

Amphipathic Solvation of Indole: Implications for the Role of Tryptophan in Membrane Proteins

Andrew J. Johnston,[†] Yapei (Rosie) Zhang,[‡] Sebastian Busch,[¶]

Luis Carlos Pardo,[§] Silvia Imberti,^{||} and Sylvia E. McLain^{*,†}

Department of Biochemistry, University of Oxford, Oxford, United Kingdom, Department of Biochemistry, Princeton University, Princeton, NJ USA, German Engineering Materials Science Centre (GEMS) at Heinz Maier-Leibnitz Zentrum (MLZ), Helmholtz-Zentrum Geesthacht GmbH Lichtenbergstr. 1, 85747 Garching bei Mnchen, Germany, Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Barcelona, Catalonia (Spain), and ISIS Facility, Rutherford Appleton Laboratory, Didcot, Chilton, UK

E-mail: sylvia.mclain@bioch.ox.ac.uk

*To whom correspondence should be addressed

[†]University of Oxford

[‡]University of Oxford

[¶]German Engineering Materials Science Centre

[§]Universitat Politècnica de Catalunya

^{||}ISIS Facility

Abstract

The microscopic structure of the tryptophan side chain, indole, in an amphiphilic environment has been investigated using a combination of neutron diffraction measurements and simulations in solution. The results show that indole is preferentially solvated by hydrogen bonding interactions between water and alcohol -OH groups rather than the interaction being dominated by indole-methyl interactions. This has implications for understanding how tryptophan interacts with the amphipathic membrane environment to anchor proteins into membranes, where the results here suggest that the benzene ring of tryptophan interacts directly with the interfacial water at the membrane surface rather than being buried into the hydrophobic regions of the membrane bilayer.

KEYWORDS: Hydration in solution, Neutron diffraction with isotopic substitution, EPSR, membrane-peptide interactions

INTRODUCTION

How amphiphilic proteins are anchored into cellular membranes is dependent upon the complex environment of the membrane-bilayer. In transmembrane proteins, the amino acid tryptophan (Trp) has been found to cluster preferentially at the interfacial region^{1,2} where it is thought to anchor proteins at the bilayer-solution interface.³⁻⁶ Previous studies⁷⁻⁹ indicate that Trp favors the interfacial region over the aqueous solvent more strongly than any other natural amino acid, where it is suggested that Trp is mostly bound at the interfacial region between the aqueous solvent and the headgroups of the lipid bilayer.^{10,11} Despite a supposed fundamental role in protein structure and thereby function, the details of the interactions of Trp with lipids and the aqueous solution at the membrane-water interface remain unknown.

Interactions of Trp with the membrane-water interface are governed by its indole side chain (Fig.1), which has an quadrupole moment from the aromatic, delocalized π electrons, creating regions of partial negative charge above and below the benzene ring.^{3,12} This is significant, as it has been suggested that charged and aromatic residues, in general, act as interfacial anchoring residues for membrane proteins.¹³ Considering its predominantly hydrophobic structure, the Trp side chain is expected to have an inherent affinity for the lipid bilayer.¹³ Although it has been long known that Trp locates preferentially at the membrane-water interface, the mechanism by which this binding occurs is less clear. Previous binding studies have shown that indole occupies a location partly exposed to a polar environment in bilayers but have not specified its exact location.¹⁴

Previous research has yielded mixed hypotheses about the dominant factors that dictate indole, and hence tryptophan, binding in membranes. While some investigations have suggested that cation- π interaction is the dominant factor governing Trp-lipid specificity,¹⁵⁻¹⁷ or even that the aromatic ring acts as a hydrogen bond acceptor either with water or hydroxyl groups present at the lipid-water interface,¹⁸ indole N-H hydrogen bond donation to the surrounding lipid head groups has also been proposed as the dominant factor governing Trp specificity.^{10,19} In contrast, some theories suggest that Trp residues are spatially oriented

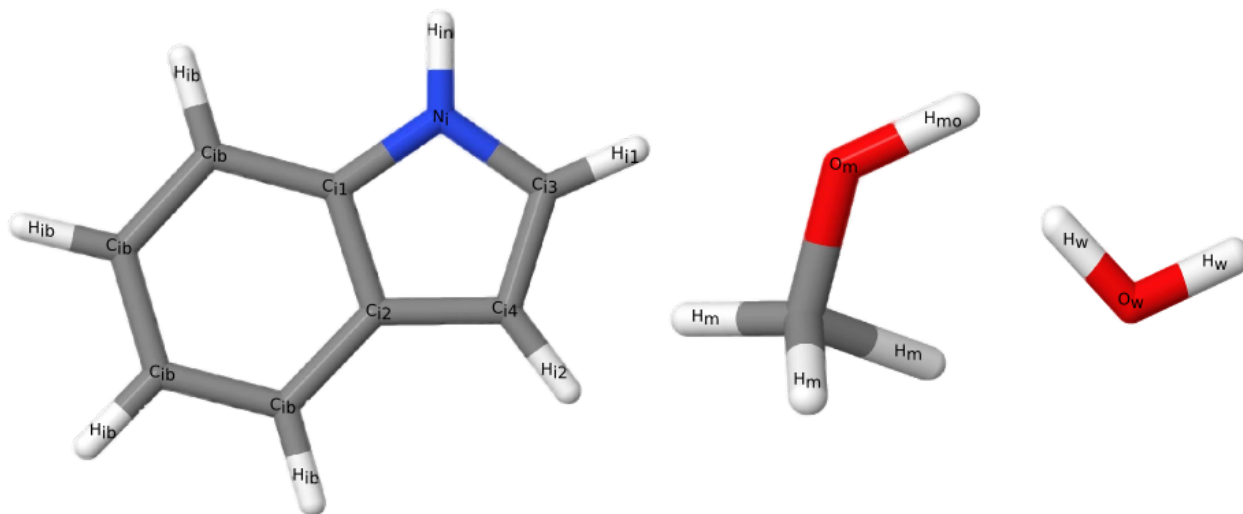


Figure 1: The molecular structure of indole, methanol and water molecules with appropriate atomic labels.

with the benzene ring sequestered into the hydrophobic domain with the indole -NH group aligned so as to hydrogen bond to the membrane surface.²⁰⁻²² Yet other investigations claim that hydrophobic interactions with the lipid tails also govern indole orientation, suggesting that hydrophobic rings are immersed in the lipid part of the bilayer while the NH moieties are free to bind with the lipid head groups.³

Given the multiplicity of theories, further detailed investigation at the atomic level is necessary to discern the physical basis underlying the preference of Trp for the interface.²³ In order to provide atomic resolution structural data to inform this debate, the microscopic solvation structure of indole (the functional group of tryptophan) in methanol/water solutions has been measured using a combination of neutron diffraction with isotopic substitution (NDIS) and computer simulation - Empirical Potential Structural Refinement (EPSR). Methanol/water solutions were used, as this amphiphilic solution shares some important characteristics of amphiphilic biological interfaces. Previous measurements on methanol/water solutions have revealed that on the atomic length scale, water molecules and hydroxyl groups are segregated from the close-packed methyl groups,²⁴ yet this solution remains fully miscible in the bulk phase. Indole has a low solubility in water,²⁵ yet it is extremely soluble in methanol/water

solutions with a high water content. While clearly there is some interplay between molecular structures in this ternary solution these are not yet understood. Understanding indole in these solutions is essential to recognizing how indole is partitioned into an amphiphilic environment and indeed interacts with biological membranes in solution.

MATERIALS AND METHODS

Theory

Neutron Diffraction with Isotopic Substitution

Neutron diffraction by isotopic substitution (NDIS) is a method by which the structure of biomolecular solutions can be determined. As opposed to X-ray scattering, where the intensity of the scattering signal is proportional to the number of electrons, intensity in a neutron diffraction experiment is dictated by the neutron-nucleus interaction where scattering intensity is roughly similar across the periodic table. In addition, the use of neutrons as a probe is particularly useful when determining the structure of hydrogen bonding solutes in water, as the interactions between neutrons and H and D yields scattering of different intensity.²⁶ Measurement of chemically identical yet isotopically unique samples allows for multiple measurements of the same system which only differ with respect to their H and D concentrations.

Neutron diffraction provides a measure of the static structure factor, $F(Q)$, which is a direct measure of the structure in solution. Scattered neutron intensity as a function of the momentum transfer, Q , between the incident neutron and atoms in the sample ($Q = 4\pi/\lambda \cdot \sin(2\theta/2)$) with the neutron wavelength, λ , and the scattering angle 2θ) is related to $F(Q)$ via

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

where c and b are the concentrations and the scattering lengths, respectively, of all the

atom types α and β in the sample. $\delta_{\alpha\beta}$ is the Kronecker delta function. Importantly, the scattering lengths for H and D are -3.373 fm and 6.671 fm,²⁶ respectively; where this large difference gives rise to distintively different diffraction patterns (*vide infra*). As can be seen from 1, $F(Q)$ is the weighted sum of the partial structure factors $S_{\alpha\beta}(Q)$ for the different atomic pairs present in the sample. These partial structure factors in reciprocal space are directly related to the radial distribution functions (RDFs) $g_{\alpha\beta}(r)$ in real space *via* Fourier transformation

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

where ρ is the atomic number density of the sample (atoms \AA^{-3}) and $g_{\alpha\beta}(r)$ are the atomic distances between β and α atoms as a function of distance (r). Coordination numbers ($n_{\alpha}^{\beta}(r)$), i. e. the average number of β atoms around a central α atom at a distance between r_{\min} and r_{\max} , can be calculated by integration of the $g(r)$ functions *via*

$$n_{\alpha}^{\beta} = 4\pi\rho c_{\beta} \int_{r_{\min}}^{r_{\max}} r^2 g(r) dr \quad . \quad (3)$$

Empirical Potential Structure Refinement

Empirical Potential Structure Refinement (EPSR) is a simulation technique which has been explicitly written for the determination of the local order present in disordered materials, such as liquids and glasses.²⁷ EPSR is a reverse Monte Carlo technique where the atomic conformation of the system in question is constrained to fit a set of diffraction data. Beginning with a box of molecules at the density and temperature of the diffraction measurements, a set of seed potentials are given to each unique atomic component - which in this case consisted of Lennard-Jones potentials (σ and ϵ) as well as appropriate atomic charges. These seed potentials are subsequently refined by the EPSR process, generating a new set of potentials. While EPSR provides a three-dimensional representation of the system in question, similar to all simulation processes, this structure is not unique, in that it is not the only possible struc-

ture which could be consistent with the data. However, by using physical constraints, such as minimum bonding distances, a set of potentials as well as appropriate atomic charges an experimentally consistent 3-dimensional structural model of the system in question emerges. EPSR coupled with NDIS data has been used to determine the local structure of liquids in a wide variety of hydrogen containing solutions.^{24,28-40} Further details on EPSR can be found elsewhere in the literature.^{27,41,42}

Neutron Diffraction Measurements

Indole was purchased from Sigma and recrystallized by dissolution in ethanol for 24 hours before the solvent was removed by vacuum. All isotopomers of methanol were dried over Mg which had been previously activated with I_2 and were subsequently refluxed at $\sim 330K$ for 48 hours under vacuum. The methanol isotopomers were then cryogenically distilled onto pre-dried 3 Å molecular sieves, yielding anhydrous methanol which was verified by 1H NMR. For all neutron diffraction and NMR samples, solutions of indole in water/methanol solutions were prepared by weight under an N_2 atmosphere.

For the neutron scattering measurements several different isotopomers of indole/methanol/water solutions were prepared at a relative molar ratio of 1 indole: 29 methanol: 30 water ($\sim 1M$). The diffraction data were collected using the SANDALS instrument, located at the ISIS facility (STFC, UK). The samples were measured in flat Ti/Zr canisters with thickness of 1mm. Ti/Zr alloy cans were used as Ti and Zr have coherent scattering lengths of -3.438 fm and 7.160 fm,²⁶ respectively, and can be combined to produce negligible scattering from the sample containers. Diffraction data were collected from the samples themselves as well as the empty canisters and empty instrument to ensure a suitable background subtraction. The data were corrected for absorption, multiple scattering and inelasticity effects and then subsequently to $F(Q)$ using the program GUDRUN, which is based on the ATLAS suite of program to correct neutron diffraction data and is freely available at ISIS.⁴³ The measured $F(Q)$ data is shown in Figure 2.

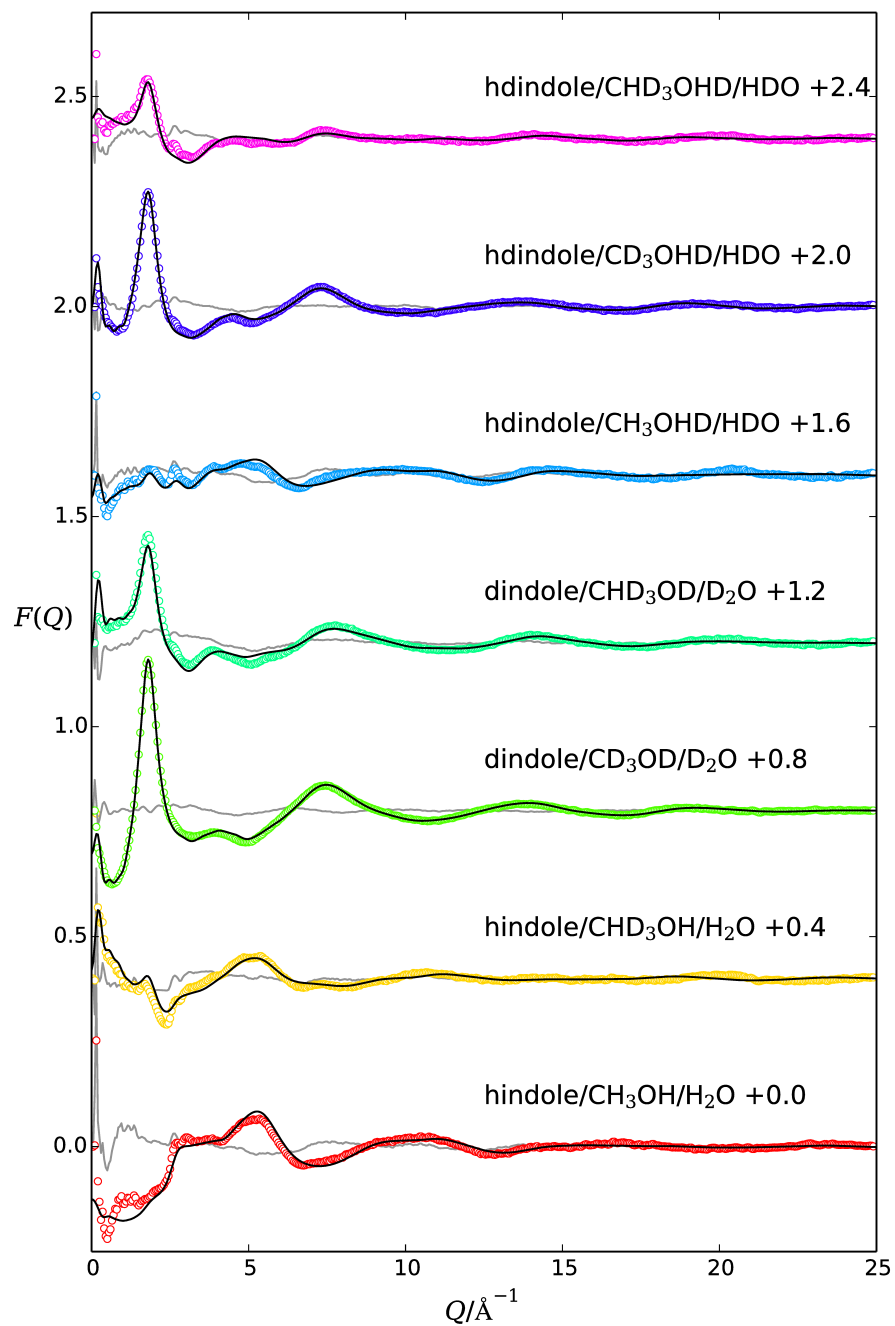


Figure 2: Measured neutron diffraction data, $F(Q)$ (coloured lines), EPSR fits to the data (black) and the difference (grey). The data have been translated by 0.4 for clarity.

Emperical Potential Structure Refinement Simulations

An EPSR simulation box was constructed to correspond to the measured solution density of $0.0912 \text{ atoms } \text{\AA}^{-3}$ at 298K and contained 25 molecules of indole, 725 of methanol and 750 of water. As mentioned above, the seed potentials used in the EPSR simulations are listed in the SI. Briefly, atoms of indole, depending on the similarity of their charges, were grouped together and given potentials defined by the average of the parameters from the OPLS all-atom (OPLS-AA) force field. In order to keep indole planar, uncharged dummy atoms with epsilon and sigma values of zero were positioned above and below the geometrical centers of both the pyrrole and benzene rings of indole in the EPSR simulation. These atoms restricted any movement of the indole molecule away from being planar. EPSR simulations were performed which confirmed that the addition of these dummy atoms did not influence the fits to the diffraction data but nevertheless effected a more physically realistic model. This comparison is shown in the SI. For water, SPC/E parameters were used,⁴⁴ with the addition of σ and ϵ values for hydrogen atoms (see SI). Methanol parameters were from the OPLS-AA force field and are also shown in the SI.⁴⁵ Molecular structures of indole and methanol were created and energetically optimized using Avogadro⁴⁶ before being included in the EPSR simulation box. The measured neutron diffraction data ($F(Q)$) along with the EPSR fits to the data and the difference between are shown in the SI.

The three-dimensional arrangements of solvent molecules around indole were determined using the program ANGULA^{47,48} from the EPSR simulation box. A coordinate system was first assigned to the indole molecules, with the origin taken as the mid-point of the carbon atoms joining the benzene and pyrrole rings. Coordinate systems were similarly assigned to the solvent molecules, such that the angles between any pair of coordinate systems could be calculated in addition to the distribution of the solvent molecules around the solute. In this instance, only molecules within 5\AA were used for further analysis. This procedure was performed for multiple snapshots of the simulation box, such that the aggregate molecule distributions are plotted as spatial density map (SDMs; *vide infra*).

RESULTS AND DISCUSSION

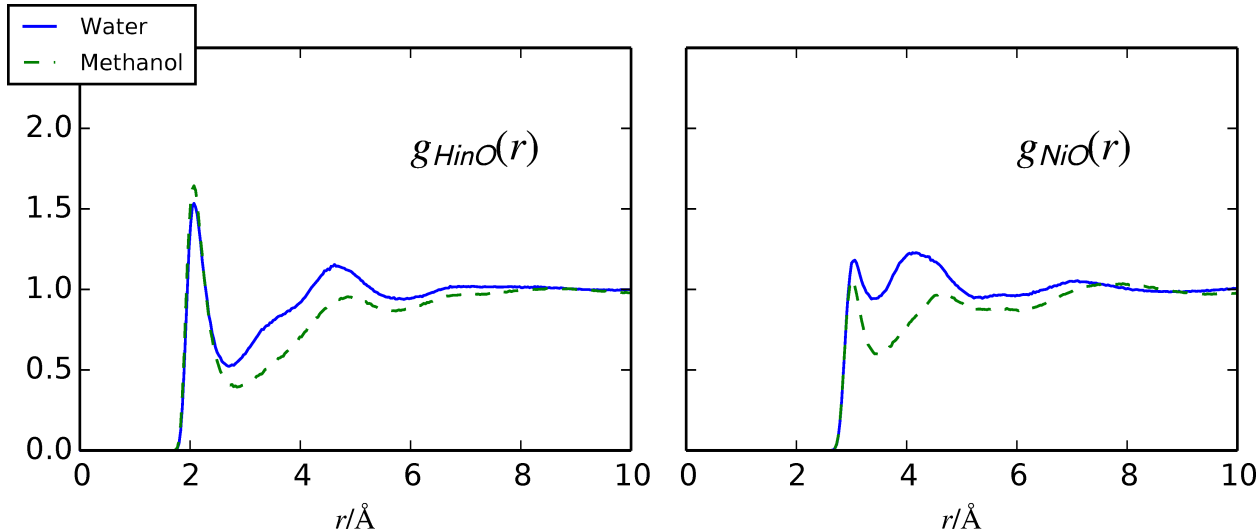


Figure 3: RDFs for interactions of the N-H indole hydrogen (H_{in}) and indole nitrogen (N_i) with the water oxygen (O_w ; blue, solid line) and the methanol oxygen (O_m ; green, dashed line).

Fig. 3 shows RDFs for indole N-H hydrogen (H_{in}) and indole nitrogen (N_i) interactions with the surrounding water and methanol molecules (Fig. 1). Both $g_{H_{in}O_w}(r)$ and $g_{H_{in}O_m}(r)$ show a sharp peak at ~ 2.1 Å, indicating the presence of a hydrogen bond between the indole H_{in} and the solvent molecules. The coordination numbers for these RDFs (Table 1) indicate that roughly 50% of H_{in} donate hydrogen bonds to water ($n_{H_{in}}^{O_w} = 0.51$) while the other half occurs between H_{in} and methanol ($n_{H_{in}}^{O_m} = 0.48$), with no clear binding preference to either solvent. Even though indole-water hydrogen bonding here is somewhat larger than might be expected for indole amine-water hydrogen bonds, it seems unlikely that this bonding is responsible for indole solubility in the present solutions, given that indole is sparingly soluble in pure water.

Similar to $g_{H_{in}O_w}(r)$ and $g_{H_{in}O_m}(r)$, both $g_{N_iO_w}(r)$ and $g_{N_iO_m}(r)$ show two peaks. The first of these peaks is at ~ 3.0 Å for both solvents, and the second peak is at ~ 4.2 Å and ~ 4.7 Å in water and methanol, respectively. The first peak is related to the aforementioned

$H_{in}-O_w$ and $H_{in}-O_m$ bonds, yet the second peaks for both O_w and O_m oxygens in the $g_{N_iO}(r)$ s are at much shorter distances than would be expected for a second hydration shell around the indole NH group. If this second peak were merely reflective of a second solvent layer around the hydrogen bound solvent, this distance would be at a distance $\sim 1.0 \text{ \AA}$ greater than the second peak in the $g_{H_{in}O}(r)$ s. However, for both methanol and water $g_{N_iO}(r)$ s this distance is smaller. This indicates that the solvent molecules are coordinated around indole by means of interaction unassociated with the N-H group hydrogen-bond, where this behavior is more pronounced for water compared with methanol. Consideration of the corresponding coordination numbers (Table 1) underlines this observation with water molecules showing a higher coordination ($n_{N_i}^{O_w} = 0.77$) compared with methanol ($n_{N_i}^{O_m} = 0.59$).

Table 1: Coordination numbers for the RDFs shown in Fig. 3.

n_{α}^{β}	EPSR	$r/\text{\AA}$
$n_{H_{in}}^{O_w}$	0.51	2.70
$n_{N_i}^{O_w}$	0.77	3.42
$n_{H_{in}}^{O_m}$	0.48	2.70
$n_{N_i}^{O_m}$	0.59	3.42

Fig. 4 shows the RDFs for interactions between indole’s benzene ring carbon atoms (C_B ; where $C_B = 4C_{ib} + C_{i1} + C_{i2}$; Fig. 1) and both water and methanol, and Table 2 lists the coordination numbers for these functions. The large peak for the C_B -water oxygen interactions indicate that waters are preferentially located around the indole aromatic ring; the relative size and sharpness of this peak, further suggests that these waters are interacting *directly* with the benzene ring. Conversely, for the methanol hydroxyl oxygen O_m , while there is some density of methanol hydroxyl groups present around the benzene ring at slightly less than 4 \AA , the more prominent interactions occur over a larger distance range, evidenced by the peak at around 5 \AA . Further, the methyl carbon - benzene contacts in Fig. 4 ($g_{C_B C_m}(r)$), which are a measure of the hydrophobic-hydrophobic distances between methanol and the indole benzene ring, have a peak position of $\sim 4.1 \text{ \AA}$, showing that these non-polar interactions take place over a larger length scale compared with water-benzene interactions. Inspection of

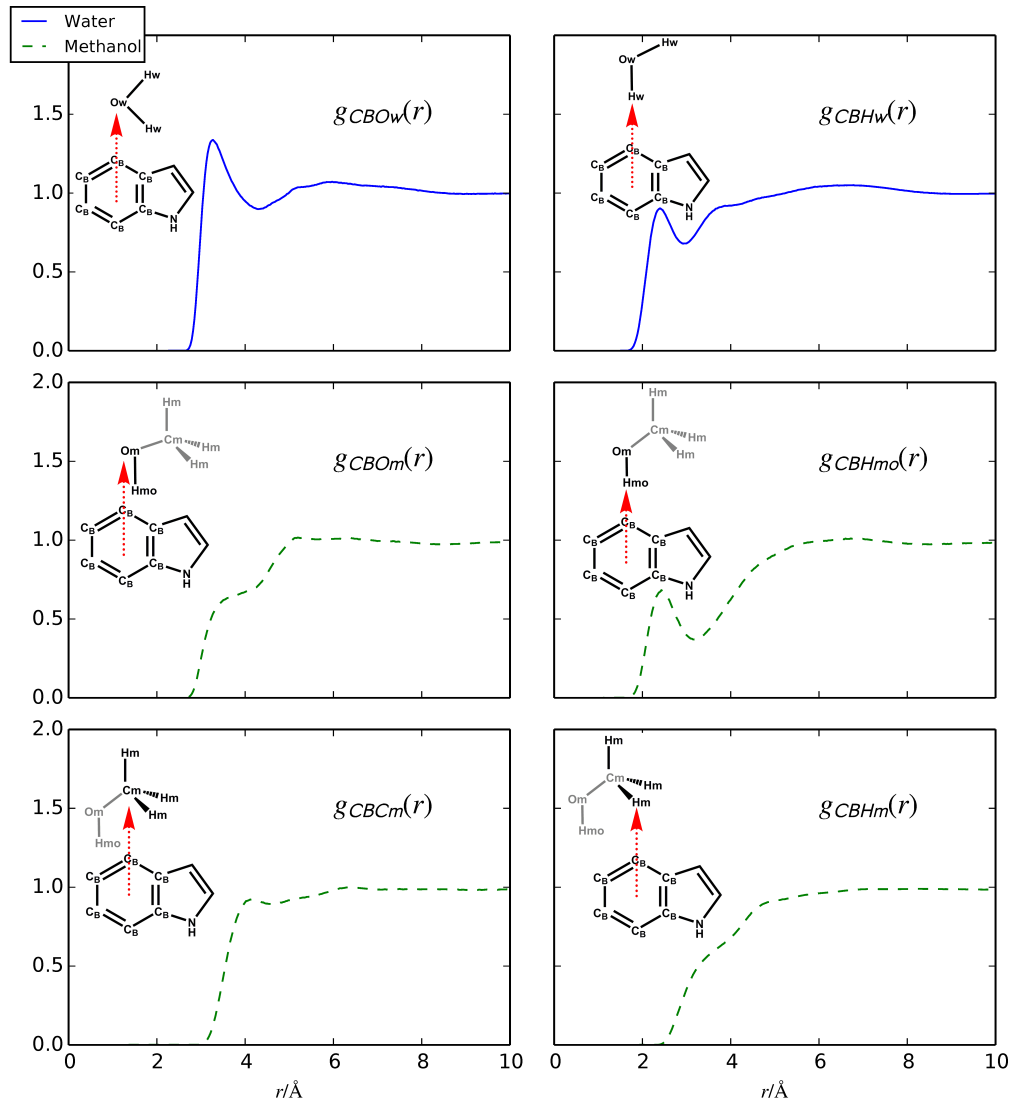


Figure 4: RDFs for interactions between the benzene ring (C_B - the weighted average of C_{i1} , C_{i1} and C_{i2}) and solvent atoms of water (O_w , H_w ; blue, solid line) and methanol (O_m , H_{mo} , H_m , C_m ; green, dashed line).

the coordination numbers for the C_B -functions at 4.3 Å for the O_w , O_m and C_B atoms shows that for these nearest neighbor solvent molecules the total number of hydroxyl interactions ($O_m + O_w$) with indole is ~ 4 times greater within this distance range than the indole-methyl interactions, as the C_B - C_m coordination number at this same distance is ~ 1.3 .

Table 2: Coordination numbers of the RDFs in Fig. 4.

n_{α}^{β}	EPSR	$r/\text{Å}$
$n_{C_B}^{O_w}$	2.50	4.32
$n_{C_B}^{H_w}$	1.14	2.94
$n_{C_B}^{O_m}$	1.36	4.32
$n_{C_B}^{H_{mo}}$	0.49	3.18
$n_{C_B}^{C_m}$	1.34	4.32

Given the surprisingly high probability density of O_w and O_m around indole’s benzene ring, the presence (or absence) of hydrogen bonding configurations can be determined by inspection of the solvent hydrogen-indole RDFs which are also shown in Fig. 4. There is a clear indication of a H-bond-like interaction between the water hydrogen (H_w) indole at around 2.4 Å where the coordination number of this peak of ~ 1.1 indicates that, on the average, at least one water molecule is bound with this distance range. Interestingly, the methanol -OH hydrogen (H_{mo}) - indole RDFs also show hydrogen bonding from the hydroxyl group to the benzene carbons. The presence of peaks for both water and hydroxyl hydrogen RDFs around indole’s C_6 ring at these short distances in Fig. 4 suggests that a large proportion of the electrostatic interactions in this solution are by virtue of benzene-hydrogen bonds. Conversely, the methyl hydrogen (H_m) - indole RDFs show no distinct peaks, indicating relatively disordered interactions between the -CH₃ hydrogens and the benzene ring. Moreover, these interactions, similar to the indole- C_m interactions, take place over a longer distance range compared to the electrostatic hydrogen-bond interactions from water or methanol hydrogen groups.

Fig. 5 shows the average orientations of methanol and water molecules around indole’s benzene. The benzene- H_w - O_w angle is almost linear (Fig. 5a, while a second peak in the

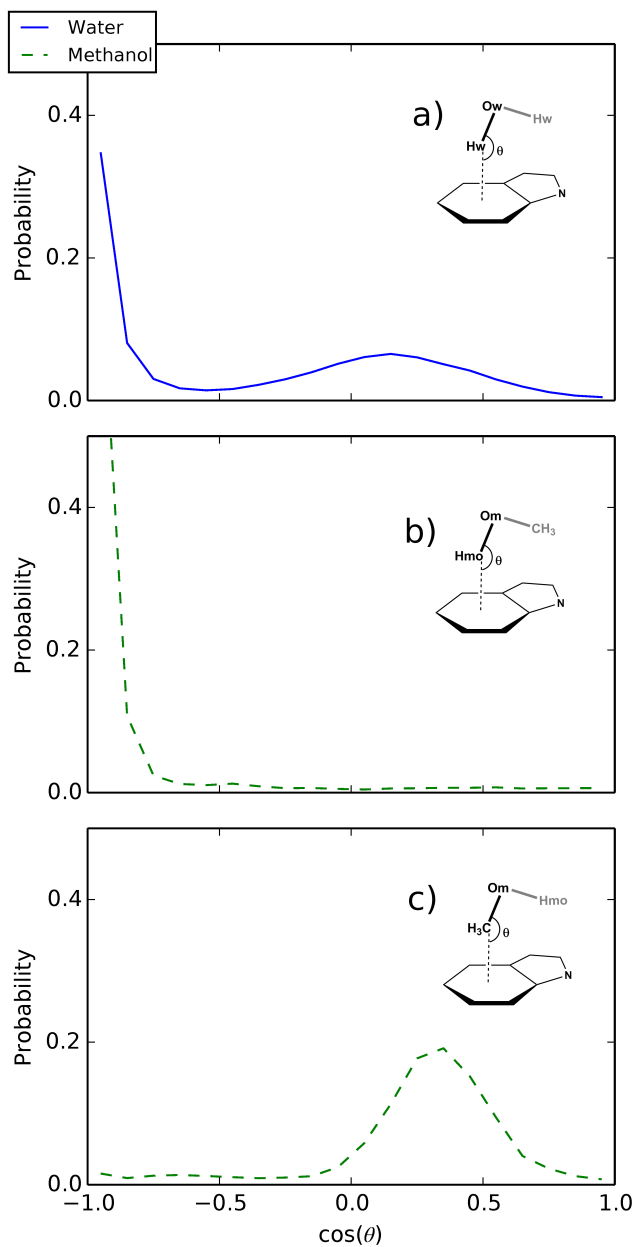


Figure 5: Average angular orientation ($\cos(\theta)$) from the center of the indole ring to the surrounding solvent molecules where a) depicts the benzene- O_w - H_w angle, b) the benzene- H_m - O_m angle and c) the benzene- O_m - C_m angle from both EPSR.

benzene-H_w-O_w reaches a maximum at a value of around $\cos(0.25)$ ($\sim 75^\circ$). This broad peak shows the second O_w-H_w from the nearest neighbor waters and suggests that this second covalently bound hydrogen is freely rotating around the hydrogen-bond formed by the benzene \cdots H_w-O_w motif. The benzene-H_m-O_m and benzene-C_m-O_m angles are broadly similar to the benzene-H_w-O_w orientation, that is the benzene-H_m-O_m shows at $\cos=-1$ ($\sim 180^\circ$) and the benzene-C_m-O_m has a broad density distribution at $\cos=0.75$.

Solvation information in three dimensions can also be extracted from the simulations *via* Spatial Density Maps (SDMs)^{47,48} which show the most probable locations of water and methanol molecules around indole in the solutions. Fig. 6 shows the SDMs for water (O_w-indole) and methanol molecules both with respect to the methanol O_m-indole correlations and to the methanol C_m-indole correlations, where the density of the solvent molecules is illustrated by an isocontour surface that envelops the top 50% of solvent molecules and by cuts through the orthogonal xy, xz and yz planes, where these slices were taken through the origin.⁴⁸

As might be anticipated from Fig. 3, the hydration SDMs show a high probability of water binding to the N-H nitrogen on the pyrrole ring by virtue of an N-H \cdots O_w-H_w hydrogen bond. Perhaps more surprisingly, but consistent with the $g(r)$ s shown in Fig. 4, these SDMs also clearly show that water is preferentially located around the benzene ring - specifically in the middle of the benzene ring. The analogous SDM for methanol O_m around indole similarly shows a high propensity for the hydroxyl group of methanol to be located either around the benzene ring in addition to forming N-H \cdots O_m hydrogen bonding motifs. The methyl groups, on the other hand, show a broad distribution of preferred locations around indole where the density of -CH₃ groups above the benzene ring is much lower compared with either O_w or O_m.

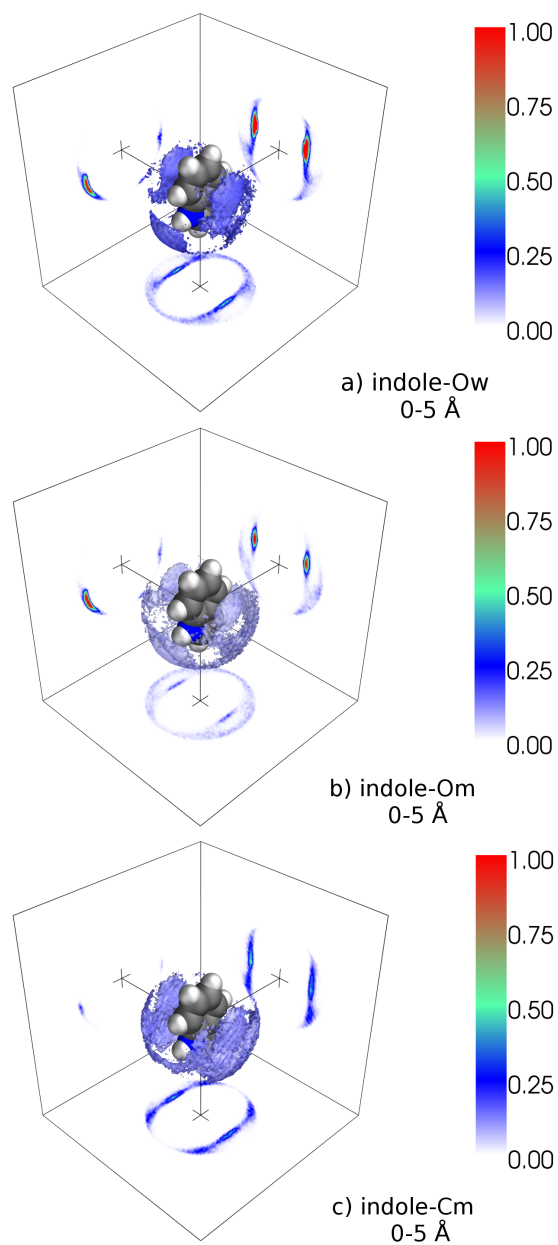


Figure 6: Spatial density map for a) water O_w, b) methanol O_m, and c) methanol C_m around indole for EPSR enclosing 0-5 Å from the indole mid point (see SI).

These maps show the density of neighboring molecules around central indole molecule per Å³ and contour surface encloses 50% of molecules within the highest density.

CONCLUSIONS

In the current indole/water/methanol solution electrostatic configurations appear to be preferred over hydrophobic-hydrophobic interactions for indole in this model amphiphilic solution. Even though it might be expected that some water-benzene or OH-benzene interactions would be present, that the solution structure of indole is dominated by these electrostatic interactions is somewhat surprising, given that Trp is sometimes ranked as the most hydrophobic of amino acids.⁴⁹ The standard theory of the hydrophobic effect would predict indole to be preferentially solvated by hydrophobic motifs, as methanol microsegregates in solution,²⁴ offering a somewhat continuous surface of hydrophobicity in which the indole could, in principle, be partitioned. On the other hand, it could be that the hydrophobic surfaces in the present solutions are ‘too small’, as according to Chandler, a large enough hydrophobic cavity must be present for the ‘hydrophobic effect’ to drive assembly.⁵⁰ Whatever the case may be, the data here show that electrostatic interactions between the benzene ring from either methanol or water comprises most of the nearest-neighbor solvation around indole. To show this, a representative snapshot of both water and methanol solvating indole from the EPSR fits to the data is shown in Fig. 7.

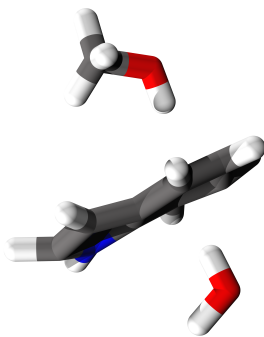


Figure 7: A representative snapshot of the hydrogen bonding configurations to indole from methanol and water solvent molecules taken from the EPSR simulation box.

Taken overall these data imply that electrostatic forces have a larger influence than might

be expected on tryptophan side chain interactions within the membrane bilayer. Rather than merely being excluded from the hydrocarbon core due to entropic reasons,⁸ the atomic scale structural data here show that indole is likely to interact with charges present at the interface of a bilayer, predominantly through its benzene ring motif. It could be that the benzene ring binds with water at this interface similar to the interactions shown in Fig. 7 and that this water may in turn be bound to the polar regions of the lipid head groups, aligning the Trp in such a way that it can easily anchor membrane proteins into the bilayer. From the present work, it is not possible to make any suggestions as to the orientation of Trp residues relative to the membrane bilayer itself.

Our results however do emphasize the importance of electrostatic interactions in the behavior of proteins and their interactions with amphiphilic membrane interfaces. There are many previous studies which show hydrogen bonding interactions with π electrons from aromatic systems,^{51,52} where the X-H $\cdots\pi$ interactions can be considered as weak hydrogen-bonding interactions.⁵³ Although cation- π interactions are rare in proteins compared with other electrostatic interactions - such as from salt bridges - such interactions have been proposed as important factors in determining the three-dimensional structure and recognition properties of proteins and have been observed previously in crystal structures of small molecules and proteins as well as in gas-phase spectroscopic measurements.^{51,52,54-57}

In a biological membrane, a steep polarity gradient exists across the interfacial region from the highly polar aqueous phase to the highly apolar hydrocarbon core,⁴ where it is not clear how molecules - traditionally classed as polar or non-polar - will interact with a membrane interface. The classification of certain molecular motifs as polar or non-polar does not necessarily dictate their specific interactions in protein structures or, by analogy, in specific stabilizing interactions at the protein membrane interface.^{29,58,59} It has rather been noted that molecular association in amphiphilic environments is more complex at the atomic scale.²⁹

Here, it appears that the amphiphilic solution is indeed needed for indole to interact with

water; indole is sparingly soluble in water yet highly soluble in methanol/water solutions. It is clear that it is not merely the electrostatics themselves but the interplay between hydrophobic and electrostatic interactions which give rise to the resultant structure observed here. Interestingly, it has been noted that structures which are deficient in hydrogen-bond acceptors will make use of the π -systems of aromatic rings.⁵³ This appears to be pertinent to the results presented here as water, in the absence of methanol, does not bind to indole strongly enough to allow for mixing between the solvent and the solute. Perhaps the electrostatic interactions between benzene and the -OH groups in methanol place the -CH₃ groups close enough to benzene-ring to participate in stabilizing van der Waals interactions between indole and methanol.

This mechanism would also allow for the binding of water, as water is microsegregated from the methyl groups in the solution²⁴ and as a result would be naturally placed near the benzene ring in indole. The propensity for indole interact with water molecules only when in an amphiphilic environment is consistent with experimental data showing a greater percentage of Trp in membrane proteins compared with water soluble proteins.^{3,20} In the more complex membrane environment, it is possible that the interplay of both electrostatic interactions and hydrophobic interactions between the indole and water in conjunction with the amphiphilic lipid molecules may allow for tryptophan to be stabilized at the membrane bilayer interface. For instance, the PO₄⁻ oxygens the phosphocholine lipid head group form highly directed hydrogen bonds with water in solution.³⁰ This would leave the water hydrogen that is not directly participating in hydrogen bonding to the P=O group available to bind with the benzene ring in Trp residue to help stabilize the protein into the bilayer. This view is consistent with energetic calculations which show that Trp has a global energetic minimum when it is located in the head group region of a membrane.⁶⁰ Whether or not this mechanism is correct, what is clear from the present work is that electrostatic- π interactions play a significant role in the solvation of indole and these interactions are likely to dictate the role Trp plays in anchoring membrane proteins into their native environments.

Acknowledgement

We thank the ISIS Facility (Rutherford Appleton Laboratories, STFC, UK) for the allocation of neutron beam time, the UK Engineering and Physical Sciences Research Council (EP/J002615/1) for funding and Dr. Richard Gillams (Oxford) for assistance with data visualization and Prof. Mark Sansom and Dr. Phillip Stansfeld (both of Oxford) and Prof. Alan Soper (STFC, UK) for useful discussions.

Supporting Information Available

Additional tables and figures showing seed potentials for EPSR simulations, the Fourier Transformation of the data compared with the EPSR fits in real space, details of the charges used in the simulations for indole as well as RDFs for the solvent solvent interactions and EPSR with and without dummy atoms for indole are shown in the supporting information. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) von Heijne, G. Membrane Proteins: From Sequence to Structure. *Annual Review of Biophys. and Biomol. Structure* **1990**, *23*, 167–192.
- (2) Persson, S.; Killian, J. A.; Lindblom, G. Molecular Ordering of Interfacially Localized Tryptophan Analogs in Ester- and Ether-Lipid Bilayers Studied by ^2H -NMR. *Biophys. J.* **1998**, *75*, 1365–1371.
- (3) Schiffer, M.; Chang, C. H.; Stevens, F. J. The Functions of Tryptophan Residues in Membrane Proteins. *Protein Engineering* **1992**, *5*, 213–214.
- (4) Killian, J. A.; von Heijne, G. How Proteins Adapt to a Membrane-water Interface. *Trends in Biochemical Sciences* **2000**, *25*, 429–434.

- (5) O'Connell, A. M.; Koeppe, R. E.; Andersen, O. S. Kinetics of Gramicidin Channel Formation in Lipid Bilayers: Transmembrane Monomer Association. *Science (New York, N. Y.)* **1990**, *250*, 1256–1259.
- (6) Ridder, A. N. J. A.; Morein, S.; Stam, J. G.; Kuhn, A.; de Kruijff, B.; Killian, J. A. Analysis of the Role of Interfacial Tryptophan Residues in Controlling the Topology of Membrane Proteins. *Biochem.* **2000**, *39*, 6521–6528.
- (7) Jacobs, R. E.; White, S. H. The Nature of the Hydrophobic Binding of Small Peptides at the Bilayer Interface: Implications for the Insertion of Transbilayer Helices. *Biochem.* **1989**, *28*, 3421–3437.
- (8) Yau, W. M.; Wimley, W. C.; Gawrisch, K.; White, S. H. The Preference of Tryptophan for Membrane Interfaces. *Biochem.* **1998**, *37*, 14713–14718.
- (9) White, S. H.; Wimley, W. C. Membrane Protein Folding and Stability: Physical Principles. *Annual Review of Biophys. and Biomol. Structure* **1999**, *28*, 319–365.
- (10) de Planque, M. R.; Kruijtzter, J. A.; Liskamp, R. M.; Marsh, D.; Greathouse, D. V.; Koeppe, R. E.; de Kruijff, B.; Killian, J. A. Different Membrane Anchoring Positions of Tryptophan and Lysine in Synthetic Transmembrane Alpha-helical Peptides. *J. Biol. Chem.* **1999**, *274*, 20839–20846.
- (11) Petersen, F. N. R.; Jensen, M. Ø.; Nielsen, C. H. Interfacial Tryptophan Residues: A Role for the Cation- π Effect? *Biophys. J.* **2005**, *89*, 3985–3996.
- (12) Dougherty, D. Cation- π Interactions in Chemistry and Biology : A New View of Benzene, Phe, Tyr, and Trp. *Science* **1996**, *271*, 163–168.
- (13) van der Wel, P. C. A.; Reed, N. D.; Greathouse, D. V.; Koeppe, R. E. Orientation and Motion of Tryptophan Interfacial Anchors in Membrane-Spanning Peptides. *Biochem.* **2007**, *46*, 7514–7524.

- (14) Kachel, K.; Asuncion-Punzaland, E.; London, E. Anchoring of Tryptophan and Tyrosine Analogs at the Hydrocarbon-Polar Boundary in Model Membrane Vesicles: Parallax Analysis of Fluorescence Quenching Induced by Nitroxide-Labeled Phospholipids. *Biochem.* **1995**, *34*, 15475–15479.
- (15) Woolf, T. B.; Grossfield, A.; Pearson, J. G. Indoles at Interfaces : Calculations of Electrostatic Effects with Density. *International J. Quantum Chem.* **1998**, *75*, 197–206.
- (16) Blaser, G.; Sanderson, J. M.; Wilson, M. R. Free-energy Relationships for the Interactions of Tryptophan with Phosphocholines. *Organic & Biomol. Chem.* **2009**, *7*, 5119–5128.
- (17) Sanchez, K. M.; Kang, G.; Wu, B.; Kim, J. E. Tryptophan-Lipid Interactions in Membrane Protein Folding Probed by Ultraviolet Resonance Raman and Fluorescence Spectroscopy. *Biophys. J.* **2011**, *100*, 2121–2130.
- (18) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A.; Blake, G. A. Benzene Forms Hydrogen Bonds with Water. *Science* **1992**, *257*, 942–945.
- (19) de Jesus, A. J.; Allen, T. W. The Role of Tryptophan Side Chains in Membrane Protein Anchoring and Hydrophobic Mismatch. *Biochimica et Biophysica Acta* **2012**, *1828*, 864–876.
- (20) Hu, W.; Cross, T. A. Tryptophan Hydrogen Bonding and Electric Dipole Moments: Functional Roles in the Gramicidin Channel and Implications for Membrane Proteins. *Biochem.* **1995**, *34*, 14147–14155.
- (21) Hu, W.; Lee, K.-C.; Cross, T. A. Tryptophans in Membrane Proteins: Indole Ring Orientations and Functional Implications in the Gramicidin Channel. *Biochem.* **1993**, *32*, 7035–7047.

- (22) Esbjörner, E.; Caesar, C. E.; Albinsson, B.; Lincoln, P.; Nordeñ, B. Tryptophan Orientation in Model Lipid Membranes. *Biochemical and Biophysical Research Communications* **2007**, *361*, 645–650.
- (23) Koeppe, R. E. Concerning Tryptophan and Protein-Bilayer Interactions. *J. General Physiology* **2007**, *130*, 223–224.
- (24) Dixit, S.; Crain, J.; Poon, W. C. K.; Finney, J. L.; Soper, A. K. Molecular Segregation Observed in a Concentrated Alcohol-Water Solution. *Nature* **2002**, *416*, 42–57.
- (25) Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics, Internet Version 2005*; CRC Press: Boca Raton, FL, 2005.
- (26) Sears, V. F. Neutron Scattering Lengths and Cross Sections. *Neutron News* **1992**, *3*, 26–37.
- (27) Soper, A. K. Computer Simulation as a Tool for the Interpretation of Total Scattering Data from Glasses and Liquids. *Mol. Sim.* **2012**, *38*, 1171–1185.
- (28) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. Solvation of Hydroxyl Ions in Water. *J. Phys. Chem.* **2003**, *119*, 5001–5004.
- (29) Bowron, D. T.; Finney, J. L. Association and Dissociation of an Aqueous Amphiphile at Elevated Temperatures. *J. Phys. Chem. B* **2007**, *111*, 9838–9852.
- (30) Foglia, F.; Lawrence, M. J.; Lorenz, C. D.; McLain, S. E. On the Hydration of the Phosphocholine Headgroup in Aqueous Solution. *J. Chem. Phys.* **2010**, *133*, 145103.
- (31) Headen, T. F.; Howard, C. A.; Skipper, N. T.; Wilkinson, M. A.; Bowron, D. T.; Soper, A. K. Structure of Interactions in Aromatic Liquids. *J. Am. Chem. Soc.* **2010**, *132*, 5735–5742.

- (32) Hargreaves, R.; Bowron, D. T.; Edler, K. Atomistic Structure of a Micelle in Solution Determined by Wide Q-Range Neutron Diffraction. *J. Am. Chem. Soc.* **2011**, *133*, 16524–16536.
- (33) Rhys, N. H.; Soper, A. K.; Dougan, L. The Hydrogen-Bonding Ability of the Amino Acid Glutamine Revealed by Neutron Diffraction Experiments. *J. Phys. Chem. B* **2012**, *116*, 13308–13319.
- (34) Busch, S.; Bruce, C. D.; Redfield, C.; Lorenz, C. D.; McLain, S. E. Water Mediation Essential to Nucleation of β -Turn Formation in Peptide Folding Motifs. *Angew. Chem. Int. Ed.* **2013**, *49*, 13091–13095.
- (35) Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. The Nature of Hydrogen Bonding in Protic Ionic Liquids. *Angew. Chem. Int. Ed.* **2013**, *52*, 4623–4627.
- (36) McLain, S.; Soper, A.; Daidone, I.; Smith, J.; Watts, A. Charge-Based Interactions between Peptides Observed as the Dominant Force for Association in Aqueous Solution. *Angew. Chem. Int. Ed.* *47*, 9059–9062.
- (37) Scoppola, E.; Sodo, A.; McLain, S.; Ricci, M.; Bruni, F. Water-Peptide Site-Specific Interactions: A Structural Study on the Hydration of Glutathione. *Biophys. J.* **2014**, *106*, 1701–1709.
- (38) Mason, P. E.; Neilson, G. W.; Saboungi, M.; Brady, J. W. The Conformation of a Ribose Derivative in Aqueous Solution: A Neutron-Scattering and Molecular Dynamics Study. *Biopolymers* **2013**, *99*, 739–745.
- (39) Gillams, R. J.; Busto, J. V.; Busch, S.; Goi, F. M.; Lorenz, C. D.; McLain, S. E. Solvation and Hydration of the Ceramide Headgroup in a Non-Polar Solution. *J. Phys. Chem. B* **2015**, *119*, 128–139.

- (40) Towey, J. J.; Soper, A. K.; Dougan, L. What Happens to the Structure of Water in Cryoprotectant Solutions? *Faraday Discuss.* **2013**, *167*, 159–176.
- (41) Soper, A. K. Empirical Potential Monte Carlo Simulation of Fluid Structure. *Chem. Phys.* **1996**, *202*, 295–306.
- (42) Soper, A. K. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Phys. Chem.* **2013**, *2013*, 1–67.
- (43) Hannon, A. C.; Howells, W.; Soper, A. ATLAS : A Suite of Programs for the Analysis of Time-of-flight Neutron Diffraction Data from Liquid and Amorphous Samples. *Inst. Phys. Conf. Ser.* **1990**, 193–211.
- (44) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (45) Kaminski, G. A.; Friesner, R. A. Evaluation and Reparametrization of the OPLS-AA Forde Field for Proteins via Comparison with Accurate Qunatum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (46) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchinson, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualisation and Analysis Platform. *J. Cheminformatics* **2012**, *4*, 17.
- (47) Pardo, L. C.; Rovira-Esteva, M.; Tamarit, J. L.; Veglio, N.; Bermejo, J. F.; Cuello, G. J. In *Metastable Systems under Pressure*; Sylwester, R., Aleksandra, D., Victor, M., Eds.; NATO Science for Peace and Security Series A: Chemistry and Biology; Springer Netherlands, 2010; pp 79–91.
- (48) Busch, S.; Lorenz, C. D.; Taylor, J. W.; Pardo, L. C.; McLain, S. E. On the Structure of Concentrated Proline in Solution. *J. Phys. Chem. B.* **2014**, *118*, 14267–14277.

- (49) Nozaki, Y.; Tanford, C. The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions. *J. Bio. Chem.* **1971**, *246*, 2211–2217.
- (50) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature* **2005**, *437*, 640–647.
- (51) Steiner, T.; Koellner, G. Hydrogen Bonds with π -Acceptors in Proteins: Frequences and Role in Stabilizing Local 3D Structures. *J. Mol. Biol.* **2001**, *305*, 535–557.
- (52) Steiner, T. Hydrogen Bonds from Water Molecules to Aromatic Acceptors in Very High-Resolution Protein Crystal Structures. *Biophys. Chem.* **2002**, *95*, 195–201.
- (53) Malone, J. F.; Murray, C. M.; Charlton, M. H.; Dochertyb, R.; Laveryb, A. J. X–H $\cdots\pi$ (Phenyl) Interactions. *J. Chem. Soc., Faraday Trans.* **1997**, *437*, 640–647.
- (54) Perutz, M. F. The Role of Aromatic Rings as Hydrogen-bond Acceptors in Molecular Recognition. *Phil. Trans. R. Soc.* **1993**, *345*, 105–112.
- (55) Adams, H.; Carver, F. J.; Hunter, C. A.; Osborne, N. J. Amide-Aromatic Hydrogen-Bonds in Host-Guest Recognition. *Chem. Commun.* **1996**, 2529–2530.
- (56) Levitt, M.; Perutz, M. F. Aromatic Rings Act as Hydrogen Bond Acceptors. *J. Mol. Biol.* **1988**, *201*, 751–754.
- (57) Loewenthal, R.; Sancho, J.; Fersht, A. R. Histidine-Aromatic Interactions in Barnase. *J. Mol. Bio.* **1991**, *224*, 759–770.
- (58) Finney, J. L.; Gellatly, B. J.; Golton, I. C.; Goodfellow, J. Solvent Effects and Polar Interactions in the Structural Stability and Dynamics of Globular Proteins. *Biophys. J.* **1980**, *32*, 17–33.
- (59) Pace, C. N.; Trevino, S.; Prabhakaran, E.; Scholtz, J. M. Protein Structure, Stability and Solubility in Water and Other Solvents. *Royal Soc.* **2011**, *359*, 1225–1235.

- (60) Senes, A.; Chadi, D. C.; Law, P. B.; Walters, R. F.; Nanda, V.; DeGrado, W. F. Ez, a Depth-dependent Potential for Assessing the Energies of Insertion of Amino Acid Side-chains into Membranes: Derivation and Applications to Determining the Orientation of Transmembrane and Interfacial Helices. *Journal of Molecular Biology* **2007**, *366*, 436 – 448.

Graphical TOC Entry

