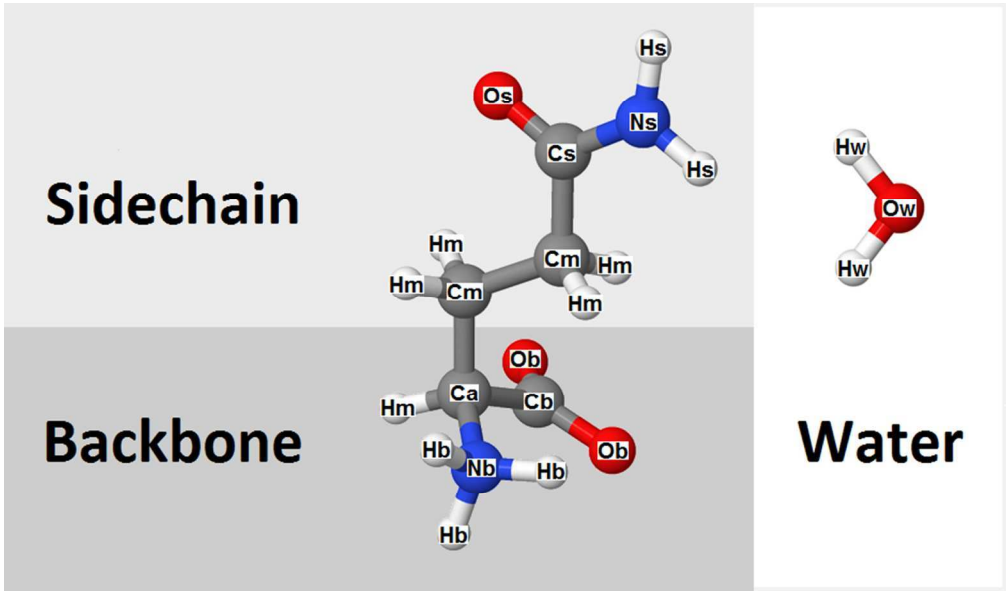
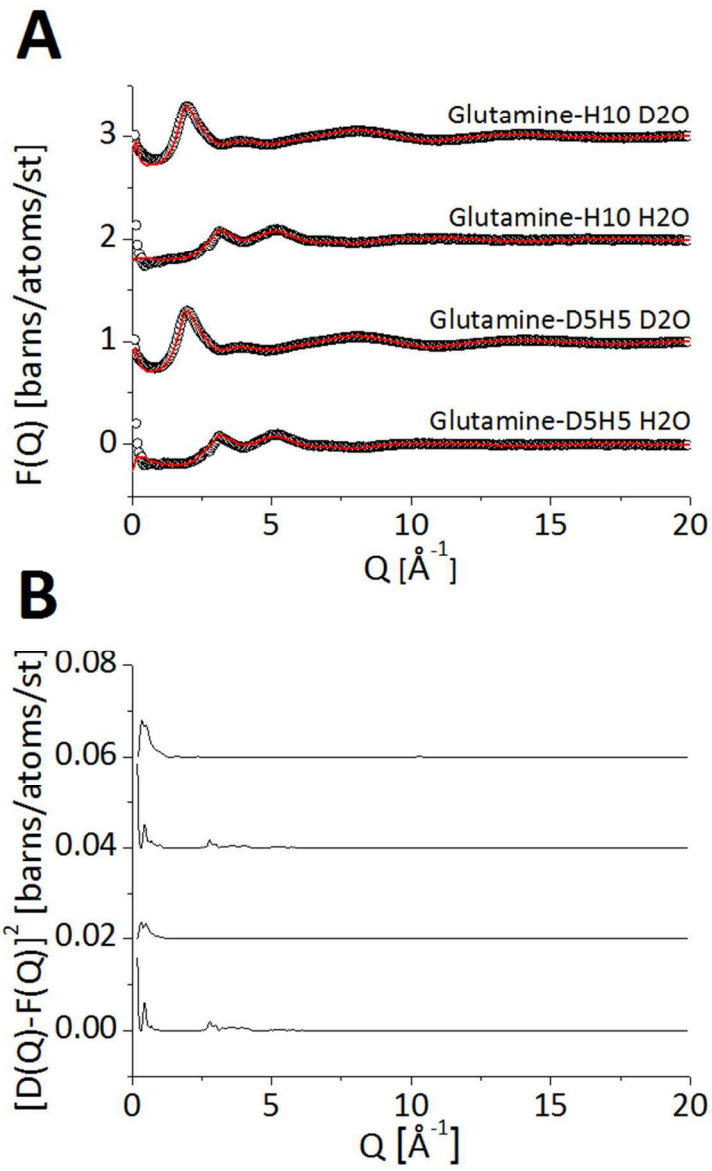
Article:
Rhys, NH, Soper, AK and Dougan, L (2015) Hydrophilic Association in a Dilute Glutamine Solution Persists Independent of Increasing Temperature. Journal of Physical Chemistry B, 119 (51). 51. pp. 15644-15651. ISSN 1520-6106

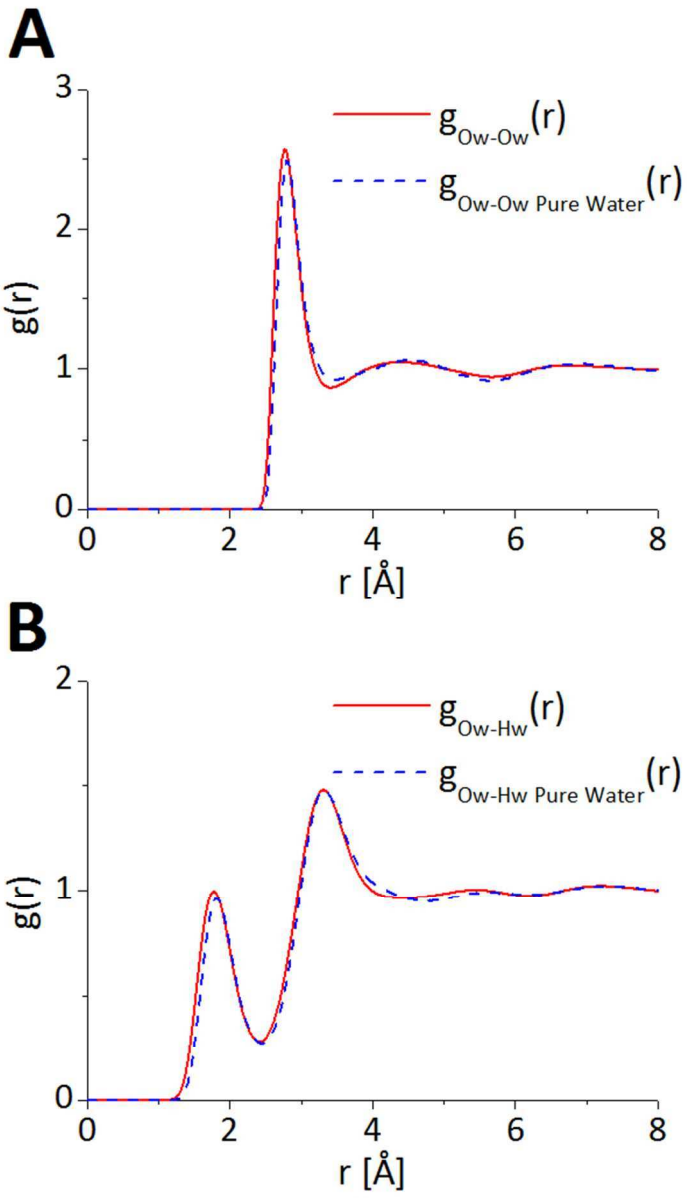
https://doi.org/10.1021/acs.jpcb.5b07413



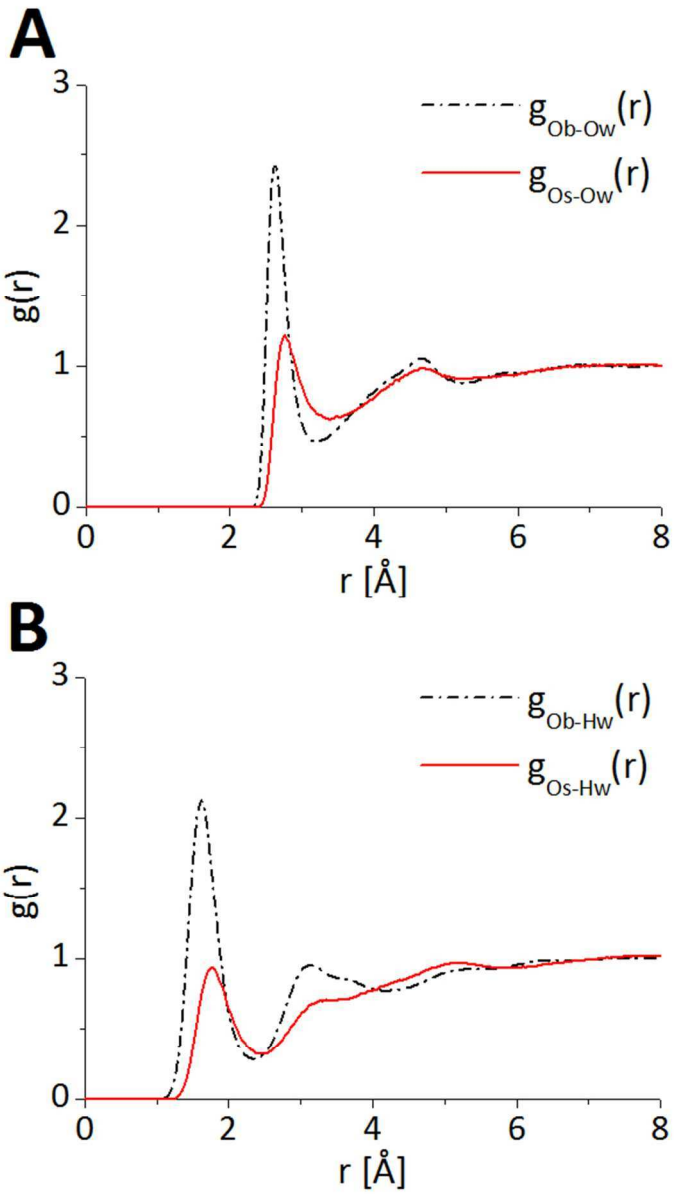
118x69mm (200 x 200 DPI)



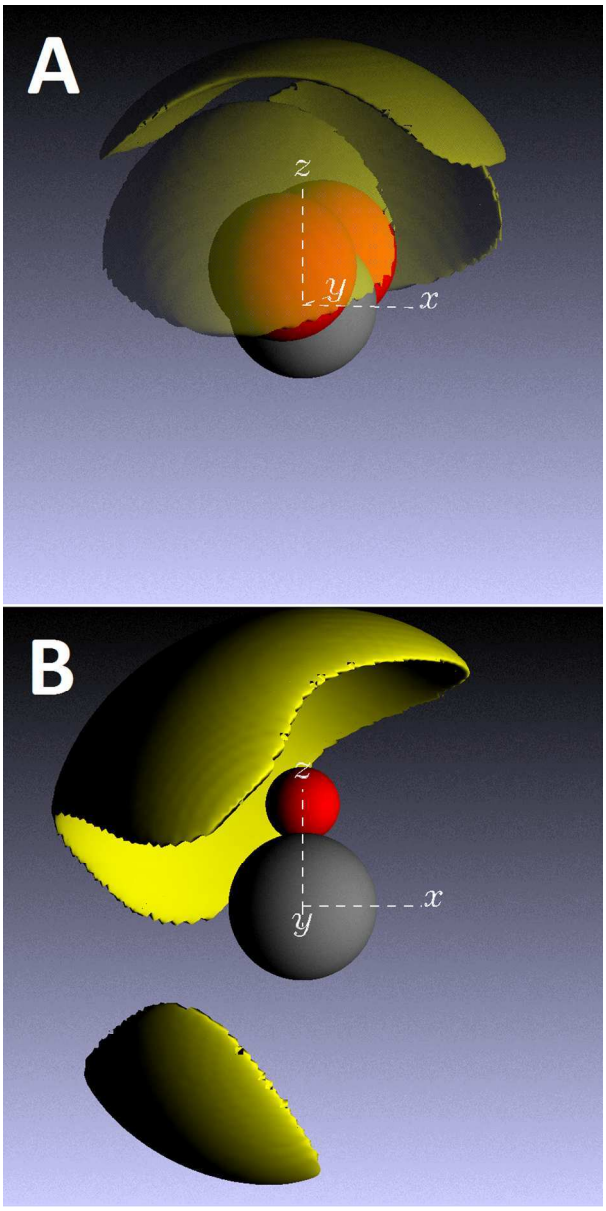
66x107mm (250 x 250 DPI)



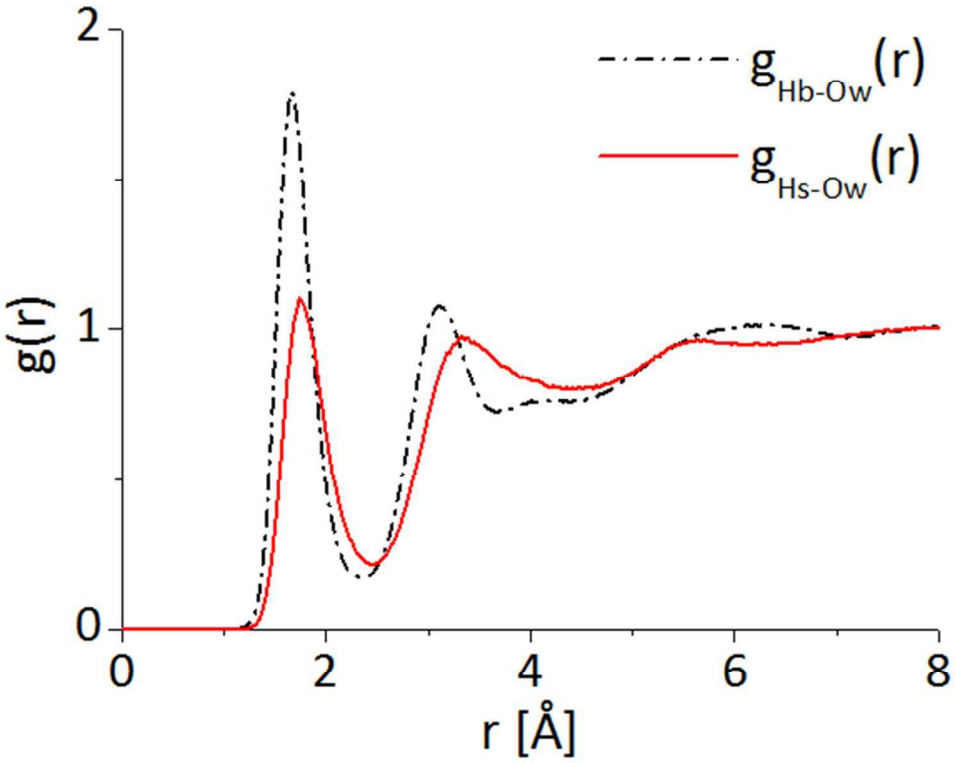
76x131mm (200 x 200 DPI)



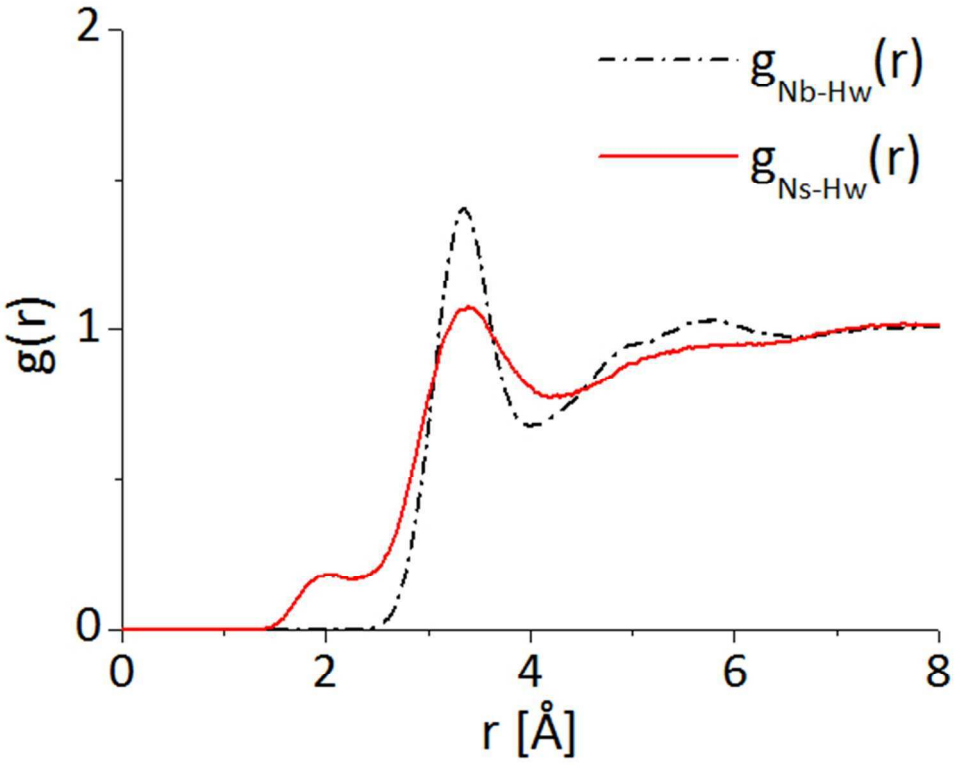
76x131mm (200 x 200 DPI)



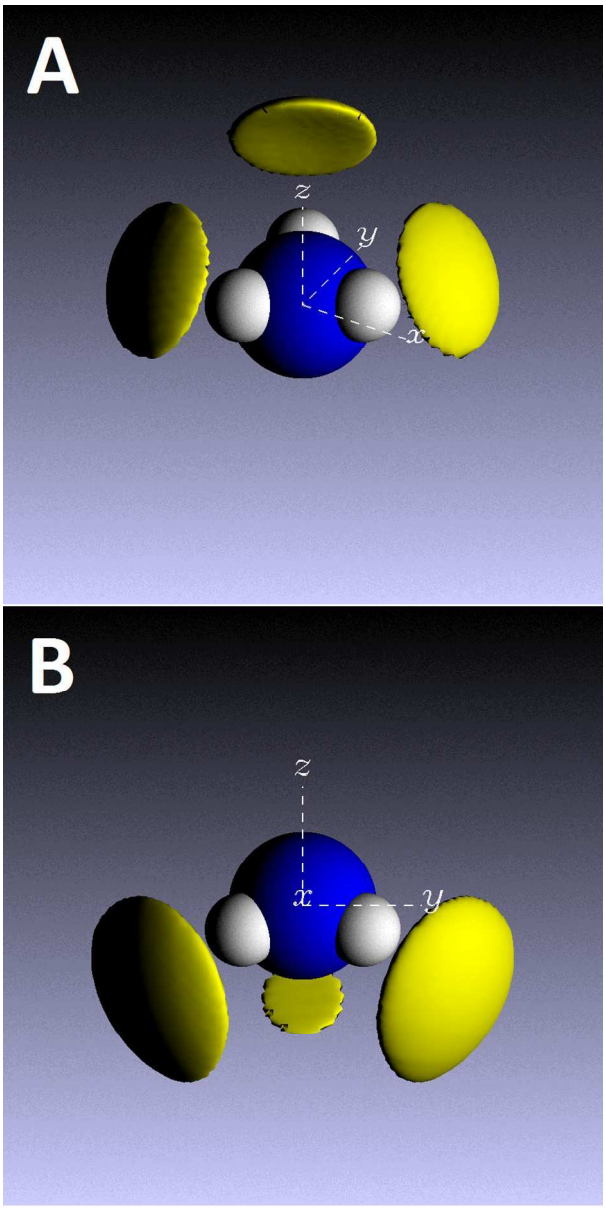
91x183mm (250 x 250 DPI)



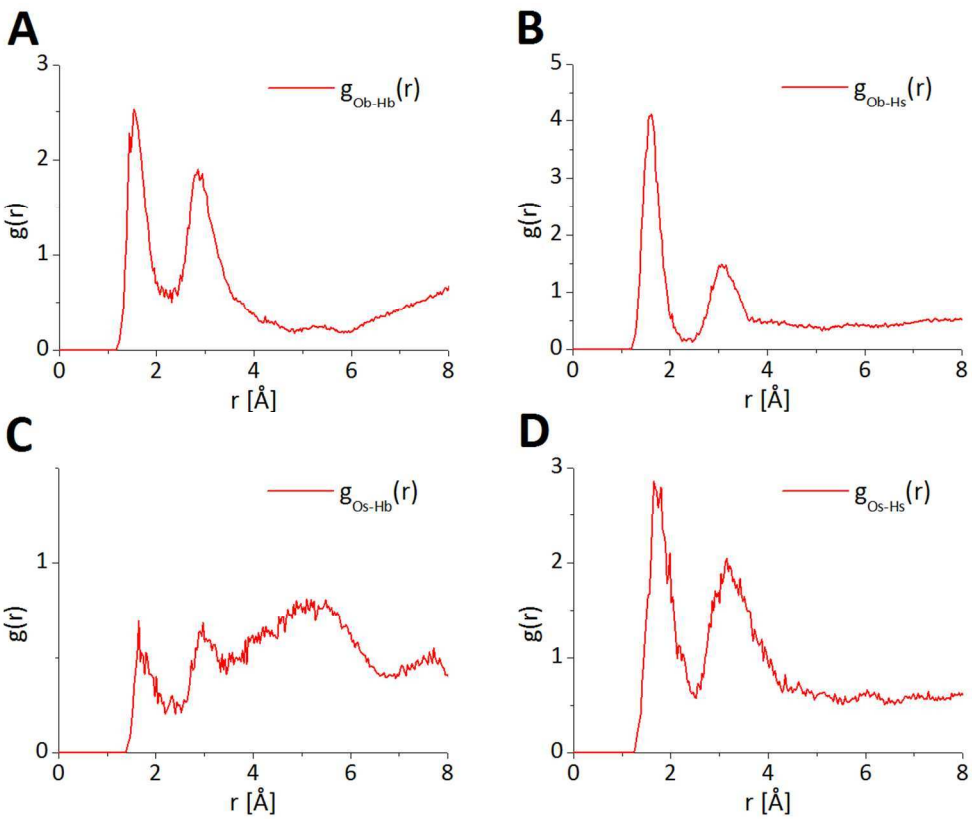
76x59mm (200 x 200 DPI)



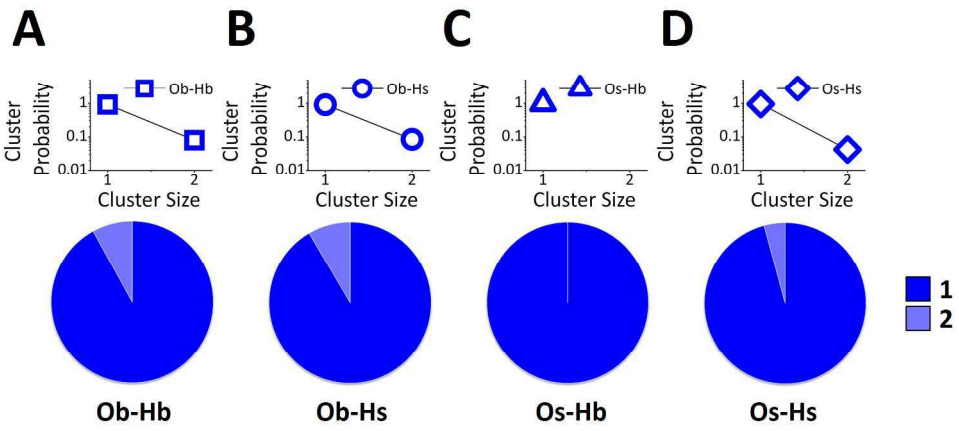
76x59mm (200 x 200 DPI)



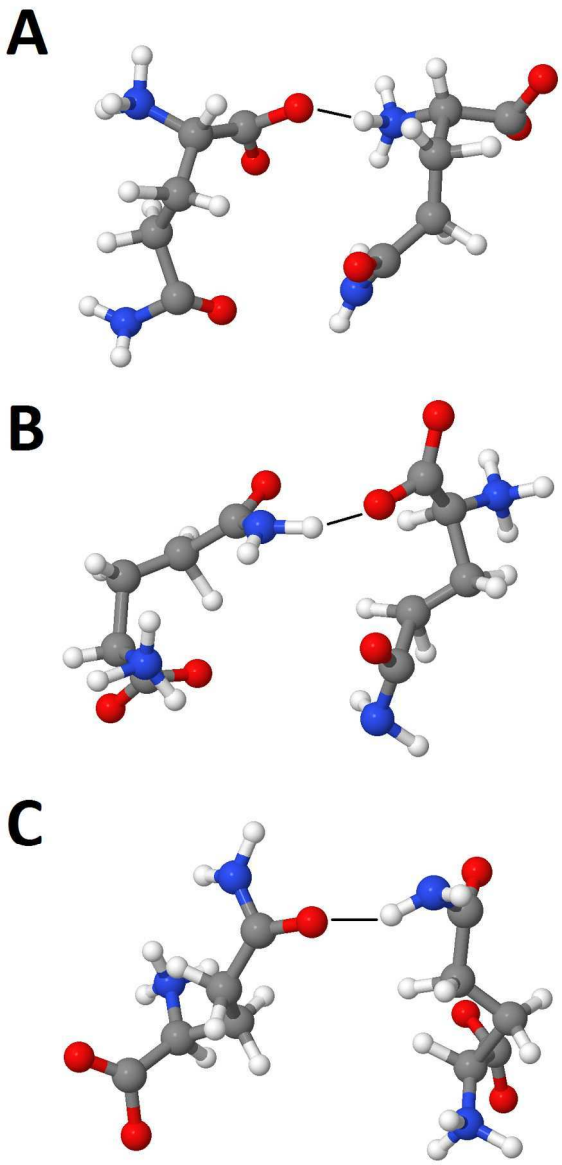
91x183mm (250 x 250 DPI)



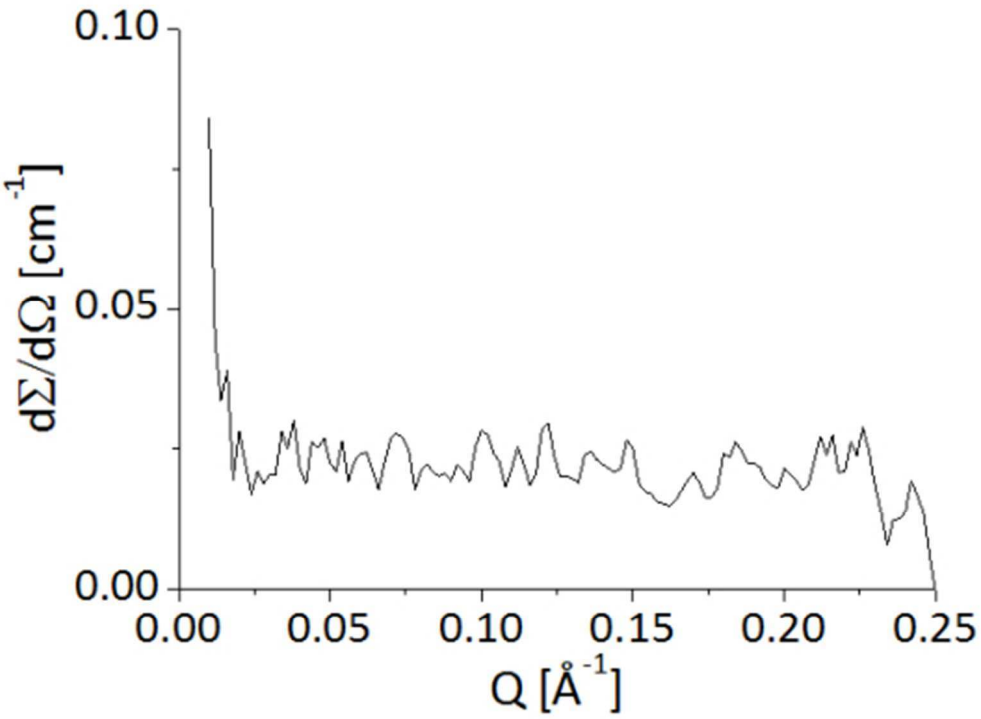
161x131mm (200 x 200 DPI)



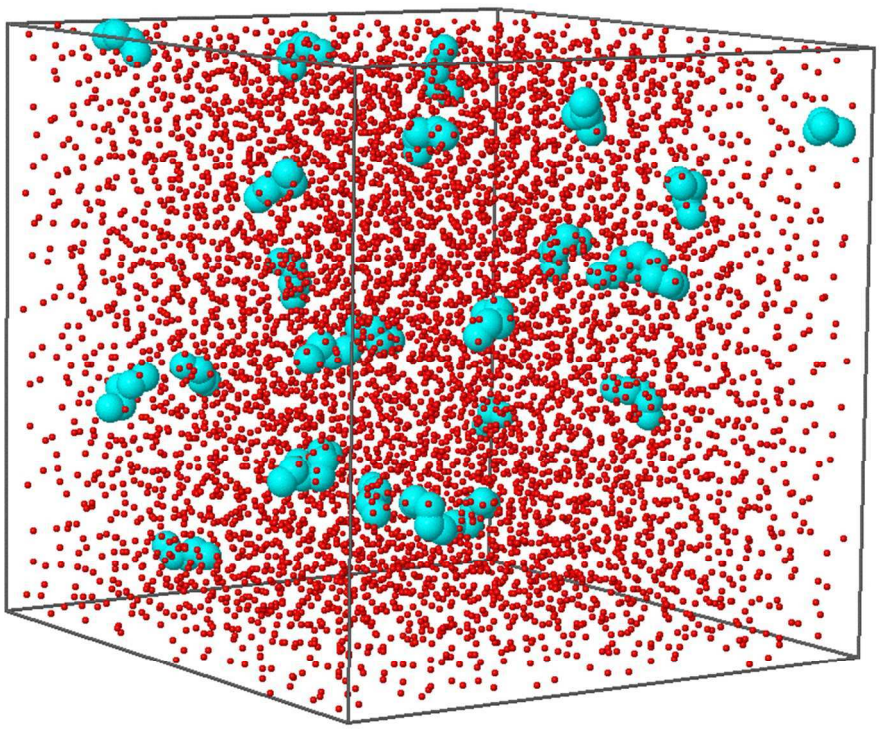
243x108mm (250 x 250 DPI)



118x229mm (250 x 250 DPI)



54x40mm (250 x 250 DPI)



91x73mm (250 x 250 DPI)

The hydrogen bonding ability of the amino acid glutamine revealed by neutron diffraction experiments

N. H. Rhys[*], A. K. Soper[†] and L. Dougan[*]

[*] *School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK*

[†] *ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 OQX, UK*

Abstract

Hydrogen bonding between glutamine residues has been identified as playing an important role in the intermolecular association and aggregation of proteins. To establish the molecular mechanisms of glutamine interactions, neutron diffraction coupled with hydrogen/deuterium isotopic substitution in combination with computational modelling has been used to investigate the structure and hydration of glutamine in aqueous solution. The final structures obtained are consistent with the experimental data and provide insight into the hydrogen bonding ability of glutamine. We find that the backbone of glutamine is able to coordinate more water molecules than the side chain, suggesting that charged groups on the glutamine molecule are more successful in attracting water than the dipole in the side chain. In both the backbone and the side chain we find that the carbonyl groups interact more readily with water molecules than the amine groups. We find that glutamine-glutamine interactions are present, despite their low concentration in this dilute solution. This is evidenced through the occurrence of dimers of glutamine molecules in the solution, demonstrating the effective propensity of this molecule to associate through backbone-backbone, backbone-side chain and side chain-side chain hydrogen bond interactions. The formation of dimers of glutamine molecules in such a dilute solution (30 mg/ml glutamine) may have implications in the aggregation of glutamine-rich proteins in neurological diseases where aggregation is prevalent.

Keywords

Water
Neutron diffraction
Structure
Polyglutamine
Hydration
Hydrogen bond

1. Introduction

L-glutamine[1] is the most abundant naturally occurring, non-essential amino acid in the human body where it is found circulating in the blood as well as in the brain [1]. It is synthesised in the body by the enzyme glutamine synthetase, from glutamate and ammonia [2], and plays an important role in metabolism [3]. In proteins homopolypeptide (HPP) repeat regions of glutamine can exist, which are comprised of neighbouring, repeating glutamine residues. HPP regions of glutamine residues can be found in the amino acid sequence of many proteins, and are most commonly found in proteins responsible for DNA and RNA synthesis [1]. While their importance is yet to be elucidated, the uncontrolled genetic expansion of HPP regions has been associated with the development of a number of serious and debilitating human diseases [1,4]. In particular, polyglutamine (polyQ) regions have been associated with nine neurodegenerative diseases. For example, Huntington's disease (HD) is linked to the insertion of extra glutamines in the protein huntingtin, resulting in a longer HPP region, known as a polyQ expansion [5]. PolyQ expansions in diseased proteins are thought to self-associate to form ordered aggregates and postmortem examinations of HD patients identify large aggregates deposited within brain neurons [6, 7]. Products of proteolysis are rich in polyQ and their aggregation appears to be essential for toxicity [8, 9]. Given its presence in neurological diseases, polyQ is currently the subject of intense scientific research [10-16].

While the aggregation propensity of glutamine-rich proteins has been the subject of many studies [11, 15, 17-19], less is known about the specific interactions between glutamine molecules which might be important for polyQ association. One approach to understanding the stability and aggregation propensity of polyQ is to uncover the molecular mechanisms of the glutamine molecule's interactions with itself, both as an individual molecule and as part of a polyQ chain or protein. Experimental and computational studies have shown that single polyQ chains form collapsed globules in aqueous solutions [20, 21]. Recent experimental studies have also demonstrated that disease associated polyQ stretched preferentially adopt compact conformations [22]. This is surprising given glutamine's ability to hydrogen bond through both the backbone and side chain of the molecule, which might suggest that glutamine-water interactions would be favourable. However, the presence of a collapsed polyQ chain suggests that water is a poor solvent for this HPP chain and glutamine would prefer to minimise any glutamine-water interactions [20]. Indeed, studies have suggested that the stability of collapsed, globular polyQ chains results from an extensive hydrogen bonding network between glutamine molecules, where glutamine-glutamine hydrogen bond interactions outweigh interactions with the surrounding solvent environment [21]. As well as studies on single polyQ chains work has focused on polyQ disease-related proteins. A recent small angle neutron scattering study on a truncated version of the Huntingtin protein containing a polyQ chain with 42 residues found that dimers and trimers formed by the truncated protein were smaller than would be expected if they were random coils. Instead they were spherical and compact, again suggesting water is a poor solvent for a polyQ chain [16].

Given a number of studies suggest that water is a poor solvent for polyQ chains and glutamine-glutamine interactions are more favourable than glutamine-water interactions, it is interesting to consider how glutamine interacts with other glutamine molecules. Recent work has implicated a major role for hydrogen bonding in the side chains of glutamine in the irreversible aggregation of a protein [23]. The protein Ataxin-3 consists of a globular domain and a stretch of consecutive glutamines that are thought to trigger the neurodegenerative

[1] In this paper we will refer to L-glutamine as glutamine

disorder spinocerebellar ataxia type 3, when it is expanded beyond a critical threshold. This work suggested that side chain-side chain hydrogen bonding between glutamine residues was responsible for irreversible aggregation of the protein Ataxin-3. The importance of glutamine side chain hydrogen bonding has also been identified in a density functional theory study where the formation of glutamine-rich aggregates were examined [24]. This work proposed that hydrogen bonding between side chains of glutamines contribute to the stabilising energy and formation of β-sheets in the aggregates and the lowering of distortion energy. This study also proposed that side chain – side chain hydrogen bonds between glutamine are more cooperative than backbone -backbone hydrogen bonds. This is in agreement with molecular dynamics simulations of polyQ chains have demonstrated that the association of glutamine residues is driven by glutamine side chain hydrogen bond interactions, resulting in the formation of larger aggregates. [25]

The prevalence and importance of glutamine-glutamine interactions in all of these studies highlights the need to obtain direct, structural insight into the hydrogen bonding ability of glutamine. An extensive study of hydrogen bonding motifs involving glutamine in the solid state has been completed, using a data set of 1370 protein crystal structures, revealing common structural motifs determined by the hydrogen bonding between the glutamine residue side chain and backbone [26]. However, no such experimental, structural study has been completed in the liquid state for the single glutamine molecule in aqueous solution. Further insight is needed into the structural conformation of the glutamine molecule in the liquid state and its ability to hydrogen bond both with itself and the surrounding liquid environment. This information is important for evaluating the length scales and magnitude of intra- and inter-molecular hydrogen bonds in the system. An understanding of the solvation properties of glutamine in the liquid state is the necessary first step in building a molecular level framework to structurally characterize polyQ expansion diseases. The availability of new and detailed experimental structural information on glutamine will provide a fundamental benchmark for spectroscopic and computational studies aimed at characterising the properties of systems involving glutamine.

In the present study we use experimental and computational methods to allow for the determination of a complete set of partial radial distribution functions for glutamine molecules in the liquid state. Neutron diffraction is a suitable probe for the structural study of aqueous solutions due to the large scattering cross section of deuterium and the high contrast achievable using H/D selective substitution on specific hydrogen sites in the molecule. Previous work on small biological molecules in solution has demonstrated that this is a powerful approach for elucidating hydration and association behaviour [27-35]. The main goal of this work was to obtain high quality structural data of glutamine in aqueous solution at 297 K to allow determination of the conformation of the molecule and investigation of the hydration of glutamine by water. The combination of isotopic substitution neutron diffraction and computational modelling was employed to obtain the average structural interactions of glutamine in solution on a local length scale (1-10 Å). Specifically, the modelling is a three-dimensional molecular reconstruction using the Empirical Potential Structure Refinement (EPSR) method which is constrained by experimental neutron diffraction data, described in the following section. In addition, we have completed small-angle neutron scattering experiments of glutamine in solution to determine if any long-range structures (25 – 800 Å) are formed in the solutions.

2. Methods

2.1 Neutron diffraction experiments

Neutron diffraction experiments were completed on the SANDALS time-of-flight diffractometer at the ISIS pulsed neutron facility at the Rutherford Appleton Laboratory, United Kingdom. The instrument SANDALS is a total scattering neutron diffractometer optimised for the study of liquids and amorphous samples containing hydrogen [34, 35]. The physical quantity measured by the diffractometer is the differential scattering cross section $d\sigma/d\Omega$ as a function of the exchanged wave vector Q (defined as the difference between the incident and the scattered neutron wave vectors). The SANDALS instrument collects data with a Q range of 0.1 – 50 $Å^{-1}$. Through the theory of neutron scattering [36], it is possible to relate $d\sigma/d\Omega$ to the static structure factor $S(Q)$, which is the Fourier transform of the atomic pair distribution function $g(r)$. The function $g(r)$ provides information on how atomic densities vary as a function of radial distance, r, from any particular atom [37]. Neutron diffraction experiments are combined with isotopic substitution to allow labelling of individual atomic sites in a molecule and the extraction of $g(r)$ [38].

Deuteriated samples of water and protiated samples of L-glutamine were obtained from Sigma-Aldrich. The deuterated water (> 99.99 % purity) was molecular biology grade and the protiated glutamine sample was molecular biology grade (≥ 99.5 % purity). Partially deuterated samples of L-glutamine, where the methyl hydrogens are deuterated, were obtained from Cambridge Isotope Laboratories (> 97 % purity). Ultrapure H_2O was obtained from a Millipore purification system. A total of 4 isotopically distinct samples were measured at standard temperature and pressure (297 K and 1 bar) and are listed in Table 1. Each sample had a concentration of 30 mg/ml, close to but below the solubility limit of glutamine [39]. The samples were prepared by weight and then placed in flat cells made from a titanium-zirconium alloy which gives negligible coherent scattering signal. The cells were mounted on an automated sample changer to cycle through the samples. The typical collection time for each sample is ~ 8 hours. The differential scattering cross-section for each sample was obtained by normalising to a vanadium standard. Corrections for attenuation and multiple scattering were made using the Gudrun program suite which has been detailed previously [40]. A further correction for inelastic scattering was made and has been described in detail elsewhere [41]. Neutron diffraction on solutions yields the quantity, $F(Q)$, which is the total interference differential scattering cross section, and which is the sum of all partial structure factors $S_{\alpha\beta}(Q)$ present in the sample each weighted by their composition c and scattering length b,

$$F(Q) = \Sigma_{\alpha\beta} c_\alpha c_\beta b_\alpha b_\beta (S_{\alpha\beta}(Q) - 1)$$

where Q is the magnitude of the change in momentum vector by the scattered neutrons ($Q = (4\pi//\lambda) \sin\theta$). Fourier transform of $S_{\alpha\beta}(Q)$ gives the respective atom-atom radial distribution functions (RDFs) $g_{\alpha\beta}(r)$, and integration of $g_{\alpha\beta}(r)$, gives the coordination numbers of atoms α around atoms β between two distance r_1 and r_2.

2.2 Computational modelling

In this paper we refer to 2 distinct atomic components in water and 11 distinct atomic components in glutamine molecules (see Figure 1). For glutamine the carbon atoms are labelled Cb, Cs, Cm and Ca, the oxygen atoms Ob and Os, the nitrogen atoms Nb and Ns and amine hydrogen atoms, Hb and Hs, and the methyl hydrogen atoms, H. Glutamine is a polar amino acid and contains the backbone (CO_2CHNH_3) and an uncharged side chain ($CH_2CH_2CONH_2$). The backbone of glutamine consists of charged groups, whilst the side chain amide group is a dipole system. In aqueous solution glutamine is in the zwitterionic form with the amine group deprotonating the carboxylic acid group on the backbone, resulting in a charged backbone. In Figure 1 the glutamine molecule is shown in its zwitterionic form where the amine group is protonated to form a NH_3^+ group and the carboxylic acid group has been deprotonated to form a COO^- or carboxylate group to be consistent with the average structure of glutamine in solution. The side chain contains an amide group, constituting a carbonyl (CO) and amine portion (NH_2). This amide group can hydrogen bond via two lone pairs on the carbonyl oxygen, and one on the amine nitrogen, as well as two hydrogen atoms on the amine section. The amine and carbonyl parts of the side chain amide group will be analysed separately in this work. The water oxygen (Figure 1) is labelled Ow and the water hydrogen Hw. A full structural characterization of the system requires the determination of 91 RDFs, which is well beyond the capability of any existing diffraction techniques by themselves.

To build a model of glutamine and water liquid structure, the experimental data are used to constrain a computer simulation. In the simulation the empirical potential is obtained directly from the diffraction data. This potential drives the structure of the three-dimensional model toward molecular configurations that are consistent with the measured partial structure factors from the neutron diffraction experiments. The diffraction data were interpreted via a computer simulation procedure, Empirical Potential Structure Refinement (EPSR) [42]. EPSR aims to produce a model with a simulated differential scattering cross section ($D_i(Q)$) which fits the experiment results as closely as possible. Given that in this case there are more site-site radial distribution functions than diffraction data sets, extra information is required to define the structure. This is achieved by forcing the glutamine molecules in the simulation box to adopt the expected molecular geometries. A reference interaction potential is used which serves to generate hydrogen bonding between the relevant atom sites and to prevent atomic overlap at unrealistic distance ranges. The combined empirical and reference potentials do not guarantee the final reconstruction of the structure is unique, but they do ensure it is consistent with the diffraction data as well as being physically plausible.

For the simulations, a total of 25 glutamine molecules and 6725 water molecules were contained in a cubic box of the appropriate dimension to give the measured atomic number density of 0.099511 at 297 K. The intra-molecular structure of the glutamine molecule has been defined using the crystal structure determination of glutamine [43]. Table 2 shows the intra-molecular geometry for the glutamine molecules in the EPSR simulation. A three-dimensional computer model of the solution is constructed and equilibrated using relevant interaction potentials. The water molecules were produced using the SPC/E model [44] and the OPLS potentials have been used for glutamine [45]. The charges and Lennard-Jones constants are shown in Table 3. Periodic boundary conditions were imposed and the Coulomb interactions are truncated by means of a derivative of the reaction field method [46], and other interactions are truncated as described previously, using a radial cut-off of 12 Å in both cases [47]. Information from the diffraction data is then introduced as a constraint whereby the

difference between observed and calculated partial structure factors enters as a perturbation potential to drive the computer model (via Monte Carlo updates of atomic positions) towards agreement with the measured data. The perturbation is refined in successive iterations of the procedure until a satisfactory fit is obtained. The ESPR simulations were run under the same conditions and equilibrated for ~10 000 Monte Carlo cycles. In this way an ensemble of three-dimensional molecular configurations of the mixture is generated which exhibit average structural correlations that are consistent with the diffraction data.

Small angle neutron (SANS) measurements were also taken on protiated glutamine in D_2O at standard temperature and pressure (298 K and 1 bar) using the LOQ instrument at the ISIS pulsed neutron facility at the Rutherford Appleton Laboratory, UK [48]. Experiments were completed at the same concentration measured by neutron diffraction (30 mg/ml). Samples were prepared by weight and transferred into Hellma fused silica spectrophotometry cuvettes with a 1 mm path length. The samples were measured for ~ 2 hours on LOQ. As a background, a similar cuvette filled with pure D_2O was measured and was subsequently subtracted from the glutamine-water data. After accounting for detector efficiency and pixel solid angles, sample transmission, illuminated volume and the incident flux, the differential cross-section was determined [49].

3. Results

The experimental structure factors $F_i(Q)$ (black circles) and the fits obtained with the EPSR method $D_i(Q)$ (solid line) for all solutions are shown in Figure 2. Minor discrepancies are observed in the low Q region and are caused by difficulties in removing the effect of nuclear recoil from the measured data. However, this recoil effect is expected to have a monotonic dependence on Q and so does not influence the model structure to any significant extent. The simulated structure factors in Figures 2 show very good agreement with the experimental data. For a simulation where M data sets have been refined simultaneously, the quality of the fit R is defined as

$$R = \frac{1}{M}\sum_i \frac{1}{n_Q(i)}\sum_Q \left[D_i(Q) - F_i(Q)\right]^2$$

where $D_i(Q)$ is the experimental interference differential cross section, $F_i(Q)$ is the simulated structure factor the i^{th} dataset and $n_Q(i)$ is the number of Q values within the i^{th} dataset. An R-factor close to zero denotes a good fit. The R-factor we obtain is 2.2×10^{-3}. The quality of fit for the EPSR simulation can be compared visually by plotting residuals squared $([D_i(Q)-F_i(Q)]^2)$ for all datasets as a function of Q (Figure 2 B).

Below we focus on some of the RDFs associated with water structure, the glutamine-water interactions with both the backbone and side chain of glutamine, as well as the interactions between glutamine molecules. Figure 3 shows the RDFs for the water-water interactions for the aqueous glutamine solution taken from the EPSR analysis of neutron diffraction data. The site-site partial radial distribution functions shown are $g_{Ow\text{-}Ow}$ (A) and $g_{Ow\text{-}Hw}$ (B) (solid lines). Using previously published neutron diffraction data[50] we have completed an EPSR simulation of pure water (dashed lines) for comparison. In each case the RDFs are similar to those for pure water indicating that there is no dramatic change in the bulk water structure upon the addition of small quantities of glutamine. The position of the 1^{st} peak in the RDFs for each function shows little change between aqueous glutamine and pure water (see Table

4). Inspection of the coordination numbers for the first shell for aqueous glutamine solution and pure water, also shown in Table 4, indicate a small reduction in coordination number in aqueous glutamine. For $g_{Ow\text{-}Ow}(r)$ the coordination number is 4.8 for aqueous glutamine as compared with 5.2 for pure water. While larger changes to the first coordination shell in water have been observed for other amino acids and dipeptides in solution at higher concentrations [29-32], there is no strong effect on water structure in the present study. The second peak in the $g_{Ow\text{-}Ow}(r)$ function of aqueous glutamine is slightly flattened compared to the same peak in pure water (Figure 3A). This peak corresponds to the second nearest neighbour distance for water molecules and its flattening in aqueous glutamine may indicate a reduction in available water molecules available to form a second coordination shell, although this is not seen in the measured coordination numbers in Table 5, where the coordination number slightly increases from 18.9 in pure water to 19.4 in aqueous glutamine. Interestingly, similar studies of glutamic acid showed that the addition of glutamate to solution resulted in a marked disruption to the bulk water structure at concentrations of 1:28 glutamic acid-water [31]. However, it is worth noting that in this study the solution also contained Na^+ ions, which are known to have a marked effect on water structure [51, 52].

Next we consider hydrogen bonding between glutamine and water. The glutamine molecule is capable of forming hydrogen bonds through the carbonyl groups and the amine groups. The carbonyl groups include the backbone carboxylate anion COO^-, labelled atoms Cb and Ob in Figure 1, and the amide carbonyl on the side chain CO, labelled atoms Cs and Os. To understand the hydrogen bonding ability of glutamine with water we first examine the site-site partial radial distribution function $g_{O\text{-}Ow}$. Figure 4 shows the RDF for the backbone $g_{Ob\text{-}Ow}$ (dashed line) and side chain $g_{Os\text{-}Ow}$ (solid line). The first peak in the $g_{O\text{-}Ow}$ (r) occurs at 2.63 Å for the backbone and 2.76 Å for the side chain indicating the backbone-water hydrogen bond interactions are more strongly correlated than the side chain-water hydrogen bond interactions. The coordination numbers for the RDFs are shown in Tables 6-7 showing the coordination numbers of backbone $g_{Ob\text{-}Ow}$ is 2.9 while for the side chain $g_{Os\text{-}Ow}$ the coordination is only 2.5

The first coordination shell distances measured for glutamine-water interactions are also slightly shorter than the equivalent correlation for the water-water interaction, namely $g_{Ow\text{-}Ow}(r)$ (Figure 3). These shortened distances between water and the glutamine molecule indicate a strong correlation between water and the carbonyl groups in glutamine. However, the intensities of the peaks in $g_{O\text{-}Ow}$ (r) (Figure 4) are slightly reduced compared to the analogous water-water correlation $g_{Ow\text{-}Ow}$. This reduction in intensity is also observed in the coordination numbers for $g_{O\text{-}Ow}$ (r) (Table 6-7), where the backbone and side chain oxygen atoms accept 2- 3 hydrogen bonds from the surrounding water molecules of glutamine. When compared with the coordination number of water in this solution (Table 4) where each oxygen atom accepts 4.8 hydrogen bonds, it is clear that the geometric constraints of orientating water molecules around the glutamine carbonyl group, leads to a reduction in surrounding water molecules. To visualize the hydrogen bonding between the carbonyl group in the backbone and side chain and the water molecules in three dimensions, the spatial density function (SDF) between the groups have been determined from the EPSR simulation. The distribution of water molecules around the backbone carbonyl group was analysed by placing the COO^- group of the backbone at the centre of the laboratory axis, with oxygen atoms lying in the zy-plane, while for the side chain the CO group was central in the laboratory axis, with the z axis set to run through the oxygen. From these central axes the positions of water molecules were probed giving rise to a SDF, which depicts the location of water molecules in the three-dimensional space around each carbonyl group. Figure 5 shows

the resulting SDF for the backbone (A) and the side chain (B). For the backbone, the surface contour of the water shell encloses 25 % of the water molecules from 2 to 4.28 Å. For the side chain, 25% of the water molecules are enclosed from 2 – 4.41 Å. These distances correspond to the minimum of the first peak in the corresponding *C-O$_W$ (r)* function (Figure SI 1). For the backbone carbonyl group the preferred location of the water molecules in the first coordination shell is either directly over each oxygen atom or in the zx plane, with density found above and on each side of the group. There is absence of density below the COO^- group where this group is bound to a carbon. In the case of the side chain, one extended area of water density is found above the oxygen atom, corresponding to bonding via the two lone pairs on the oxygen. An additional area of density is found to one side of the carbon atom due to interactions with the side chain amine group. The density of the water shell for the backbone and side chains in the SDFs further illustrate that the backbone has more interactions with water than the side chain. Below we compare backbone and side chain interactions between glutamine molecules to determine whether this trend continues. This will help to shed light on the intermolecular interactions which might favour glutamine–glutamine interactions over glutamine-water interactions, which may be important in understanding the self-assembly of glutamine in larger structures.

As well as hydrogen bonding through the carbonyl groups the glutamine molecule is capable of hydrogen bonding through the amine groups to the surrounding water molecules. In the glutamine molecule the amine groups include the backbone ammonium cation NH_3^+, labelled atoms Nb and Hb in Figure 1, and the amine on the side chain NH_2, labelled atoms Ns and Hs. Water and glutamine can hydrogen bond via donation from NH_3^+ and NH_2 to the surrounding water molecules. Figure 6 shows the hydrogen bond interactions between glutamine hydrogen atoms and water oxygen atoms in the backbone $g_{Hb\text{-}Ow}$ (dashed line) and the side chain $g_{Hs\text{-}Ow}$ (solid line). The first peak in the $g_{H\text{-}Ow}$ *(r)* are similar occurring at 1.66 Å for the backbone and 1.74 Å for the side chain. Again, the backbone first coordination shells are at a smaller distance than that of the side chain, indicating the backbone-water hydrogen bond interactions are more strongly correlated. The number of hydrogen bonds between Hb and Hs and Ow is ~ 1 for both the backbone and the side chain (Table 6). Next, we consider the interactions between the nitrogen in the backbone Nb and side chain Ns, with the surrounding water molecules. The $g_{N\text{-}Hw}$ *(r)* functions are shown in Figure 7 for the backbone (dashed line) and the side chain (solid line). A peak can be seen at 3.35 Å for $g_{Nb\text{-}Hw}$ *(r)* for the backbone, with a corresponding coordination number of 10.7 (Table 8). In the case of the side chain – water interactions a peak in the g_{NsHw} (r) is observed at 3.39 Å, with a coordination number of 13.4, corresponding to the hydrogen bond interactions between the nitrogen on glutamine and hydrogen on water. In addition to this peak, a smaller peak is observed at ~ 2 Å in $g_{Ns\text{-}Hw}$ *(r)* (Figure 7), which may be indicative of hydrogen bonding between the lone pair of electrons on the nitrogen, allowing the possibility of accepting a hydrogen bond from water.

The SDF for the water molecules surrounding the NH_3^+ group of the backbone and the NH_2 group of the side chain are shown in Figure 8. For the backbone, the surface contour of the water shell encloses 25 % of the water molecules from 2 to 3.42 Å. For the side chain, 25% of the water molecules are enclosed from 2 – 3.45 Å. These distances correspond to the minimum of the first peak in the corresponding *N-Ow (r)* function (Figure SI 2). In each SDF plot the nitrogen atom is located at the origin of the central axes, with one of the amine hydrogen atoms in the zy plane. From Figure 8 the most likely location of water molecules around the NH_3^+ group in the backbone is over each hydrogen atom. This suggests that the hydrogen atoms successfully donate a hydrogen bond to surrounding water molecules. In the

case of the NH_2 side chain the most likely location of water molecules is over each of the hydrogen atoms. In this SDF the distribution is reflective of the fact that there are only two hydrogen bonds available to donate a hydrogen bond to the surrounding water molecules. Due to the presence of a lone pair of electrons on the nitrogen there is also the possibility of accepting a hydrogen bond, shown as a small lobe of density in the centre, back of the SDF (Figure 8B). This hydrogen bonding can also be observed as a small peak in the $g_{Ns\text{-}Hw}(r)$ function in Figure 7.

Although the concentration of glutamine in the solution is small it is interesting to examine the RDFs for glutamine-glutamine interactions to fully understand the hydrogen bonding ability of this molecule. We consider the interactions between glutamine oxygen and glutamine hydrogen atoms, for which there are 4 pair distributions namely $g_{Ob\text{-}Hb}(r)$, $g_{Ob\text{-}Hs}(r)$, $g_{Os\text{-}Hb}(r)$ and $g_{Os\text{-}Hs}(r)$ (Figure 9). Tables 9 and 10 show the corresponding coordination numbers for these functions. In each of the RDFs shown in Figure 9 there are prominent intermolecular correlations between carbonyl and amine groups in the solution. Three of the oxygen – hydrogen RDFs in Figure 9 display two prominent peaks, indicating either two different hydrogen bonding distances present or one bonded hydrogen and one non-bonded hydrogen interaction. In the case of the backbone – backbone interactions (Figure 9A) the first peak in the RDF $g_{Ob\text{-}Hb}(r)$ occurs at 1.53 Å, while for side chain –side chain interactions (Figure 9D) the first peak in the RDF $g_{Os\text{-}Hs}(r)$ occurs at 1.65 Å (see Table 9). We can compare this with the first peak position in the RDF of oxygen hydrogen interactions for water in aqueous glutamine, where $g_{OW\text{-}HW}(r)$ has a first peak position at 1.77 Å. Thus the backbone-backbone RDF first coordination shells are at a smaller distance than that of both the side chain – side chain and water-water RDF first coordination shells, indicating the backbone-backbone hydrogen bond interactions are more strongly correlated. This trend continues for the second coordination shell where the second peak position in the RDF $g_{Ob\text{-}Hb}(r)$ is at a smaller distance for the backbone than for the side chain interactions $g_{Os\text{-}Hs}(r)$ (see Table 10). As well as the radial positions of the first and second coordination shell we can examine the peak intensities of the first and second peaks in the RDF $g_{O\text{-}H}(r)$ for both the backbone and side chain. There are interesting differences between the peak intensities of the glutamine – glutamine RDFs in Figure 9. The peak intensity is highest for backbone –side chain interactions (see $g_{Os\text{-}Hs}(r)$ Figure 9C) than for backbone-backbone interactions (see $g_{Ob\text{-}Hb}(r)$ in Figure 9A) and side chain-side chain interactions (see $g_{Os\text{-}Hs}(r)$ in Figure 9D). Inspection of the coordination numbers in Tables 9 and 10 indicate that hydrogen bonding between glutamine molecule is present, but very limited at this dilute concentration.

To determine whether glutamine interacted strongly with itself in the solution, despite the dilute concentration, a cluster analysis was completed. Glutamine molecules were considered to be involved in clusters through hydrogen bonded interactions if the distance between carbonyl oxygen atoms and amine hydrogen atoms is within the range covered by the first coordination shell of the corresponding RDF. The glutamine clusters were defined in this way because there was a prominent correlation between carbonyl oxygen and amine hydrogen atoms in the RDFs (Figure 9). The results of the cluster analysis are shown in Figure 10 for the aqueous glutamine solution. From Figure 10 it can be seen that most of the glutamine molecules exist as isolated molecules in this dilute solution. However, there is evidence for dimers of glutamine molecules in the solution. This is an interesting result given there are only a small number of glutamine molecules in the EPSR simulation (25 glutamine molecules and 6725 water molecules). Glutamine clusters of size two are seen for the glutamine-glutamine interactions involving backbone-backbone, backbone-side chain and side chain-side chain. Representative glutamine-glutamine interactions obtained from the EPSR

simulation are shown in Figure 11 for (A) backbone-backbone, (B) backbone-side chain and (C) side chain-side chain. The glutamine-glutamine interactions shown in Figure 9 and the resulting dimers of glutamine in Figure 10 demonstrate the ability of glutamine to hydrogen bond with itself in the solution. Inspection of the position and width of the first coordination shell for all glutamine – glutamine RDFs in Figure 9 show that the most strongly correlated hydrogen bond interaction is between backbone-backbone interactions.

To determine whether the dimers observed in Figure 10 are indicative of larger-range structures being present in the solution or whether it is simply the result of glutamine molecules randomly packing (with no attractive interactions) we completed further EPSR simulations. In the first case, a simulation was run at the same concentration of glutamine with no atomic charges and the empirical potential set to zero, thus removing all of the attractive interactions between the molecules other than Van der Waals interactions and removing the empirical potential refinement to the data. If the clustering was a consequence of attractive interactions between molecules and refinement to the experimental data then the clustering would not occur in this 'randomly packed' simulation. The 'randomly packed' simulation then 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 glutamine molecules (or indeed water molecules). The results of this test are shown in Figure SI 3B, with the results of the original EPSR simulation shown in Figure SI 3A for comparison. The cluster distributions demonstrate that only isolated glutamine monomers are present in the 'randomly packed' simulation. In the second case a simulation was run which included atomic charges but no empirical potential, allowing us to determine whether the neutron diffraction data was important in the refinement of the empirical potential. The results of these tests are shown in Figure SI 3C. The cluster distributions demonstrate that some glutamine clusters of size two, three and four are found, when considering glutamine interactions between the backbones of glutamine molecules. However, no dimers or larger clusters are observed for glutamine molecules, when considering interactions between glutamine side chains. Importantly, the glutamine dimers formed by interactions between side chains in glutamine are only observed in the EPSR simulation which includes refinement (Figure 10), demonstrating that these dimers are a result of refinement to the neutron diffraction data. Interestingly it is side chain – side chain glutamine interactions which have been identified as playing an important role in the interaction and association of glutamine containing polypeptide chains and proteins [23-26].

The largest cluster size observed in Figure 10 is two glutamine molecules and may be a limit of the number of glutamine molecules in EPSR simulation. As a control we completed small angle scattering experiments to determine whether larger clusters were found in this dilute solution. Figure 12 shows the differential scattering cross-section for aqueous glutamine after subtraction of the scattering from D_2O and the cuvettes from the samples. Inspection of the data shows that there is no large scale structure present in the solution at the concentration studied, namely 30 mg/ml. This is evidenced by the lack of any strong signal arising at any place in the spectra in Figure 12. An exception occurs at very low Q due to the detector cut-off rather than the sample scattering itself. Given the LOQ instrument is capable of measuring long-range structures in the length range 25 – 800 Å in solution, the results demonstrate that there are no long range glutamine structures > 25 Å. If glutamine is aggregating in the solution it is either forming small aggregates of less than 25 Å in diameter or forming clusters with a larger number of molecules with very compact conformations, less than 25 Å in diameter.

Conclusion

We have used neutron diffraction and computational modelling to examine the hydrogen bonding present in a glutamine water solution. At the dilute concentration studied here glutamine is capable of forming hydrogen bonds with itself, as evidenced by the presence of prominent hydrogen bonding peaks in the relevant RDFs (Figure 9A-D) as well as the presence of dimers of glutamine molecules (Figure 10). Glutamine's ability to hydrogen bond through both the backbone and side chain provides a range of possible hydrogen bonding interactions between glutamine. The presence of dimers in the solution is an interesting result given there are only a small number of glutamine molecules in the EPSR simulation, demonstrating the propensity for this molecule to associate. Glutamine dimers are seen for glutamine-glutamine interactions involving backbone-backbone, backbone-side chain and side chain-side chain. Examination of all of the RDFs $g_{O\text{-}H}(r)$ for glutamine–glutamine, water – glutamine and water –water interactions allows comparison between first neighbour level hydrogen bond interaction. The shortest distance measured in the first peak position of $g_{O\text{-}H}(r)$ is for the glutamine-glutamine interactions between backbone oxygen and backbone hydrogen (1.53 Å), suggesting that hydrogen bonding between glutamine backbones is most strongly correlated. While side chain –side chain interactions were observed, there is no indication that they are more correlated or cooperative than backbone-backbone interactions, as has been suggested by previous studies on polyQ [24, 25]. These differences may reflect the subtle difference in hydrogen bond ability of the glutamine monomer and glutamine within a polypeptide chain, highlighting the importance of understanding the intermolecular interactions of both the individual building block, glutamine, and its assembly into larger structures, such a polyQ chains.

The dominant interaction at the glutamine concentration studied here is between glutamine and the surrounding water, without significantly perturbing the bulk water structure. On average, single carbonyl oxygen atoms can coordinate more water molecules, compared to amine hydrogen atoms in agreement with previous work on aqueous proline [30] and glutamic acid [31]. Interestingly, in the study of aqueous glutamic acid [31] the carbonyl oxygen atom coordinated more water molecules (~4.5 water molecules) than that observed for glutamine or proline (~3 water molecules) [30]. This may be due to the presence of sodium hydroxide in the solution in the study [31]. Examination of the RDFs and corresponding coordination numbers reveals that the backbone of glutamine has more interactions with water than the side chain, demonstrating the powerful propensity of this molecule to hydrogen bond. This is in agreement with previous studies of a glutamic acid solution which showed that both the backbone and side chain of glutamic acid hydrogen bonded with water [31]. Given the strong interaction between glutamine and water it might be expected that water would be a good solvent for a polypeptide chain made up of glutamine, polyQ. In a good solvent a polypeptide chain is hydrated by water and forms an expanded structure, maximizing the contact between chain and water [53]. However, recent experimental and computational studies have made the discovery that water is a poor solvent for polyQ chains [20, 21, 54]. In these studies polyQ chains, varying in length from 10 glutamine repeats to 75 repeats, where found to form collapsed disordered globules in aqueous solution. The observed behavior of the collapsed polyQ chain is therefore at odds with the notion of glutamine-water interactions being favorable and instead suggests that the glutamine molecules in the chain would prefer to interact with themselves rather than with water. This is in agreement with molecular dynamics simulations which suggested that the stability of collapsed, globular polyQ chains results from an extensive hydrogen bonding network between glutamine molecules, with glutamine-glutamine hydrogen bond interactions being

more favourable than glutamine – water interactions [21]. The importance of inter-amino acid interactions has previously been observed in a study examining a series of dipeptides of differing hydrophobicity and hydrophilicity [29]. This study found that dipeptides containing more hydrophilic groups interacted more strongly with each other than those containing hydrophobic groups, suggesting that it is the hydrophilic nature of a peptide with drives association [29]. This hydrophilic association has also been observed in proline solutions, where backbone-backbone interactions resulted in clustering [30]. While the present study shows extensive glutamine-water interactions, the presence of prominent glutamine-glutamine interactions at this dilute concentration points to the importance of glutamine hydrogen bonding, which will be more prevalent in a polyQ chain. Further studies on polyQ in water are essential to learn more about the intermolecular interactions which might favour glutamine–glutamine interactions over glutamine-water interactions, which may be important in understanding the self-assembly of glutamine in larger structures. The work presented here forms only the first step in an exhaustive study of the structure of glutamine and polyQ chains of different length at different concentrations, temperatures and solvent environment.

Acknowledgments

This work was supported by the European Research Council through a grant to Lorna Dougan. Experiments at the ISIS Pulsed Neutron and Muon Source were supported by a beam time allocation from the Science and Technology Facilities Council. We are grateful to Dr Silvia Imberti, Dr Stephen King and Dr Ann Terry at the ISIS Facility, Rutherford Appleton Laboratory for many useful discussions.

Supporting Information

Additional figures are available in the supporting information. This material is available free of charge via the Internet at http://pubs.acs.org.

Sample No.	Sample Name	Description
1	Glutamine-H_{10} D_2O	Fully protiated Glutamine(99.5%) with D_2O (Purity 99.99 % D)
2	Glutamine-H_{10} H_2O	Fully protiated (99.5%) with milliQ water
3	Glutamine-D_5H_5 D_2O	Deuterated methyl atoms (97 atom % D) with D_2O (Purity 99.99 % D)
4	Glutamine-D_5H_5 H_2O	Deuterated methyl atoms (97 atom % D) with milliQ water

Table 1. Glutamine-water samples for which the structure factor has been measured with neutron diffraction on the SANDALS instrument.

Intra-molecular Bond Lengths	
Bond Pair	Bond Length (Å)
Cb-Ca	1.5400
Cb-Ob	1.2510
Ca-Cm	1.5280
Ca-Nb	1.4980
Ca-Hm	1.0970
Cm-Cm	1.5200
Cm-Hm	1.1048
Cs-Cm	1.5110
Cs-Os	1.2310
Cs-Ns	1.3340
Nb-Hb	1.0400
Ns-Hs	1.0070

Intra-molecular Bond Angles	
Atom Types	Bond Angle (°)
Ca-Cb-Ob	116.65
Ob-Cb-Ob	126.70
Cb-Ca-Cm	110.30
Cb-Ca-Nb	110.20

Cm-Ca-Nb	111.10
Cb-Ca-Hm	108.70
Cm-Ca-Hm	109.70
Nb-Ca-Hm	106.60
Ca-Cm-Cm	114.00
Ca-Cm-Hm	108.00
Cm-Cm-Hm	110.28
Hm-Cm-Hm	106.90
Cm-Cm-Cs	113.10
Cs-Cm-Hm	107.55
Cm-Cs-Os	122.10
Cm-Cs-Ns	115.20
Ns-Cs-Os	122.70
Ca-Nb-Hb	110.10
Hb-Nb-Hb	108.83
Cs-Ns-Hs	120.95
Hs-Ns-Hs	117.70

Intra-molecular Dihedral Angles	
Atom Types	Dihedral Angles (°)
Cb-Ca-Cm-Cm	-61.50
Nb-Ca-Cm-Cm	66.10
Ca-Cm-Cm-Cs	175.50
Cm-Cm-Cs-Ns	167.20

Table 2. Geometry for glutamine molecules used in the EPSR simulations: representative intra-molecular bond distances, angles and dihedral angles.

Atom Name	ε (kJ/mol)	σ (Å)	m (a.m.u.)	q (e)
Cb	0.43932	3.750	12	0.7000
Ca	0.33472	3.800	12	0.2100
Cm	0.49731	3.905	12	0.0000
Cs	0.43932	3.750	12	0.5000
Ob	0.87864	2.960	16	-0.8000
Os	0.87864	2.960	16	-0.5000
Nb	0.71128	3.250	14	-0.3000
Ns	0.71128	3.250	14	-0.8500
Hb	0.0	0.0	2	0.3300
Hs	0.0	0.0	2	0.4250
Hm	0.0	0.0	2	0.0
Ow	0.65000	3.166	16	-0.8476
Hw	0.0	0.0	2	0.4238

Table 3. Lennard-Jones parameters, masses and Coulomb charges defining the potentials used for EPSR simulations of glutamine and water at 297 K. The SPC/E model potentials have been used for water, whilst OPLS potentials have been used for glutamine.

Bond	r_1 (Å)	r_2 (Å)	1st peak position (Å)	Coordination number	s.d.
Ow-Ow	0	3.41	2.77	4.8	1.1
Ow-Ow Pure Water	0	3.51	2.80	5.2	1.2
Ow-Hw	0	2.41	1.77	1.8	0.7
Ow-Hw Pure Water	0	2.45	1.80	1.8	0.6

Table 4. Water-Water Interactions: Peak positions and coordination numbers for the first coordination shell of the radial distribution function $g_{Ow\text{-}Ow}$ for water oxygen atoms (Ow), shown in Figure 3. The values are shown for both the EPSR analysis of neutron diffraction data of the glutamine– water mixture and the same functions for pure water, where previously published neutron diffraction data [50] has been used to complete an EPSR simulation. The standard deviation s.d. has been calculated over all configurations (~10 000).

Bond	r_1 (Å)	r_2 (Å)	2nd peak position (Å)	Coordination number	s.d.
Ow-Ow	3.41	5.68	4.38	19.4	2.4
Ow-Ow Pure Water	3.51	5.64	4.56	18.9	2.1
Ow-Hw	2.41	4.42	3.30	20.5	2.9
Ow-Hw Pure Water	2.45	4.80	3.31	27.5	2.6

Table 5: Water-Water Interactions: Peak positions and coordination numbers for the second coordination shell of the radial distribution function $g_{Ow\text{-}Ow}$ for water oxygen atoms (Ow), shown in Figure 3. The standard deviation s.d. has been calculated over all configurations (~10 000).

Bond	r_1 (Å)	r_2 (Å)	1st peak position (Å)	Coordination number	s.d.
Ob-Ow	0	3.21	2.63	2.9	0.7
Os-Ow	0	3.36	2.76	2.5	0.9
Hb-Ow	0	2.35	1.66	1.0	0.3
Hs-Ow	0	2.45	1.74	0.9	0.4

Table 6: Glutamine– Water Interactions: Peak positions and coordination numbers for the first coordination shell of the radial distribution functions $g_{O\text{-}Ow}$ and $g_{H\text{-}Ow}$, shown in Figures 4 and 6. The glutamine backbone atoms are denoted Ob and Hb and the glutamine side chain atoms are denoted Os and Hs. The standard deviation s.d. has been calculated over all configurations (~10 000).

Bond	r_1 (Å)	r_2 (Å)	2nd peak position (Å)	Coordination number	s.d.
Ob-Ow	3.21	5.27	4.67	13.2	2.7
Os-Ow	3.36	5.18	4.67	11.8	2.1
Hb-Ow	2.35	3.67	3.11	3.6	1.1
Hs-Ow	2.45	4.38	3.32	7.4	1.7

Table 7: Glutamine– Water Interactions: Peak positions and coordination numbers for the second coordination shell of the radial distribution functions $g_{O\text{-}Ow}$ and $g_{H\text{-}Ow}$, shown in Figures 4 and 6. The glutamine backbone atoms are denoted Ob and Hb and the glutamine side chain atoms are denoted Os and Hs. The standard deviation s.d. has been calculated over all configurations (~10 000).

Bond	Interaction Type	r_1 (Å)	r_2 (Å)	Peak position (Å)	Coordination number	s.d.
Nb-Hw	Amine hydrogen atoms	0	3.99	3.35	10.7	1.8
Ns-Hw	Lone pairs	0	2.23	2.03	0.3	0.5
Ns-Hw	Amine hydrogen atoms	2.23	4.31	3.39	14.5	2.8

Table 8: Amine Nitrogen- Water Interactions Peak positions and coordination numbers for the coordination shells of the radial distribution function $g_{Nb\text{-}Hw}$ and $g_{Ns\text{-}Hw}$, shown in Figure 7. The glutamine backbone atoms are denoted Nb and the glutamine side chain atoms are denoted Ns. The standard deviation s.d. has been calculated over all configurations (~10 000).

Bond	r_1 (Å)	r_2 (Å)	1st peak position (Å)	Coordination number	s.d.
Ob-Hb	0	2.31	1.53	0.0	0.2
Ob-Hs	0	2.46	1.62	0.0	0.1
Os-Hb	0	2.21	1.65	0.1	0.0
Os-Hs	0	2.51	1.65	0.0	0.0

Table 9: Glutamine – Glutamine Interactions Peak positions and coordination numbers for the first coordination shell of the radial distribution function $g_{O\text{-}H,}$ shown in Figure 9. The glutamine backbone atoms are denoted Ob and Hb and the glutamine side chain atoms are denoted Os and Hs. The standard deviation s.d. has been calculated over all configurations (~10 000).

Bond	r_1 (Å)	r_2 (Å)	2nd peak position (Å)	Coordination number	s.d.
Ob-Hb	2.31	4.81	2.82	0.1	0.5
Ob-Hs	2.46	3.60	3.06	0.0	0.7
Os-Hb	2.21	3.46	2.97	0.1	0.1
Os-Hs	2.51	5.51	3.15	0.2	0.6

Table 10: Glutamine – Glutamine Interactions Peak positions and coordination numbers for the second coordination shell of the radial distribution function $g_{O\text{-}H,}$ shown in Figure 9. The glutamine backbone atoms are denoted Ob and Hb and the glutamine side chain atoms are denoted Os and Hs. The standard deviation s.d. has been calculated over all configurations (~10 000).

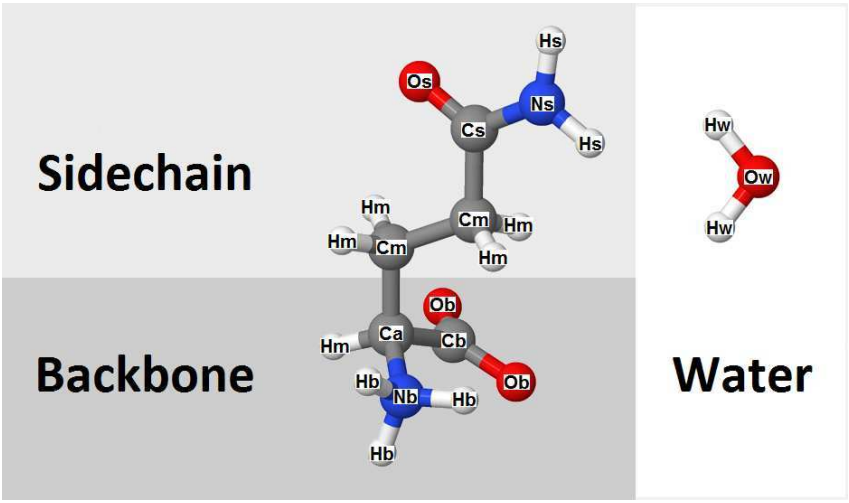


Figure 1. Representation of the glutamine and water molecules. The single atoms have been labelled according to the symbols used in the simulation and throughout this paper. For glutamine the carbon atoms are labelled Ca, Cb and Cm, the carbonyl oxygen atoms Ob and Os, the amine hydrogen atoms Hb and Hs, and the methyl hydrogen atoms Hm. The atoms Cb, Ob and Hb are in the glutamine backbone while the atoms Cs, Os and Hs are in the side chain. For water the oxygen atom is labelled OW and the water hydrogen atom HW.

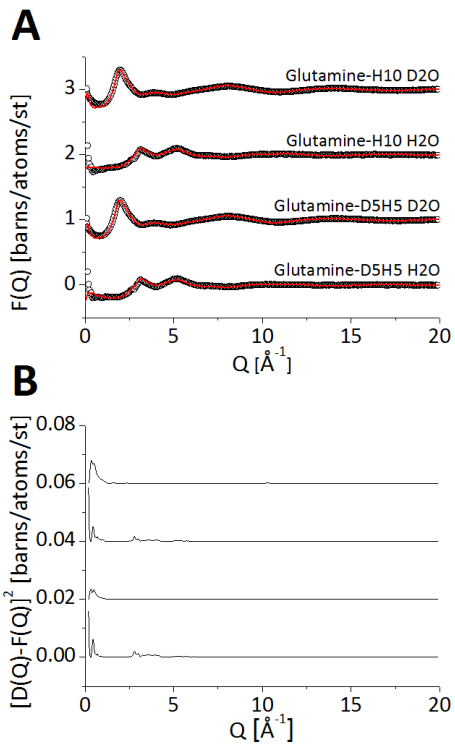


Figure 2. (A) Measured neutron diffraction data (black circles) compared to fits obtained by EPSR analysis (red lines), for a dilute aqueous glutamine solution at 297 K for all four samples. (B)The residual difference between the data and the simulation, calculated as the residual squares ($[D(Q)-F(Q)]^2$) as a function of Q. The data and fits are labelled according to Table 1 and have been shifted vertically for improved clarity.

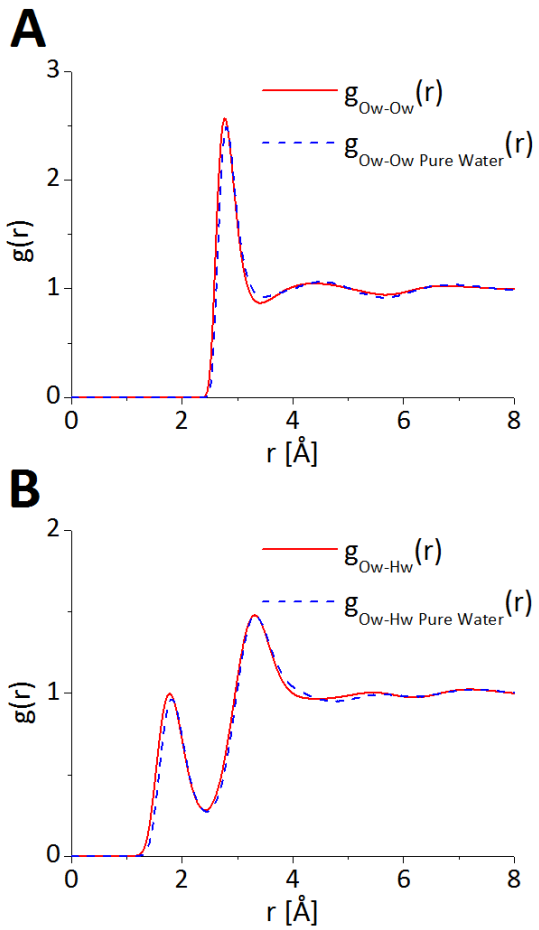


Figure 3. Water – water interactions: Water-water site-site partial radial distribution functions (RDFs) in a dilute aqueous glutamine solution. These include (A) $g_{Ow\text{-}Ow}$ and (B) $g_{Ow\text{-}Hw}$. The RDFs are taken from the EPSR analysis of neutron diffraction data of the glutamine– water mixture (solid lines) and are compared with the same functions for pure water (dashed lines), where previously published neutron diffraction data [50] has been used to complete an EPSR simulation.

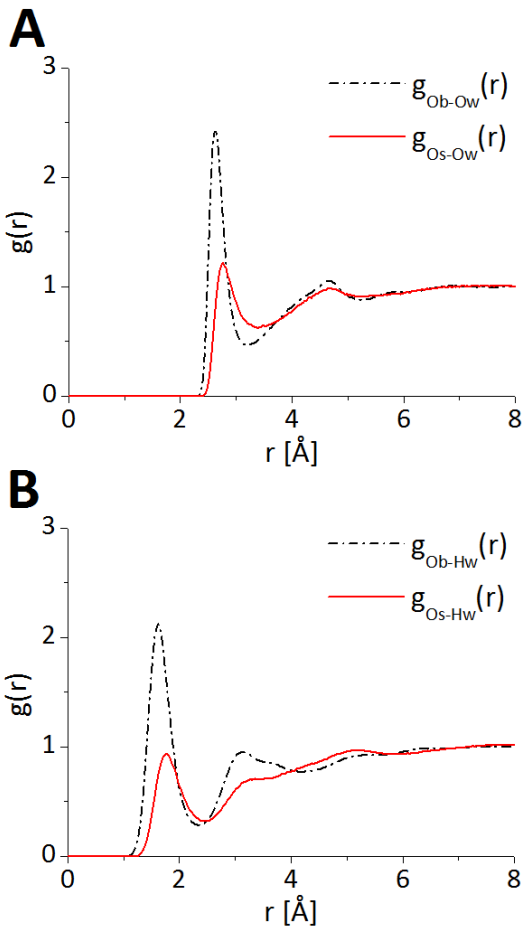


Figure 4. **Glutamine – water interactions**: Carbonyl-water site-site partial radial distribution functions (RDFs) from the EPSR analysis of neutron diffraction data of the glutamine– water mixture. The $g_{Ob\text{-}Ow}$ (dashed, black line) shows the ability of oxygen atoms on the backbone (Ob) to coordinate water molecules (O_W). The $g_{Os\text{-}Ow}$ (solid, red line) shows the ability of oxygen atoms on the side chain (Os) to coordinate water molecules (O_W).

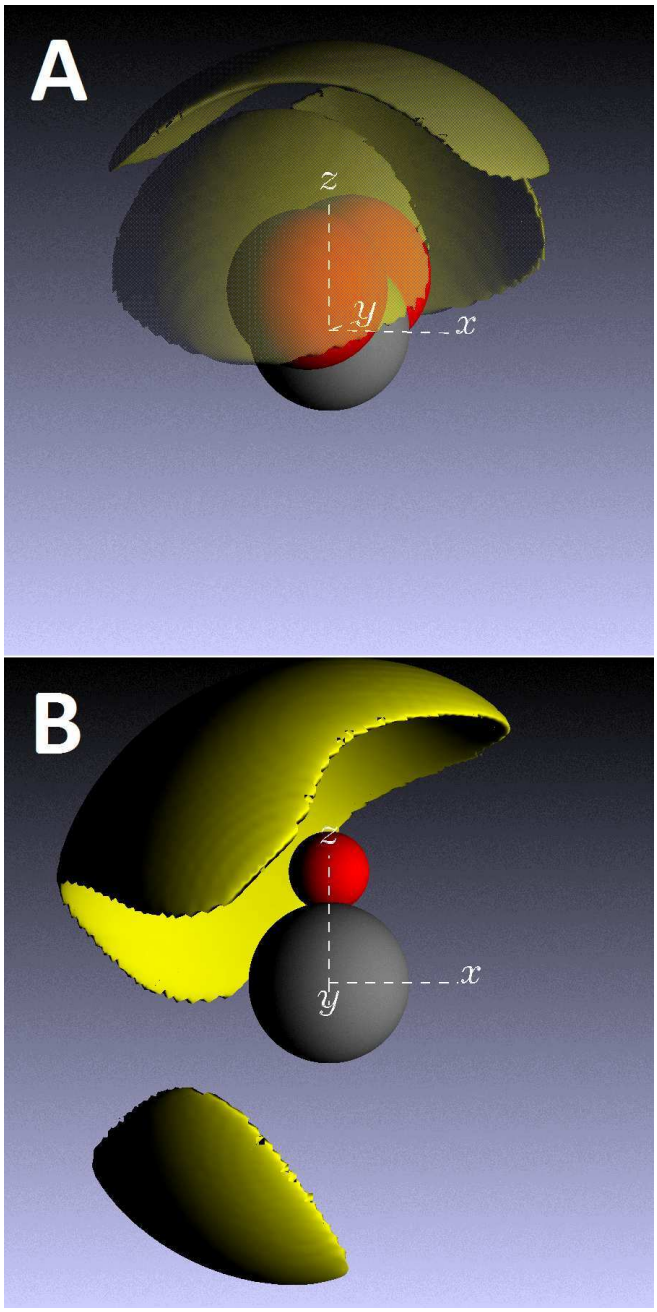


Figure 5: Spatial density function (SDFs) showing the distribution of water molecules around glutamine taken from the EPSR analysis of neutron diffraction data of aqueous glutamine. The yellow shaded areas represent the regions where there is a probability of finding a water molecule surrounding (A) the backbone carbonyl group of glutamine at a distance range of 2 – 4.28 Å and (B) the side chain carbonyl group of glutamine at a distance range of 2 – 4.41 Å. The distance ranges correspond to the first coordination shell of the corresponding C-Ow radial distribution function (Figure SI 1). The SDFs show 25% of the water molecules enclosed in this region.

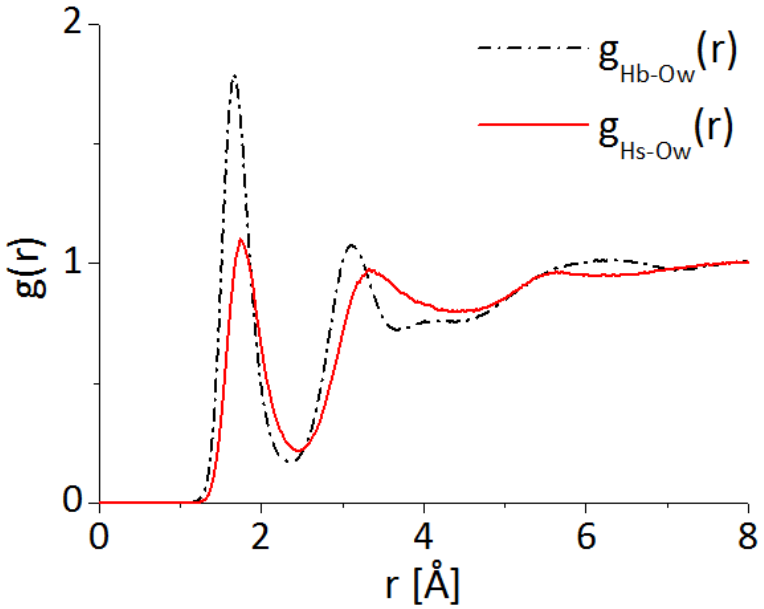


Figure 6. **Glutamine – water interactions**: Amine-water site-site partial radial distribution functions (RDFs) from the EPSR analysis of neutron diffraction data of the glutamine– water mixture. The $g_{Hb\text{-}Ow}$ (dashed, black line) shows the ability of hydrogen atoms on the backbone (Hb) to coordinate water molecules (O_W). The $g_{Hs\text{-}Ow}$ (solid, red line) shows the ability of hydrogen atoms on the side chain (Hs) to coordinate water molecules (O_W).

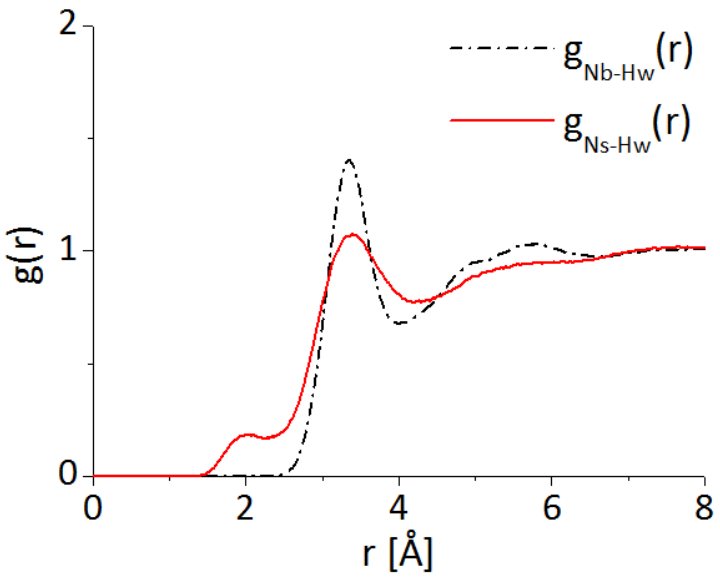


Figure 7. **Glutamine – water interactions**: Amine-water site-site partial radial distribution functions (RDFs) from the EPSR analysis of neutron diffraction data of the glutamine– water mixture. The $g_{Nb\text{-}Hw}$ (dashed, black line) shows the ability of nitrogen atoms on the backbone (Nb) to coordinate water molecules (H_W). The $g_{Ns\text{-}Hw}$ (solid, red line) shows the ability of nitrogen atoms on the side chain (Ns) to coordinate water molecules (H_W).

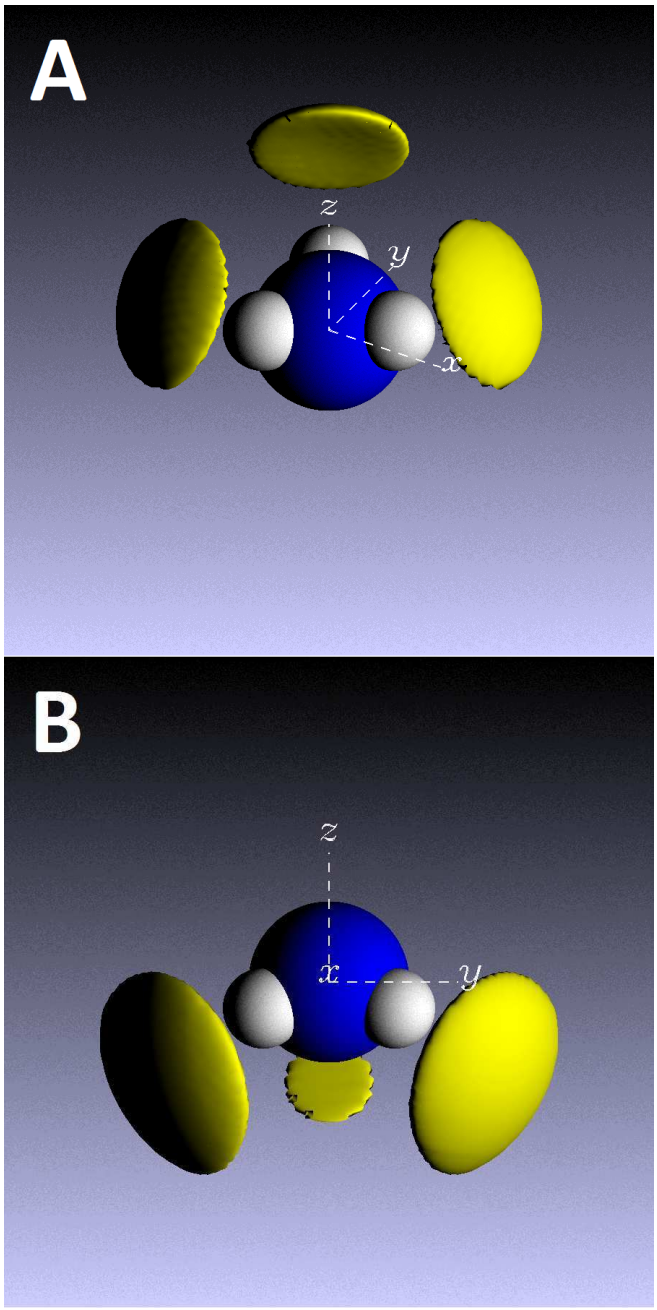


Figure 8: Spatial density function (SDFs) showing the distribution of water molecules around glutamine taken from the EPSR analysis of neutron diffraction data of aqueous glutamine surrounding (A) the backbone amine group of glutamine at a distance range of 2 – 3.42 Å (B) the side chain amine group of glutamine at a distance range of 2 – 3.45 Å. The distance ranges correspond to the first coordination shell of the corresponding N-Ow radial distribution function (Figure SI 2). The SDFs show 25% of the water molecules enclosed in this region.

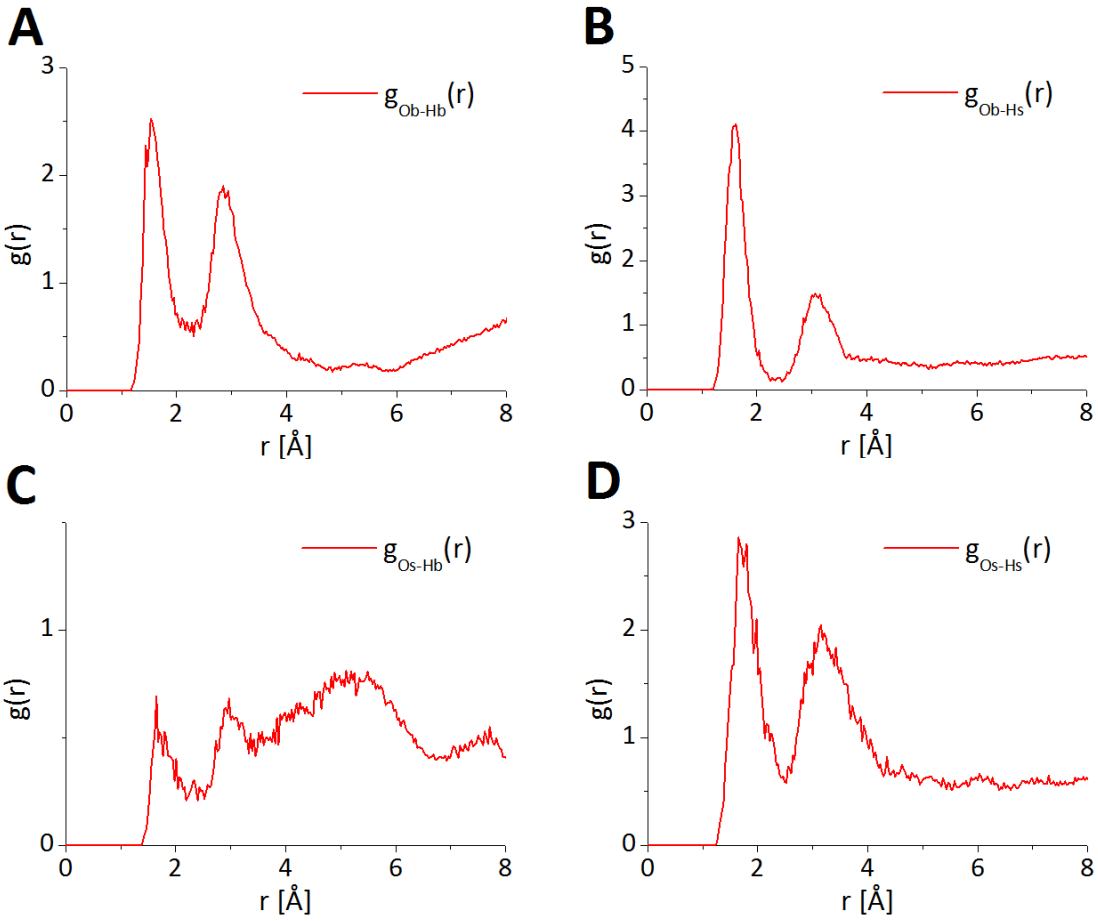


Figure 9: Glutamine – glutamine interactions: Glutamine-glutamine site-site partial radial distribution functions (RDFs) from the EPSR analysis of neutron diffraction data of the glutamine– water mixture. (A) The g_{Ob-Hb} shows the ability of oxygen atoms on the backbone (Ob) to coordinate hydrogen atoms on the backbone (Hb). (B) The g_{Ob-Hs} shows the ability of oxygen atoms on the backbone (Ob) to coordinate hydrogen atoms on the side chain (Hs). (C) The g_{Os-Hb} shows the ability of oxygen atoms on the side chain (Os) to coordinate hydrogen atoms on the backbone (Hb). (D) The g_{Os-Hs} shows the ability of oxygen atoms on the side chain (Os) to coordinate hydrogen atoms on the side chain (Hs).

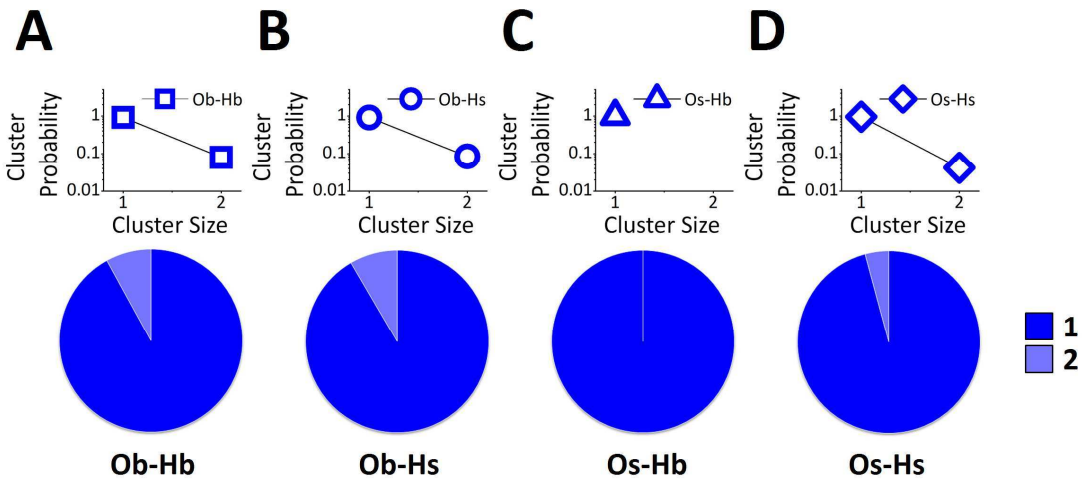


Figure 10: Glutamine –glutamine cluster analysis: (A)-(D) Upper, the probability of having a glutamine-glutamine molecule cluster of a given size is shown for a dilute glutamine-water solution. A glutamine molecule is defined as being in the same cluster as another if the oxygen atom of one glutamine molecule is within a specified distance of the hydrogen atom on another glutamine molecule. The specified distance is taken from the position of the relevant trough within the glutamine oxygen –glutamine hydrogen RDFs (Figure 9). The EPSR cluster distribution shows a maximum cluster size of two molecules for some of the interactions. **(A)-(D) Lower**, pie chart showing the proportion of glutamine molecules that are found within clusters of different sizes in a dilute glutamine-water solution. In the pie charts monomers (clusters size of one) are shown with the darkest colors and lighter shades are used for clusters with increasing size. Ob-Hb shows the ability of oxygen atoms on the backbone (Ob) to cluster with hydrogen atoms on the backbone (Hb). Ob-Hs shows the ability of oxygen atoms on the backbone (Ob) to cluster with hydrogen atoms on the side chain (Hs). Os-Hb shows the ability of oxygen atoms on the side chain (Os) to cluster with hydrogen atoms on the backbone (Hb). Os-Hs shows the ability of oxygen atoms on the side chain (Os) to cluster with hydrogen atoms on the side chain (Hs).

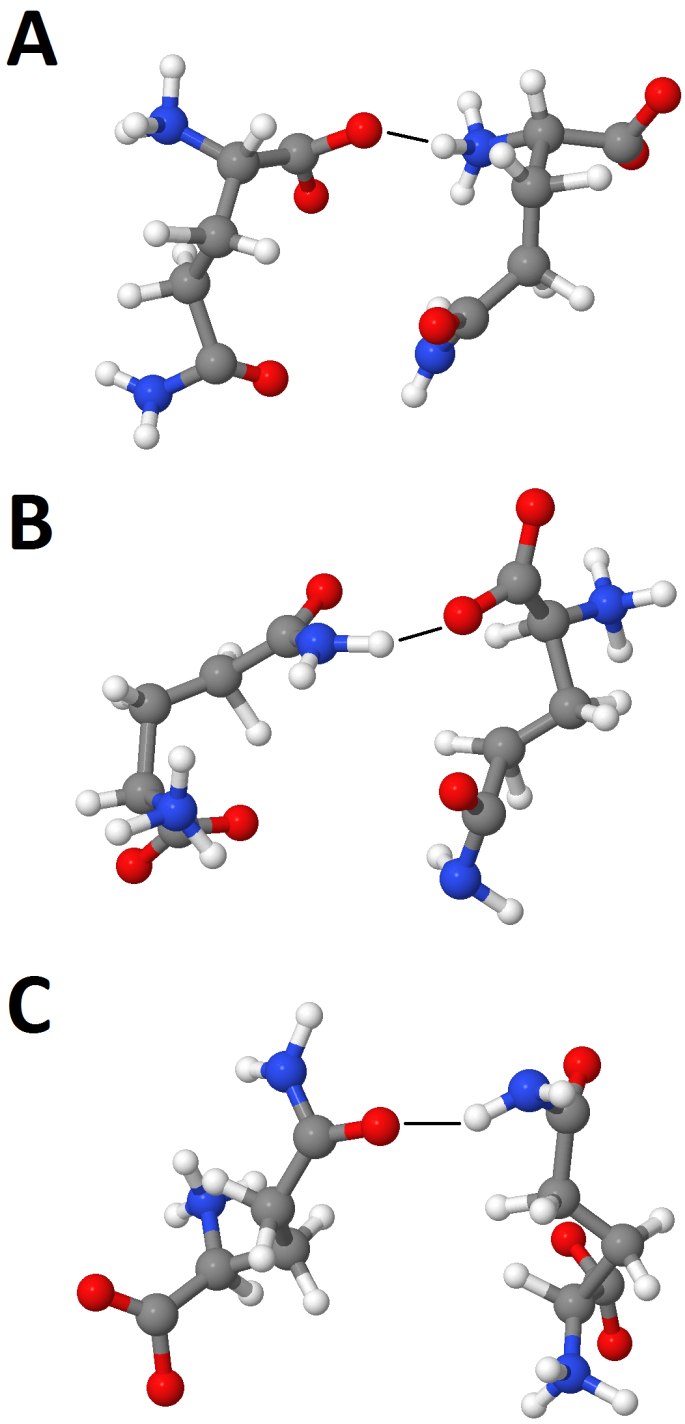


Figure 11: Representative snapshots from the EPSR simulation showing glutamine dimers which have formed from hydrogen bond interactions between (A) the backbone of one glutamine molecule and the backbone of another glutamine molecule (B) the backbone of one glutamine molecule and the side-chain of another glutamine molecule and (C) the side-chain of one glutamine molecule and the side-chain of another glutamine molecule

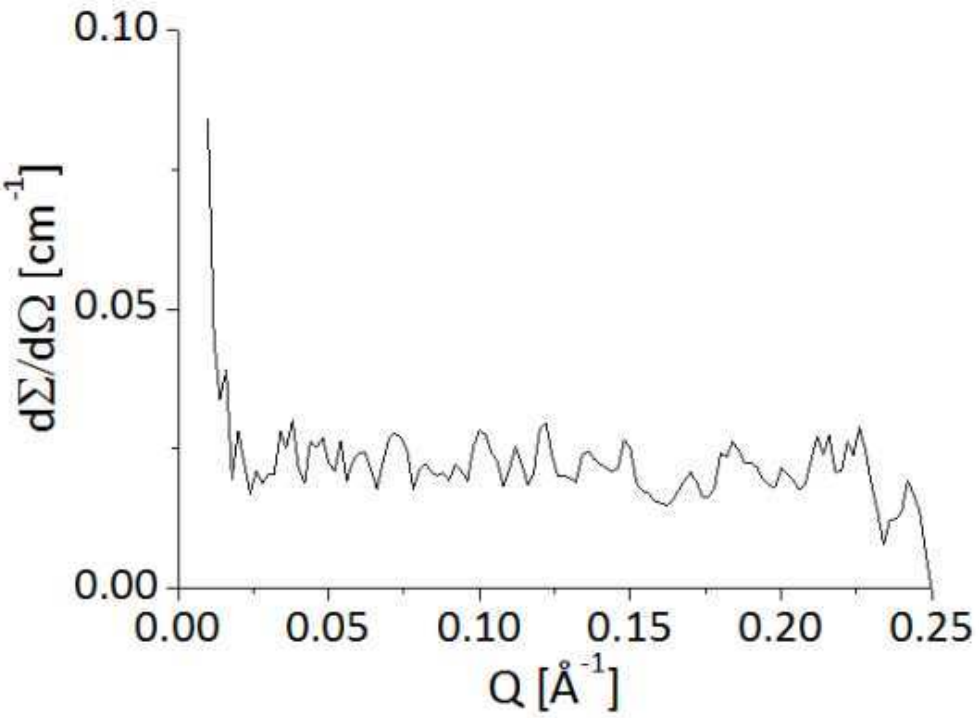


Figure 12: The differential scattering cross section from the instrument LOQ for the glutamine water solution at 298 K.

References

1. Faux, N. G.; Bottomley, S. P.; Lesk, A. M.; Irving, J. A.; Morrison, J. R.; de la Banda, M. C.; Whisstock, J. C., Functional insights from the distribution and role of homopeptide repeat-containing proteins. *Genome Research* **2005,** *15* (4), 537-551.
2. Krebs, H. A., Metabolism of amino-acids: The synthesis of glutamine from glutamic acid and ammonia, and the enzymic hydrolysis of glutamine in animal tissues. . *Biochemical Journal* **1935,** *29*, 1951-1969
3. Neu, J.; Shenoy, V.; Chakrabarti, R., Glutamine nutrition and metabolism: Where do we go from here? *Faseb Journal* **1996,** *10* (8), 829-837.
4. Wetzel, R., Physical Chemistry of Polyglutamine: Intriguing Tales of a Monotonous Sequence. *Journal of Molecular Biology* **2012,** *421* (4-5), 466-490.
5. Bates, G. P.; Benn, C., *The polyglutamine diseases. In Huntington's Diseases*. Oxford University Press: Oxford, UK, 2002; p 429-472.
6. Mangiarini, L.; Sathasivam, K.; Seller, M.; Cozens, B.; Harper, A.; Hetherington, C.; Lawton, M.; Trottier, Y.; Lehrach, H.; Davies, S. W.; Bates, G. P., Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* **1996,** *87* (3), 493-506.
7. Osmand, A. P.; Berthelier, V.; Wetzel, R., Imaging polyglutamine deposits in brain tissue. *Amyloid, Prions, and Other Protein Aggregates, Pt B* **2006,** *412*, 106-122.
8. Bocharova, N.; Chave-Cox, R.; Sokolov, S.; Knorre, D.; Severin, F., Protein aggregation and neurodegeneration: clues from a yeast model of Huntington's disease. *Biochemistry (Mosc)* **2009,** *74* (2), 231-4.
9. Sanchez, I.; Mahlke, C.; Yuan, J., Pivotal role of oligomerization in expanded polyglutamine neurodegenerative disorders. *Nature* **2003,** *421* (6921), 373-9.
10. Digambaranath, J. L.; Campbell, T. V.; Chung, A.; McPhail, M. J.; Stevenson, K. E.; Zohdy, M. A.; Finke, J. M., An accurate model of polyglutamine. *Proteins-Structure Function and Bioinformatics* **2011,** *79* (5), 1427-1440.
11. Kar, K.; Jayaraman, M.; Sahoo, B.; Kodali, R.; Wetzel, R., Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nature Structural & Molecular Biology* **2011,** *18* (3), 328-336.
12. Miller, J.; Arrasate, M.; Shaby, B. A.; Mitra, S.; Masliah, E.; Finkbeiner, S., Quantitative Relationships between Huntingtin Levels, Polyglutamine Length, Inclusion Body Formation, and Neuronal Death Provide Novel Insight into Huntington's Disease Molecular Pathogenesis. *Journal of Neuroscience* **2010,** *30* (31), 10541-10550.
13. Nakano, M.; Watanabe, H.; Rothstein, S. M.; Tanaka, S., Comparative Characterization of Short Monomeric Polyglutamine Peptides by Replica Exchange Molecular Dynamics Simulation. *Journal of Physical Chemistry B* **2010,** *114* (20), 7056-7061.
14. Wang, X. L.; Vitalis, A.; Wyczalkowski, M. A.; Pappu, R. V., Characterizing the conformational ensemble of monomeric polyglutamine. *Proteins-Structure Function and Bioinformatics* **2006,** *63* (2), 297-311.
15. Williamson, T. E.; Vitalis, A.; Crick, S. L.; Pappu, R. V., Modulation of Polyglutamine Conformations and Dimer Formation by the N-Terminus of Huntingtin. *Journal of Molecular Biology* **2010,** *396* (5), 1295-1309.
16. Stanley, C. B.; Perevozchikova, T.; Berthelier, V., Structural Formation of Huntingtin Exon 1 Aggregates Probed by Small-Angle Neutron Scattering. *Biophysical Journal* **2011,** *100* (10), 2504-2512.

17. Wang, Y. T.; Voth, G. A., Molecular Dynamics Simulations of Polyglutamine Aggregation Using Solvent-Free Multiscale Coarse-Grained Models. *Journal of Physical Chemistry B* **2010,** *114* (26), 8735-8743.
18. Khare, S. D.; Ding, F.; Gwanmesia, K. N.; Dokholyan, N. V., Molecular origin of polyglutamine aggregation in neurodegenerative diseases. *Plos Computational Biology* **2005,** *1* (3), 230-235.
19. Chen, S. M.; Ferrone, F. A.; Wetzel, R., Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** *99* (18), 11884-11889.
20. Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V., Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** *103* (45), 16764-16769.
21. Dougan, L.; Li, J. Y.; Badilla, C. L.; Berne, B. J.; Fernandez, J. M., Single homopolypeptide chains collapse into mechanically rigid conformations. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (31), 12605-12610.
22. Peters-Libeu, C.; Miller, J.; Rutenber, E.; Newhouse, Y.; Krishnan, P.; Cheung, K.; Hatters, D.; Brooks, E.; Widjaja, K.; Tran, T.; Mitra, S.; Arrasate, M.; Mosquera, L. A.; Taylor, D.; Weisgraber, K. H.; Finkbeiner, S., Disease-Associated Polyglutamine Stretches in Monomeric Huntingtin Adopt a Compact Structure. *Journal of Molecular Biology* **2012,** *421* (4-5), 587-600.
23. Natalello, A.; Frana, A. M.; Relini, A.; Apicella, A.; Invernizzi, G.; Casari, C.; Gliozzi, A.; Doglia, S. M.; Tortora, P.; Regonesi, M. E., A Major Role for Side-Chain Polyglutamine Hydrogen Bonding in Irreversible Ataxin-3 Aggregation. *Plos One* **2011,** *6* (4), 1-10.
24. Plumley, J. A.; Dannenberg, J. J., The Importance of Hydrogen Bonding between the Glutamine Side Chains to the Formation of Amyloid VQIVYK Parallel beta-Sheets: An ONIOM DFT/AM1 Study. *Journal of the American Chemical Society* **2010,** *132* (6), 1758-1759.
25. Marchut, A. J.; Hall, C. K., Side-chain interactions determine amyloid formation by model polyglutamine peptides in molecular dynamics simulations. *Biophysical Journal* **2006,** *90* (12), 4574-4584.
26. Vasudev, P. G.; Banerjee, M.; Ramakrishnan, C.; Balaram, P., Asparagine and glutamine differ in their propensity to form specific side chain-backbone hydrogen bonded motifs in proteins. *Proteins-Structure Function and Bioinformatics* **2011,** *80* (4), 991-1002.
27. Hargreaves, R.; Bowron, D. T.; Edler, K., Atomistic Structure of a Micelle in Solution Determined by Wide Q-Range Neutron Diffraction. *Journal of the American Chemical Society* **2011,** *133* (41), 16524-16536.
28. Malardier-Jugroot, C.; Bowron, D. T.; Soper, A. K.; Johnson, M. E.; Head-Gordon, T., Structure and water dynamics of aqueous peptide solutions in the presence of co-solvents. *Physical Chemistry Chemical Physics* **2010,** *12* (2), 382-392.
29. McLain, S. E.; Soper, A. K.; Daidone, I.; Smith, J. C.; Watts, A., Charge-Based Interactions between Peptides Observed as the Dominant Force for Association in Aqueous Solution. *Angewandte Chemie-International Edition* **2008,** *47* (47), 9059-9062.
30. McLain, S. E.; Soper, A. K.; Terry, A. E.; Watts, A., Structure and hydration of L-proline in aqueous solutions. *Journal of Physical Chemistry B* **2007,** *111* (17), 4568-4580.
31. McLain, S. E.; Soper, A. K.; Watts, A., Structural studies on the hydration of L-glutamic acid in solution. *Journal of Physical Chemistry B* **2006,** *110* (42), 21251-21258.

32. McLain, S. E.; Soper, A. K.; Watts, A., Water structure around dipeptides in aqueous solutions. *European Biophysics Journal with Biophysics Letters* **2008,** *37* (5), 647-655.
33. Towey, J. J.; Dougan, L., Structural Examination of the Impact of Glycerol on Water Structure. *Journal of Physical Chemistry B* **2012,** *116* (5), 1633-1641.
34. Towey, J. J.; Soper, A. K.; Dougan, L., The structure of glycerol in the liquid state: A neutron diffraction study
Physical Chemistry Chemical Physics **2011,** *12*, 9397-9347.
35. Towey, J. J.; Soper, A. K.; Dougan, L., Preference for Isolated Water Molecules in a Concentrated Glycerol-Water Mixture. *Journal of Physical Chemistry B* **2011,** *115* (24), 7799-7807.
36. Fisher, H. E.; Barnes, A. C.; Salmon, P. S., Neutron and x-ray diffraction studies of liquids and glasses. *Reports on Progress in Physics* **2006,** *69*, 233-299.
37. Keen, D. A., A comparison of various commonly used correlation functions for describing total scattering. *Journal of Applied Crystallography* **2001,** *34* (172-177), 172 - 177.
38. Finney, J. L.; Soper, A. K., Solvent Structure and Perturbations in Solutions of Chemical and Biological Importance. *Chemical Society Reviews* **1994,** *23* (1), 1-10.
39. CRC, *Handbook of Chemistry and Physics*. 55th ed.; CRS Press: Cleveland, Ohio.
40. Soper, A. K., GudrunN and GudrunX : programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section. *RAL Technical Reports RAL-TR-2011-013* **2011**.
41. Soper, A. K., Inelasticity corrections for time-of-flight and fixed wavelength neutron diffraction experiments. *Molecular Physics* **2009,** *107* (16), 1667-1684.
42. Soper, A. K., Partial structure factors from disordered materials diffraction data: An approach using empirical potential structure refinement. *Physical Review B* **2005,** *72* (10), 104204-104211.
43. Koetzle, T. F.; Frey, M. N.; Lehmann, M. S.; Hamilton, W. C., Precision Neutron-Diffraction Structure Determination of Protein and Nucleic-Acid Components .XIII. Molecular and Crystal-Structure of Amino-Acid L-Glutamine. *Acta Crystallographica Section B-Structural Science* **1973,** *29* (Nov15), 2571-2575.
44. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The Missing Term in Effective Pair Potentials. *Journal of Physical Chemistry* **1987,** *91* (24), 6269-6271.
45. Jorgensen, W. L.; Tiradorives, J., The Opls Potential Functions for Proteins - Energy Minimizations for Crystals of Cyclic-Peptides and Crambin. *Journal of the American Chemical Society* **1988,** *110* (6), 1657-1666.
46. Hummer, G.; Soumpasis, D. M.; Neumann, M., Computer-Simulation of Aqueous Na-Cl Electrolytes. *Journal of Physics-Condensed Matter* **1994,** *6*, A141-A144.
47. Soper, A. K., Empirical potential Monte Carlo simulation of fluid structure. *Chemical Physics* **1996,** *202* (2-3), 295-306.
48. Heenan, R. K.; Penfold, J.; King, S. M., SANS at pulsed neutron sources: Present and future prospects. *Journal of Applied Crystallography* **1997,** *30*, 1140-1147.
49. Wignall, G. D.; Bates, F. S., Absolute Calibration of Small-Angle Neutron-Scattering Data. *Journal of Applied Crystallography* **1987,** *20*, 28-40.
50. Soper, A. K., Joint structure refinement of x-ray and neutron
diffraction data on disordered materials: application to
liquid water. *Journal of Physics-Condensed Matter* **2007,** *19*, 335206-335224.
51. Leberman, R.; Soper, A. K., Effect of High-Salt Concentrations on Water-Structure. *Nature* **1995,** *378* (6555), 364-366.

52. Soper, A. K.; Weckstrom, K., Ion solvation and water structure in potassium halide aqueous solutions. *Biophysical Chemistry* **2006,** *124* (3), 180-191.
53. Flory, P. J., *Principles of Polymer Chemistry*. Cornell University Press: 1953.
54. Vitalis, A.; Wang, X. L.; Pappu, R. V., Quantitative characterization of intrinsic disorder in polyglutamine: Insights from analysis based on polymer theories. *Biophysical Journal* **2007,** *93* (6), 1923-1937.