

Interaction of Urea with Amino Acids: Implications for Urea-Induced Protein Denaturation

Martin C. Stumpe and Helmut Grubmüller*

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

Received August 17, 2007; E-mail: hgrubmu@gwdg.de

Abstract: The molecular mechanism of urea-induced protein denaturation is not yet fully understood. Mainly two opposing mechanisms are controversially discussed, according to which either hydrophobic, or polar interactions are the dominant driving force. To resolve this question, we have investigated the interactions between urea and all 20 amino acids by comprehensive molecular dynamics simulations of 22 tripeptides. Calculation of atomic contact frequencies between the amino acids and solvent molecules revealed a clear profile of solvation preferences by either water or urea. Almost all amino acids showed preference for contacts with urea molecules, whereas charged and polar amino acids were found to have slight preferences for contact with water molecules. Particularly strong preference for contacts to urea were seen for aromatic and apolar side-chains, as well as for the protein backbone of all amino acids. Further, protein–urea hydrogen bonds were found to be significantly weaker than protein–water or water–water hydrogen bonds. Our results suggest that hydrophobic interactions are the dominant driving force, while hydrogen bonds between urea and the protein backbone contribute markedly to the overall energetics by avoiding unfavorable unsatisfied hydrogen bond sites on the backbone. In summary, we suggest a combined mechanism that unifies the two current and seemingly opposing views.

1. Introduction

Urea is a widely used protein denaturant. Despite its widespread use, however, the molecular mechanism underlying urea-induced denaturation is not well understood. Two classes of interaction models are distinguished in the literature. In the first, direct interactions between urea and the protein are considered the main denaturation driving force.^{1–5} In the second, urea-induced changes in the water structure are suggested as indirect interactions that drive unfolding.^{6–8} While several recent studies support the direct interaction model,^{5,9–13} it is still unclear whether polar or apolar residues or the peptide backbone constitute the main interaction sites for urea. That the peptide backbone is an important interaction site for urea is now widely

accepted.^{14–16} However, some studies^{4,10,11,13,17,18} stress the importance of urea–protein hydrogen bonds to polar residues. Other studies^{12,19–28} support the importance of apolar urea–protein contacts weakening the hydrophobic effect. Hence, more detailed insights into the interactions of a denaturant with amino acids is imperative to understand how denaturants work.

This study aims to elucidate and quantify by extended molecular dynamics simulations the interactions of urea with each of the natural 20 amino acids. To this aim, interaction frequencies between urea and the individual amino acids are investigated to decide whether urea interacts preferentially with polar or apolar residues or with the peptide backbone. To quantify residue interaction with urea, a contact coefficient C_{UW}

- (1) Astrand, P.-O.; Wallqvist, A. and Karlstrom, G. *J. Phys. Chem.* **1994**, *98*, 8224–8233.
- (2) Tirado-Rives, J.; Orozco, M.; Jorgensen, W. L. *Biochemistry* **1997**, *36* (24), 7313–7329.
- (3) Grdadolnik J.; Maréchal, Y. *J. Mol. Struct.* **2002**, *615* (1–3), 177–189.
- (4) Mountain R. D.; Thirumalai, D. *J. Am. Chem. Soc.* **2003**, *125* (7), 1950–1957.
- (5) Klimov, D. K.; Straub, J. E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (41), 14760–14765.
- (6) Frank, H. S.; Franks, F. *J. Chem. Phys.* **1968**, *48* (10), 4746–4757.
- (7) Finer, E. G.; Franks, F.; Tait, M. J. *J. Am. Chem. Soc.* **1972**, *94* (13), 4424–4429.
- (8) Hoccart, X.; Turrell, G. *J. Chem. Phys.* **1993**, *99* (11), 8498–8503.
- (9) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2002**, *106*, 1491–1500.
- (10) Caballero-Herrera, A.; Nordstrand, K.; Berndt, K. D.; Nilsson, L. *Biophys. J.* **2005**, *89* (8), 842–857.
- (11) Oostenbrink, C.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2005**, *7* (1), 53–58.
- (12) Lee, M.-E.; van der Vegt, N. F. A. *J. Am. Chem. Soc.* **2006**, *128* (15), 4948–4949.
- (13) O'Brien, E. P.; Dima, R. I.; Brooks, B.; Thirumalai, D. *J. Am. Chem. Soc.* **2007**, *129* (23), 7346–7353.

- (14) Courtenay, E. S.; Capp, M. W.; Record, T. M., Jr. *Protein Sci.* **2001**, *10*, 2485–2497.
- (15) Auton, M.; Bolen, W. D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (42), 15065–15068.
- (16) Möglich, A.; Krieger, F.; Kiefhaber, T. *J. Mol. Biol.* **2005**, *345* (1), 153–162.
- (17) Makhataadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1992**, *226* (2), 491–505.
- (18) Bennion, B. J.; Daggett, V. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (9), 5142–5147.
- (19) Nozaki, Y.; Tanford, C. *J. Biol. Chem.* **1963**, *238* (12), 4074–4081.
- (20) Tanford, C. *Adv. Protein Chem.* **1970**, *24*, 1–95.
- (21) Prakash, V.; Loucheux, S.; Scheuffele, C.; Gorbunoff, M. J.; Timasheff, S. N. *Archives Biochem. Biophys.* **1981**, *210* (2), 455–464.
- (22) Müller, N. *J. Phys. Chem.* **1990**, *3856–3859* (94), 3856–3859.
- (23) Alonso, D. O. V.; Dill, K. A. *Biochemistry* **1991**, *30* (24), 5974–5985.
- (24) Duffy, E. M.; Kowalczyk, P. J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1993**, *115* (20), 9271–9275.
- (25) Tsai, J.; Gerstein, M.; Levitt, M. *J. Chem. Phys.* **1996**, *104* (23), 9417–9430.
- (26) Zou, Q.; Habermann-Rottinghaus, S. M.; Murphy, K. P. *Proteins: Structure, Function, Bioinform.* **1998**, *31* (2), 107–115.
- (27) Ikeguchi, M.; Nakamura, S.; Shimizu, K. *J. Am. Chem. Soc.* **2001**, *123* (4), 677–682.
- (28) Timasheff S. N.; Xie, G. *Biophys. Chem.* **2003**, *105*, 421–448.

is introduced as a measure for preferential interaction with urea relative to that with water. The C_{UW} analysis will also provide detailed insights into urea–peptide interactions on the atomic level. Additionally, the role of hydrogen bonds between urea or water and the peptide residues is investigated and hydrogen bond energies are estimated.

To separate sequence dependence and secondary or tertiary structure effects from the immediate interaction between urea and the respective amino acids, all 20 amino acids were investigated by simulations of glycine-capped tripeptides (GXG). The influence of sequence and structure on the immediate or direct interactions of the amino acids are discussed in the Conclusions.

2. Methods

Simulation Setup. Each of the 20 natural amino acids (X) was simulated in a glycine-capped tripeptide (GXG). The glycine termini with NH_2 and COOH were kept uncharged. For histidine, all three protonation states were considered, resulting in 22 simulations in total, each of 100 ns length. The initial backbone angles of the tripeptides were set in β -sheet conformation ($\phi = -135^\circ$, $\psi = 128^\circ$). Since the autocorrelation time of (ϕ , ψ) was found to be below 1 ns in our simulations, these starting conditions did not impose significant bias. All tripeptides were simulated individually in aqueous urea solution with 1250 water molecules and 250 urea molecules, corresponding to a urea mole fraction of 0.17 and a concentration of 6.5 M. An appropriate counterion (Na^+ or Cl^-) was added for each charged amino acid. Ion concentration might affect the results.²⁹ Here, however, we are interested in the interactions without salt effects, and therefore, only one counterion was added where necessary to obtain an electrically neutral simulation system. Additionally, a second set of simulations was performed with all electrostatic interactions switched off to estimate steric effects on the calculated contact coefficients, which may arise from different volumes of urea and water molecules.

All simulations were performed using the Gromacs^{30,31} program package, version 3.3, with the OPLS-all-atom force field,³² the TIP4P water model,³³ and the urea model of Smith et al.³⁴ Particle Mesh Ewald summation (PME) was used to calculate the long-range electrostatic interactions with a grid-spacing of 0.12 nm and an interpolation order of 4. A cutoff of 1.0 nm was used for the short-range Coulomb and the Lennard–Jones interactions. All simulations were performed in the NpT -ensemble using Berendsen-type temperature-coupling³⁵ with a coupling coefficient of $\tau_T = 0.1$ ps and Berendsen-type pressure-coupling³⁵ at 1 bar with a coupling coefficient of $\tau_p = 1$ ps. An integration time step of 2 fs was chosen. The initial size of the periodic cubic box was set to $(3.9 \text{ nm})^3$ to accommodate 1250 water and 250 urea molecules in addition to the tripeptide.

To setup the simulation systems, 250 (non-overlapping) urea molecules were placed at random positions in the simulation box containing the tripeptide. Subsequently, the box was filled up with 1250 water molecules. Each simulation was preceded by a 200-step steepest-descent energy minimization, 500 ps equilibration with position restraints on the tripeptide, and finally a 1 ns equilibration without position restraints. The total simulation time for data collecting was 4.4 μs .

Contact Coefficient. To quantify the frequency of interactions between urea and the amino acids, we define the contact coefficient C_{UW} for a particular amino acid X:

$$C_{UW_x} = \frac{N_{X-U} M_W}{N_{X-W} M_U} \quad (1)$$

where N_{X-U} and N_{X-W} are the numbers of atomic contacts of amino acid X with urea and water molecules, respectively. Atoms were defined to be in contact if they are closer than 0.35 nm. C_{UW} is normalized using the total numbers of urea atoms (M_U) and water atoms (M_W). Hence, a residue with a contact coefficient of $C_{UW} = 1.0$ has no interaction preference for either urea or water. Values above 1.0 indicate preferential interaction with urea; values below 1.0 indicate preferential interaction with water. Since C_{UW} relates to interaction frequencies, it can be regarded as a measure for the free energy of contact formation which gives rise to the first peak in the respective radial distribution function.

The autocorrelation time of the instantaneous contact coefficient (determined from single snapshots) was found to be about 100 ps for the analysis on the residue level and about 10 ps for the analysis on the atomic level. Lifetimes of contacts were distributed exponentially with a similar time-constant. The correlation time was used to calculate the number of statistically independent frames within the 100 ns simulation time for the statistical error estimate of the average contact coefficients.

Note that the contact coefficient can easily be extended to quantify interaction preferences of solute molecules X in solvents consisting of more than two components S_i :

$$C_{S_i} = \frac{N_{X-S_i} \sum_{j=1}^k M_{S_j}}{M_{S_i} \sum_{j=1}^k N_{X-S_j}} \quad (2)$$

where k is the number of solvent components, N_{X-S_i} is the number of atomic contacts between the solute X and the solvent molecules S_i , and M_{S_i} denotes the total number of atoms of all solvent molecules i .

Hydrogen Bonds. The number of hydrogen bonds per molecule was calculated using the standard Gromacs tool `g_hbond` with a cutoff radius of 0.35 nm between donor and acceptor and a cutoff angle of 30° as geometric criteria for the existence of a hydrogen bond.³⁶ Energies of hydrogen bonds were estimated using the empirical function,³⁷ where d denotes the distance between hydrogen and acceptor atom in nm.

The estimated energies are for isolated hydrogen bonds and are certainly not identical with the free energy contribution of these hydrogen bonds for proteins in solution.³⁸ We note that we here use this formula only for a (semiquantitative) measure of the hydrogen bond strength.^{39–43} In particular, our conclusions are only based on the monotonic dependence of hydrogen bond energy on distance (in the considered distance range) and not on accurate numbers, such that this simple treatment should suffice.

- (29) Donnini, S.; Mark, A. E.; Juffer, A. H.; Villa, A. *J. Comp. Chem.* **2005**, *26* (2), 115–122.
 (30) Berendsen, H. J.; van der Spoel, C. D.; van Drunen, R. *Comp. Phys. Comm.* **1995**, *91*, 43–56.
 (31) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
 (32) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
 (33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
 (34) Smith, L. J.; Berendsen, H. J. C.; van Gunsteren, H. F. *J. Phys. Chem. B* **2004**, *108* (3), 1065–1071.
 (35) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.

- (36) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; and Berendsen, H. J. C. www.gromacs.org 2004.
 (37) Espinosa, E.; Molins, E.; Lecomte, C. *Chem. Phys. Lett.* **1998**, *285* (3–4), 170–173.
 (38) Sheu, S.-Y.; Yang, D.-Y.; Selzle, H. L.; Schlag, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (22), 12683–12687.
 (39) Arnold, W. D.; Eric Oldfield, E. *J. Am. Chem. Soc.* **2000**, *122*, 12835–12841.
 (40) Jenkins, S.; Morrison, I. *Chem. Phys. Lett.* **2000**, *317*, 97–102.
 (41) Rozenberg, M.; Loewenschuss, A.; Marcus, Y. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2699–2702.
 (42) Alkorta, I.; Elguero, J. *J. Phys. Chem. A* **1999**, *103*, 272–279.
 (43) Gálvez, O.; Gómez, P. C.; Pacios, L. F. *J. Chem. Phys.* **2001**, *115* (24), 11166–11184.

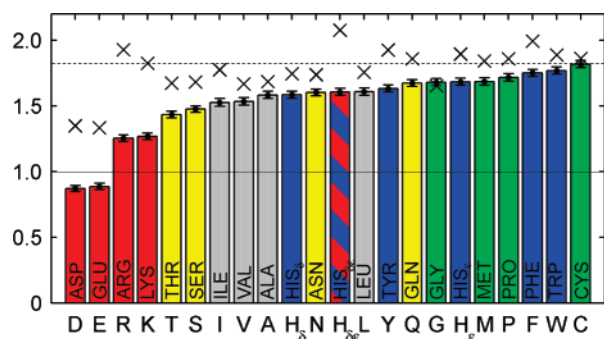


Figure 1. Contact coefficient C_{UW} for each amino acid. High values above 1 indicate preferential interactions with urea; a value of 1 corresponds to equal probability to interact with urea or with water. The color characterizes the amino acids. Red, charged; yellow, polar; gray, aliphatic; blue, aromatic; green, apolar. Crosses denote the C_{UW} of the backbone alone; the dotted line at 1.82 marks the backbone average over all amino acids.

Force Field Energies. Force field energies (Lennard–Jones and Coulomb) between amino acid X and urea (E_{X-U}) or water (E_{X-W}) were calculated for all atoms in atomic contact using the same distance cutoff criterion as for the contact coefficients. Energies per atom were then defined as

$$E_X^{\text{norm}} = \frac{E_{X-W}}{N_X} - \frac{E_{X-U}}{N_U} \quad (3)$$

where N_X and N_U denote the number of atomic contacts of residue X with water or urea, respectively.

3. Results

Contact Coefficients. We first focus on contact coefficients C_{UW} (Figure 1). As can be seen, C_{UW} is higher than 1, indicating preferential contacts to urea, for all amino acids except the anionic ASP and GLU, which have a significantly lower C_{UW} of about 0.89 each. The cationic ARG and LYS have the second lowest C_{UW} with 1.26 each. The other amino acids exhibit a C_{UW} between 1.44 (THR, SER) and 1.82 (CYS). For each amino acid the C_{UW} of the backbone alone is higher than for the complete amino acid. The C_{UW} of the backbone alone, averaged over all residues, is 1.78 ± 0.18 .

In summary, urea interacts mainly with aromatic and nonpolar residues, as well as with the protein backbone. Polar and especially charged residues interact less frequently with urea, the charged amino acids ASP and GLU show even more interactions with water than with urea. Note that $C_{UW} < 1$ does not necessarily imply a positive free energy of transfer from water to urea solution, which has been found to be negative for hydrophobic, as well as for hydrophilic, residues.^{44–47}

To elucidate which parts of the amino acids show contact preferences for either urea or water, we calculated C_{UW} atomwise. This in-depth analysis was further motivated by the difference in average C_{UW} for the backbones and the complete residues. Figure 2 shows atomic interaction sites for urea and water for all amino acids. Blue indicates preferential interaction with water (low C_{UW}), green indicates preferential interaction with urea (high C_{UW}), and white corresponds to no interaction preference ($C_{UW} = 1$).

Again, clear differences in the C_{UW} are seen for the different the amino acids. In particular, the carboxyl groups of ASP and GLU have very low C_{UW} values and represent the main interaction sites for water. For both amino acids, also the side-chain CH_2 groups and even the backbone show reduced interactions with urea due to the charge of the carboxyl group. This effect is slightly more pronounced for ASP than for GLU, since the backbone is separated from the carboxyl group by one CH_2 group less in ASP. ARG and LYS show both, preferential interaction sites for water, as well as for urea. The amino groups are the main interaction sites for water, whereas the backbones exhibit high C_{UW} values and are not significantly affected by the charged amino groups due to the long apolar side-chains of both amino acids. Further interaction sites for water are the hydroxyl group of TYR and the two nitrogen atoms in the $\text{HIS}_{\delta\epsilon}$ -ring.

Pronounced interactions with urea are seen for the aromatic rings of PHE, TRP, and TYR, as well as for the whole CYS, and the C_{β} and C_{γ} atoms of ILE, LEU, and MET. These atoms have been assigned a small charge (between $-0.06e$ and $-0.12e$) in the OPLS-AA force field. The peptide backbone shows preferential interaction with urea for all amino acids.

Overall, we observe for the residue, as well as for the atomic level, that polar parts (with large partial charges) mainly interact with water while less polar parts (with small partial charges) interact mainly with urea.

Hydrogen Bonding. Hydrogen bonds between water, urea, and the amino acids were analyzed, and their strength was estimated via the Espinosa formula.³⁷ The average hydrogen bond energies are given in Figure 2 for all side-chain donors and acceptors. As can be seen, all hydrogen bonds to water (blue numbers) are stronger than the corresponding hydrogen bonds to urea (green numbers). Therefore, hydrogen bond sites on the side-chains are preferentially solvated by water rather than by urea molecules. The strongest hydrogen bonds (≈ 46 kJ/mol) are seen between the carboxyl groups of ASP or GLU and water. Since the energies of hydrogen bonds to the backbone do not significantly vary between the amino acids, their average energies is given in Figure 3 and they are not shown individually in Figure 2. Figure 3 shows the hydrogen bond energies between protein (backbone), water, and urea. The larger numbers in Figure 3a denote the average hydrogen bond energies for all donor/acceptor combinations. The small number pairs next to the arrows in panel (a) refer to hydrogen bonds with donor/acceptor in the direction of the respective arrow. For instance, hydrogen bonds between water and the peptide backbone have an energy of 18.0 kJ/mol with the peptide as donor and water as acceptor and an energy of 29.8 kJ/mol with water as donor and the peptide backbone as acceptor. On average, intrawater hydrogen bonds are the strongest with an energy of 27.8 kJ/mol, followed by hydrogen bonds between water and the peptide backbone with an energy of 25.8 kJ/mol. For urea, the energy difference between hydrogen bonds to protein or water is very small. Judging from the hydrogen bond energies alone, water seems to be a significantly more favorable hydrogen bond partner than urea for the peptide backbone. That this difference in hydrogen bond energies between protein–urea and protein–water does not lead to preferential solvation of the backbone by water, as one might expect, shows that optimization of peptide–solvent hydrogen bonds is not the determinant for

(44) Nandi, P. K.; Robinson, D. R. *Biochemistry* **1984**, *23*, 6661–6668.

(45) Wetlaufer, D. B.; Malik, S. K.; Stoller, L.; Coffin, R. L. *J. Am. Chem. Soc.* **1964**, *86*, 508–514.

(46) Kresheck, G. C.; Benjamin, L. *J. Phys. Chem.* **1964**, *68*, 2476–2486.

(47) Roseman, M.; Jencks, W. P. *J. Am. Chem. Soc.* **1975**, *97*, 631–640.

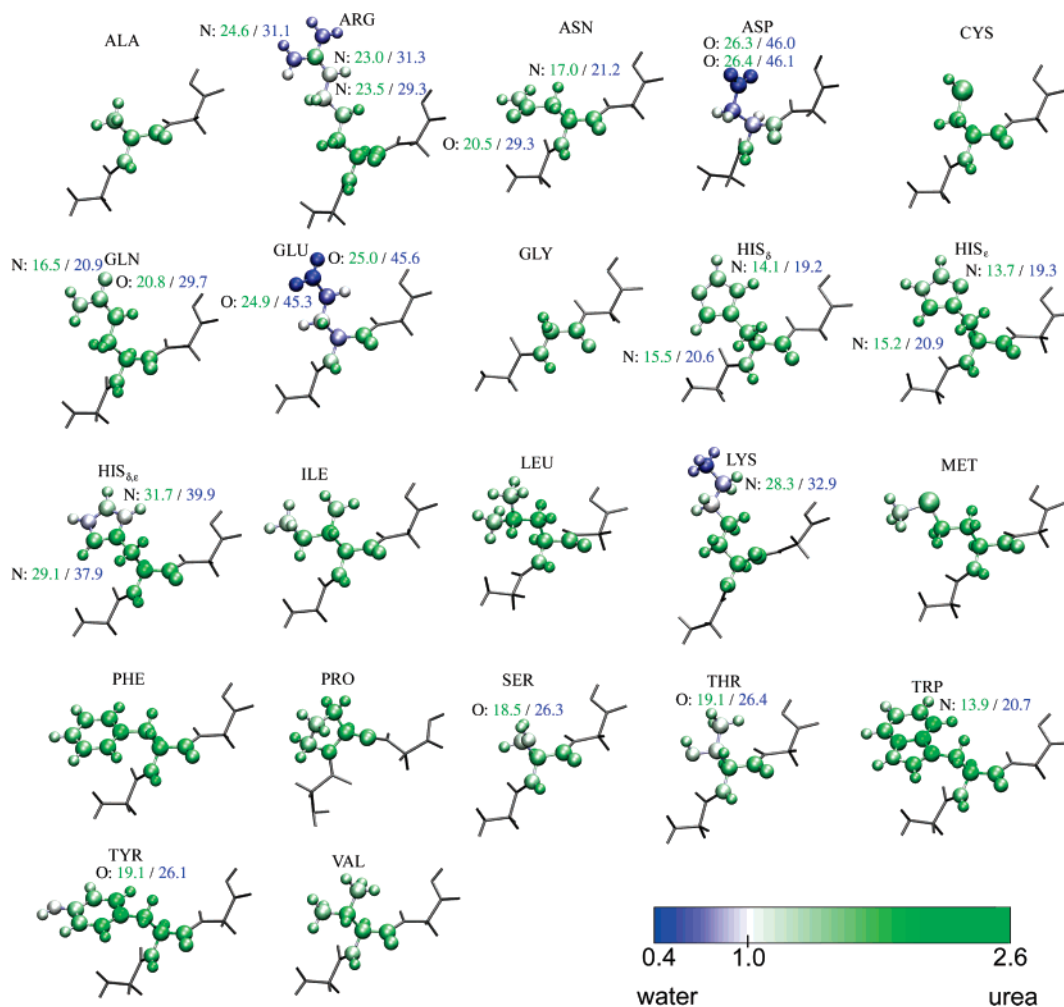


Figure 2. Atomic interaction sites for urea and water. Blue indicates low C_{UW} , green high C_{UW} , and white corresponds to $C_{UW} = 1$. The numbers denote average energies of hydrogen bonds to water (blue) or urea (green) for the respective atoms in kJ/mol.

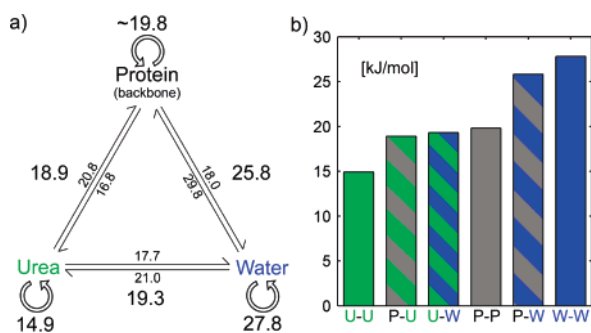


Figure 3. Average hydrogen bond energies (in kJ/mol) between the protein backbone, urea, and water. Statistical errors are below 0.1 kJ/mol. The small number pairs next to the arrows in panel (a) refer to hydrogen bonds with donor/acceptor in the direction of the respective arrow. The large numbers denote the weighted average. The energy of protein-protein hydrogen bonds was taken from simulations with the CI2 protein (to be published elsewhere).

backbone solvation by water or urea in this simple view. This issue will be discussed further below.

Intraprotein hydrogen bond energies were calculated from simulations of the CI2 protein (to be published elsewhere). Statistical errors are below 0.1 kJ/mol. Note that the empirical formula used to estimate the hydrogen bond energies is certainly less accurate. In particular, the energies calculated here are estimates for isolated hydrogen bond energies which are known to be larger than those in solution.³⁸ However, as discussed in

the Methods section, our conclusions are only based on the relative strengths of the hydrogen bonds and not on the absolute numbers, such that this simple treatment will suffice.

Driving Forces for Preferential Interactions. In order to explore the driving forces of the preferential interactions, we calculated the force field energies (Coulomb + Lennard-Jones) between the amino acids X and urea/water (Figure 4a). Charged and polar residues (low C_{UW}) have the largest force field energies with both water and urea. The Lennard-Jones contributions to the potential energies were all positive and much smaller than the respective Coulomb part.

The enthalpy gain upon substituting urea with water in the solvation shell, measured by the energy difference between X-water and X-urea, is largest for these residues. For residues with a high C_{UW} , the enthalpy gain is much smaller.

Figure 4b correlates the potential energy differences E_X^{norm} (eq 3) between X-water and X-urea with the contact coefficients C_{UW} . Indeed, the linear fit indicates a significant correlation (regression coefficient $r^2 = 0.8$). In particular, more negative (i.e., stronger) interaction energies imply low C_{UW} values. From this correlation we conclude that the interaction of water molecules with charged and polar residues (low C_{UW}) is dominated by enthalpy contributions, in particular electrostatic ones. Those residues with only a small enthalpy gain are solvated preferentially by urea (high C_{UW}) because displacement

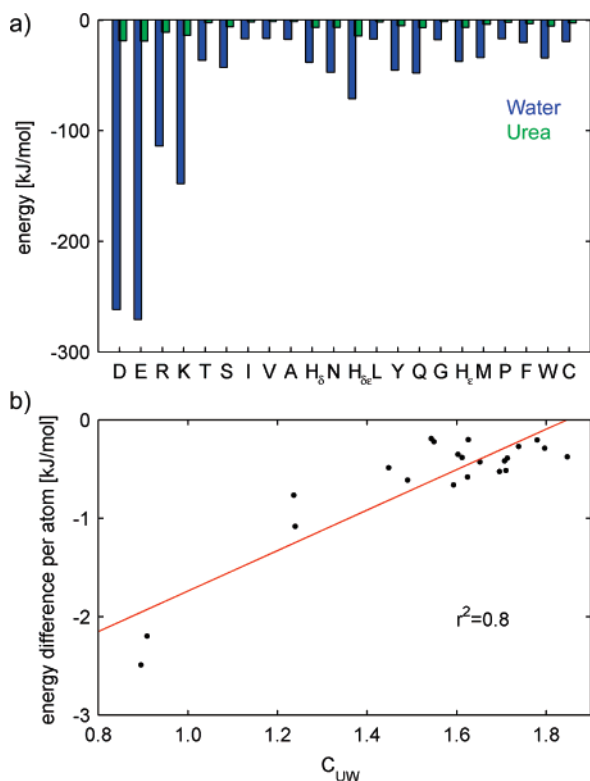


Figure 4. Force field energies between urea and water. (a) Total potential energy within cutoff radius (blue, water; green, urea). (b) Energy difference between residue–water and residue–urea per atom versus contact coefficient C_{UW} . The linear fit with a regression coefficient of $r^2 = 0.8$ is drawn in red.

of water from less polar protein surface into bulk is enthalpically and entropically favorable. This picture agrees with the results from the hydrogen bond energies discussed above.

We also carried out a set of 22 simulations with all electrostatic interactions switched off to estimate the contribution of Coulomb and Lennard–Jones energies to the contact preferences and to extract purely steric contributions. Contact coefficients from these simulations were all quite similar to each other (1.25 ± 0.06). This result further supports the view that electrostatic interactions, as opposed to Lennard–Jones interactions, are (directly via enthalpy or indirectly via the hydrophobic effect) the main determinants of the observed contact preference profile.

4. Discussion

Our simulations suggest that polarity/apolarity is the main determinant of the interaction preference C_{UW} of amino acids with urea or water. One would therefore expect that C_{UW} correlates to hydrophobicity scales. Indeed, we find qualitative agreement with common scales.^{48–53} In particular, amino acids ranked very hydrophobic by these scales (CYS, PHE, TRP, etc.) interact preferentially with urea in our simulations (high C_{UW}), whereas amino acids ranked as very hydrophilic (ASP, GLU, ARG, LYS, etc.) interact preferentially with water (low C_{UW}).

On the atomic level, we find good agreement with atomic hydrophilicities reported by Kuhn et al.,⁵³ who found oxygen atoms to be most hydrophilic, followed by nitrogen, followed by carbon and sulfur ($O \approx O^- > N^+ > N \approx C \approx S$). As the correlations between these different hydrophobicity scales are in the range of $r^2 = 0.69–0.95$ (problems involved in defining and interpreting such scales are well known^{54,55}), one would not expect correlations with C_{UW} higher than this value. The observed correlation coefficient of $r^2 = 0.35–0.53$ between C_{UW} and the different hydrophobicity scales therefore suggests that the hydrophobic effect is a key determinant for the contact preferences C_{UW} but certainly not the only one.

Indeed, comparison with free energies of transfer from water to urea solution (ΔG_{W-U}) shows a similar correlation. The C_{UW} values reported here agree with main features from the early study of Nozaki and Tanford¹⁹ and from more recent studies by Auton and Bolen.¹⁵ In particular, less polar residues like PHE, TRP, and TYR have a large ΔG_{W-U} , as well as a high C_{UW} . The correlation of C_{UW} with all ΔG_{W-U} reported in ref 19 is $r^2 = 0.32$, and $r^2 = 0.31$ for the data reported in ref 15. Hence, transfer free energies ΔG_{W-U} are obviously also related to C_{UW} . Although we certainly cannot exclude further effects contributing to the observed C_{UW} , our results suggest these two to be the main determinants for the observed contact preferences.

We have also analyzed the relation to m -value contributions, which measure the variation of the free energy of unfolding with denaturant concentration.⁵⁶ Indeed, very good agreement is seen with the m -value contributions reported by Auton and Bolen.¹⁵ Their data are particularly suitable for comparison with C_{UW} values, since both are normalized to size or surface area. The regression coefficient with their data (Figure 2a in ref 15) is $r^2 = 0.57$ (see Supporting Information for the correlation diagram). We note that CYS is a unique and puzzling outlier, with completely opposite tendencies, and was excluded from the regression fit. Large m -value contributions are reported for less polar residues like TRP, PHE, and LEU, as well as for the backbone, while the contribution of ASP, GLU, ARG, and LYS is very small. This large correlation shows that the contact coefficient C_{UW} introduced here can obviously be related to the m -value contribution per surface reported in ref 15.

By determining contact coefficients from tripeptides (GXG) with neutralized termini, we have aimed at providing values for individual amino acids. In particular, this approach excludes effects from sequence, secondary or tertiary structure interactions, and effects from the N and C termini. In a protein, in contrast, the interaction of each amino acid with the surrounding water and urea solvent will additionally depend on local neighboring residues. For example, we expect C_{UW} to be lower than in the GXG tripeptides when X is flanked by a polar residue like GLU or higher when X is flanked by an apolar residue like TRP. The influence of flanking peptide bonds on hydrophilicity has previously been discussed by Roseman et al.⁵⁷

(48) Nozaki, Y.; Tanford, C. *J. Biol. Chem.* **1971**, *246* (7), 2211–2217.

(49) Janin, J. *Nature* **1979**, *277*, 491–492.

(50) Guy, H. R. *Biophys. J.* **1985**, *47*, 61–70.

(51) Rose, G. D.; Geselowitz, A. R.; Lesser, G. J.; Lee, R. H.; Zehfus, M. H. *Science* **1985**, *229* (4716), 834–838.

(52) Bordo, D.; Argos, P. *J. Mol. Biol.* **1991**, *217* (4), 721–729.

(53) Kuhn, L. A.; Swanson, C. A.; Pique, M. E.; Trainer, J. A.; Getzoff, E. D. *Proteins: Structure, Function, Genetics* **1995**, *23*, 536–547.

(54) Charton, M.; Charton, B. I. *J. Theor. Biol.* **1982**, *99* (4), 629–644.

(55) Cornette, J. L.; Cease, K. B.; Margalit, H.; Spouge, J. L.; Berzofsky, J. A.; DeLisi, C. J. *Mol. Biol.* **1987**, *195* (3), 659–685.

(56) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Science* **1995**, *4*, 2138–2148.

(57) Roseman, M. A. *J. Mol. Biol.* **1985**, *200*, 513–522.

We note that the choice of water and urea model may affect our results and in particular the obtained contact coefficients and hydrogen bond energies. Hence, the contact coefficient can also be used to investigate the effect of force field choice on urea/water interactions with proteins. While we do not expect other models to yield qualitatively different results, particularly regarding the spectrum of contact coefficients for polar or apolar residues, the quantitative numbers may differ. Future work should therefore compare contact coefficients for different choices of urea and water models. Such studies might also address effects of urea concentration and temperature, which would help to further discriminate between enthalpic and entropic contributions to interaction preferences.

5. Conclusions

Aiming at a comprehensive characterization of the interactions of amino acids with aqueous urea solutions, we have calculated contact preferences for urea with individual amino acids from molecular dynamics simulations of 22 tripeptides. All amino acids (except ASP and GLU) were found to interact preferentially with urea. A clear spectrum of contact preferences has emerged, ranging from slight preferences for water contacts (ASP and GLU) to high preferences for urea contacts (TRP and CYS). In summary, urea was found to preferentially solvate apolar and aromatic residues, as well as the peptide backbone.

These findings suggest a number of important implications for the mechanism of protein denaturation by urea. Under native conditions in water and mainly due to the hydrophobic effect, apolar and aromatic residues are typically buried within the hydrophobic core of the folded state and not exposed to the solvent. Further, the protein backbone is largely shielded from solvent contact via formation of secondary structure elements. In aqueous urea solution, our contact analysis shows that urea molecules are located at the protein surface, in particular accumulating close to less polar residues and the backbone. The resulting displacement of water molecules from the solvation shell of less polar residues and the backbone into bulk water is favorable both entropically and enthalpically. The former because the translational and rotational entropy is increased, as one urea molecule frees about three water molecules; the latter because these three water molecules can form strong water–water hydrogen bonds.

The resulting overall weakening of the hydrophobic effect renders unfolding of the protein by exposure of the hydrophobic core and dissociation of secondary structure elements energetically favorable. Taken together, our results, as discussed so far, strongly support the view that hydrophobic interactions are the main determinant of urea-induced protein denaturation^{19,26} rather than interactions with polar residues.^{4,10,11,17,18}

This result seems to be at variance with previous findings,^{4,18,58} however, which reported hydrogen bonds to be essential. The results of our hydrogen bond energy analysis now allow to resolve this puzzle by quantifying the different

contributions discussed controversially by different authors. In particular, in our simulations urea was seen to form hydrogen bonds to the peptide backbone which are of similar strength as backbone–backbone hydrogen bonds but are significantly weaker than backbone–water or water–water hydrogen bonds. Thus, these results confirm the widely held view that hydrogen bonds between urea and the protein backbone are essential^{4,10,18,58} but do not support the suggestion¹⁸ that a competition of urea with native interactions is the driving force for urea-induced protein denaturation. In particular, it has already been argued⁵⁹ that in such competition between native protein–protein hydrogen bonds and protein–solvent hydrogen bonds, protein–urea hydrogen bonds would have to be stronger than protein–water hydrogen bonds to explain denaturation. Indeed, we found the opposite. Therefore, rather than being the driving force, we suggest that urea–protein hydrogen bonds only serve to avoid highly unfavorable unsatisfied hydrogen bond sites of the backbone,^{60,61} while at the same time shielding it from entropically unfavorable water contact. Note that this finding is not at variance with our simulation result that the weaker urea–protein hydrogen bonds actually tend to replace the stronger water–protein hydrogen bonds, which we suggest to be due to the entropic effect described above.

In summary, our simulation study suggests a synthesis of seemingly opposing viewpoints. Whereas urea–protein hydrogen bonds do not seem to drive the denaturation, they do contribute to the overall energetics. According to the mechanism proposed here, the denaturation power of urea rests on its tradeoff between two essential but conflicting features. First, it is apolar enough to solvate apolar groups; second, it is polar enough to form weak hydrogen bonds to the backbone and to incorporate well into the water hydrogen bond network.^{3,62,63} Hence, urea can be regarded to denature proteins by interfacing between water and natively buried parts of the protein.

The suggested dominance of the entropic effect leads us to speculate that the mere fact that urea molecules are significantly larger than water molecules substantially contributes to its denaturation power.

Acknowledgment. We thank Ira Tremmel for carefully reading the manuscript. M.C.S. gratefully acknowledges support from the Deutsche Volkswagen Stiftung, Grant No. I/78 839.

Supporting Information Available: Correlation diagram of contact coefficients with *m*-value contributions reported by Auton and Bolen.¹⁵ This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA076216J

- (59) Dill, K. A. *Biochemistry* **1990**, *29* (31), 7133–7155.
(60) Krantz, B. A.; Srivastava, A. K.; Nauli, S.; Baker, D.; Sauer, R. T.; Sosnick, T. R. *Nature Struct. Biol.* **2002**, *9* (6), 458–463.
(61) Fleming, P. J.; Rose, G. D. *Protein, Science* **2005**, *14*, 1911–1917.
(62) Sharp, K. A.; Madan, B.; Manas, E.; Vanderkooi, J. M. *J. Chem. Phys.* **2001**, *114* (4), 1791–1796.
(63) Stumpe, M. C.; Grubmüller, H. *J. Phys. Chem. B* **2007**, *111* (22), 6220–6228.

(58) Robinson, D. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87* (11), 2462–2470.