The simulated change in density due to urea agrees well with the experimental data.

**B. Energy.** The average potential interaction energies per molecule of mixture,  $E_{\text{pot}}(x)$ , and the excess interaction energies  $E_{\text{pot}}^{\text{ex}}(x)$  with respect to pure water are reported in Table 4 as a function of urea mole fraction  $x$ . The simulated values for the potential energy comprise the total nonbonded interaction energy between all molecules of the mixture. The excess interaction energies are defined as

$$E_{\text{pot}}^{\text{ex}}(x) = E_{\text{pot}}(x) - E_{\text{pot}}(x=0) \quad (2)$$

where  $E_{\text{pot}}(x=0)$  is obtained from simulation. The reason to use excess values for comparison with experiment is that we wish to test the performance of the urea model with respect to SPC water, rather than testing the SPC model itself.

Experimental values for the excess interaction energies were obtained as follows. Consider the vaporization process

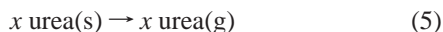
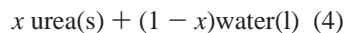


Assuming that the contributions to the energy due to internal motions of the water and urea molecules are the same in the liquid (l) and gas (g) phases (i.e., we neglect changes in

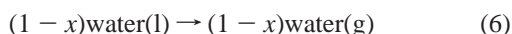


vibrational frequencies that could influence zero-point energies), and assuming ideal gas conditions and classical behavior of rotational and translational motions, the change in internal energy,  $\Delta U$ , of this process equals the negative of the potential interaction energy  $E_{\text{pot}}(x)$  we are looking for. This vaporization process is not experimentally accessible, but it can be built up from the following processes, for which the change in internal energy is experimentally accessible:

solution ( $x$  urea,  $(1 - x)$  water)  $\rightarrow$



and



Process (4) is the inverse of the solution process of  $x$  moles of solid (s) urea into  $(1 - x)$  moles of pure water. The heat of solvation of crystalline urea in water at various concentrations is given in the NBS tables<sup>34</sup> and varies between  $15.1 \text{ kJ mol}^{-1}$  at  $x = 0$  to  $12.4 \text{ kJ mol}^{-1}$  at  $x = 0.25$ . These are differential heats, dissolving a small amount of urea at a given concentration. For the integral heat of solvation to reach a mole fraction  $x$  from solid urea and pure water, we derived the values given in the last column of Table 4.

Process (5) is the sublimation of solid urea. The heat of sublimation of urea is not accurately known and causes the largest uncertainty in the result. We use the value of  $93.5 \pm 1.3 \text{ kJ mol}^{-1}$  given by Ferro.<sup>36</sup> Because the heat of sublimation equals the enthalpy change, a value of  $RT$  must be subtracted to obtain the change in internal energy upon sublimation.

Process (6) is the vaporization of  $(1 - x)$  mole of water. The energy change per mole is  $-E_{\text{pot}}$  of SPC water, or  $41.60 \text{ kJ mol}^{-1}$  (simulation A).

Adding the internal energy changes of these three processes and subtracting the energy change of SPC water yields the excess values reported in Table 4.

We estimate the error in the simulated numbers to be smaller than  $0.1 \text{ kJ mol}^{-1}$ . The error in the experimental excess values is proportional to  $x$  and reaches about  $0.4 \text{ kJ mol}^{-1}$  for the saturated solution (system E). The simulated and experimental values agree quite well, only at the highest concentration is the simulated interaction slightly stronger than the experimental one.

**C. Free Enthalpy.** When calculating the free enthalpy of desolvation of a particular urea molecule by thermodynamic integration, the intermolecular interaction of the designated urea molecule with the rest of the solution,  $E_{\text{pot}}^{\text{inter}}(1 \text{ urea}^*)$  is switched off as a function of the coupling parameter  $\lambda$ , and the free enthalpy of desolvation  $\Delta G^*$  of the designated urea molecule (symbolized by the  $*$ ) is calculated as

$$\Delta G^* = \int_0^1 \left\langle \frac{\partial E_{\text{pot}}^{\text{inter}}(1 \text{ urea}^*)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (7)$$

This expression is an approximation when using MD simulations that generate an  $(N, p, T)$  ensemble. Because in MD simulations the three degrees of freedom of the center of mass of the simulated system are decoupled from the rest of the degrees of freedom of the system and their motion does not participate in thermal equilibration, these three degrees of freedom are not properly sampled in the simulation and must be separately integrated in the expression for the ensemble average of a given

quantity. This integration yields the volume  $V$  of the computational box. Therefore, strictly speaking, when a constant pressure simulation is used to evaluate the ensemble average  $\langle \dots \rangle_{\lambda}$  in eq 7, the fluctuations in the volume  $V$  should be taken into account.<sup>37</sup> However, because the effect is negligible for almost incompressible systems with small volume fluctuations, we have omitted the dependence of  $V$  on time when eq 7 is used.

Expression 7 corresponds to the process of the transfer of a designated urea molecule from the solution to the gas phase: a solution of  $n_u - 1$  urea molecules plus one designated urea molecule (indicated by the symbol  $*$ ) and  $n_w$  water molecules is changed into a solution of  $n_u - 1$  urea molecules and  $n_w$  water molecules, and one urea molecule in the gas phase. In terms of free enthalpies  $G$  and chemical potentials  $\mu$  we can write

$$\Delta G^* = G(\text{sol}(n_u - 1, n_w)) + \mu_u(g, p=RT/V) - G(\text{sol}(n_u - 1, 1 \text{ urea}^*, n_w)) \quad (8)$$

Here the first term on the right-hand side is the free enthalpy of a mixture of  $n_u - 1$  urea molecules and  $n_w$  water molecules, the last term is the free enthalpy of a mixture of  $n_u - 1$  urea molecules plus one designated or tagged urea molecule and  $n_w$  water molecules, and the middle term is the gas-phase chemical potential of urea at pressure  $p$ . The relation between  $G(\text{sol}(n_u - 1, 1 \text{ urea}^*, n_w))$  and  $G(\text{sol}(n_u, n_w))$  can be found by considering the corresponding partition functions,

$$Z_{NpT}(n_u - 1, 1 \text{ urea}^*, n_w) = \frac{1}{(n_u - 1)! n_w!} Q_{NpT} \quad (9)$$

and

$$Z_{NpT}(n_u, n_w) = \frac{1}{n_u! n_w!} Q_{NpT} \quad (10)$$

where  $Q_{NpT}$  is the partition function for distinguishable particles. The combinatorial factors in eqs 9 and 10 account for the indistinguishability of the  $n_w$  water molecules and of the  $n_u - 1$  or  $n_u$  urea molecules. Eliminating  $Q_{NpT}$  from eqs 9 and 10, we find

$$Z_{NpT}(n_u - 1, 1 \text{ urea}^*, n_w) = n_u Z_{NpT}(n_u, n_w) \quad (11)$$

or in terms of free enthalpies

$$\begin{aligned} G(\text{sol}(n_u - 1, 1 \text{ urea}^*, n_w)) &= -RT \ln Z_{NpT}(n_u - 1, 1 \text{ urea}^*, n_w) \\ &= -RT \ln Z_{NpT}(n_u, n_w) - RT \ln n_u \\ &= G(\text{sol}(n_u, n_w)) - RT \ln n_u \end{aligned} \quad (12)$$

The process of designating or tagging one urea molecule corresponds to a free enthalpy change of  $-RT \ln n_u$ , which is a purely entropic term. Using the relation between  $G$  and  $\mu$ ,

$$G(\text{sol}(n_u, n_w)) = n_u \mu_u \left( \frac{n_u}{n_u + n_w} \right) + n_w \mu_w \left( \frac{n_w}{n_u + n_w} \right) \quad (13)$$

and inserting (12) into (8) we obtain

$$\Delta G^* = (n_u - 1)\mu_u\left(\frac{n_u - 1}{n_u - 1 + n_w}\right) + n_w\mu_w\left(\frac{n_w}{n_u - 1 + n_w}\right) - n_u\mu_u\left(\frac{n_u}{n_u + n_w}\right) - n_w\mu_w\left(\frac{n_w}{n_u + n_w}\right) + \mu_u(g, p=RT/V) + RT \ln n_u \quad (14)$$

This expression can be further simplified by assuming that the solution behaves ideally, i.e.

$$\mu_u(x) = \mu_u^{\text{id}}(x) = \mu_u^0 + RT \ln x \quad (15)$$

and

$$\mu_w(x) = \mu_w^{\text{id}}(x) = \mu_w^0 + RT \ln x \quad (16)$$

A correction for nonideal behavior is derived in the Appendix. Substituting (15) and (16) in eq 14 yields

$$\Delta G^* = RT \left\{ (n_u - 1) \ln \left( \frac{n_u - 1}{n_u - 1 + n_w} \right) + n_w \ln \left( \frac{n_u + n_w}{n_u - 1 + n_w} \right) - n_u \ln \left( \frac{n_u}{n_u + n_w} \right) + \ln n_u \right\} - \mu_u^0 + \mu_u(g, p=RT/V) \quad (17)$$

Because  $\mu_u^0$  is unknown, we must connect it to the gas phase  $\mu_u(g, p=RT/V)$ . Urea gas with vapor pressure  $p_{\text{sat}}$  in equilibrium with crystalline urea has the same chemical potential as urea in a saturated solution,

$$\mu_u(g, p=p_{\text{sat}}) = \mu_u(x=x_{\text{sat}}) \quad (18)$$

or using (15)

$$\mu_u(g, p=p_{\text{sat}}) = \mu_u^0 + RT \ln x_{\text{sat}} \quad (19)$$

The dependence of  $\mu_u$  on  $p$  is given by

$$\frac{p}{p_{\text{sat}}} = \frac{e^{-\mu_u(g, p)/RT}}{e^{-\mu_u(g, p_{\text{sat}})/RT}} \quad (20)$$

Using  $p = RT/V$ , we find from (20)

$$\mu_u(g, p=RT/V) = \mu_u(g, p=p_{\text{sat}}) + RT \ln \left( \frac{RT}{p_{\text{sat}} V} \right) \quad (21)$$

Using (19) in (21) we find

$$\mu_u(g, p=RT/V) = \mu_u^0 + RT \ln \left( \frac{x_{\text{sat}} RT}{p_{\text{sat}} V} \right) \quad (22)$$

Inserting (22) in (17) we find our final expression for  $\Delta G^*$ ,

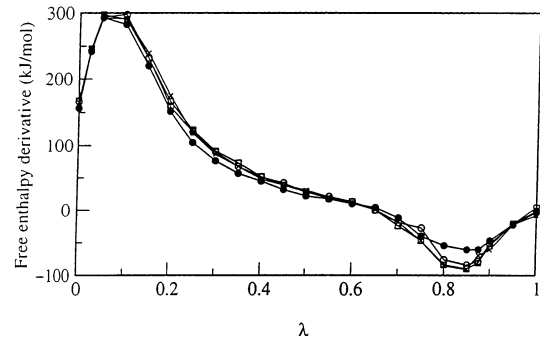
$$\Delta G^* = RT \left\{ (n_u - 1) \ln \left( \frac{n_u - 1}{n_u - 1 + n_w} \right) - n_u \ln \left( \frac{n_u}{n_u + n_w} \right) + n_w \ln \left( \frac{n_u + n_w}{n_u - 1 + n_w} \right) + \ln(n_u) + \ln \left( \frac{x_{\text{sat}} RT}{p_{\text{sat}} V} \right) \right\} \quad (23)$$

To use this expression to obtain an experimental value for  $\Delta G^*$ , we need the mole fraction  $x_{\text{sat}}$  of the saturated solution and the vapor pressure  $p_{\text{sat}}$  of crystalline urea. The *Encyclopedia of Chemical Technology*,<sup>38</sup> gives the solubility of urea for various

**TABLE 5: Thermodynamic Properties, Free Enthalpy of Desolvating One Urea Molecule<sup>a</sup>**

mixture name	mole fraction $x$	free enthalpy simulation	$\Delta G^*$ (kJ mol <sup>-1</sup> ) experiment	
			ideal	corrected
B	0.001517	58.5	60.9	60.1 ± 2.5
C	0.018519	56.7	58.4	57.8 ± 2.5
D	0.153846	60.0	58.1	58.0 ± 2.5
E	0.250000	60.1	57.9	57.9 ± 2.5

<sup>a</sup> The simulated values have been obtained through thermodynamic integration, eq 7. The precision of the calculated values is about 1 kJ mol<sup>-1</sup> as inferred from comparing the values obtained when different urea molecules are evaporated. The experimental value indicated as ideal has been obtained through eq 23, assuming ideal behavior of the solution, eqs 15 and 16. The corrected value includes a correction, eq 29, for nonideal behavior of the solution; see Figure 4.



**Figure 2.** Value of  $\langle \partial E_{\text{pot}}^{\text{inter}}(1 \text{ urea}^*) / \partial \lambda \rangle_{\lambda}$  in kJ mol<sup>-1</sup> as function of the thermodynamic integration coupling parameter  $\lambda$ . Results for the desolvation of the single urea molecule of system B at urea mole fraction  $x = 0.001517$  are indicated by the solid circles. Results for the desolvation of three different urea molecules of system E at urea mole fraction  $x = 0.25$  are indicated by the open circles, open squares, and crosses. The ensemble averages were obtained from 100 ps periods after 5 ps equilibration.

temperatures. Interpolation gives  $x_{\text{sat}} = 0.262$  for 298 K. The accuracy of this number is about 0.01 if temperatures of 298 and 300 K are not distinguished. For the saturated vapor pressure  $p_{\text{sat}}$  there is some confusion in the literature, as measurements are mostly done at higher temperatures and values are extrapolated downward. De Wit et al.<sup>39</sup> give the following equation

$$\ln(p_{\text{sat}} \text{ in Pa}) = -\frac{\Delta G_m}{R\Theta} + \frac{\Delta H_m}{R} \left[ \frac{1}{\Theta} - \frac{1}{T} \right] \quad (24)$$

with  $\Theta = 351.32$  K,  $\Delta G_m = 2.676$  kJ mol<sup>-1</sup>, and  $\Delta H_m = 96.9$  kJ mol<sup>-1</sup>, yielding  $p_{\text{sat}} = 1.0 \times 10^{-3}$  Pa at 298 K. Ferro et al.<sup>36</sup> give the equation

$$\log_{10}(p_{\text{sat}} \text{ in kPa}) = (10.30 \pm 0.30) - \frac{4750 \pm 100}{(T \text{ in K})} \quad (25)$$

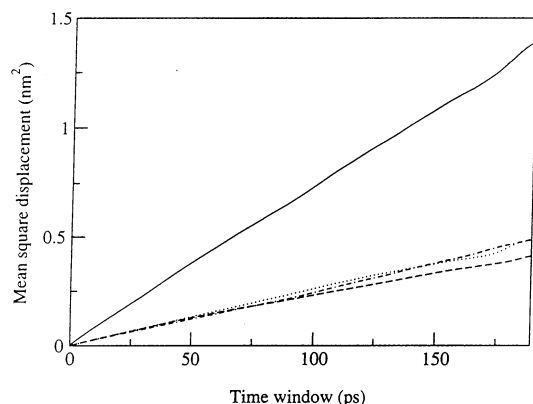
which yields  $p_{\text{sat}} = 2.3 \times 10^{-3}$  Pa at 298 K with an error of a factor of 3. We shall use the value  $p_{\text{sat}} = 2 \times 10^{-3}$  Pa in eq 23. The uncertainty in this value causes a dominating inaccuracy of  $\pm 2.5$  kJ mol<sup>-1</sup> in the experimental value for  $\Delta G^*$ . Using the mentioned values for  $x_{\text{sat}}$  and  $p_{\text{sat}}$  in (23) the ideal experimental values for  $\Delta G^*$  listed in Table 5 were obtained. To these values a nonideality correction must be added, which is worked out in the Appendix.

The simulated values for  $\Delta G^*$  as given in Table 5 were obtained by thermodynamic integration using eq 7. In Figure 2 the values of the integrand of eq 7 are shown as a function of  $\lambda$  for systems B and E. The three curves for system E represent

**TABLE 6: Dynamic Properties, Self-Diffusion of Urea<sup>a</sup>**

mixture name	mole fraction $x$	diffusion constant ( $10^{-3} \text{ nm}^2 \text{ ps}^{-1}$ )	
		simulation	experiment <sup>40</sup>
B	0.001517	3.0	1.37
C	0.018519	2.1	1.31
D	0.153846	1.7	
E	0.250000	1.2	

<sup>a</sup> The simulated values have been obtained using an Einstein relation by averaging over urea molecules.



**Figure 3.** Mean-square displacement (solid line) of urea molecules for system E at urea mole fraction  $x = 0.25$  as a function of time. The values are averages over  $n_u = 432$  urea molecules. The dotted, dashed and dot-dashed lines represent the  $x$ ,  $y$ , and  $z$  components of the mean-square displacement.

calculations in which the interactions of three different urea molecules were switched off. The simulated  $\Delta G^*$  values listed in Table 5 do agree within experimental error with the experimental values.

**D. Diffusion.** The diffusion constant of urea as function of mole fraction  $x$  is given in Table 6. The values have been obtained from the slope of the mean-square displacement of urea molecules as a function of time; see Figure 3. The agreement between simulation and experiment<sup>40</sup> improves with urea concentration. For very dilute solutions the simulated values are too high. This is due to the too high diffusivity of the SPC water model. Its diffusion constant is  $4.2 \times 10^{-3} \text{ nm}^2 \text{ ps}^{-1}$  at 301 K to be compared with an experimental value of  $2.3 \times 10^{-3} \text{ nm}^2 \text{ ps}^{-1}$  at 298 K.

#### IV. Discussion

The results from the simulations agree surprisingly well with experimental values. This applies to density and energy as well as to free enthalpies. The small density discrepancy is completely due to the density of SPC water, which is too low. The potential energy of the simulated saturated solution is slightly too low. The free enthalpies of urea solvation are correct but can only be judged within the uncertainty of the experimental values, which is larger than the uncertainty of the simulated values. The experimental uncertainty is dominated by the inaccuracy of the experimental saturated vapor pressure of urea crystals.

The results lend confidence to the practical use of the evaluated force field parameters in biomolecular simulations and their incorporation into the GROMOS96 force field. At least the interaction of urea with water is reliable and apparently sufficiently accurate even in the absence of explicit polarizability terms in the force field. The interaction of urea with protein atoms, for example, in protein denaturation studies, is equally

important for the simulation of protein behavior in urea solutions but can unfortunately not be validated with similar precision.

Nonideality corrections to the free enthalpy of urea solution are small and fall within the error of the simulations. This shows that deviations from ideality cannot be studied accurately by simulation, at least for urea–water mixtures, which do not deviate strongly from ideality.

**Acknowledgment.** Financial support was obtained through the National Center of Competence in Research (NCCR) in Structural Biology of the Swiss National Science Foundation, which is gratefully acknowledged. L.J.S. thanks the Oxford Supercomputing Center for access to its facilities and INTAS for financial support (project 01-2126).

#### Appendix. Nonideality Corrections

For the evaluation of the chemical potentials of urea and water in solution, we have used (15) and (16), valid for an ideal mixture. For the almost saturated solution at  $x = 0.25$  the chemical potential is given by the equilibrium with solid state and saturated vapor, and no corrections to the free enthalpy  $\Delta G^*$  of desolvation of one urea molecule from the solution are expected for system E. For the dilute system B, however, a nonideality correction may be required.

The real chemical potential of the solute urea cannot be determined directly, e.g., from the urea vapor pressure but can be derived from the water activity by means of the Gibbs–Duhem relation. The water activity is given by the osmotic pressure, e.g., expressed by the molal osmotic coefficient  $\Phi(m)$ , which has been discussed by Gucker and Pickard,<sup>41</sup> or by the osmolality  $O = m\Phi(m)$  listed in the *Handbook of Chemistry and Physics*.<sup>42</sup> The molality  $m$  of the solution is defined as

$$m = \frac{x}{(1-x)M_w} \quad (26)$$

in which  $x$  is the mole fraction urea and  $M_w$  is the molecular weight of water ( $18.0154 \text{ g mol}^{-1}$ ). The excess chemical potential of water above the ideal value is given by

$$\mu_w^{\text{ex}}(x) = -RT\{M_w m \Phi(m) + \ln(1-x)\} \quad (27)$$

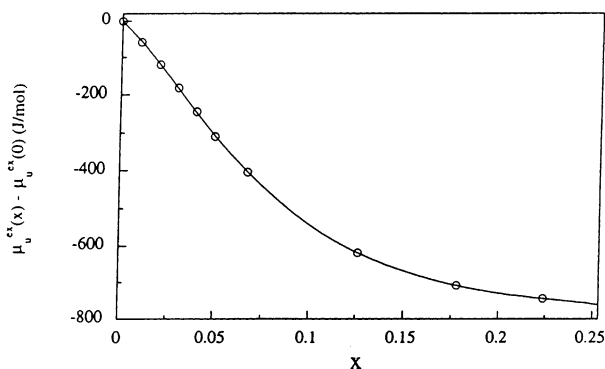
From the Gibbs–Duhem relation

$$x \frac{\partial \mu_u}{\partial x} + (1-x) \frac{\partial \mu_w}{\partial x} = 0 \quad (28)$$

which is also valid for the excess chemical potentials, we find through integration from zero to  $x$  that

$$\begin{aligned} \mu_u^{\text{ex}}(x) &= \mu_u^{\text{ex}}(0) + \int_0^x \left[1 - \frac{1}{x'}\right] \frac{\partial \mu_w^{\text{ex}}}{\partial x'} dx' \\ &= \mu_u^{\text{ex}}(0) + \left[1 - \frac{1}{x}\right] \mu_w^{\text{ex}}(x) - \int_0^x \frac{\mu_w^{\text{ex}}(x')}{(x')^2} dx' \end{aligned} \quad (29)$$

From the values of  $\Phi(m)$  given in Figure 5 of Gucker and Pickard,<sup>41</sup> we used the following smoothed values:  $\Phi(m) = (1 - 0.03m)$  for  $m < 4$  molal, and  $\Phi(4) = 0.888$ ,  $\Phi(8) = 0.834$ ,  $\Phi(12) = 0.810$ , and  $\Phi(16) = 0.795$ . Using these values for  $\Phi(m)$  in eqs 27 and 29 yields the excess chemical potential of urea  $\mu_u^{\text{ex}}(x)$  as shown in Figure 4. We see that  $\mu_u^{\text{ex}}(0)$  at zero concentration is  $760 \text{ J mol}^{-1}$  higher than computed from ideality, taking the value for a saturated solution as reference. The



**Figure 4.** Experimental excess chemical potential  $\mu_{\text{pot}}^{\text{ex}}(x) - \mu_u^{\text{ex}}(0)$  for urea (in  $\text{kJ mol}^{-1}$ ) as a function of the urea mole fraction  $x$  as obtained from eqs 29 and 27 with the experimental values for the osmolality  $m\Phi(m)$  taken from Gucker and Pickard.<sup>41</sup> The value for  $\mu_u^{\text{ex}}(x_{\text{sat}})$  of the saturated solution is  $760 \text{ J mol}^{-1}$  lower than the value of  $\mu_u^{\text{ex}}(x = 0)$  at infinite dilution.

corrections to  $\Delta G^*$  for nonideality behavior as obtained from Figure 4 are  $-0.8$ ,  $-0.6$ ,  $-0.1$ , and  $0.0 \text{ kJ mol}^{-1}$  for systems B–E, respectively. The experimental values for  $\Delta G^*$ , the free enthalpy of desolvating one urea molecule from the solution, that include the given corrections for nonideality are given in the last column of Table 5.

## References and Notes

- (1) Tirado-Rives, J.; Orozco, M.; Jorgensen, W. L. *Biochemistry* **1997**, *36*, 7313–7329.
- (2) Caffisch, A.; Karplus, M. *Structure* **1999**, *7*, 477–488.
- (3) Zhang, Z. Y.; Zhu, Y. J.; Shi, Y. Y. *Biophys. Chem.* **2001**, *89*, 145–162.
- (4) Wallqvist, A.; Covell, D. G.; Thirumalai, D. *J. Am. Chem. Soc.* **1998**, *120*, 427–428.
- (5) Ikeguchi, M.; Nakamura, S.; Shimizu, K. *J. Am. Chem. Soc.* **2001**, *123*, 677–682.
- (6) Duffy, E. M.; Kowalczyk, P. J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1993**, *115*, 9271–9275.
- (7) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331–342.
- (8) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (9) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (10) Glättli, A.; Daura, X.; van Gunsteren, W. F. *J. Chem. Phys.* **2002**, *116*, 9811–9828.
- (11) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *Isr. J. Chem.* **1993**, *33*, 323–330.
- (12) Boek, E. S.; Briels, W. J.; van Eerden, J.; Feil, D. *J. Chem. Phys.* **1992**, *96*, 7010–7018.
- (13) Boek, E. S.; Briels, W. J. *J. Chem. Phys.* **1993**, *98*, 1422–1427.
- (14) Hernandez-Cobos, J.; Ortega-Blake, I.; Bonilla-Marin, M.; Moreno-Ballo, M. *J. Chem. Phys.* **1993**, *99*, 9122–9134.
- (15) Astrand, P.-O.; Wallqvist, A.; Karlström, G. *J. Chem. Phys.* **1994**, *100*, 1262–1273.
- (16) Astrand, P.-O.; Wallqvist, A.; Karlström, G. *J. Chem. Phys.* **1994**, *98*, 8224–8233.
- (17) Tsai, J.; Gerstein, M.; Levitt, M. *J. Chem. Phys.* **1996**, *104*, 9417–9430.
- (18) Levitt, M.; Hirshberg, M.; Sharon, R.; Daggett, V. *Comput. Phys. Commun.* **1995**, *91*, 215–231.
- (19) Levitt, M.; Hirshberg, M.; Sharon, R.; Laidig, K. E.; Daggett, V. *J. Phys. Chem. B* **1997**, *101*, 5051–5061.
- (20) Vanzi, F.; Madan, B.; Sharp, K. *J. Am. Chem. Soc.* **1998**, *120*, 10748–10753.
- (21) Idrissi, A.; Sokolic, F. *J. Chem. Phys.* **2000**, *112*, 9479–9488.
- (22) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2000**, *104*, 5854–5864.
- (23) Cirino, J. J. V.; Bertran, C. A. *Quim. Nova* **2002**, *25*, 358–363.
- (24) Bertran, C. A.; Cirino, J. J. V.; Freitas, L. C. G. *J. Braz. Chem. Soc.* **2002**, *13*, 238–244.
- (25) Dias, L. G.; Florenzano, F. H.; Reed, W. F.; Baptista, M. S.; Souza, S. M. B.; Alvarez, E. B.; Chaimovich, N.; Cuccovia, I. M.; Amaral, C. L. C.; Brasil, C. R.; Romsted, L. S.; Politi, M. *J. Langmuir* **2002**, *18*, 319–324.
- (26) Sokolic, F.; Idrissi, A.; Perera, A. *J. Chem. Phys.* **2002**, *116*, 1636–1646.
- (27) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, 1996.
- (28) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535–547.
- (29) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
- (30) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (31) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (32) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (33) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169–3174.
- (34) *Handbook of Chemistry and Physics*, 81st ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, 2000–2001; pp 8–83.
- (35) Kawahara, K.; Tanford, C. *J. Biol. Chem.* **1966**, *241*, 3228–3232.
- (36) Ferro, D.; Barone, G.; Della Gatta, G.; Piacente, V. *J. Chem. Thermodyn.* **1987**, *19*, 915–923.
- (37) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: New York, 1996.
- (38) *Encyclopedia of Chemical Technology*, 4th ed.; Wiley: New York, 1998.
- (39) de Wit, H. G. M.; van Miltenburg, J. C.; de Kruijff, C. G. *J. Chem. Thermodyn.* **1983**, *15*, 651–663.
- (40) Gostling, L. J.; Akeley, D. F. *J. Am. Chem. Soc.* **1952**, *74*, 2058–2060.
- (41) Gucker, F. T.; Pickard, H. B. *J. Am. Chem. Soc.* **1940**, *62*, 1464–1472.
- (42) *Handbook of Chemistry and Physics*, 56th ed.; Weast, R. C., Ed.; CRC Press: Boca Raton, FL, 1976.