Structure and Hydration of L-Proline in Aqueous Solutions

Sylvia E. McLain,*,†,‡ Alan K. Soper,† Ann E. Terry,† and Anthony Watts‡

ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom, and Biochemistry Department, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3QU, United Kingdom

Received: December 5, 2006; In Final Form: February 18, 2007

The structure and hydration of L-proline in aqueous solution have been investigated using a combination of neutron diffraction with isotopic substitution, empirical potential structure refinement modeling, and small-angle neutron scattering at three concentrations, 1:10, 1:15, and 1:20 proline/water mole ratios. In each solution the carboxylate oxygen atoms from proline accept less than two hydrogen bonds from the surrounding water solvent and the amine hydrogen atoms donate less than one hydrogen bond to the surrounding water molecules. The solute−solute radial distribution functions indicate relatively weak interactions between proline molecules, and significant clustering or aggregation of proline is absent at all these concentrations. The spatial density distributions for the hydration of the COO^- group in proline show a similar shape to that found previously in L-glutamic acid in aqueous solution but with a reduced coordination number.

I. Introduction

Over 40 years ago it was proposed that there is an evolutionary basis for the ubiquity of certain elements and many small molecules found in living systems.[1] This view was subsequently supported in a study by Yancey et al.[2] who found that certain biological molecules—the free amino acids proline, glycine, and arginine to name a few—were found to accumulate in the cells of a large variety of water-stressed organisms, where they purportedly serve the same biological function in many plant, animal, and bacterial systems. In general this class of molecules, known as osmolytes, has the ability to counter environmental water stress and to protect cellular proteins against denaturation under these extreme conditions.[2−6] These molecules are synthesized and accumulate in cells as a response to osmotically induced dehydration in the presence of high salt environments,[7] where proline in particular has been shown to stimulate growth and respiration in bacteria in the presence of low water activities. Proline, in essence, acts as a "chaperone" to protein formation in these conditions.[8] It has been proposed that osmolytes, sometimes termed osmoprotectants, are effective because of their ability to promote protein refolding due to their exclusion from the protein surface itself.[9,10] Although it has long been known that proline has this biological function, details of this interaction and how the free amino acid interacts with a surrounding water environment to date remain unknown.

Proline has a unique structure among the amino acids; the amino nitrogen binds with the side chain, leading to the formation of a pyrolidine ring with the α-carbon. This structure necessarily restricts the conformations that proline can adopt within a peptide or protein, giving proline a unique role in the secondary and tertiary structures of proline-containing proteins. Given this conformation proline is often found in the turns in complex proteins[11] and plays a binding role in collagen, which has a high proline content.[12] Proline-rich regions in proteins are also found in many kinase binding sites[13] and are linked to many other cellular processes.[14] Proline-rich regions can adopt a left-handed helical conformation, polyproline II (PP_{II}), and this conformation has been linked to ligands for signaling proteins and is important for the structural integrity of many proteins.[14−16] That proline is a very rigid amino acid is exemplified by the fact that it is one of the largest components of the protein collagen, the main protein that forms connective tissue in humans and other higher organisms.

In living systems, proline, as a free amino acid, is invariably in an aqueous solution of some description. Additionally, as peptide and protein constituents, many amino acids interact with water molecules in these configurations as well. To understand more about how the aqueous environment hydrates the proline molecule, we have investigated the structure of L-proline in solution at several concentrations using a combination of neutron diffraction measurements and empirical potential structure refinement modeling. The combination of these techniques is employed to obtain the average structural interactions of this amino acid in solution on a local length scale (∼1−10 Å). Additionally we have performed small-angle neutron scattering (SANS) experiments on each of the concentrations in solution to determine if any longer-range structures (∼25−800 Å) are formed in the solutions.

The purpose of this investigation is 2-fold. In the first instance, given that its role as an osmoprotectant is most likely linked to the interaction between proline and the bulk water environment, this investigation of the free amino acid in solution provides a direct measurement of the hydrogen bonding between proline molecules and the water solvent. Additionally, at the concentrations measured, it is likely that there will be a number of proline−proline interactions, which can provide insight into the ability of proline to act as a chaperone during protein formation under environmentally challenged conditions. Secondly, the hydration of a protein in solution can be partly described by the interactions of its constituent amino acids with the surrounding water environment.[17] Understanding the hydrogen-bonding interactions in an amino acid/water system is a

* Author to whom correspondence should be addressed. Present e-mail: mclainse@ornl.gov.
† Rutherford Appleton Laboratory.
‡ University of Oxford.

10.1021/jp068340f CCC: $37.00
Published on Web 04/10/2007

necessary first step toward understanding the development of higher structure in proteins.

II. Theoretical Background

A. Neutron Diffraction. Neutron diffraction augmented by isotopic substitution has been widely used to study the structure of hydrogen-containing liquids as well as the structure of solutes in aqueous and other solvent systems.[18−24] Given that different isotopes give rise to different scattering intensities, multiple measurements on the same system using different isotopomers give rise to different diffraction patterns, allowing for the interpretation of the structure and coordination around different sites within the liquid.

The quantity obtained in a neutron diffraction experiment after appropriate corrections[25] is the interference differential cross-section, $F(Q)$, which is defined as

$$F(Q) = \sum_{\alpha\beta\geq\alpha} (2 - \delta_{\alpha b}) c_\alpha c_\beta b_\alpha b_\beta (S_{\alpha b}(Q) - 1) \qquad (1)$$

where c_α is the atomic fraction and b_α is the scattering length of isotope α. Q, the magnitude of the change in the wave vector by the scattered neutrons, is defined as $Q = 4\pi \sin\theta/\lambda$, where θ represents the scattering angle and λ is the wavelength of the scattered radiation. $F(Q)$ is the sum of all partial structure factors, $S_{\alpha b}(Q)$, present in the sample each weighted by their composition and scattering intensity. For each system measured there are $m(m + 1)/2$ partial structure factors for m distinct atom types.

The Fourier transform of any structure factor ($S_{\alpha b}(Q)$) yields the associated radial distribution function (RDF), $g_{\alpha b}(r)$, where these two functions are related by

$$S_{\alpha b}(Q) = 1 + \frac{4\pi\rho}{Q} \int r[g_{\alpha b}(r) - 1]\sin(Qr)\, dr \qquad (2)$$

where ρ corresponds to the atomic number density of the sample. To understand the average local structure of a liquid, integration of $g_{\alpha b}(r)$ gives the coordination number of atoms of type β around an α atom at the origin over the distance range from r_1 to r_2

$$n_\alpha^\beta(r) = 4\pi c_\beta \rho \int_{r_1}^{r_2} g_{\alpha b}(r) r^2\, dr \qquad (3)$$

The coordination number is usually taken by integration up to the first minimum (r_2) after the first obvious peak in $g_{\alpha b}(r)$.

B. Empirical Potential Structure Refinement. In practice, because of the limitations imposed by the availability of isotopes, it is usually not feasible to directly measure all of the partial structure factors and thereby all of the site−site RDFs present in multicomponent systems. To obtain a full set of correlations for the systems studied here, empirical potential structure refinement (EPSR), a computational method for disordered materials, was used to model the diffraction data. EPSR is a method that begins with a standard Monte Carlo simulation using a set of reference potentials. Subsequently EPSR iteratively perturbs these potentials, giving rise to new potentials that aim to give the best possible agreement with the structural data. While EPSR does not necessarily provide the only possible interpretation of the structural data it does provide a model that is consistent with the measured diffraction data. More detailed descriptions of EPSR are given elsewhere in the literature.[24,26,27] In addition to determining the RDFs from the EPSR model, some of the three-dimensional spatial density functions (SDFs), which show the location of molecules or portions of molecules relative to one another, were also determined (vide infra). The same structural model also allows the orientational correlation functions (OCFs) to be extracted simultaneously. These OCFs show the relative orientation of one molecule to another at a specific location in three-dimensional space. Details of both the spherical harmonic expansion as well as the orientational correlation function calculation are given in detail elsewhere.[24,28,29]

C. Small-Angle Neutron Scattering. SANS is widely used to determine the size and shape as well as the polydispersity of aggregates in solution.[30] In a similar manner to neutron diffraction, a SANS experiment measures samples as a function of Q by determining the differential cross-section, $d\Sigma/d\Omega$. This function is analogous to $F(Q)$ for the neutron diffraction measurements except the scattering is not normalized to the number density of the sample and multiple scattering and inelasticity effects are not normally corrected for in a SANS experiment. One other notable difference between SANS and neutron diffraction is that the Q-range over which the sample is measured is much lower in SANS than that for the wider-angle diffraction experiment. In the present instance, the Q-range for the SANS experiments was $0.008 < Q < 0.24$ Å^{-1}. This corresponds to a distance range up to 1000 Å and is significantly larger than the distances (<30 Å) probed by the diffraction experiment alone.

III. Experimental Section

A. Sample Preparation. Fully protonated L-proline ($C_5H_9NO_2$) and D_2O (99.8% D) were purchased from Sigma-Aldrich Chemical Co., and d_7-L-proline ($C_5D_7H_2NO_2$) was purchased from Cambridge Isotopic Laboratories. Ultrapure H_2O was obtained from a Millipore purification system. Fully deuterated L-proline was prepared by dissolving d_7-L-proline in a surplus of D_2O in a borosilicate glass ampule to deuterate the two exchangeable hydrogens. The mixture was subsequently freeze-dried using an all-glass vacuum apparatus (~10^{-3} mbar).

B. Neutron Diffraction Measurements. Three different concentrations of proline solutions were measured, namely, concentrations of 1:20 proline/water (~2.8 M), 1:15 proline/water (~3.7 M), and 1:10 proline/water (~5.5 M) at standard temperature and pressure (298 K and 1 bar). The samples were prepared by weight and then transferred into vanadium flat plate containers with wall thicknesses of ~1 mm each coated with a ~0.1 mm layer of poly(tetrafluoroethylene) (PTFE). Vanadium containers were used because the scattering from this metal is predominantly incoherent and leads to a more tractable analysis of the scattering from the sample itself. The PTFE coating was used to prevent interactions between the amino acid in solution and the vanadium metal surface.

The diffraction data were obtained using SANDALS (Small-Angle Neutron Diffractometer for Amorphous and Liquid Samples) located at the ISIS pulsed neutron facility at Rutherford Appleton Laboratory in the U. K. where the typical collection time for each sample is ~6−8 h at 1500 μA. SANDALS is an instrument well-suited for structural measurements of liquids containing hydrogen with detectors that range from 3.9° to 39°, giving a Q-range from 0.15 to 50 Å^{-1}. For each measurement the raw data for each sample were converted to $F(Q)$, after correcting for absorption, multiple scattering, container scattering, and inelasticity effects, using the program Gudrun derived from the ATLAS suite of programs available at ISIS.[25] SANDALS is further equipped with a transmission monitor that measures the total cross-section of the sample being measured, σ_{trans}, relative to the incident beam. Data were collected for each of the samples as well as the empty sample containers to ensure

TABLE 1: Proline/Water Solutions Measured by Neutron Diffraction

sample	composition	sample thickness
	20:1 proline/water	
I	h_9-proline/H_2O	0.115
II	d_7-proline/H_2O	0.113
III	h_7-proline/HDO	0.115
IV	h_7-proline/D_2O	0.138
V	d_9-proline/D_2O	0.113
	15:1 proline/water	
VI	h_9-proline/H_2O	0.138
VII	h_7-proline/HDO	0.138
VIII	h_7-proline/D_2O	0.115
IX	d_9-proline/D_2O	0.113
	10:1 proline/water	
X	h_9-proline/H_2O	0.138
XI	h_7-proline/HDO	0.115
XII	h_7-proline/D_2O	0.115

an effective background subtraction. Finally to ensure an accurate container size each sample container was measured filled with D_2O to assess the appropriate sample thickness for each sample container.

For each of the three concentrations of proline investigated, several isotopomers of the solutions were measured. The samples measured are listed in Table 1 along with the corresponding thickness of the sample for each measurement. In each case the level of scatter was reasonable compared with the expected theoretical values,[31] with the one exception of sample X (h_9-proline/10 H_2O), which was below the expected level, most likely from insufficiently filling the sample can. In this case the diffraction pattern was multiplied by the appropriate factor to account for the discrepancy between expected and measured diffraction levels and to ensure correct normalization of the collected data.

C. EPSR Modeling of the Neutron Diffraction Data. For the three EPSR models, one model for each concentration, a box of molecules was constructed at the appropriate density for each measurement, 0.10408, 0.10518, and 0.10658 atoms $Å^{-3}$ for 1:20, 1:15, and 1:10 proline/water concentrations, respectively.[32] Each EPSR model contained 1000 water molecules and varied with respect to the number of proline molecules present in the solution, 50 for the 1:20 concentration, 67 for the 1:15 concentration, and 100 for the 1:10 concentration. The intra-atomic distances and angles for the proline molecules were taken from the crystal structure determination of hydroxyl-L-proline.[33] The molecular structure for a typical proline molecule used for the EPSR modeling box is shown in Figure 1. In this figure proline is shown in its zwitterionic form; that is, the amide group is protonated to form a NH_2^+ group, and the carboxylic acid group has been deprotonated to form a COO^- or carboxylate group to be consistent with the average structure of proline in solution. At each of the concentrations measured, the pH of the solution is ~6.5, either with proline dissolved in H_2O or D_2O, and at this pH proline is most likely to be found as a zwitterion in a time-averaged measurement. The protonated amine group is technically termed an ammonium group as it is formally charged; however in this work we refer to the ammonium group as an amine group to avoid confusion.

Each potential for the EPSR model is listed in Table 2 where the atoms correspond to the labeling scheme shown in Figure 1 and Ow and Hw are the oxygen and hydrogen atoms on the water molecule, respectively. The reference potentials used for the proline molecules are from OPLS potentials developed by Jorgenson et al. for proteins and cyclic peptides,[34] and the SPC/E potential was used for the water reference potential.[35] The

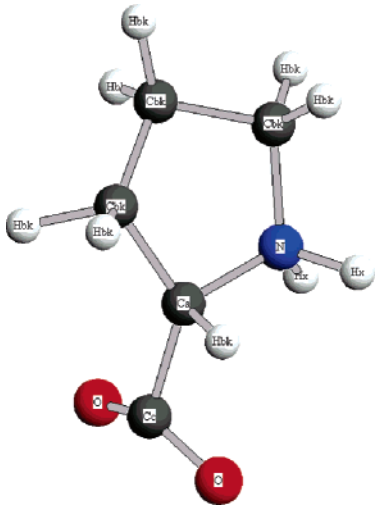


Figure 1. Molecular structure of L-proline used in EPSR fits to the diffraction data.

TABLE 2: Parameters for the EPSR Reference Potentials Used in the Fits to the Neutron Diffraction Data

atom	ϵ/kJ mol^{-1}	σ/Å	q_e
O_w	0.65000	3.166	−0.8476
H_w	0.05	0.0	0.4238
Hbk	0.0075	0.0	−0.0125
Cbk	0.42937	3.905	0.125
Ca	0.33472	3.800	0.2125
Cc	0.43932	3.750	0.3
O	0.87864	2.960	−0.5
N	0.71128	3.250	−0.3
Hx	0.0075	0.0	0.25

charges were adjusted slightly from the original potentials for proline to obtain the appropriate charge balance for the system. However it has been noted that the exact choice of potentials in similar systems is not critical but is dependent upon adjusting the charges to achieve electroneutrality.[36] Also it should be noted that, as indicated above and shown in Table 1, several different isotopomers at each of the three concentrations have been measured. The measurement of chemically equivalent isotopomeric samples gives several neutronically unique measurements of the same system and as such ensures a more reliable model as all of the data sets are fit in one EPSR model simultaneously.

D. Small-Angle Neutron Scattering Measurement. SANS measurements were taken on h_7-proline in D_2O at standard temperature and pressure (298 K and 1 bar) using the LOQ instrument located at the ISIS facility at the same concentrations measured by neutron diffraction, 1:20, 1:15, and 1:10 proline/water. Similar to the preparation of the neutron diffraction samples, the samples were prepared by weight and then transferred into Hellma fused silica spectrophotometry cuvettes with a 1 mm path length. As a background, a similar can filled with pure D_2O was measured and was subsequently subtracted from the measured proline/water data for each concentration. After accounting for detector efficiency and pixel solid angles, sample transmission, illuminated volume, and the incident flux, $d\Sigma/d\Omega$ was determined. Additionally, because LOQ is calibrated with a high-quality standard (such as a polyblend mix copolymer), the differential cross-section can be determined on an absolute scale; although as opposed to the diffraction data, the data have not been normalized by the number density or corrected for multiple scattering or inelastic effects. The samples were each measured for ~2 h at 170 μA on LOQ.

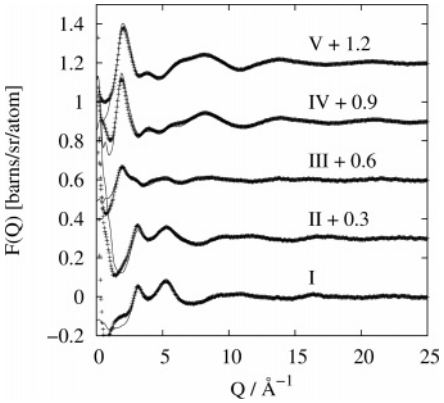


Figure 2. Measured diffraction data (crosses) and the EPSR fits to the data (crosses) for 1:20 proline/water solutions. The data and the fits have been shifted as indicated for clarity.

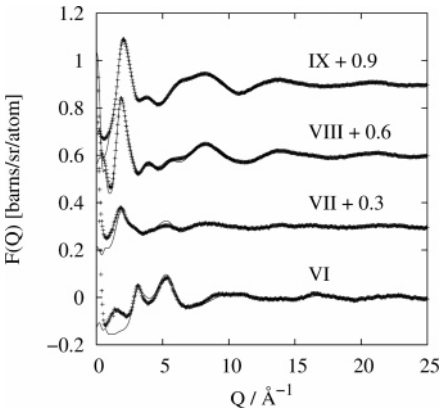


Figure 3. Measured diffraction data (crosses) and the EPSR fits to the data (solid lines) for 1:15 proline/water solutions. The data and the fits have been shifted as indicated for clarity.

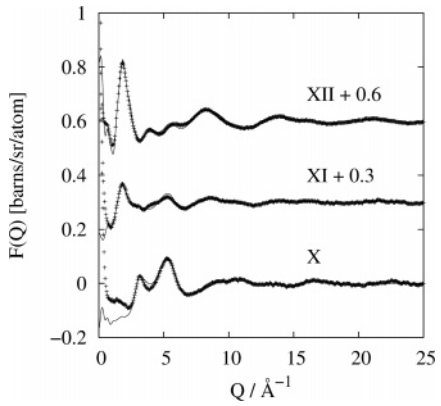


Figure 4. Measured diffraction data (crosses) and the EPSR fits to the data (solid lines) for 1:10 proline/water solutions. The data and the fits have been shifted as indicated for clarity.

IV. Results and Discussion

The measured diffraction data, $F(Q)$, along with the EPSR fits to the data are shown in Figures 2−4 for the 1:20, 1:15, and 1:10 proline/water solutions, respectively, where each sample is labeled according to Table 1 and the data have been shifted vertically for clarity. The agreement between EPSR fits and the experimentally obtained structure factors is good in each data set, with the exception of the low-Q region ($Q < 3$ Å^{-1}) in the data sets that contain hydrogen. In this region the background and inelasticity corrections to the data are most difficult to remove when light hydrogen is present in the sample. Given that $F(Q)$ is a measurement of all of the partial structure factors, it is not possible to observe directly each site−site interaction in the diffraction pattern or in the corresponding RDF. However, as mentioned above, it is possible to extract each individual site−site RDF from the EPSR model. By inspection of Table 2 and Figures 1−3 there are a total of 9 unique atoms in the measured diffraction pattern, which gives rise to 45 individual RDFs. Here we show only the RDFs associated with the bulk water structure, the water−proline interactions with the COO^- and NH_2^+ groups as well as the proline−proline interactions between COO^- and NH_2^+ groups on different proline molecules, and the ring center−center distribution function.

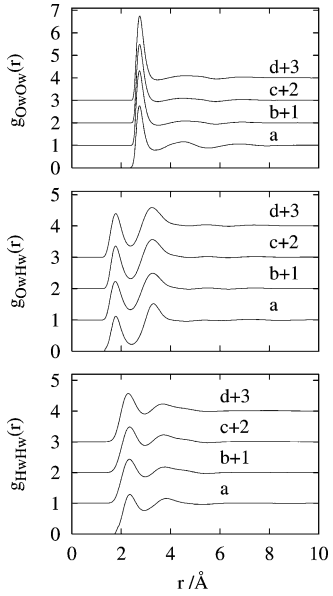


Figure 5. Comparison of water−water RDFs from EPSR fits to the diffraction data for proline/water solutions compared with pure water: (a) pure water; (b) 1:20 proline/water solution; (c) 1:15 proline/water solution; (d) 1:10 proline/water solution.

TABLE 3: Coordination Numbers for Water−Water Correlations Shown in Figure 5 from Proline/Water Solutions Compared with Pure Wate

$n_\alpha^\beta(r)$	$g_{OwOw}(r)$ ($r_2 = 3.54$ Å)	$g_{OwHw}(r)$ ($r_2 = 2.40$ Å)	$g_{HwHw}(r)$ ($r_2 = 2.94$ Å)
pure water	∼4.5−5	∼1.8	∼4−5
1:10 proline/water	4.64	1.64	4.42
1:15 proline/water	4.95	1.69	4.70
1:20 proline/water	5.10	1.73	4.83

A. Water−Water Interactions. Figure 5 shows the RDFs for the water−water interactions from the proline solutions along with previously published RDFs from EPSR simulations on pure water total structure factor measurements.[18] In each case these functions are similar to those for pure water indicating that there is no radical change in the bulk water structure upon the addition of proline. In each panel of Figure 5, the positions of the first peak are the same for each of the proline solutions in comparison to those of pure water. The comparison between coordination numbers ($n_\alpha^\beta(r)$) for this system and for pure water, which are listed in Table 3 and follow the nomenclature of eq 4, shows that the coordination between water molecules in the proline solutions is only slightly reduced from pure water. The $g_{OwHw}(r)$ coordination number is ∼1.7 in both the 1:15 and the 1:20 proline/water solutions, only slightly decreased from pure water where the number of hydrogen bonds is 1.8, while in the 1:10 proline/water solution this number has decreased somewhat further to 1.64. This deviation from pure water can be attributed

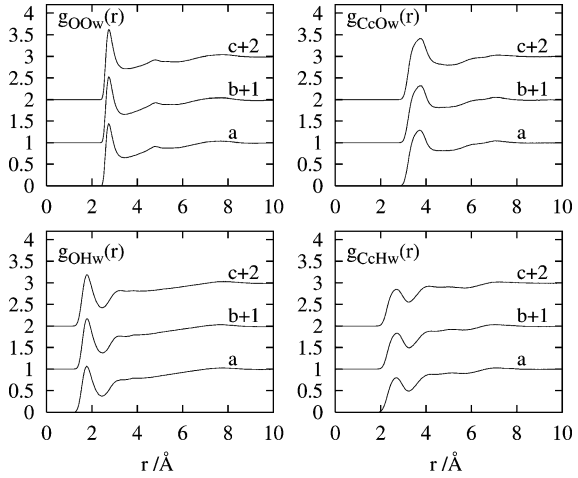


Figure 6. RDFs for water−carboxylate group correlations from EPSR fits to the diffraction data for proline/water solutions compared with pure water: (a) 1:20 proline/water solution; (b) 1:15 proline/water solution; (c) 1:10 proline/water solution.

TABLE 4: Coordination Numbers for RDFs Shown in Figure 6

$n_\alpha^\beta(r)$	1:20	1:15	1:10	r_2/Å
$g_{OOw}(r)$	2.81	2.72	2.64	3.54
$g_{OHw}(r)$	1.67	1.65	1.57	2.40
$g_{CcOw}(r)$	7.10	6.77	6.48	4.53
$g_{CcHw}(r)$	3.40	3.31	3.12	3.24

to the fact that the volume of a proline molecule is quite large compared to water ($r \approx 5$ Å) and that the solutions are all fairly concentrated, with the highest concentration (1:10) showing the largest decrease in water−water coordination. From these RDFs it appears that in each of the solutions water essentially makes a "hole" for the proline with only a small disruption to the water hydrogen-bonding network in the first hydration shell. This interpretation is furthermore supported by the fact that the only slight difference in Figure 5 is in the $g_{OwOw}(r)$ function in the second peak where there is a "flattening" of the peak in the proline/water solutions relative to the same peak in pure water. This peak corresponds to the second nearest neighbor distance for water molecules in the bulk water solvent. This flattening is most likely due to an insufficient number of water molecules being present in these solutions to form a second coordination shell. The absence of perturbation to the bulk water network in solutions also occurs with other large solute molecules at relatively high concentrations in aqueous solution such as *tert*-butyl alcohol,[37] which is a protein denaturing agent,[9] and the neurotransmitter acetylcholine.[22] Interestingly, this phenomenon was not observed in similar studies of glutamic acid where the addition of glutamate to solution resulted in a marked disruption to the bulk water structure even at concentrations lower than those measured here (1:28).[24] However, in the latter case, the solution contained Na^+ ions in addition to the glutamic acid, and spherical ions are well-known to have marked structure-disturbing properties on water.[38,39]

B. Water−Carboxylate Interactions. *1. Radial Distribution Functions.* Although the bulk water network has been largely preserved in the each of the proline water solutions, proline will have some interaction with the surrounding solvent given its high solubility in water. Proline can act both as a hydrogen bond donor via the amine group and as a hydrogen bond acceptor via the carboxylate group. Figure 6 shows the water−proline carboxylate group (Cc and O) RDFs, while Figure 7 shows the water−amine group correlations (N and Hx). The

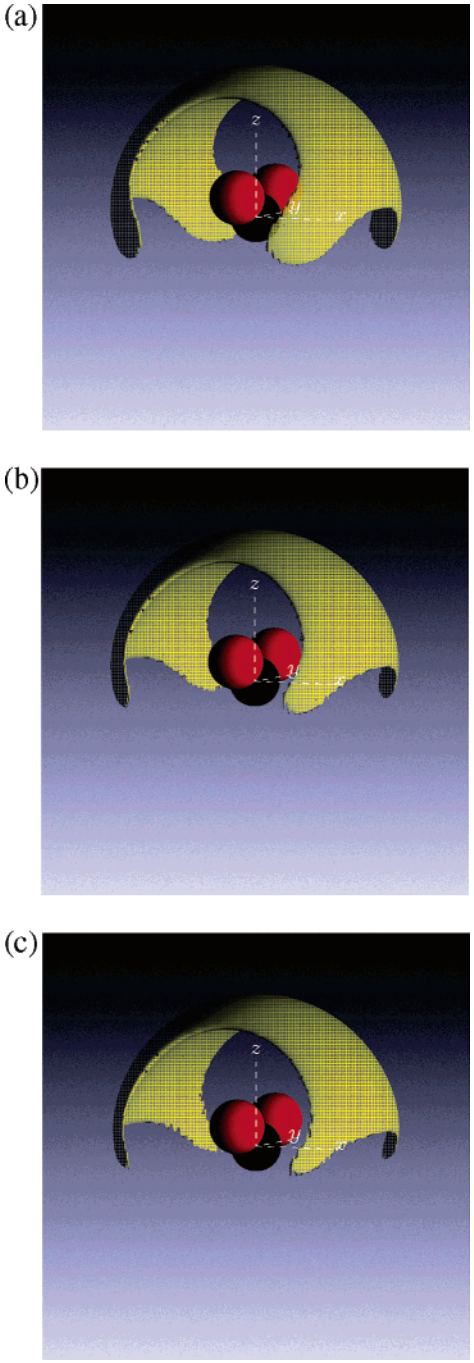


Figure 7. SDFs for the closest hydration shell of water molecules surrounding the COO^- group in proline for (a) 1:20 proline/water, (b) 1:15 proline/water, and (c) 1:10 proline/water concentrations. In each case the SDF is shown from 2 to 3.5 Å where the contour surface encloses 75% of the water molecules shown in the surrounding shell, and the radius of the plotting box is 5 Å.

TABLE 5: Coordination Numbers for RDFs Shown in Figure 9

$n_\alpha^b(r)$	1:20	1:15	1:10	r_2/Å
$g_{HxOw}(r)$	0.63	0.62	0.62	2.52
$g_{HxHw}(r)$	2.11	2.00	1.97	2.94
$g_{NOw}(r)$	1.70	1.63	1.63	3.48
$g_{NHw}(r)$	6.77	6.40	6.20	3.99

coordination numbers for the correlations shown in Figures 6 and 7 are listed in Tables 4 and 5, respectively.

From Figure 6, the hydrogen bonding between the COO^- and water molecules can be assessed. In each solution, the first

peak maxima of both the $g_{OHw}(r)$ ($r_{max} = 1.8 \pm 0.03$ Å) and the $g_{OOw}(r)$ ($r_{max} = 2.8 \pm 0.05$ Å) functions (Figure 6) occur at the same position as in the corresponding functions for the water−water correlations, namely, $g_{OwHw}(r)$ and $g_{OwOw}(r)$ (Figures 5a and 5b). These distances indicate a fairly strong correlation between water and the carboxylate group at each measured concentration. However, while the intensities of the peaks in $g_{OHw}(r)$ are nearly the same as in the corresponding $g_{OwHw}(r)$ function, the intensity of the peaks in $g_{OOw}(r)$ is markedly reduced compared to the analogous water−water correlation ($g_{OwOw}(r)$). These trends are also seen in the corresponding coordination numbers (Table 4 compared to Table 3) and are reflective of the fact that each carboxylate oxygen atom cannot be surrounded on all sides by water molecules. The coordination numbers between water and the COO^- group (Table 4), specifically for the $g_{CcOw}(r)$ functions, show that there are approximately seven water molecules surrounding this group in the nearest neighbor shell at each concentration. Moreover these coordination numbers also show that each oxygen atom accepts, on the average, less than two hydrogen bonds from the surrounding water solvent and the number of hydrogen bonds decreases with increasing proline concentration: There are 1.67 hydrogen bonds from the water donor hydrogens to the proline oxygen acceptor in the 1:20 concentration, and this number decreases to 1.57 in the 1:10 concentration. The number of hydrogen bonds from water to the carboxylate groups are only slightly lower than the coordination between water molecules in pure water where the number of hydrogen bonds is 1.8 (Table 3).[18] This coordination is also consistent with the observation that the bulk water network is not significantly disrupted upon the addition of proline at these concentrations. Interestingly, the coordination number about the carboxylate group is significantly lower than was observed for carboxylate groups in aqueous glutamate, where the coordination number of the $g_{OHw}(r)$ function is approximately three hydrogen bonds per oxygen atom.[24] There are several explanations for this difference between these two amino acids in solution. First, glutamate is a molecular ion in solution and carries a charge of −1 and as such causes a significant disruption to the bulk water network in solution as is the case with other ions in solution. Conversely, proline is electrically neutral and does not significantly perturb the water−water correlations in the surrounding solvent. Furthermore, the proline solutions presented here are at a higher concentrations than that measured for glutamic acid, and as the coordination numbers indicate in Table 4, the number of Hw−O hydrogen bonds are reduced with an increasing concentration of amino acid. The proline solutions were measured at the present concentrations to understand not only the behavior of the amino acid with respect to the bulk water environment but also to understand the interactions that may occur between proline molecules in solutions, as this may provide some insight into proline as an osmolyte. At lower concentrations, as was measured for glutamic acid, the solute−solute contacts are minimized and therefore provide less information concerning amino acid contacts.

2. Spatial Density Functions. To visualize the hydrogen bonding between the carboxylate group and the water molecules in three dimensions, the SDFs and orientational correlation functions (OCFs) between these two groups have been determined from the EPSR modeling box for each solution. The distribution of water molecules around the carboxylate group was analyzed by placing the COO^- group at the center of the standard laboratory axis, with the oxygen atoms lying in the *zy*-plane. From this central axis the positions of the water molecules were probed giving rise to a SDF, which depicts the location of the water molecules in three-dimensional space around the COO^- group. Figure 7 shows the resultant SDFs from this operation at each concentration. In each panel the oxygen atoms bisect the *z*-axis of the coordinate system with the O−C−O group lying flat in the *zy*-plane, and the shell surrounding the COO^- on the central axes shows the location of water molecules around this group. In each case, the surface contour of the shell encloses 75% of the water molecules from 2 to 3.5 Å, corresponding approximately with the maximum of the first peak in the $g_{CcOw}(r)$ function (Figure 6a). It should be noted that this not the entire first coordination shell with respect to the center of the COO^- group (Cc), which lies at the origin of the central axes, but rather represents the leading edge of the Cc−Ow coordination shell, namely, the closest hydration sphere around the carboxylate group. Furthermore, this is the same distance range probed for the carboxylate−water interactions in glutamic acid and as such provides a direct comparison of carboxylate hydration with respect to the nearest neighbor water molecules between these two amino acids in solution. Additionally in each panel in Figure 7, the central axis has been rotated by 65° to provide a better visualization of the SDF.

At each concentration, the preferred location of water molecules in this closest hydration sphere is either directly above the COO^- group where the *z*-axis bisects the two oxygen atoms or directly in front or behind O−C−O group in the *zx*-plane. Additionally, there is an absence of density in the *xy*-plane below the COO^- group where this group is bound to the α-carbon and in the *zy*-plane to the sides of the oxygen atoms. The SDFs for the 1:20 and the 1:15 concentrations (Figure 7a and 7b) both contain lobes located in front of the *zx*-plane ∼20% below the *xy*-plane. The presence of water molecules in this region in the 1:15 concentration is slightly reduced compared with the 1:20 concentration, while in the 1:10 concentration (Figure 7c) these lobes are reduced further, and although vestiges of water molecules are present in this location, here the density distribution does not extend below the *xy*-plane. The clear trend is that with increasing proline concentration water molecules are removed from the region occupied by these lobes, indicating that water molecules will preferentially coordinate above the COO^- group when there are an insufficient number of water molecules present to fully hydrate this portion of the amino acid in the first hydration sphere.

The coordination number for the center−center distribution between the carboxylate group and water molecules ($g_{CcOw}(r)$, Table 4) gives an indication of the number of water molecules present in the SDFs shown in Figure 7, although the minimum distance used for the Cc−Ow coordination number in Table 4 is longer than the distance probed for the SDFs, which only represent the closest hydration sphere. Inspection of these numbers shows that the change in the carboxylate−water SDFs upon increasing proline concentration is not merely due to a lowering of the number of molecules present in the surrounding shell at the higher concentrations as these numbers are quite similar ranging from 7.25 in the 1:20 concentration to 6.69 in the 1:10 concentration. Given that the number of water molecules in this region does not vary considerably between concentrations, the change in the shape of the SDFs in Figure 7 may be attributable due to the presence of other proline molecules. (The solute−solute correlations are described in detail below.) Interestingly, former studies on glutamate show a similar hydration of the carboxylate group in this distance range as that of the 1:20 concentration, implying that COO^- hydration is similar in structure among the amino acids, even

though the coordination number itself varies, depending on the local charge distribution.[24]

3. Orientational Correlation Functions. Figure 7 shows that there is a preferred location for the water molecules surrounding the carboxylate group in the closest sphere of hydration in each solution rather than a random distribution of water molecules, which would result in an isotropic distribution of water molecules about the central axes in this figure. As a result, there must be an alignment of the water dipole moment with the COO^- group where, in water, this dipole moment lies along the bisector of the two Ow−Hw bonds. The orientations of the dipole moment vectors from the surrounding water molecules in the SDF were determined by extracting the appropriate OCFs. This orientation of the water molecules was probed at three locations in the shell surrounding the carboxylate group from the 1:10 proline/water solutions (Figure 7c) and are shown in Figures 8 a−c. In each of these figures the water molecule is now placed on the central axis with the water oxygen atom (Ow) at the origin. The representative water molecule on the central axis and the surrounding yellow shell shows the orientational probability distribution of the water molecule's dipole moment vector. In each case the central water is placed in its most likely orientation at the specified direction relative to the SDF of Figure 7c. These directions are indicated in the text and subsequently in each panel of Figures 8a−c. The central axes in all of the panels in Figure 8 have been rotated by 65° for clarity. The location of the water molecules relative to the carboxylate group is given by the spherical polar coordinates (θ_L,φ_L), while the orientation of the water molecule is denoted by the Euler angles (φ_M,θ_M,χ_M), both sets of coordinates being measured relative to the coordinate axes of the central carboxylate group at the origin. In this coordinate system when (φ_M,θ_M,χ_M) = (0,0,0) where all angles are measured in degrees, the water molecule lies with its dipole moment vector pointing along the *z*-axis, with the two hydrogen atoms symmetrically placed either side of the *zx*-plane in the *zy*-plane.

Figure 8a shows the orientations of water molecules directly above the *z*-axis relative to the SDF in Figure 7c at (θ_L,φ_L) = (0,0) in a distance range from r = 2 to 3.5 Å from the carboxylate carbon atom. In this case the OCF of the water molecules at this location was extracted by fixing φ_M at 0 and probing the dipole vector of water as a function of θ_M and χ_M,[28] giving the function shown in Figure 8a. With the same process, Figure 8b shows the water orientation in the direction (θ_L,φ_L) = (20,0) in the *zx*-plane of the SDF shown in Figure 7c again over the distance range r = 2−3.5 Å probing the dipole moment direction as a function of θ_M and χ_M. From these two figures it is evident that water, on the average, is oriented so that one hydrogen bond from the water donor to the oxygen acceptor sites is "shared" or split between the two carboxylate oxygen atoms.

Figure 8c shows the most probable orientation at a position of (θ_L,φ_L) = (70,0) relative to the carboxylate group at the central axis where the lobes are located in the 1:10 carboxylate water SDF (Figure 7c). As opposed to the OCFs shown in Figures 8a and 8b, the orientation of water molecules at this location was extracted by fixing χ_M at 0 and probing the dipole moment vector of water as a function of φ_M and θ_M.[28] In this figure, the orientation shell surrounding the oriented water molecule is much larger than was seen at the other locations. This indicates that the water molecule is more highly oriented above the carboxylate group and becomes less strongly oriented moving into the *xy*-plane of the carboxylate−water SDF (Figure 7c). However it is notable that in all three cases one of the

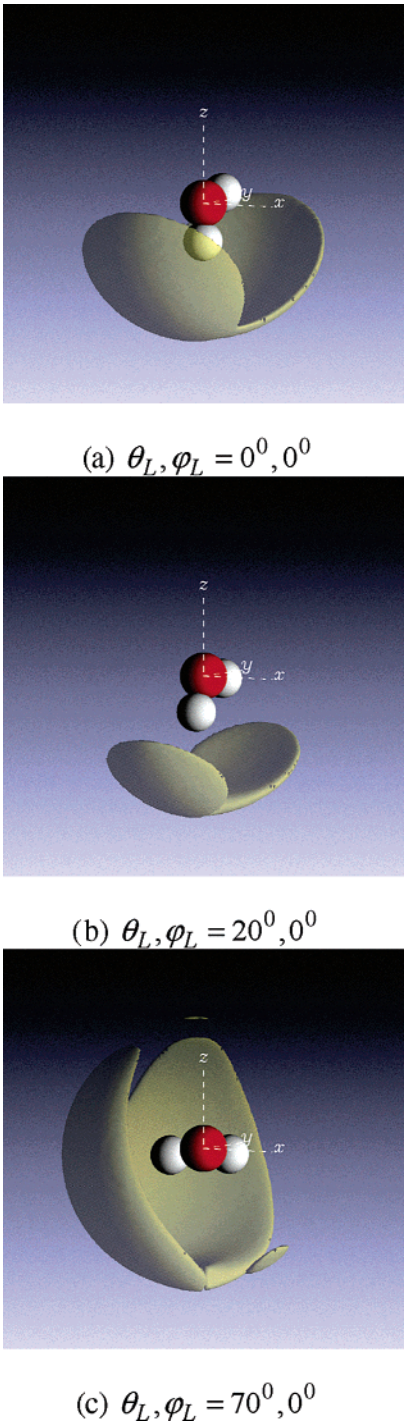


Figure 8. OCFs for water molecule surrounding the COO^- depicted in Figure 7c for the 1:10 proline/water concentration. In each of the panels the distance range probed is r = 2−3.5 Å, the contour level of the shell is 75%, and the water molecule is pointing to the most likely orientation at the location indicated. The length of each plotting box is 10 Å.

O−Hw bonds on the coordinating water molecule is always pointing toward the midpoint of the two carboxylate oxygen atoms. This orientation of water relative to the oxygen atom has been shown to be a low-energy configuration in density functional theory studies of water−proline clusters in the gas phase.[40] However these calculations do not assume a carboxylate group but rather a carboxylic acid group (COOH) so that the water molecule is not located at the midpoint between the two oxygens but rather is oriented pointing toward the C═O oxygen.

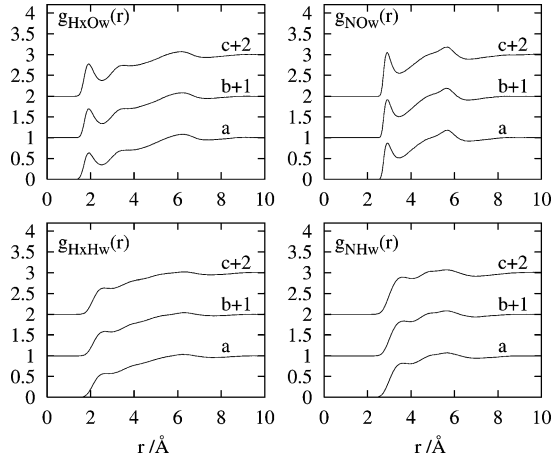


Figure 9. RDFs for water−amine group correlations from EPSR fits to the diffraction data for proline/water solutions compared with pure water: (a) 1:20 proline/water solution; (b) 1:15 proline/water solution; (c) 1:10 proline/water solution.

Only the OCFs from the 1:10 concentration have been shown here for brevity as the other concentrations measured here gave similar orientational distributions as those shown in Figure 8. Comparable orientations were also seen in the glutamic acid/water system where again the hydrogen bond was split between the two oxygen atoms, although the number hydrogen bonds between water and the COO^- are greater in glutamic acid.[24] This again supports the view that the hydration of carboxylate groups are similar among the amino acids; moreover the orientation of the hydration shell may also be similar among this group of biomolecules.

C. Water−Amide Interactions. *1. Radial Distribution Functions.* The other potential hydrogen-bonding interaction between water and proline is via donation from the amine group to the recipient water oxygen sites. Figure 9 shows the hydrogen bond interactions between the proline amine hydrogen sites (Hx) and the water oxygen sites (Ow). The number of hydrogen bonds between the water oxygen and the amine hydrogens is less than one in each solution, with virtually the same coordination number for each concentration, ∼0.6 hydrogen bonds per amine hydrogen (Table 5). Moreover the hydrogen bonds from the amine group are donated to more than one water molecule, as is demonstrated by the $g_{NOw}(r)$ coordination number, which ranges from 1.63 water molecules in the 1:10 solution to ∼1.70 in the 1:20 solution. These coordination numbers are lower than might be expected in each solution, given that there are a sufficient number of water molecules present to completely hydrate a single proline molecule in the first coordination sphere, as indicated by the water−water correlations (Figure 5). There is also lower hydrogen bond donation compared to hydrogen bonding in pure water where each Hw atom, on the average, forms ∼0.9 hydrogen bonds to other water molecules. However, given that there is no significant disruption to the bulk water network in these solutions (Figure 5), it is reasonable to expect that this coordination would be decreased, albeit slightly, from what might be expected based on the charge of an amine group. As was the case with the COO^- function, the hydrogen bond donation from the NH_2^+ group on proline to the water solvent is decreased when compared to measurements on glutamate where the amide hydrogens formed exactly 1.0 Hx···Ow bonds per amine hydrogen.[24] Again, it should be emphasized that in addition to being a molecular ion glutamate was measured at a lower concentration, and it appears that, similar to carboxylate−water hydrogen bonds, amine−hydrogen bonds decrease with an increasing concentration of amino acid present in solution.

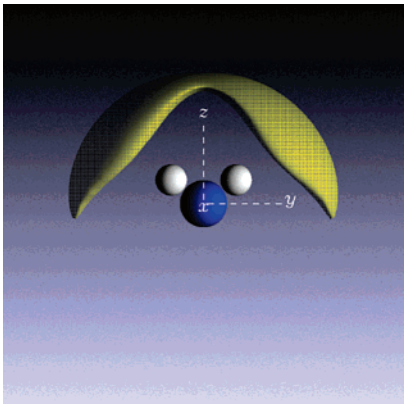
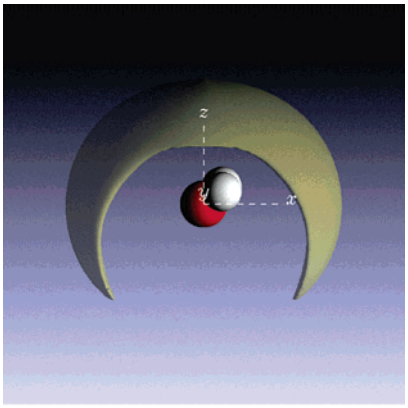


Figure 10. SDFs for water distribution around the amine group in proline from the 1:10 proline/water solution. The contour surface of the shell encloses 75% of the water molecules in a distance range of $r = 2-3.5$ Å, and in each case the length of the plotting box is 10 Å.

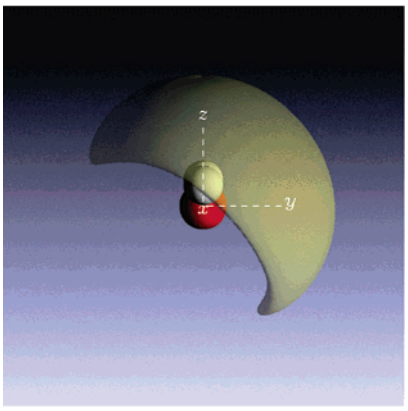
2. Spatial Density Functions. Figure 10 shows the SDF for water molecules surrounding the amine group for the 1:10 proline water solution. In this figure, the surface contour encloses 75% of the water molecules in the distance range from $r = 2$ to 3.5 Å corresponding to the minimum in the $g_{NOw}(r)$ function shown in Figure 9. Similar to the carboxylate groups in Figure 7, in each plot the nitrogen atom of the amine group is located at the origin of the central axes while the two amine hydrogen atoms (Hx) bisect the *z*-axis of the coordinate system with the Hx−N−Hx triangle lying flat in the *zy*-plane. However in opposition to the carboxylate−water SDFs (Figure 7), these SDFs show the entire first hydration shell about the amine group instead of just the closest hydration shell in this region. The entire first hydration sphere is shown here because the total number of water molecules is lower than was the case for the water molecule coordination around the carboxylate groups; therefore in this case it is easier to directly observe the distribution for all of the water molecules present in this coordination sphere. From Figure 10, the most likely location of water molecules about the amine group is in the *zy*-plane in bands directly above the hydrogen atoms with an absence of density below the *xy*-plane where the pyrolidine ring is located. Only the 1:10 proline/water concentration is shown as the amine−water SDFs at the more dilute concentrations were virtually identical, although the number of molecules occupying this shell is slightly different at each concentration (Table 5). In this SDF the distribution is reflective of the fact that there are only two hydrogens available to donate a hydrogen bond to the surrounding water molecules. Also the amine group is sterically constrained by the pyrolidine ring such that no water molecules could easily approach from below the *xy*-plane.

3. Orientational Correlation Functions. Figure 11 shows the water dipole orientation (OCF) around the amine group at two specific locations of the lobes depicted in Figures 10a and 10c corresponding to the 1:10 concentrations. Again, these functions are only shown for the highest proline/water concentrations as the other solutions showed virtually identical OCFs. As was the case for the carboxylate−water OCFs (Figure 8) the water molecule is now located on the central axis, and in both Figures 11a and 11c, the shell contour encloses 75% of the surrounding water molecules in the distance range from $r = 2$ to 3.5 Å. The water molecule is again shown pointing toward the most probable orientation.

Figure 11a shows the OCFs for water around proline at the position directly above the *z*-axis in Figure 10, specifically at, $(\theta_L,\varphi_L) = (0,0)$ relative to the central axis occupied by the Hx−



(a) $\theta_L, \varphi_L = 0^0, 0^0$



(b) $\theta_L, \varphi_L = 45^0, 90^0$

Figure 11. OCFs for water distribution around the amine group for 1:10 proline/water solution. (a) The OCF was extracted at the position $\omega_L = (\varphi_L, \theta_L) = (0,0)$ relative to Figure 10, and the central axis has been rotated by 90° for clarity. (b) The OCF was extracted at the position $\omega_L = (\varphi_L, \theta_L) = (90,45)$ again relative to Figure 10. In each case, the surface contour of the shell encloses 75% of the molecules in the distance range from $r = 2$ to 3.5 Å, and the length of each plotting box is 10 Å.

N−Hx triangle. Additionally, the OCF has been rotated by 90° relative to the central axis to provide a better visualization of the OCFs. Here, the most likely orientation is with the dipole moment of water roughly tangential to the dipole orientation of the amine group. The orientation at this position is similar to that seen in tetramethyl ammonium and the onium head group in acetylcholine where the water dipole moment is also tangential to the methyl groups in both of these molecules.[22,41] More specifically, in acetylcholine the orientation of water molecules at the analogous location is virtually identical to the orientation of water molecules shown in Figure 11a, with the water dipole moment lying at approximately 60° to the N−O_W axis. However, the amine−water RDFs (Figure 9), when compared with the corresponding RDFs for acetylcholine, show that the water molecules are much closer to the amine group than the water molecules around the corresponding onium head group in acetylcholine.[22] This is interesting, given that there is no clear hydrogen bonding in either TMA or acetylcholine between the methyl groups and water. It is possible therefore that an angled orientation of water molecules at this position is indicative of decreased hydrogen bond strength between the amine group in proline and water, compared to the corresponding water−water bond.

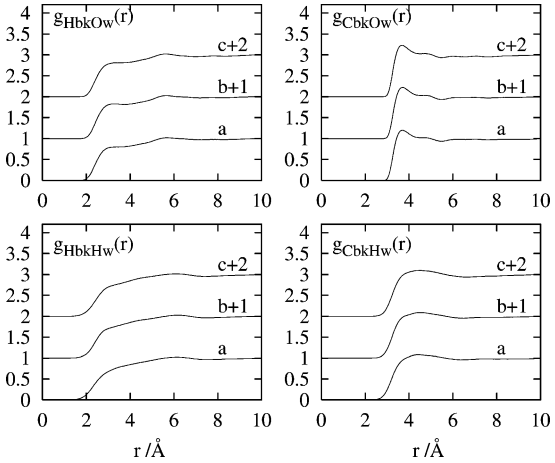


Figure 12. RDFs for water−pyrolidine ring correlations from EPSR fits to the diffraction data for proline/water solutions compared with pure water: (a) 1:20 proline/water solution; (b) 1:15 proline/water solution; (c) 1:10 proline/water solution.

Figure 11b shows the OCFs for water around proline at the position $(\theta_L, \varphi_L) = (45,90)$ relative to the central axes in Figure 10, corresponding to the large lobes that extend into the *zy*-plane of the SDF. Figure 11b shows that the most probable orientation of the water molecule has rotated relative to the water molecule in Figure 11a, but it is still lying at an angle of ∼60° relative to the N−O_W axis, and there is still a broad distribution of water molecule orientations in this direction.

D. Water−Pyrolidine Ring Interactions. The RDFs corresponding to the interactions between the pyrolidine ring in proline (Cbk and Hbk from Figure 1 and Table 3) and water are shown in Figure 12. In each solution this interaction is very slight, the most prominent interaction being between the water oxygen and the pyrolidine carbons. Although the interaction between these two sites is small, these correlations are indeed indicative of a hydrophobic interaction, which is expected given that proline is an amphiphlic molecule, namely, it has polar and nonpolar regions. The coordination numbers for these RDFs have not been listed as there is not a clear minimum present for any of the peaks at any of the concentrations. In fact the RDFs are indicative of a nearly isotropic distribution of water−ring interactions showing that there are only Van der Waals interactions between this portion of the proline molecule and water even at the most dilute concentration measured.

E. Carboxylate−Amide Interactions. *1. Radial Distribution Functions.* Prominent proline−proline interactions are most likely between the portions of the molecules that carry the largest charges, namely, the amine group and the carboxylate group. Figure 13 shows the RDFs for the interactions between these two groups at each concentration, and Table 6 shows the corresponding coordination numbers for these functions. In each of the RDFs shown in Figure 13, there are prominent intermolecular correlations between carboxylate and amine sites in each of the solutions. The $g_{CcN}(r)$ that corresponds to the center−center distribution between these two groups shows a broad peak at ∼4.5 Å, and the corresponding coordination number shows that at the highest concentration each proline amine group is coordinated by 0.79 carboxylate groups in the average bulk structure (Table 6). This number decreases to 0.66 in the 1:15 solution and 0.59 in the 1:20 solution, indicating that there are fewer proline−proline interactions in the more dilute solutions. That this number is less than 1 at all concentrations shows that a single proline molecule does not always form an amine−carboxylate hydrogen bond with another proline molecule in

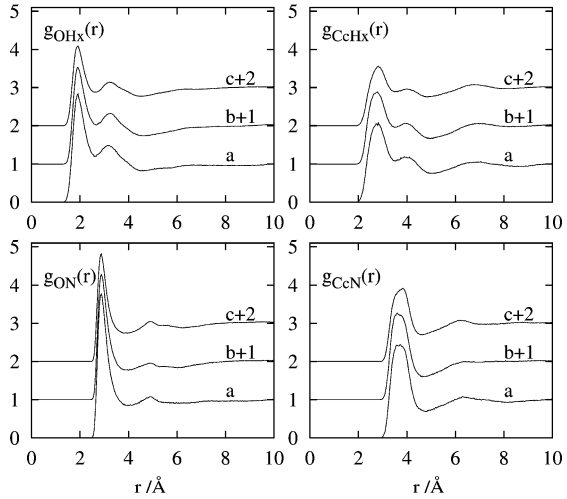


Figure 13. RDFs for carboxylate−amine group correlations from EPSR fits to the diffraction data for proline/water solutions compared with pure water: (a) 1:20 proline/water solution; (b) 1:15 proline/water solution; (c) 1:10 proline/water solution.

TABLE 6: Coordination Numbers for Amine−Carboxylate Group Interactions as Shown in Figure 13

$n_\alpha^\beta(r)$	1:20	1:15	1:10	r_2/Å
$g_{OHx}(r)$	0.29	0.33	0.36	2.61
$g_{OHx}(r)$	1.31	1.46	1.78	4.62
$g_{ON}(r)$	0.42	0.46	0.54	3.96
$g_{CcHx}(r)$	0.54	0.61	0.68	3.45
$g_{CcHx}(r)$	1.36	1.52	1.85	4.80
$g_{CcN}(r)$	0.59	0.66	0.79	4.56

solution. For instance, at the lowest concentration this only occurs with a probability of roughly 50% (Table 6); that is, on the average only half of the proline molecules in the 1:20 solution are hydrogen bound to another proline molecule in that solution. Both the $g_{OHx}(r)$ and the $g_{CcHx}(r)$ functions in Figure 13 show double peaks, indicating either two different hydrogen-bonding distances present from the amine hydrogens to the oxygen atoms of the carboxylate group or one bonded hydrogen and a non-hydrogen-bonding interaction. The latter interpretation is the most reasonable given that the distance of the second peak in $g_{OHx}(r)$ is ∼3.5 Å, which is longer than generally accepted for even a weak hydrogen bond.[42] The most prominent peak of the $g_{OHx}(r)$ function is ∼1.9 Å at each concentration, which is identical to the water $g_{OwHw}(r)$ first peak position in each measured solution (Figure 5c) and to that of pure water.[18] The number of hydrogen bonds at this location (Table 6) range from 0.29 in the most dilute case to 0.36 at the highest concentration, indicating that hydrogen bonding between proline molecules is limited. This peak is followed by a second broader peak at ∼3.5 Å at each concentration (Figure 13), and the coordination numbers for the combination of these two peaks range from 1.31 to 1.78 for the 1:20 and 1:10 concentrations, respectively. Although at first glance the $g_{OHx}(r)$ function indicates that there are two hydrogen-bonding distances between proline molecules, it again is more likely that the second peak merely reflects the fact that each amine group contains two hydrogens and that only one hydrogen bond occurs between these groups. This view is also supported by the $g_{CcN}(r)$ coordination numbers, discussed above, which show that less than one amine group surrounds the carboxylate groups in each solution measured.

2. Spatial Distribution Functions. Similar to the proline−water SDFs described above, the amine−carboxylate interactions

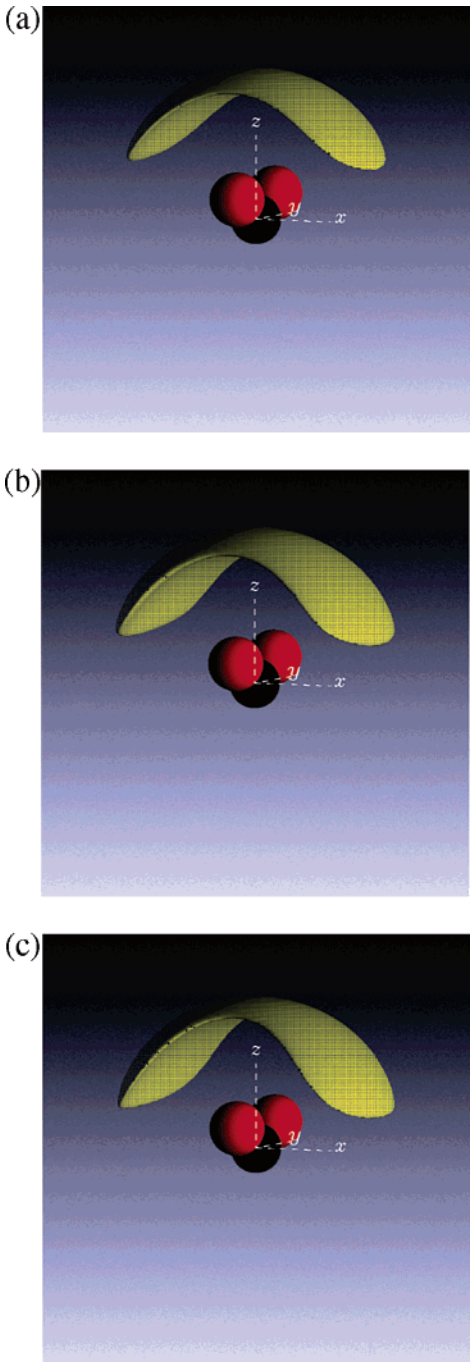


Figure 14. SDFs for amine groups surrounding the COO^- group for (a) 1:20 proline/water, (b) 1:15 proline/water, and (c) 1:10 proline/water concentrations. In each case the SDF is shown from 3 to 3.5 Å with the contour level of the shell enclosing 75% of the surrounding amine groups in this distance range, and the length of the plotting box is 10 Å.

were extracted in three-dimensional space by evaluating the SDFs between these two groups on different proline molecules. Here the carboxylate group was placed on the central laboratory axes (in the same configuration as the carboxylate group in SDFs in Figure 7) and the distribution of amine groups around the COO^- group was probed. Figures 14a−c show the resultant SDFs for each of the measured concentrations. In each case, the surrounding shell encloses 75% of the amine groups surrounding the central carboxylate group at a distance range from $r = 3$ to 3.5 Å. The minimum distance range probed was set at 3 Å to distinguish from the intramolecular proline−proline interactions in the modeling box. The maximum distance of 3.5

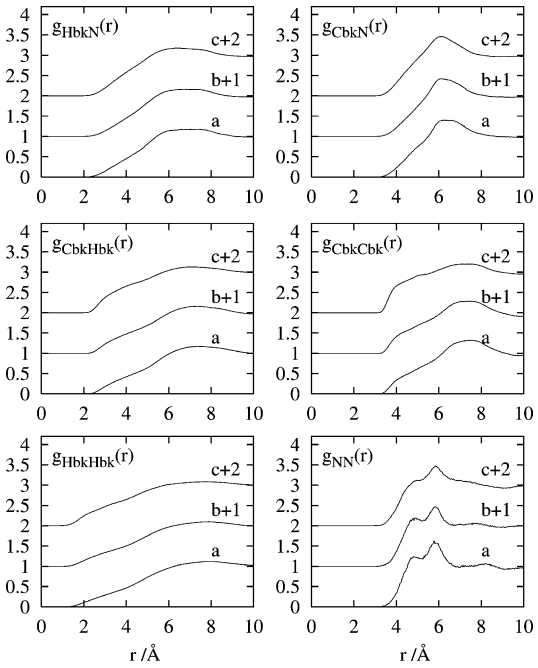


Figure 15. RDFs for water−pyrolidine ring correlations from EPSR fits to the diffraction data for proline/water solutions compared with pure water: (a) 1:20 proline/water solution; (b) 1:15 proline/water solution; (c) 1:10 proline/water solution.

Å was chosen as this distance corresponds approximately with the first peak maximum in the $g_{NCc}(r)$ function. Additionally, using this distance range allows for a direct comparison with the carboxylate−water SDFs (Figure 7) and shows the closest sphere of association between amine and carboxylate groups.

The amine−carboxylate SDFs are similar for each concentration, namely, there is a large band of density directly above the carboxylate group that extends from both sides of the central *z*-axis into the *zx*-plane. This shell extends slightly further in the *x*-direction with increasing concentrations of proline in solution with the 1:10 concentration showing the most density in this region (Figure 14c). The presence of amine groups about the carboxylate group in this region is similar to that seen in the SDFs for water molecules surrounding the carboxylate groups (Figures 7a−c), with the absence of the lobes that appear for water around COO^- in the *zx*-plane. This suggests that there is a competition between water and other proline molecules for the coordination of the carboxylate group at these locations in the SDF, thus preventing a greater number of hydrogen bonds from water to the carboxylate oxygen atoms as discussed previously (Table 3). However, not all water molecules will be excluded by amine groups given the absence of density in the *zx*-plane that occurs in the carboxylate−water SDFs (Figure 7). This indicates that in addition to competitive bonding about the carboxylate group that there is also simultaneous bonding of water molecules, even in the presence of amine−carboxylate association. Because the overall number of amine−carboxylate interactions are small (Table 6), the corresponding OCFs are not shown as in these cases it is difficult to extract a reliable orientations.

F. Pyrolidine Ring−Ring Interactions. *1. Radial Distribution Functions.* Clustering between proline molecules, in particular by a stacking of pyrolidine rings, has been suggested to account for proline's osmoprotectant properties.[43] Through the EPSR model it is also possible to ascertain if there is evidence of ring−ring contacts present in the solutions. Figure 15 shows the salient RDFs for these correlations for each solution. From inspecting the Hbk- and Cbk-containing RDFs,

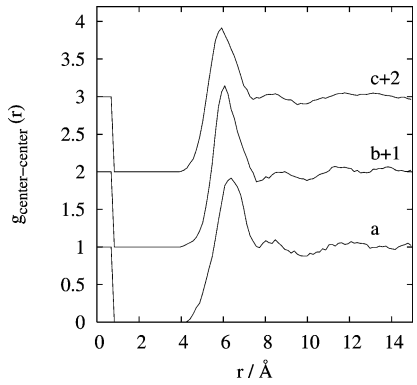


Figure 16. Pyrolidine ring center−center RDF for (a) 1:20 proline/water, (b) 1:15 proline/water, and (c) 1:10 proline/water solutions.

contact between the pyrolidine rings is not evident. There are no obvious peaks in any of these functions, but rather there appears to be only a random distribution of ring−ring contacts at each concentration measured. The only exception to this is the $g_{NN}(r)$ function that shows a prominent peak at ∼6 Å in the 1:10 and 1:20 concentrations while at the intermediate 1:15 concentration this peak is double-humped (Figure 15). This may be evidence of a ring−ring interaction, or also it may be that this correlation is related to the amine−carboxylate interactions (Figure 13). However, ring−ring interactions may indeed occur even in the absence of a noticeable correlation from the ring−ring RDFs (Figure 15) given that there are three backbone carbons (Cbk's, Figure 1) and seven backbone hydrogens (Hbk's), and each individual RDF for these functions is difficult to distinguish.

2. Center−Center Radial Distribution Functions. In the absence of clear information on ring−ring contacts from the atom−atom RDFs, the ring−ring correlations between the pyrolidine ring centers were extracted from the EPSR modeling box to determine if there was any possible evidence for ring−ring stacking that cannot be distinguished by the atom−atom RDFs alone. The function was generated by defining the center of the pyrolidine ring as a site and generating the center−center $g(r)$ at each of the concentrations; the results are shown in Figure 16. At each concentration there is a peak that has a maximum at ∼6 Å where the coordination numbers are 3.9, 2.9, and 2.3 for the 1:10, 1:15, and 1:20 concentrations, respectively, at r = 7.35 Å. While these numbers show some association between proline molecules, the coordination numbers in each case are quite low, indicating that the result of this center−center distribution may be due to simple packing effects and not clustering per se. Also this distance is large (6 Å) and may be only a reflection of the proline amine−carboxylate acceptor donor coordination discussed, as it is not possible to disentangle the ring−ring orientations with respect to each other by the radial center−center RDF given the intramolecular structure of the proline molecule.

G. Analysis of Proline Cluster formation. It is possible that proline could form clusters in an alternative manner to the stacking of pyrolidine rings, as is indicated by the amine−carboxylate donor−receptor behavior discussed above. For this reason a cluster analysis was generated from each of the proline solutions for both the proline molecules and the water molecules in the EPSR simulations.[44] Water molecules were considered to be assigned to a cluster if they were involved in a continuous hydrogen-bonding network. A hydrogen bond between water molecules was defined with distance constraints from r_{min} = 1.4 Å to r_{max} = 2.5 Å between Ow and Hw atoms on different water molecules. Proline molecules, however, were considered

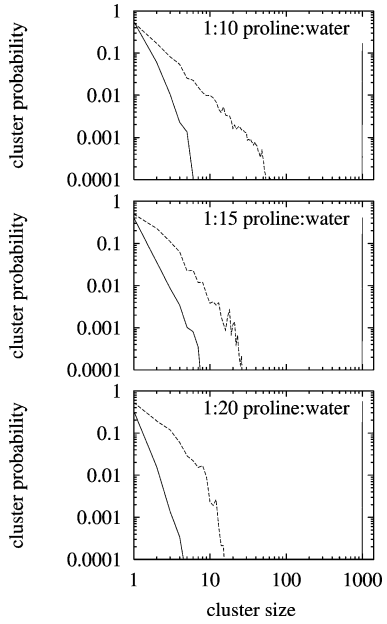


Figure 17. Cluster analysis for the proline/water solutions. In each panel the solid line represents the water clusters, and the dashed line represents the proline clusters.

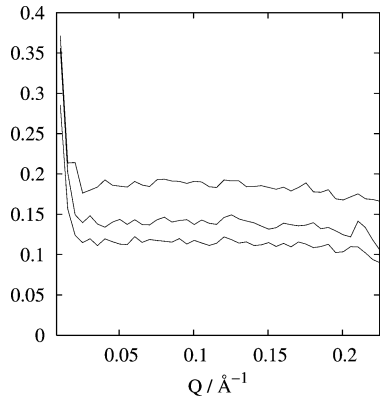


Figure 18. $d\Sigma/d\Omega$ from LOQ for proline water solutions. The top curve is the 1:10 proline/water solution, the middle curve is the the 1:15 solution, and the bottom curve is the 1:20 solution.

to be involved in clusters by hydrogen bonding between the carboxylate and amine groups, which were defined between N and Cc atoms with $r_{min} = 2.5$ Å and $r_{max} = 5.0$ Å. The potential proline clusters were defined as such on the assumption that the only prominent proline−proline contacts in the solutions were from this hydrogen bond donor−acceptor contact (Figure 11). The results of the cluster analysis are shown in Figure 17 for each of the solutions measured here. From this figure it is clear that there is indeed some very slight clustering of the proline molecules, with proline−proline clusters more prominent in 1:10 proline/water solution (Figure 17c). However, this most likely is merely a consequence of the higher concentration of proline in the 1:10 solution, given that the volume of a proline molecule, as indicated above, is quite large. The question is whether the degree of clustering shown in Figure 17 is indicative of longer-range structures being present in the solution or is simply reflective of the fact that the proline molecules are randomly packed with no attractive interactions. To establish whether the clustering observed here is significant the simulations were also performed with no atomic charges and the empirical potential set to zero, thus removing all of the attractive interactions between molecules other than the simple Van der Waals (dispersion) interactions. If the clustering is a consequence of the attractive interactions between molecules, then it should not occur in the simulation in the absence of atomic charges.[44] Without refining the empirical potential, it is not possible to fit the data using the EPSR method, but this simulation provides a Monte Carlo simulation that approximates a random distribution of molecules in solution where there is no driving force for hydrogen-bonding interactions between the molecules. The results of this test (not shown) gave a similar clustering to that shown in Figure 17, supporting the view that proline−proline interactions are not any more pronounced than those seen in a random packing of molecules. It can therefore be concluded here that the proline/water solutions investigated show, at best, limited amine−carboxylate clustering between proline molecules.

H. Small-Angle Neutron Scattering. Figure 18 shows the $d\Sigma/d\Omega$ determined for 1:10, 1:15, and 1:20 proline/water solutions using SANS after subtraction of the scatter from the D_2O and the cuvettes from the samples themselves. It should be noted that in Figure 18 the samples have not been displaced for clarity but rather represent the level of scatter measured for the differing concentrations of proline. What is clear from this figure is that no large scale structures are formed in any of the aqueous proline solutions since there is no strong signal arising at any place in the spectra in Figure 18. An exception occurs at very low Q; however the sharp rise at the smallest Q values is due to the detector cutoff rather than the sample scattering itself. Although a small degree of aggregation cannot be completely eliminated by the SANS experiment, which can only detect longer-range structures (>30 Å), there is a Q-range overlap with the diffraction data that have a Q_{min} of 0.1 Å^{-1}. From inspection of Figure 18 and Figures 4−6 there is no clear signal for significant aggregation from any of the low-Q regions (<0.2 Å^{-1}) in either the SANS data or the neutron diffraction data for any of the isopomers measured. Hence it appears such aggregation is highly unlikely. This again supports the findings above that aggregation in proline, at the concentrations measured here, does not occur and is likely not a mechanism by which proline acts as an osmoprotectant in solution.

V. Conclusions

Although it has been suggested that proline acts as an osmoprotectant through clustering, by virtue of pyrolidine rings stacking in solution,[43] the results presented here show no evidence of such interactions, given the absence of any clear intensity in the ring−ring RDFs or any obvious clustering of the proline molecules over and above what might be expected for a random distribution of molecules. The center−center RDF does show a pronounced peak at ∼6 Å, but the coordination number of this peak is small, ranging from 3.9 at the highest concentration to 2.3 at the lowest. This low coordination number does not indicate that any enhanced correlation between the rings occurs other than what is seen by random packing of proline molecules at the concentrations measured here. This view is also supported by a recent molecular dynamics study on proline in water that also finds no self-aggregation between proline molecules.[9] In its role as an osmoprotectant, it appears that proline is excluded from interacting with the protein itself[45] but rather acts as a "chaperone" to prevent denaturation and refolding of the protein in question.[8]

From the results presented here, it appears that while proline has a somewhat weak interaction with other proline molecules in solution it has a fairly strong interaction with the surrounding water environment without significantly perturbing the bulk water structure. Put another way, the tetrahedral structure of

water remains intact even though the polar portions of the proline molecule are sufficiently hydrated to ensure the high solubility of proline in aqueous solution. It appears that while in solution even at relatively high concentrations the proline carboxylate groups are hydrated by relatively strong hydrogen bonds as witnessed by the peak maximum (~2 Å) in the $g_{OHw}(r)$ function (Figure 6).

Interestingly there are obvious interactions between the polar groups in proline, the amine and carboxylate groups. Though the coordination numbers are small (Table 6) the interactions are still obviously present (Figure 11) while at the same time there are no pronounced interactions between the pyrolidine rings in the solutions measured here. This may be an effect of concentration, e.g., that the solutions do not contain a sufficient number of proline molecules to give rise to strong ring−ring correlations. Moreover, in an osmolytically challenged environment, where proline serves its role as a "chaperone", there will be ions present, and it is possible ring−ring correlations are only seen with the addition of ions to the solution. However, the amine−carboxylate interactions could be linked to the role that proline has in osmolytic conditions. It is likely that some amine groups may replace water molecules in bonding to the carboxylate groups, which is evident not only in the similarity of the amine− and water−carboxylate SDFs, but also by the fact that the O−Hw bonding is less than 2 in each solution (Table 5). Also, as was shown above in the SDF for the carboxylate−water interactions (Figures 7a−c), water molecules are located in the *zy*-plane around the carboxylate group in a position not seen in any of the carboxylate−amine SDFs (Figure 7c). It is possible that this phenomenon leaves proline sufficiently hydrated to ensure solubility while being able to coordinate with other proline molecules in solution. Additionally, this concomitancy or simultaneous bonding between proline−proline and proline−water molecules around the polar sites in proline may be indeed what gives rise to the effectiveness of proline as an osmoprotectant. Specifically, because proline exhibits a strong interaction with the water without perturbing the water structure but still associates loosely with other proline molecules, it may perhaps form a protective sheath around the protein creating a microenvironment thus allowing the protein to interact with water in a normal fashion while resisting or preventing interaction of the osmotic compounds with the protein.

Acknowledgment. We thank the U. S. National Science Foundation for fellowship monies for Sylvia McLain under award OISE-0404938 and Jason Crain, Simon Bates (both of the University of Edinburgh), and Silvia Imberti (ISIS, RAL) for useful discussions. Additionally, we thank the SANS express system at ISIS for allocation of beam time.

References and Notes

(1) Wald, G. *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *52*, 595.
(2) Yancey, P. H.; Clark, M. E.; Hand, S. E.; Bowlus, R. D.; Somero, G. N. *Science* **1982**, *217*, 1214.
(3) Rosgen, J.; Pettitt, B. M.; Bolen, D. W. *Biochemistry* **2004**, *43*, 14472.
(4) Somero, G. N. *Am. J. Phys.* **1986**, *251*, R197.
(5) Sharma, S. S.; Dietz, K. J. *J. Exp. Botany* **2006**, *57*, 711.
(6) Bhargava, S. *J. Biosci.* **2006**, *31*, 265.
(7) Measures, J. C. *Nature* **1975**, *257*, 398.
(8) Kim, S. H.; Yan, Y. B.; Zhou, H. M. *Biochem. Cell Biol.* **2006**, *84*, 30.
(9) Civera, M.; Sironi, M.; Fornili, S. L. *Chem. Phys. Lett.* **2005**, *416*, 274.
(10) Courtenay, E. S.; Capp, M. W.; Anderson, C. F.; Record, M. T. *Biochemistry* **2000**, *39*, 4455.
(11) Rai, R.; Aravinda, S.; Kanagarajadurai, K.; Raghothama, S.; Shamala, N.; Balaram, P. *J. Am. Chem. Soc.* **2006**, *126*, 7916.
(12) Bronco, S.; Cappelli, C.; Monti, S. *J. Phys. Chem. B* **2004**, *108*, 10101.
(13) Ferreon, J. C.; Hilser, V. J. *Biochemistry* **2004**, *43*, 7787.
(14) Stapley, B. J.; Creamer, T. P. *Protein Sci.* **1999**, *8*, 587.
(15) Kelly, M. A.; Chellgren, B. W.; Rucker, A. L.; Troutman, J. M.; Fried, M. G.; Miller, A.-F.; Creamer, T. P. *Biochemistry* **2001**, *40*, 14376.
(16) Ostrander, D. B.; Ernst, E. G.; Lavoie, T. B.; Gorman, J. A. *Eur. J. Biochem.* **1999**, *262*, 26.
(17) Kuntz, I. D., Jr. *J. Am. Chem. Soc.* **1971**, *93*, 514.
(18) Soper, A. K. *Chem. Phys.* **2000**, *258*, 121.
(19) Soper, A. K. *Physica B* **2000**, *12*, 276−278.
(20) Fischer, H. E.; Barnes, A. C.; Salmon, P. S. *Rep. Prog. Phys.* **2006**, *69*, 223.
(21) Turner, J. F. C.; Soper, A. K. *Polyhedron* **2004**, *23*, 2975.
(22) Hulme, E. C.; Soper, A. K.; McLain, S. E.; Finney, J. L. *Biophys. J.* **2006**, *91*, 2371.
(23) Thompson, H.; Wasse, J. C.; Skipper, N. T.; Hayama, S.; Bowron, D. T.; Soper, A. K. *J. Am. Chem. Soc.* **2003**, *125*, 2572.
(24) McLain, S. E.; Soper, A. K.; Watts, A. *J. Phys. Chem. B* **2006**, *110*, 21251.
(25) Soper, A. K.; Howells, W. S.; Hannon, A. C. *ATLAS: Analysis of Time-of-Flight Diffraction Data from Liquid and Amorphous Samples*; RAL Report No. RAL-89.046; 1989.
(26) Soper, A. K. *Mol. Phys.* **2001**, *99*, 1503.
(27) Soper, A. K. *Phys. Rev. B* **2005**, *72*.
(28) Gray, C. G.; Gubbins, K. E. *Theory of Molecular Liquids: Fundamentals*; Oxford University Press: New York, 1984; Vol. 1.
(29) McLain, S. E.; Soper, A. K.; Luzar, A. *J. Chem. Phys.* **2006**, *124*, 074502.
(30) King, S. M. In *Modern Techniques for Polymer Characterization*; John Wiley and Sons: New York, 1999; p 171.
(31) Sears, V. F. *Neutron News* **1992**, *3*, 29.
(32) Ninni, L.; Meirelles, A. J. A. *Biotechnol. Prog.* **2001**, *17*, 703.
(33) Donohue, J.; Trueblood, K. N. *Acta Crystallogr.* **1952**, *5*, 419.
(34) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(35) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(36) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569.
(37) Bowron, D. T.; Soper, A. K.; Finney, J. L. *J. Chem. Phys.* **2001**, *114*, 6203.
(38) Leberman, R.; Soper, A. K. *Nature* **1995**, *378*, 364.
(39) Soper, A. K.; Weckstrom, K. *Biophys. Chem.* **2006**, *124*, 180.
(40) Lee, K.-M.; Park, S.-W.; Jeon, I.-S.; Lee, B.-R.; Ahn, D.-S.; Lee, S. *Bull. Korean Chem. Soc.* **2005**, *26*, 909.
(41) Turner, J. Z.; Soper, A. K.; Finney, J. L. *J. Chem. Phys.* **1995**, *102*, 5438.
(42) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*, 2nd ed.; Oxford University Press: Oxford, U. K., 1999; Vol. 9.
(43) Schobert, B.; Tschesche, H. *Biochim. Biophys. Acta* **1978**, *541*, 270.
(44) Dougan, L.; Bates, S. P.; Hargreaves, R.; Fox, J. P.; Crain, J.; Finney, J. L.; Reat, V.; Soper, A. K. *J. Chem. Phys.* **2004**, *121*, 6456.
(45) Arakawa, T.; Timasheff, S. N. *Biophys. J.* **1985**, *47*, 411.