

Structural Studies on the Hydration of L-Glutamic Acid in Solution

Sylvia E. McLain,^{*,†,‡} Alan K. Soper,[†] and Anthony Watts^{†,‡}

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

Received: April 18, 2006; In Final Form: August 16, 2006

A combination of neutron diffraction augmented with isotopic substitution and computer modeling using empirical potential structure refinement has been used to extract detailed structural information for L-glutamic acid dissolved in 2 M NaOH solution. This work shows that the tetrahedral hydrogen bonding network in water is severely disrupted by the addition of glutamic acid and NaOH, with the number of water–water hydrogen bonds being reduced from 1.8 bonds per water molecule in pure water to 1.4 bonds per water molecule in the present solution. In the glutamic acid molecule, each carboxylate oxygen atom forms an average of three hydrogen bonds with the surrounding water solvent with one of these hydrogens being shared between the two oxygen atoms on each carboxylate group, while each amine hydrogen forms a single hydrogen bond with the surrounding water solvent. Additionally, the average conformation of the glutamic acid molecules in these solutions is extracted.

I. Introduction

Understanding how the self-organization of a protein in solution is affected by its aqueous environment is a major challenge that has a long history.^{1,2} It is well-known that the higher structures of proteins are linked by noncovalent bonding interactions that include hydrogen bonds between the constituent amino acids as well as with the surrounding aqueous environment; therefore changes to the aqueous environment will naturally affect all these bonds, causing a protein to either fold or denature in solution. Currently there are several techniques by which the structure and dynamics of proteins in solution are investigated. Among these methods, large protein structures in solution can be determined by NMR studies,^{3–10} and the determination of structure and solvation via small angle neutron and X-ray scattering is becoming an increasingly viable option.^{11–14} While it is true that partially hydrated proteins can also be observed by crystallography, it has been noted in X-ray solution scattering studies that the protein hydration shell in solution is not necessarily the same as the hydration shell in the solid state.¹³ Many aspects of folding and structure can be determined by macromolecular NMR in solution,^{7–10,15} but these techniques necessarily give structures that are averaged on the NMR time scale, and as such it is not always possible to observe hydrogen bonding in these systems. Hence, NMR experiments are complementary to the more direct techniques such as X-ray and neutron small angle scattering and diffraction, which provide a direct view of local order on an atomic length scale.

Unlike recent studies that have determined atomic-scale structural information in solid-state proteins,¹⁶ the application of diffraction techniques to the study of protein hydration and conformation in solution at atomic-scale resolution ($\sim 0\text{--}20$ Å) has been limited. With the exception of a large body of small

angle scattering experiments as mentioned above, there has been little direct investigation at the atomic scale either of proteins themselves or of their constituent amino acids in solution.^{17,18} Even less is known about how amino acids affect the structure of bulk water in solution. Investigation over a wide length scale from 0.1 to >300 Å of a protein in solution would ideally provide a direct visualization not only of protein conformation of a fully hydrated protein but would also provide information about the structure of the water around the protein. However, at present, the size, complexity, and diversity of real proteins in solution render atomic-scale investigations impractical by diffraction techniques. However, because the hydration of a protein can be partially described by the interactions of its constituent amino acids with the surrounding water environment,⁴ understanding the hydrogen bonding interactions in an amino acid/water system is a necessary step toward understanding the development of higher structure in proteins. Specifically, understanding how water interacts with different amino acid functional groups in solution and how the water structure itself is affected by those headgroups will provide greater insight into the structures of more complex molecules in solution.

Here, L-glutamic acid is characterized in solution by a combination of experimental diffraction data and computer modeling techniques. L-Glutamic acid was chosen for this study because it is a constituent of most proteins but also because it functions as a neurotransmitter and has recently been linked with brain dysfunction.¹⁹ As one of the most abundant neurotransmitters in higher life forms, the action of glutamic acid on binding to a receptor site is still unknown, in no small part due to the lack of details concerning its mode and site of binding and hence activation of the receptor, all of which are coupled to its hydration. Additionally, L-glutamate is used in the food industry as a flavor enhancer, where it previously has been noted that high-quality, detailed, and unambiguous atomic-scale

* Author to whom correspondence should be addressed. E-mail: s.mclain@rl.ac.uk.

[†] Rutherford Appleton Laboratory.

[‡] University of Oxford.

information is needed to provide a greater understanding of the molecule itself.²⁰

By using a combination of neutron diffraction measurements augmented with isotopic substitution and subsequent computer simulations using the modeling program, Empirical Potential Structure Refinement (EPSR),^{21,22} we have determined the interactions between L-glutamic acid and its aqueous environment on an atomic scale and investigated how the water structure is affected in the presence of this ubiquitous molecule.

II. Theoretical Background

A. Neutron Diffraction. Neutron diffraction is the premier technique by which the structures of hydrogen-bound liquids such as water,^{23–25} alcohols,^{26,27} and simple acids^{28–31} have been determined. In addition to the structural determination of pure liquids, there have been a large number of studies which focus on the structure of solutes in aqueous systems.^{17,32–38} Neutrons are the most appropriate probe of hydrogen-containing samples for several reasons. First, there is no correlation between the size of an atom and the scattering intensity. Neutrons are scattered from the nucleus itself and not the electron density of the atoms in question, as is the case with X-ray scattering. The nuclear composition, i.e., the isotopic nature, of the scattering system determines the strength of the interaction, and as a result the scattering intensity is “decoupled” from the chemical size of the atoms in the system. Second, different isotopes exhibit different scattering intensities; therefore, by changing the isotopic composition of the system, the scattering intensity is perturbed while the structure is conserved. This allows for multiple measurements on the same system, giving rise to several contrasting diffraction patterns that can be interpreted to give the local environment around any site in the liquid.

The quantity obtained, after appropriate corrections,³⁹ in a neutron diffraction experiment is the structure factor, $F(Q)$, which is defined as

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

where $\delta_{\alpha\beta}$ is the Kronecker δ -function, c is the atomic fraction, and b is the scattering length the atoms in the system, α and β . Q , the magnitude of the change in the momentum vector by the scattered neutrons, is defined as $Q = 4\pi \sin \theta/\lambda$, where 2θ represents the scattering angle and λ is the wavelength of the scattered radiation. $S_{\alpha\beta}(Q)$ is the partial structure factor between atom types α and β . For each system measured there are $m(m + 1)/2$ partial structure factors for m distinct atom types. For example, pure water has three distinct $S_{\alpha\beta}(Q)$ values, namely, $S_{\text{HH}}(Q)$, $S_{\text{OH}}(Q)$, and $S_{\text{OO}}(Q)$.

The Fourier transform of any structure factor yields the associated radial distribution function, $F(r)$, which is the sum of the respective atom–atom radial distribution functions (RDFs), $g_{\alpha\beta}(r)$'s, each weighted by concentrations and scattering lengths of atomic species (α and β) present in the sample analogous to eq 1. The partial structure factors, $S_{\alpha\beta}(Q)$, are related to the RDFs, $g_{\alpha\beta}(r)$, via the Fourier transform

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

To understand the average local structure of a liquid, integration of $g_{\alpha\beta}(r)$ gives the coordination number of β atoms around α atoms over a distance range from r_1 to r_2 given as

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

where ρ corresponds to the atomic number density of the sample and c_{β} is the concentration of atom β . The coordination number is taken by integration up to the first minimum (r_{min}) usually after the first obvious peak in the RDF.

B. Empirical Potential Structure Refinement. EPSR begins with a standard Monte Carlo simulation using an initial reference potential which has an intramolecular harmonic potential to define the geometry of the molecules being modeled and an intermolecular potential, consisting of Lennard-Jones 12–6 potentials for the site–site interactions on different molecules, as well as Coloumbic interactions on some sites, namely, the water molecules, the Na^+ ions, and the polar sites on the glutamic acid molecule. After the generation of the starting configuration with the reference potential, EPSR iteratively adjusts a perturbation to this reference potential to obtain the best possible agreement between the computed $F(Q)$ and that provided by the experimental diffraction data.^{21,23}

While EPSR provides a molecular ensemble that is consistent with the diffraction data measured, it does not necessarily provide a definitive model for the structure of the liquid in question. There may be several distinct structures that give equally reasonable agreement between data and simulation. This is especially true in the present case where there are many more partial structure factors than available diffraction contrasts; specifically there are 55 $S_{\alpha\beta}(Q)$'s for the L-glutamic acid/NaOH/water system. Therefore, it is imperative that the simulation box is constrained from the outset with as much prior information regarding the properties of the water as well as the solutes in question as is possible, such as the relative atomic charges on the molecules present as well as fairly accurate starting interatomic potentials.

The purpose of the EPSR analysis in the present instance is not only to extract three-dimensional information from a model at the correct atomic number density which is consistent with the one-dimensional diffraction data but also to explore the validity of some potential models against a set of diffraction data. EPSR generates an effective site–site interaction potential that reproduces the measured diffraction data as closely as possible. This direct comparison with the diffraction data in Q -space is rarely done with conventional molecular dynamics and Monte Carlo simulations of molecular liquids. Ideally, a wide range of initial reference potentials should be explored, for example, those which include three-body or many-body forces such as polarizability. Unfortunately, such a task is still beyond most computing strategies, and as such the most likely potentials must be selected from the literature for each individual case and tested against the experimental data.

Having found, through EPSR, a model liquid structure consistent with the diffraction data, it is useful to extract structural information from the simulation box concerning the intermolecular distributions, such as the individual site–site RDFs, as this information is not directly available from the experimental data alone. Because the site–site RDFs only give a one-dimensional representation of the system in question, it is difficult to use these distances to visualize the local spatial and orientational order in three dimensions. For this reason, spatial density functions (SDFs),^{40,41} which allow a three-dimensional representation of the local liquid structure to be constructed, were used to help determine the most probable nearest neighbor positions at certain sites in the system. Although the SDFs show the most probable location of nearest neighbor molecules, they do not necessarily give direct orien-

tational information about the surrounding molecules. In light of this, some aspects of the orientational pair correlation functions (OCFs) are also shown.⁵⁹ These tasks are achieved via spherical harmonic expansion of the full orientational pair correlation function,^{31,42} using the simulation box to derive the positional and orientational coordinates of the molecules, and are described in more detail in the following section.

C. Spatial Density and Orientational Pair Correlation Functions. The details of the spherical harmonic expansion as well as the orientational correlation function calculation using a spherical harmonic expansion are given in more detail elsewhere.^{31,42} Here a summary of these techniques that follow the notation used by Gray and Gubbins explicitly is presented.⁴²

A set of Euler angles within the laboratory reference frame $\omega_M \equiv (\varphi_M \theta_M \chi_M)$ for each molecule M is calculated using a predefined set of molecular coordinate axes. The corresponding set of generalized spherical harmonic functions, $D_{mn}^l(\omega_M)$, are calculated for each molecule and for a range of (l, m, n) values (up to $l = 4$ in the present instance). The set of such functions is then correlated taking into account the relative position $\mathbf{r} \equiv (r, \omega_L) \equiv (r, \theta_L \varphi_L)$ of the second molecule with respect to the first, yielding a set of orientational correlation function expansion coefficients, $g(l_1 l_2 l; n_1 n_2; r)$.⁴² From these coefficients the full orientational pair correlation function is obtained as an expansion of the form

$$g(\mathbf{r}, \omega_1, \omega_2) = \sum_{l_1 l_2 l} \sum_{m_1 m_2 m} \sum_{n_1 n_2} g(l_1 l_2 l; n_1 n_2; r) C(l_1 l_2 l; m_1 m_2 m) \times D_{m_1 n_1}^{l_1}(\omega_1) * D_{m_2 n_2}^{l_2}(\omega_2) * D_{m_0}^l(\omega_L) \quad (4)$$

where $C(l_1 l_2 l; m_1 m_2 m)$ are the Clebsch–Gordan coefficients, ω_1 represents the Euler angles of molecule 1, ω_2 represents the Euler angles of molecule 2, and $\mathbf{r} = (r, \omega_L)$ represents the position of molecule 2 relative to molecule 1 in the laboratory coordinate frame.

To reconstruct the orientational correlation function it is convenient to set molecule 1 at the origin and orient the coordinate system so that $\omega_1 = 0$. This serves to define the coordinate system about which the spatial density and orientation of second (neighboring) molecules will be plotted. It also leads to an immediate simplification of eq 5 in that $D_{mn}^l(000) = \delta(mn)$ so that by combining this with the requirement from the Clebsch–Gordan coefficients that $m = m_1 + m_2$ the orientational pair correlation function relative to a central molecule at the origin is given by

$$g(r, \omega, \omega_M) = \sum_{l_1 l_2 l} \sum_m \sum_{n_1 n_2} g(l_1 l_2 l; n_1 n_2; r) C(l_1 l_2 l; n_1 m_2 m) \times D_{m_2 n_2}^{l_2}(\omega_M) * D_{m_0}^l(\omega_L) \quad (5)$$

where $m_2 = m - n_1$. The SDF is generated by averaging the full orientational pair correlation function over the orientations of the second molecule, $\omega_M \equiv (\varphi_M \theta_M \chi_M)$, which immediately eliminates any terms in the summation shown in eq 6 for which $l_2, m_2, n_2 \neq 0$. Hence the spatial density function is expressed as

$$g(r, \omega) = \sum_{l_1} \sum_{n_1} g(l_1 0 l_1; n_1 0; r) C(l_1 0 l_1; n_1 0 n_1) D_{n_1 0}^{l_1}(\omega_L) \quad (6)$$

from the closure relations for the Clebsch–Gordan coefficients ($l_1 + l_2 \geq l \geq |l_1 - l_2|$).

In general the full orientational pair correlation function (eq 3) is difficult to visualize because it is a function of six coordinates. To assist in this visualization, the SDF can be plotted to gauge the most likely places of finding neighboring

molecules as is done below. Whereas the SDF shows the density of the most probable *location* of molecules in the nearest neighbor shell, the OCF shows the most probable *orientation* of nearest neighbor molecules at a particular location in this shell. The orientational correlation function can be plotted for a specified ω_L , after fixing one of the remaining angular coordinates, e.g., χ_M . This leaves an orientational pair correlation function that is a function of three variables, r , φ_M , and θ_M , for a specified direction $\omega_L \equiv (\theta_L \varphi_L)$ away from the central molecule. Other orientational pair correlation functions of the angular coordinates can be obtained by fixing other terms (such as φ_M) in the full expression (eq 1).

Below, the two OCFs shown are generated from the EPSR model by probing the relative position of the water molecule to the carboxylate group set at the origin of the laboratory axes where $\mathbf{r} \equiv (r, \omega_L) \equiv (r, \theta_L \varphi_L)$ defines the position of the water molecules. These OCFs are shown at two particular locations in the nearest neighbor water shell surrounding the α -carbon carboxylate group. Both figures show a different ω_L value relative to the central axis, and in each case $r = 2\text{--}3.5$ Å. The first OCF shows the most probable orientation of water molecules surrounding the COO⁻ group at the position of $\omega_L \equiv \theta_L \varphi_L = 0^\circ$. In this case the OCF of the water molecules at this location was extracted by fixing φ_M at 0° while showing the dipole vector as a function of θ_M and χ_M . The second instance shows the most probable orientation at a position of $\omega_L \equiv \theta_L \varphi_L = 110^\circ 45'$ relative to the carboxylate group at the central axis. The orientation of water molecules at this location were extracted by fixing χ_M at 0 and probing the dipole moment vector of water as a function of φ_M and θ_M .

III. Experimental and Modeling Procedures

Fully protonated L-glutamic acid ($C_5H_9NO_4$), ultrapure NaOH (99.99%), 40 wt % NaOD/D₂O (99.9% D), and D₂O (99.8% D) were purchased from Sigma-Aldrich Chemical Co., and *d*₅-L-glutamic acid ($C_5D_5H_4NO_4$) was purchased from Cambridge Isotopic Laboratories. Ultrapure H₂O was obtained from a Millipore purification system. Fully deuterated L-glutamic acid was prepared by dissolving *d*₅-L-glutamic acid in a surplus of D₂O in a borosilicate glass ampule to deuterate the four exchangeable hydrogens. The mixture was subsequently freeze-dried using an all-glass vacuum apparatus ($\sim 10^{-3}$ mbar). This procedure was repeated twice more to ensure that the sample was fully deuterated yielding *d*₉-L-glutamic acid. Both NaOH and the NaOD solution were used without further purification, and the NaOD solution was diluted to the appropriate concentration (2 M) with D₂O.

The samples measured are listed in Table 1 where in each case the concentrations of the solution were 2 M with respect to L-glutamic acid and NaOH and the pH of each solution is ~ 5.5 . The samples were prepared and measured in Ti/Zr alloy sample containers. The use of this alloy allows for minimal coherent scattering arising from the sample container itself, leading to a more tractable data analysis for the samples. In addition each container was lined with 0.05 mm wall thickness poly(tetrafluoroethylene) (PTFE) tubing to prevent corrosion to the metal alloy from NaOH as well as to prevent L-glutamic acid from interacting with the Ti/Zr metal surface.

The neutron diffraction data were collected on the D4c diffractometer at the high flux reactor source located at the Institut Laue Langevin in Grenoble, France. Data were collected for each of the samples as well as the empty sample containers to ensure an effective background subtraction. For each

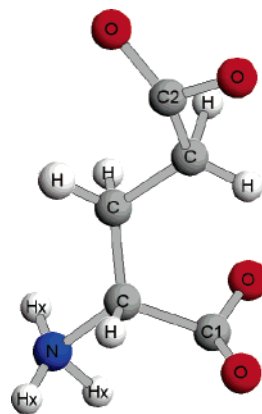


Figure 1. Molecular structure of L-glutamic acid.

TABLE 1: Samples Measured by Neutron Diffraction

sample	composition
I	C ₅ H ₉ NO ₄ /NaOH in H ₂ O
II	C ₅ H ₄ D ₅ NO ₄ /NaOH in H ₂ O
III	(C ₅ H ₉ NO ₄ /NaOH) _{1/2} /(C ₅ D ₉ NO ₄ /NaOD) _{1/2} in HDO
IV	C ₅ D ₉ NO ₄ /NaOD in D ₂ O

measurement the raw data have been converted to the structure factor using the program Gudrun available at the ISIS Facility, U.K.³⁹ In each case the level of scatter was below the expected theoretical values of the samples,⁴³ due to machining uncertainties in the sample container as well as to the presence of the PTFE liners. Therefore each diffraction pattern was adjusted to account for the discrepancies and to ensure correct normalization of the collected data.

For the EPSR model, a box of molecules was constructed at the appropriate density of the measurements (0.1 atoms Å⁻³) that contained 20 L-glutamic acid ions, 20 Na⁺ ions, and 580 water molecules. The Na⁺ ions were charged-balanced as the deprotonated glutamic acid molecule has an overall charge of -1. At the pH measured (~5.5) in solution, the L-glutamic acid is a zwitterionic form—that is the amine group is protonated to yield a NH₃⁺ group and the carboxylate group on the C_α is deprotonated to form a COO⁻ group. Also at this pH the COOH on the functional group in glutamic acid is deprotonated by combining with the OH⁻ group of NaOH to make an additional water molecule, yielding a molecular ratio of 1:1:29 Na⁺/L-glutamic acid⁻/water.

The non-hydrogen-containing intramolecular distances for the glutamic acid molecule were taken from the crystalline structure of 1:1 L-glutamic acid and L-pyroglutamic acid hydrate,⁴⁴ and the intramolecular hydrogen-containing distances were taken from a solid-state neutron diffraction study on crystalline L-glutamic acid.⁴⁵ These molecular distances were used as this structure gave the best initial fit to the neutron diffraction data. A representative L-glutamic acid molecule from the EPSR modeling box is shown in Figure 1. The reference potentials used for the model were derived from a variety of OPLS potentials—those developed by Jorgenson et al. for amines,⁴⁶ the SPC/E potentials for water,⁴⁷ and potentials from Zapatoski et al.⁴⁸ for the Na⁺ ions. Each potential for the EPSR model is listed in Table 2 where the atoms correspond to the labeling scheme shown in Figure 1.⁴⁶ Specifically, O_w and H_w are the oxygen and hydrogen atoms on the water molecule, respectively. For both the glutamic acid molecule and the Na⁺ ions, the charges were adjusted from the original potentials^{46,48} to obtain the appropriate charge balance for the system. Reported potentials for amino acids often carry charges for the backbone carbons and C_α carbon,^{46,49} which is not the case with the EPSR

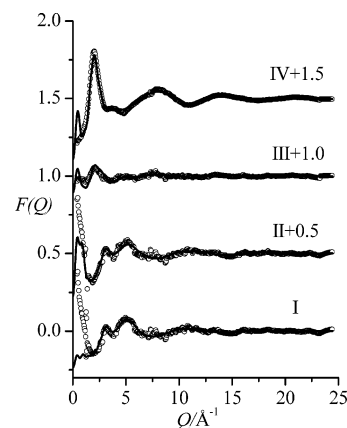


Figure 2. Measured neutron diffraction data, $F(Q)$, (circles) and subsequent EPSR fits to the data (solid line).

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

atom	$\epsilon/\text{kJ (mol}^{-1}\text{)}$	σ (Å)	q (e)
O _w	0.65000	3.166	-0.8476
H _w	0.0	0.0	0.4238
Na	0.12900	2.500	1.0000
C	0.43932	3.750	0.0
C ₁	0.43932	3.750	0.0
C ₂	0.43932	3.750	0.0
N	0.71128	3.400	-0.3000
O	0.87864	2.960	-0.5500
H _x	0.0	0.0	0.5000
H	0.0	0.0	0.0

simulations presented here. However, it has been noted that the exact choices of potentials for these sites are not critical but are dependent upon adjusting the charges to achieve electro-neutrality.⁴⁶ Also, the backbone carbons have been labeled the same (C) while the COO⁻ carbons have been labeled separately (C₁ and C₂) even though the reference potentials for these carbons are identical. This labeling was used because the COO⁻ groups have been shown to be distinguishable from one another by solid-state NMR measurements.¹⁶ Finally the COO⁻ groups and the NH₃⁺ group on each glutamic acid molecule were allowed to rotate in the model about the appropriate C-C or C-N bond.

IV. Results and Discussion

The measured diffraction data, $F(Q)$, along with the EPSR fits to the data are shown in Figure 2 where each sample is labeled with reference to Table 1 and the data have been shifted for clarity. The agreement between EPSR fits and the experimentally obtained structure factors is good in each data set. The total RDFs ($F(r)$'s) are shown in Figure 3, along with the corresponding functions from the EPSR fits to the data where again the data have been shifted for clarity. Because $F(Q)$ is a sum of all the partial structure factors, it is not possible to observe directly each individual atom-atom interaction in the diffraction pattern or in the corresponding RDF. However, from the EPSR model, it is possible to extract each individual partial structure factor and each individual site-site RDF. By inspection of Figure 1 and Table 2, there are a total of 10 unique atoms in the measured diffraction pattern, which give rise to 55 individual RDFs. Here we show the RDFs which are associated with the water molecules and the charged portions of the glutamic acid molecule and the sodium ion while the rest of these functions are shown in the Supporting Information.

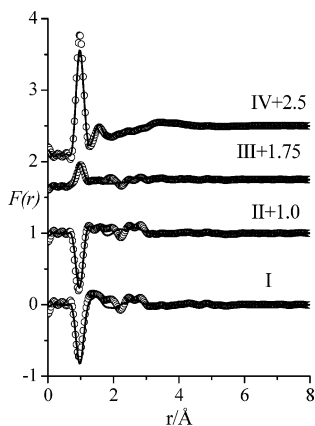


Figure 3. Total RDFs, $F(r)$, (circles) and subsequent EPSR-derived $F(r)$ (solid line).

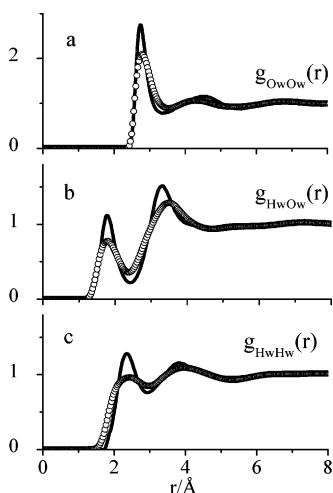


Figure 4. Water–water RDFs for pure water (solid line, ref 23) and in the glutamic acid/water system (circles).

TABLE 3: Coordination Numbers for the Water-Related RDFs for the L-Glutamic Acid/Water System and Pure Water²³

EDF	$n_{\alpha}^{\beta}(r)$ water	$n_{\alpha}^{\beta}(r)$ glutamic acid/water	r_{\min} (Å)
$g_{O_wO_w}(r)$	~4.5–5	4.3	3.51
$g_{O_wH_w}(r)$	~1.8	1.4	2.40
$g_{H_wH_w}(r)$	~4–5	4.1	2.94

A. Water Structure. Figure 4 shows the RDFs for the water–water interactions from the glutamic acid/water system, along with previously published RDFs from EPSR simulations on pure water total structure factor measurements.²³ It is clear from this figure that the water–water correlations are distorted in the glutamic acid/NaOH/water solution when compared with correlations for water in the absence of any solutes. In Figure 4, the first peak in each function is broader than the corresponding peak for pure water, indicating a significant perturbation to the water structure by the solute molecules, L-glutamic acid and Na^+ ions. This perturbation is also evident upon the comparison between coordination numbers ($n_{\alpha\beta}(r)$'s) for this system and for the case of pure water,²⁴ which are listed in Table 3 and follow the nomenclature of eq 4. The addition of glutamic acid and NaOH to water lowers the coordination numbers between the water molecules in the solvent. The most notable of these is the $g_{O_wH_w}(r)$ function (Figure 4b), which shows that the number of hydrogen bonds between water molecules has been decreased from 1.8 in pure water to 1.4 in the glutamic acid/water system. This shows a significant

TABLE 4: Coordination Numbers for Water–Glutamic Acid Radial Distribution Functions

RDF	$n_{\alpha}^{\beta}(r)$	r_{\min} (Å)
$g_{OO_w}(r)$	4.5	3.54
$g_{OH_w}(r)$	3.1	2.40
$g_{H_xO_w}(r)$	1.0	2.25
$g_{H_xH_w}(r)$	2.3	2.79
$g_{NO_w}(r)$	4.3	3.48
$g_{NH_w}(r)$	12.1	3.99

disruption by the solute molecules of the tetrahedral hydrogen-bonding network seen in pure water. This change in the solvent structure can, in part, be attributed to the relative sizes of the solute molecules, both Na^+ (~ 1.0 Å)⁵⁰ and glutamic acid (> 5 Å),⁴⁴ given that the solution is quite concentrated. However, even in the presence of large solute molecules at relatively high concentrations such as *tert*-butyl alcohol³⁸ or dimethyl sulfoxide,³⁶ the bulk water solvent is not significantly disrupted when compared to the addition of glutamic acid and NaOH. The disruption of the tetrahedral bulk water network in this study is primarily due to the charges on both solute molecules. It is notable that this disruption of the water network is similar to that seen in bulk water at high pressures, where the RDFs for water become increasingly broad with the application of pressure (up to 400 MPa).²³

B. Water–Glutamic Acid Interactions. Given that the hydrogen-bonding network in the bulk water solvent has been disrupted by glutamic acid and NaOH, examination of the solute–solvent interactions can elucidate the mechanism of this perturbation. In Figure 5, the water–glutamic acid RDFs for the oxygen and exchangeable (amine) hydrogen sites on the glutamic acid molecules are shown, and Table 4 lists the corresponding coordination numbers for these functions. The salient water–water RDFs, from Figure 4, from the present simulation are superimposed on the glutamic acid–water RDFs where appropriate. It is apparent that the glutamic acid–water interactions have peak maxima at the same positions as the water–water interactions for the analogous RDFs. This not only indicates that the coordination “missing” from the water–water interactions has been substituted by water–glutamic acid interactions but shows that water is strongly bound to the charged sites on the glutamic acid molecule. In both the $g_{OH_w}(r)$ and the $g_{OO_w}(r)$ functions, the first peaks in the RDFs are quite

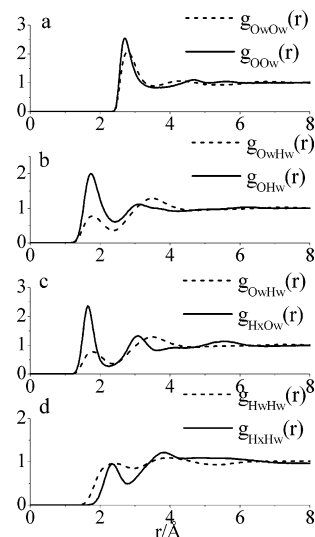


Figure 5. RDFs for glutamic acid H_x and O sites and water compared with salient water–water radial distribution functions.

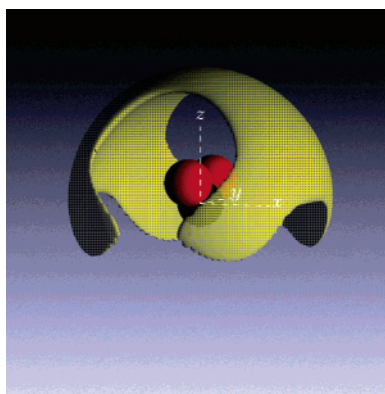
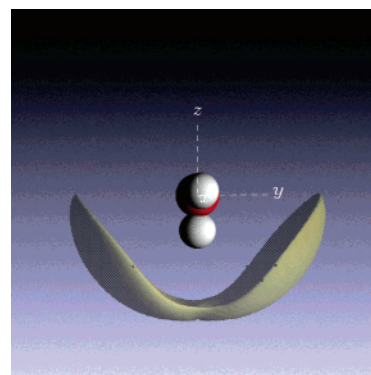


Figure 6. Distribution of water molecules around the COO^- on the α -carbon from L-glutamic acid in solution.

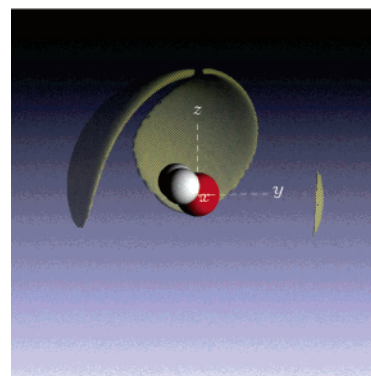
sharp when compared to the corresponding $g_{\text{O}_w\text{H}_w}(r)$ and $g_{\text{O}_w\text{O}_w}(r)$ functions (Figures 5a and 5b), showing a strong correlation between the carboxylate group oxygen atoms from the glutamic acid molecule and the surrounding water solvent. The coordination number of the $g_{\text{O}_w\text{H}_w}(r)$ function at the minimum of the first peak (r_{min}) (Figure 5b) shows, on average, that each oxygen atom from the glutamic acid molecule forms three hydrogen bonds with the surrounding water solvent. Additionally, the amine hydrogens from the glutamic acid molecule (H_x) each form one hydrogen bond with the oxygen atom (Figure 5c) on the water molecule (Table 4).

To understand the positions and the orientations of water molecules and as a result the hydrogen bonding between water molecules and the carboxylate groups in three dimensions, the correlations between these two groups have been determined from the EPSR modeling box. This is accomplished by orienting the COO^- group at the center of the standard laboratory axis, then probing the distribution of water molecules surrounding this portion of the glutamic acid molecule. The *location* of the water molecules relative to the carboxylate group gives rise to a SDF that depicts the location of water molecules around the COO^- group. Full details of this procedure are described elsewhere in the literature,⁵¹ and a brief description is provided above (section II.C). Figure 6 shows the *location* of the water molecules in the extracted SDF surrounding the C_α carboxylate group. The oxygen atoms are seen to bisect the z -axis of the coordinate system with the $\text{C}-\text{O}-\text{C}$ group lying flat in the yz -plane, and the shell surrounding the COO^- on the central axes shows the location of water molecules around this group. Here the contour level of this shell depicts 75% of these molecules at a distance range from 2 to 3.5 Å corresponding with the minimum of the first peak in the $g_{\text{O}_w\text{O}_w}(r)$ function (Figure 5a). The preferred locations of water molecules surrounding the carboxylate group (Figure 6) are either directly above the COO^- group where the z -axis bisects the two oxygen atoms or in front and behind the yz -plane of the group with an absence of distribution in the yz -plane itself directly to the sides of the COO^- group. As expected there is also an absence of density in the xy -plane below the COO^- group where it is bound to the α -carbon. Only the C_α carboxylate group is shown since the functional group carboxylate gives the same distribution.

To understand the *orientation* of the surrounding water molecules in the SDF, the orientation of the dipole moment vector in water can be determined by extracting the appropriate OCF. Figure 6 shows that there is a preferred location for the water molecules surrounding the carboxylate group rather than a random distribution of water molecules, which would result in an isotropic distribution of surrounding water molecules about



a)



b)

Figure 7. Orientation of the water molecules surrounding the α -carbon carboxylate group from L-glutamic acid in solution.

the central axes. As a result, there must be an alignment of the water dipole moment with the COO^- group where, in water, this dipole moment is located along the bisector of the two O_w-H_w bonds. To determine the hydrogen bonding from the water to the oxygen atoms, the orientation of the dipole moment vector in the water molecules relative to the carboxylate oxygen atoms was probed (the OCFs). This orientation was probed at two points in the shell surrounding the carboxylate group (Figure 6), namely, directly above the COO^- group where the z -axis bisects the two oxygen atoms and in the xz -plane directly in front of the carboxylate group in the region 20° below the xy -plane where two lobes are located. Details on the OCFs are provided in section II.C, above.

Figure 7a shows the orientation of the water molecules directly above the carboxylate group (above the z -axis in Figure 6) while Figure 7b shows the orientations of water molecules in front of the xz -plane. Each of the oxygen atoms from the carboxylate group has two hydrogen bonds provided from one water molecule in the xz -plane while the third hydrogen bond is shared between the oxygen atoms (Figure 7a). The water- COO^- coordination distribution shows the resonance nature of the COO^- group itself where the negative charge associated with this group is distributed between the two oxygen atoms. Also this coordination explains the large perturbation in the bulk water structure itself as one of the water-water hydrogen bonds in the molecules at this location has been broken from the solvent water network.

C. Na^+ -Water and Na^+ -Glutamic Acid Interactions. The bulk water structure will not only be perturbed by the presence of the amino acid but the presence of Na^+ ions will also have an effect as has been shown in studies of NaOH /water solutions.^{32,37} Figure 8 shows the RDFs for the Na^+ -glutamic acid interactions (Figures 8a-c) and for the Na^+ ion-water

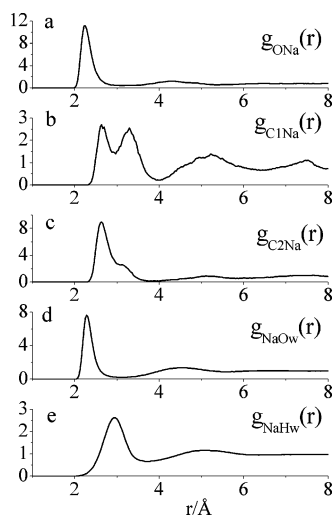


Figure 8. Sodium-containing RDFs for the L-glutamic acid/water system.

TABLE 5: Coordination Numbers for Na⁺–Glutamic Acid Radial Distribution Functions.

RDF	$n_{\alpha}^{\beta}(r)$	r_{\min} (Å)
$g_{\text{NaO}}(r)$	1.0	3.15
$g_{\text{C}_1\text{Na}}(r)$	0.1	2.97
	0.3	3.99
$g_{\text{C}_2\text{Na}}(r)$	0.4	3.82
$g_{\text{NaO}_w}(r)$	4.4	3.09

interactions (Figures 8d and 8e). The corresponding coordination numbers for these functions are shown in Table 5. The Na⁺ ion is fully coordinated from a combination of both the glutamic acid oxygen atoms as well as the oxygen atoms from the water solvent. Complete hydration of the Na⁺ ion is not possible in high concentrations of NaOH in water alone as there are an insufficient number of water molecules to fully hydrate the NaOH.³² To determine the coordination of the sodium ion by the glutamic acid molecules, the $g_{\text{ONa}}(r)$, $g_{\text{C}_1\text{Na}}(r)$, and $g_{\text{C}_2\text{Na}}(r)$ functions are shown (Figures 8a–c). The $g_{\text{ONa}}(r)$ correlation (Figure 8a) shows the closest nearest neighbor Na⁺–glutamic acid distance among the glutamic acid RDFs, which is expected given the respective charges on each of these sites. Each Na⁺ ion is coordinated by approximately four water oxygen atoms and one oxygen atom from a COO[−] group oxygen, though precisely which oxygen provides the additional coordination of Na⁺ cannot be precisely determined by the $g_{\text{NaO}}(r)$ coordination number in Table 5. The $g_{\text{C}_1\text{Na}}(r)$ (Figure 8b) and $g_{\text{C}_2\text{Na}}(r)$ (Figure 8c) functions are shown to determine if the Na⁺ ion was more likely to be found on the amino acid functional COO[−] group or the COO[−] group on the α -carbon in glutamic acid. From Figure 8 and the corresponding coordination numbers in Table 5, there is a preference for the Na⁺ ion to be associated near the functional group carboxylate group rather than the α -carbon carboxylate group. This is likely due to steric reasons where the carboxylate functional group is much less shielded than the carboxylate group at the α -carbon position. However, the α -carbon–Na⁺ RDF (Figure 8b) shows a split peak distribution showing two distinctive distances between these charged ions. It would appear that these distances arise from different glutamic acid molecules, indicating that a percentage of Na⁺ ions are coordinated by more than one amino acid in addition to being hydrated by the bulk water solvent, though the distance of this coordination is large (~ 4 Å). This phenomenon occurs to a lesser degree with the functional group carboxylate where the

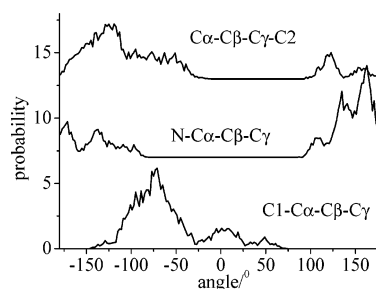


Figure 9. Glutamic acid torsional angle distribution from the L-glutamic acid/water system.

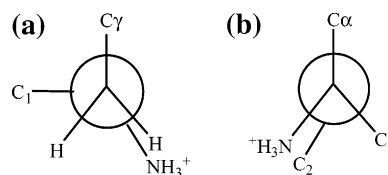


Figure 10. Newman projections from the average torsional angles in the glutamic acid molecule from the L-glutamic acid/water system.

first peak has a smaller second shoulder peak, the $g_{\text{C}_2\text{Na}}(r)$ function (Figure 8c) again showing a similar coordination to the Na⁺ ions by the amino acid itself.

D. Glutamic Acid Conformation. To understand if there is a preferred orientation present in the glutamic acid molecule in solution, three intramolecular torsional angles were extracted from the EPSR simulations, and these angles are shown in Figure 9. Here, the average nitrogen is clearly trans to the C_γ carbon as this angle shows a broad peak between centered at $\sim 150^\circ$. This is confirmed by inspection of the C₁–C_α–C_β–C_γ torsional angle where the α -carbon carboxylate group is cis to the C_γ backbone carbon with the distribution centered on the angle around -60° . However, the C_α–C_β–C_γ–C₂ torsional angle shows a more varied distribution and shows the functional group carbonyl carbon located both cis and trans relative to the α -carbon. Figure 10 shows the Newman projections for these torsional angles with the most likely configuration in Figure 10a showing this projection for both torsional angles about the C_α–C_β axis while Figure 10b shows the torsional angle about the C_β–C_γ axis where the two most likely positions for the C₂ carbon are shown.

V. Conclusions

By showing details of coordination on an atomic length scale between water, glutamic acid, and Na⁺, a detailed description of this amino acid in solution has been elucidated for the first time. This was possible only by combining neutron diffraction measurements with the modeling program EPSR, which uses the diffraction data as a necessary constraint to the resultant model. This method of extracting information from solution on an atomic scale is well proven for other liquids but has not previously been used to determine correlations in a system containing amino acids in solution. Moreover, the use of these techniques provides a time-averaged picture of how water is organized around L-glutamic acid, making this work highly complementary to current NMR and crystallographic techniques. The necessity of studying such small biological molecules in the solution phase cannot be overstated given that most of the processes in life take place in liquid water.

Acknowledgment. We thank the U. S. National Science Foundation for financial support for S.E.M. under award OISE-0404938 and G. Cuello (ILL), A. Wright, and R. Haworth (both

from University of Reading, U. K.) for assistance with the neutron diffraction measurements as well as the Institut Laue Langevin for the allocation of beam time.

Supporting Information Available: RDFs not included in the main text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- Hofmeister, F. *Arch. Exp. Pathol. Pharmacol.* **1888**, *24*, 247.
- Muta, H.; Kawauchi, S.; Satoh, M. *J. Mol. Struct. (THEOCHEM)* **2003**, *62*, 65.
- Hanson, W. M.; Beeser, S. A.; Oas, T. G.; Goldenberg, D. P. *J. Mol. Biol.* **2003**, *333*, 425.
- Kuntz, I. D., Jr. *J. Am. Chem. Soc.* **1971**, *93*, 514.
- Kupce, E.; Freeman, R. *J. Biomol. NMR* **2003**, *27*, 383.
- Kempf, J. G.; Jung, J.-Y.; Sampson, N. S.; Loria, J. P. *J. Am. Chem. Soc.* **2003**, *125*, 12064.
- Canet, D.; Last, A. M.; Tito, P.; Sunde, M.; Spencer, A.; Archer, D. B.; Redfield, C.; Robinson, C. V.; Dobson, C. M. *Nat. Struct. Biol.* **2002**, *9*, 308.
- Ruckert, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, *122*, 7793.
- Pintacuda, G.; Keniry, M. A.; Huber, T.; Park, A. Y.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2004**, *126*, 2963.
- Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Bruschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885.
- Longeville, S.; Doster, W.; Kali, G. *Chem. Phys.* **2003**, *292*, 413.
- Vertesse, B. G.; Magazu, S.; Mangione, A.; Migliardo, F.; Brandt, A. *Macromol. Biol.* **2003**, *3*, 477.
- Svergun, D. I.; Richard, S.; Koch, M. H. J.; Sayers, Z.; Kuprin, S.; Zaccai, Z. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2267.
- Svergun, D. I.; Koch, M. H. J. *Rep. Prog. Phys.* **2003**, *66*, 1735.
- Wijekinha-Betoni, R.; Dobson, C. M.; Redfield, C. *J. Mol. Biol.* **2001**, *312*, 261.
- Lemaitre, V.; de Planque, M. R. R.; Howes, A. P.; Smith, M. E.; Dupree, R.; Watts, A. *J. Am. Chem. Soc.* **2004**, *126*, 15320.
- Kameda, Y.; Sugawara, K.; Usuki, T.; Uemura, O. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 935.
- Sugawara, K.; Kameda, Y.; Usuki, T.; Uemura, O. *Bull. Chem. Soc. Jpn.* **2000**, *73*, 1967.
- Michael, N.; Erfurth, A.; Ohrmann, P.; Gossling, M.; Arolt, V.; Heindel, W.; Pfeleiderer, B. *Psychopharmacology (Berlin)* **2003**, *168*, 344.
- Lemaitre, V.; Pike, K. J.; Watts, A.; Anupold, T.; Samsom, A.; Smith, M. E.; Dupree, R. *Chem. Phys. Lett.* **2003**, *371*, 91.
- Soper, A. K. *Mol. Phys.* **2001**, *99*, 1503.
- Soper, A. K. *Phys. Rev. B* **2005**, *72*.
- Soper, A. K. *Chem. Phys.* **2000**, *258*, 121.
- Soper, A. K.; Phillips, M. G. *Chem. Phys.* **1986**, *107*, 47.
- Soper, A. K.; Silver, R. N. *Phys. Rev. Lett.* **1982**, *49*, 471.
- Benmore, C. J.; Loh, Y. L. *J. Chem. Phys.* **2000**, *112*, 5877.
- Yamaguchi, T.; Benmore, C. J.; Soper, A. K. *J. Chem. Phys.* **2000**, *112*, 8976.
- McLain, S. E.; Benmore, C. J.; Siewenie, J. E.; Urquidi, J.; Turner, J. F. C. *Angew. Chem., Int. Ed.* **2004**, *43*, 1952.
- Soper, A. K.; Egelstaff, P. A. *Mol. Phys.* **1981**, *42*, 399.
- Andreani, C.; Nardone, M.; Ricci, F. P.; Soper, A. K. *Phys. Rev. A* **1992**, *46*, 4709.
- Andreani, C.; Menzinger, F.; Ricci, M. A.; Soper, A. K.; Dreyer, J. *Phys. Rev. B* **1994**, *49*, 3811.
- Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. *J. Chem. Phys.* **2004**, *120*, 10154.
- Bowron, D. T.; Finney, J. L. *Phys. Rev. Lett.* **2002**, *89*, 215508.
- Soper, A. K.; Castner, E. W.; Luzar, A. *Biophys. Chem.* **2003**, *105*, 649.
- Soper, A. K.; Finney, J. L. *Phys. Rev. Lett.* **1993**, *71*, 4346.
- Soper, A. K.; Luzar, A. *J. Chem. Phys.* **1992**, *97*, 1320.
- Bruni, F.; Ricci, M. A.; Soper, A. K. *J. Chem. Phys.* **2001**, *114*, 8056.
- Bowron, D. T.; Soper, A. K.; Finney, J. L. *J. Chem. Phys.* **2001**, *114*, 6203.
- 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; Rutherford Appleton Laboratory: Chilton, Didcot, U.K., 1989.
- Svishchev, I. M.; Kusalik, P. G. *J. Chem. Phys.* **1993**, *99*, 3049.
- Svishchev, I. M.; Kusalik, P. G. *Chem. Phys. Lett.* **1993**, *215*, 596.
- Gray, C. G.; Gubbins, K. E. *Theory of Molecular Liquids: Fundamentals*; Oxford University Press: New York, 1984; Vol 1.
- Sears, V. F. *Neutron News* **1992**, *3*, 29.
- Taira, Z.; Watson, W. H. *Acta Crystallogr., Sect. B* **1977**, *33*, 3823.
- Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *J. Cryst. Mol. Struct.* **1972**, *2*, 225.
- Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569.
- Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
- Zapalowski, M.; Bartczak, W. M. *Comput. Chem.* **2000**, *24*, 459.
- Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
- Maslen, V. W. *Proc. Phys. Soc., London* **1967**, *90*, 259.
- McLain, S. E.; Soper, A. K.; Luzar, A. *J. Chem. Phys.* **2006**, *124*, 074502.