ELSEVIER

Biophysical Chemistry 100 (2003) 221–237

Biophysical Chemistry

www.elsevier.com/locate/bpc

A microscopic view of peptide and protein solvation

David A.C. Beck[a], Darwin O.V. Alonso[b], Valerie Daggett[a,b,*]

[a]*Biomolecular Structure & Design Program, University of Washington, Seattle, WA 98195-7610, USA*
[b]*Department of Medicinal Chemistry, University of Washington, Seattle, WA 98195-7610, USA*

Received 1 March 2002; received in revised form 13 May 2002; accepted 12 June 2002

Abstract

The structure and dynamics of the water hydrating peptides and proteins are examined here at atomic resolution via molecular dynamics simulations. Detailed solvation density and residence time data for all 20 L-amino acids in an end-capped AXA tripeptide motif are presented. In addition, the solvation of the protein chymotrypsin inhibitor 2 is investigated as a point of comparison. Residues on the surface of proteins are not isolated; they interact both locally and non-locally in sequence space, and comparison of the solvation properties of each amino acid in both the peptide and protein allow us to distinguish inherent solvation properties from context-dependent perturbations due to neighboring residues. This work moves beyond traditional radial distribution functions and presents graphical representations of preferential solvation and orientation of water by side chains and the main chain. The combination of 0.3 μs of simulation data improves the statistical sampling over previous studies and reveals the significance of bridging water molecules that stabilize and mediate side chain–side chain, side chain–main chain and main chain–main chain interactions at the solvation interface.

Keywords: Solvation; Peptide hydration; Protein hydration; Molecular dynamics; Hydrophobic effect; Water structure

1. Introduction

The structure and dynamics of pure water and the water–protein interface are critical components in the behavior of biological systems. The work of J.T. Edsall and colleagues has done much to illuminate this arena of science [1]. Advances in neutron diffraction equipment and isotopic substitution methods, as well as new methods for fitting experimental data, such as empirical potential structure refinement [2], now provide error analysis of experimentally derived pure-water radial distribution functions, which are valuable as standards for theoretical endeavors. In addition, there have been many experimental studies of the solvation of simple organic molecules and ions ([3–5] and references therein), which show that solvation of a solute induces fundamental changes in the complexity of the first-shell hydrogen-bonding networks of water with respect to the bulk [6–9]. Such information has been used to build models of solvation [10,11].

Experimental studies of protein hydration in crystal structures [12,13], in combination with simulations based on crystal hydration data [14,15], have demonstrated preferential solvation

*Corresponding author..
E-mail address: daggett@u.washington.edu (V. Daggett).

0301-4622/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0301-4622(02)00283-1

and variations in local diffusion around classes of atomic centers. NMR studies have also provided much information about protein hydration [16,17]. Theoretical studies have illuminated complex solvent-dependent processes, such as helix–coil transitions in peptides and the formation of β-hairpins [18,19], the unfolding [20] and the collapse and partial refolding of proteins in pure water [21–23] and mixed solvents [22], and the nature of inverse temperature transitions in elastin peptides [24], as well as the fundamental hydration properties of proteins (reviewed in [25,26]).

Since its conception and later publication, the flexible three-center (F3C) water model [27] has been shown to be a good classical approximation of water [7,27–29]. The model's ability to reproduce accurately experimental phenomena, both structural and dynamic, make it suitable for use in solvation studies of organic molecules, peptides and proteins, sugars and nucleic acids.

Here we present a minimal set of data to demonstrate the applicability of the F3C water model and a survey of water interaction with peptides and the protein chymotrypsin inhibitor 2 (CI2). Further information regarding the parameterization of the F3C water model and its simulation properties can be found elsewhere [27–31]. Here we supplement those previous studies by focusing on the solvation of 20 end-capped alanine-based tripeptides. These results are then compared and contrasted with the results of similar studies of a stable, well-studied protein system, CI2.

The examination of end-capped tripeptides permits a view of the conduct of primitive model systems in solution. The tripeptides are highly mobile and the central side chain is forced to interact with solvating waters. These tripeptides are then very useful in determining the range of possible solute behaviors for each residue. End-capping to reduce charge effects and the choice of the smallest possible side chain for the neighboring residues (Ala) are important in allowing the side chains to display many possible solvation states. The length of the peptide chain in these simulations severely limits the degree to which main-chain groups can directly interact, further limiting differences in solvation behavior to the side chain effects.

Radial distribution functions (RDF) are commonly used to present graphically the structural characteristics of solutions. Unfortunately, normalized RDFs are not as meaningful for protein atoms as for water, because the protein/water interface is grossly asymmetric. Another approach is to use raw pair distribution data (non-normalized RDF pair counts). These are also radially tabulated, but they are not directly comparable across residues or charge groups. The analysis of peptide and protein systems in solution, therefore, requires a method for the normalization of water density around specific charge groups. To this end, we present a method of normalizing the raw counts of water oxygens in solvation shells to produce a normalized local-solvent function, which we call a solvent density function (SDF).

Time-averaged images of local water oxygen density and orientation persistence are used to further augment information provided by these normalized average densities. These images provide a graphical representation of local water oxygen density around particular groups, as well as measures of water orientation persistence. The peaks and valleys of the solvent density function can be clearly observed in these three-dimensional images.

It is important to understand how the peptide results apply to full protein systems. To this end, the analysis methods employed for the tripeptides have also been applied to CI2. In this way, we can investigate the inherent solvation tendencies of the amino acid side chains, as well as context-dependent perturbations in the protein.

2. Methods

2.1. Molecular dynamics simulations

All simulations were performed using an in-house version of the program ENCAD [32], a previously described force field for proteins [33], and the flexible three-center (F3C) water model [27]. Molecular dynamics simulations were performed in the NVE ensemble. A non-bonded energy cut-off range of 8 Å was used. A force-

shifted truncation scheme was used to truncate smoothly the interaction energy and force values [27,33]. Systems were heated to 298 K and simulations carried out for 15 ns. The first 1 ns (5×10^5 steps) was designated for equilibration.

The pure water simulation consisted of 298 F3C water molecules at the experimental density of 0.997 gm cm^{-3} [34]. A total of 20 tripeptide simulations were conducted, each substituting a different central residue into the acetylated and amidated Ala-based tripeptide motif Ala–X–Ala. These systems were solvated with a F3C water box extending at least 8 Å from any solute atom, and the box volume was adjusted to reproduce the experimental density. The number of water molecules in these simulations ranged from 397 to 488. CI2 was similarly solvated. The detailed CI2 protocols have been documented elsewhere [30]. A total of 0.33 μs of simulation time was calculated.

2.2. Analysis of molecular dynamics simulations

Analysis of the molecular dynamics simulations utilized the final 14 ns at a sampling frequency of 0.2 ps for a total of 7×10^4 frames. For pure water simulations, the self-diffusion coefficient (D) and radial distribution function (RDF) were determined. First, solvent-accessible shell properties (water oxygen density and water oxygen residence time), solvent density function (SDF), and orientation persistence data were calculated for all simulations. Molecular dynamics data for 0.308 μs of simulation time were analyzed in this manner.

The pure-water diffusion coefficient (D) was computed from the mean square displacement of all atoms using the Einstein relation [35]:

$$D=\frac{1}{6}\lim_{t\to\infty}\left\langle\left|r(t')-r(t)\right|^2\right\rangle \quad (1)$$

where $r(t)$ is the atomic position at time t. In this study $\Delta(t,t')$ was 0.2 ps. The RDF values for the pure water simulation were calculated using Eqs. (2) and (3) [36]:

$$g\left(r+\frac{1}{2}\Delta(r,r')\right)=\frac{n(r,r')}{n_{\text{ideal}}(r,r')} \quad (2)$$

$$n_{\text{ideal}}=\frac{4\pi\rho}{3}(r'^3-r^3) \quad (3)$$

where $n(r,r')$ is the raw count in a spherical shell extending from r to r' from an atomic center factored by the number of atoms under consideration; $n_{\text{ideal}}(r,r')$ is the number of atoms expected in an ideal gas at the same density over the range r to r'.

The asymmetry in peptide and protein conformations results in the occlusion of the solvation shell volume of a given charge group by neighboring protein atoms. Therefore, in a solvation shell with radius r_s, the raw water oxygen counts $O(r_s)$ require normalization before comparisons can be made across residues and charge groups. The normalization factor is the solvent-accessible volume of the shell. This volume, $V(r_s)$, for a given charge group, was calculated by erecting a 0.2-Å grid (where each grid point represents 0.008 $Å^3$) over simulation space and counting the grid points within r_s of any of the group's atoms. Grid points within the van der Waals volume of protein atoms were marked as occluded and not included in this volume sum. The accessible volume was averaged over the final 14 ns of each simulation. The normalized value, $\rho_O(r_s)$ in Eq. (4), is the water oxygen (number) density $Å^{-3}$. Density values were multiplied by 100 to improve readability in Tables 2 and 3. For this study, the first solvation shell was considered to have r_s of 3.5 Å (the rationale for using this cut-off is explained below).

$$\rho_O(r_s)=\frac{O(r_s)}{V(r_s)} \quad (4)$$

The solvent-accessible shell methodology described above has been extended to produce a solvent density function (SDF). In this method, only non-occluded grid points in the shell from r_s to r_s' were included in the volume sum, denoted as $V(r_s,r_s')$. Similarly, only water oxygens in this range were included in the raw counts, $O(r_s,r_s')$. The shell number density for this range can be expressed as:

$$\rho_O(r_s,r'_s)=\frac{O(r_s,r'_s)}{V(r_s,r'_s)} \quad (5)$$

For these analyses, r_s was enumerated from 1.5 to 6.0 Å with $\Delta(r_s,r_s')$ of 0.1 Å.

Graphics that depict water oxygen density and water orientation persistence were computed for all 20 residues in the central position of the AXA tripeptide and for several surface residues of CI2. For a given frame in the trajectory, the periodic box was re-centered on the residue(s) of interest using the best fit method of Kearsley [37] for heavy atom positions to realign the protein to an initial reference frame. Water counts and orientations were calculated by erecting a 0.2-Å grid and sorting water oxygen atoms onto the grid. For each water oxygen, a molecular bisector was calculated extending to the midpoint of a line connecting the two water hydrogen atoms. At the end of the analysis run (7×10^4 frames) the molecular bisector vector for each grid point was averaged by the number of oxygen visits. A persistent orientation produces a longer vector, while a highly variable orientation produces a smaller, nearly point-like vector. It is important to note that the molecular-dipole orientation vector cannot be used to distinguish between two configurations with molecular bisectors that are parallel to one another. Images of these data were rendered using a locally modified version of MOLSCRIPT [38]. Water orientation vectors are colored with the hydrogens at the white end, fading to the appropriate color (based on the density scale) for the oxygen.

As a measure of side-chain conformational sampling, the χ_1 and χ_2 order parameters were calculated [39]. The form of $O(\chi_n)$ is in Eq. (6). The χ_n order parameter decreases from 1.0 as side chain mobility increases.

$$O(\chi_n)=1-\frac{1}{2}\langle\Delta\chi_n\rangle^2 \tag{6}$$

Residence times for the first solvent-accessible shell (0.0–3.5 Å) are reported for the charge groups of all 20 amino-acid side chains. The residence time was considered to be the average amount of time a given water oxygen spent in the first shell—this value is averaged over all water molecules. As with r_s in Eq. (4), 3.5 Å was used to delineate the end of the first solvent-accessible shell after examination of experimental RDFs [2,4] and SDFs developed here. It should be noted that there are other methods for the calculation of residence times, e.g. [40]. Other methods can produce higher residence times than the averages provided here. We have presented some comparisons using other methods, and, in general, the qualitative/relative values are similar to those presented in this work, although the absolute values can be sensitive to the method employed.

3. Results

3.1. Liquid water

The self-diffusion coefficient for the F3C water model, 0.23 $Å^2\ ps^{-1}$, is in excellent agreement with the experimental value of 0.23 $Å^2\ ps^{-1}$ [44]. This agreement is important, especially considered in the context of the other, widely used water models [45]. The F3C model was parameterized to reproduce several other bulk properties through a range of temperatures [27,30], but these are not strictly relevant to this study.

Critical values for the water–water RDFs from simulation and experiment are shown in Table 1. The F3C simulation data exhibit a slightly exaggerated first peak (3.18 pairs) as compared with neutron diffraction data of 2.75 ± 0.32 and 2.79 ± 0.30 pairs [2]. Previous gOO experimental data [46] determined this peak to be 3.31 pairs at 2.70 Å—to which the F3C model is in best agreement. X-Ray diffraction studies at 298 K suggest that the location of this peak may be as high as 2.86 Å [47]. Of particular importance is the offset of the second peak, which the F3C model reproduces. Considering the experimental error in data sets A and B, the F3C model overestimates the number of pairs in the first peak by fewer than 0.07 and 0.05 pairs $Å^{-1}$, respectively, for the range 2.67–2.73 Å. This slight overcounting in the first-shell gOO is also reflected in the first shell gOH. As with the gOO, this value (1.60 pairs) is slightly left-skewed. The gHH peaks and valley from simulation are well within experimental error for both sets of experiments, with the first peak (1.28 ± 0.13 for A, 1.37 ± 0.13 for B) almost optimally placed by the F3C simulation in distance from the central atom (2.35 Å). The

average water residence time for the first solvation shell (0.0–3.5 Å) was 1.25 ± 0.02 ps. Another commonly used distance range, 0.0–4.5 Å, which counts many waters in the second solvation shell, gives an average residence time of 1.53 ± 0.09 ps.

3.2. End-capped Ala-based tripeptides

Fig. 1 depicts the time-averaged water oxygen density and water molecular bisector orientations with respect to two of the tripeptide central residues, (a) Lys and (b) Leu. In this figure, higher local water density is indicated by blue points and more persistent water dipole orientation by longer vectors. The Lys NH_3^+ head group exerts a fair amount of ordering (denoted by the length of water dipole orientation vectors) on its first-shell waters. This ordering, in turn, establishes solid, but less well-ordered water density in the second shell (concentration of red points). The molecular plane of first-shell waters tends to be parallel to the hydrophobic shaft of the side chain. Although these waters are unrestricted relative to those hydrogen bonding directly with polar groups, they rarely situate themselves with their hydrogens facing the hydrophobic portion of the Lys side-chain stem—this has been observed for other aliphatic hydrocarbon chains [7,29]. The main-chain amide hydrogen is highly capable of hydrogen bonding to water oxygens, thereby inducing local solvent structure. Main-chain carbonyl oxygens favor hydrogen bonds with first-shell waters that bridge to semi-ordered second-shell waters (residue right in Fig. 1) off the NH_3^+ head group.

The central Leu of the end-capped Ala tripeptide in Fig. 1b shows similar organization of solvent waters around its amide hydrogen and carbonyl oxygen. The hydrophobic terminal-methyl groups of the side chain are solvated by two intersecting shells of disordered water. Images (a) and (b) utilize the same numerical values for color scale assignments of water oxygen density and strength of molecular dipole orientation, and can therefore be compared directly.

First-solvent-accessible-shell water oxygen density and the average water-oxygen residence time for every residue's charge groups are listed in Table 2. Boxplots of data from like charge groups across all residues are presented in Fig. 2. From the data in Fig. 2a, there is a weak relationship between a charge group's polarity and its first-shell water oxygen density. There is a stronger relationship between a group's polarity and its residence time—groups that are charged and/or highly polar have longer residence times than non-polar/neutral groups (Fig. 2b). An example of the effect of charge on residence time is the (Asp and Glu) carboxyl residence time of 5.39 ± 0.01 ps, which is higher than any other polar group and much higher than any non-polar group.

Solvent density function plots calculated from four of the tripeptide central residues are presented

Table 1
Critical values of radial distribution function for all water from 15-ns molecular dynamics simulation

RDF pair	Name	First peak		First valley		Second peak	
		g height	Å	g height	Å	g height	Å
gOO	F3C	3.18	2.70	0.82	3.30	1.04	4.45
	Exp. A	2.75 ± 0.32	2.73	0.78 ± 0.14	3.36	1.16 ± 0.13	4.50
	Exp. B	2.79 ± 0.30	2.85	0.94 ± 0.17	3.63	1.11 ± 0.13	4.26
gOH[a]	F3C	1.60	1.65	0.18	2.35	1.49	3.20
	Exp. A	1.12 ± 0.16	1.77	0.22 ± 0.05	2.40	1.51 ± 0.10	3.33
	Exp. B	1.09 ± 0.14	1.83	0.25 ± 0.05	2.43	1.55 ± 0.10	3.30
gHH[a]	F3C	1.34	2.35	0.69	2.95	1.17	3.90
	Exp. A	1.28 ± 0.13	2.34	0.76 ± 0.08	2.94	1.16 ± 0.07	3.84
	Exp. B	1.37 ± 0.13	2.34	0.82 ± 0.08	2.91	1.82 ± 0.07	3.75

There were 298 molecules of F3C water at experimental density in the periodic box. Statistics were calculated over the last 14 ns with a sampling frequency of 0.2 ps. Experimental data are from neutron diffraction with isotopic substitution [2,38].

[a] Intramolecular pairs are ignored.

in Fig. 3. Two oppositely charged residues, (a) Asp and (b) Lys, are shown as models of polar amino acids. The first solvation shell of the Asp carboxyl group is highly populated over a narrow (~0.5 Å) range. The main-chain polar groups have slightly broader first site peaks at predictable distances. Lys has somewhat similar main-chain characteristics, although the main-chain amide is much more densely solvated than that of Asp.

Also in Fig. 3 are SDF data from an aromatic residue, (c) Tyr, and an aliphatic hydrophobic residue, (d) Ile. Main-chain carbonyl groups have SDF profiles that are fairly consistent across these four residues—the same is true for Cβ and Cα groups, where the first peak has contributions from waters interacting with both the main-chain amide and carbonyl groups. Differences in the profiles for the main-chain amide group are density (peak height), but not distance (peak location) dependent. The trend of most groups by 6 Å to have similar shell water density is an indication that at this distance, most waters appear as bulk to the solute. This is not true for a few select atoms, which are within the structural core of a residue and still have a fair number of peptide atoms at this distance with their own local solvent structure.

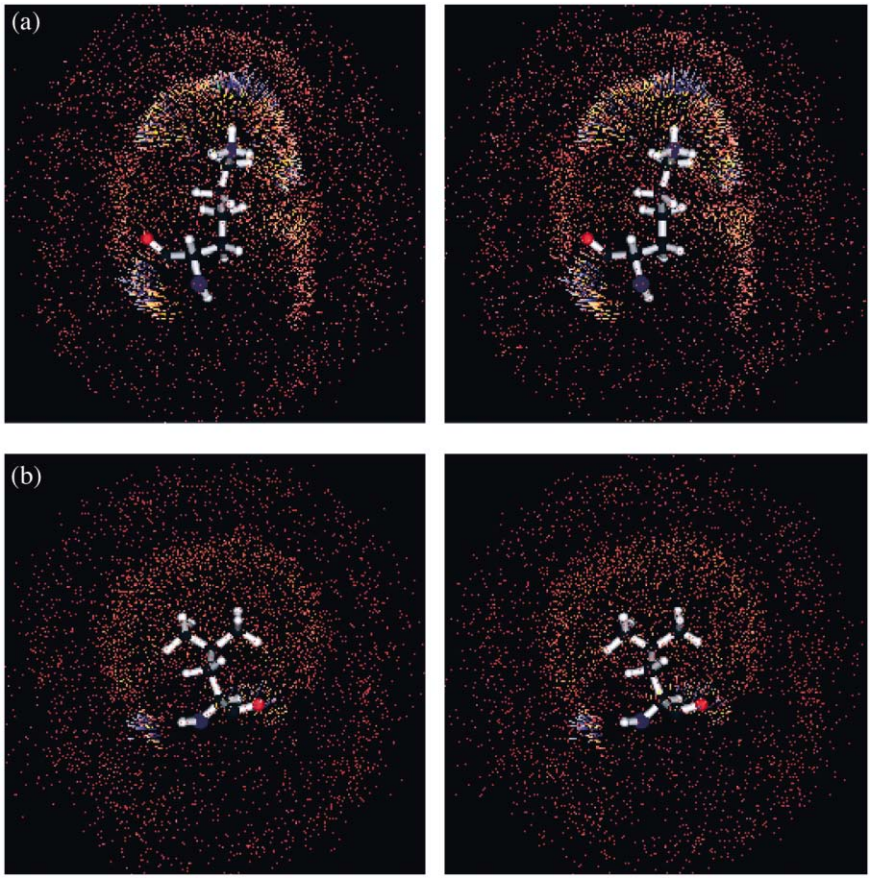

Fig. 1. Stereo images of time-averaged water oxygen density and orientation persistence. Data collected from molecular dynamics simulations of (a) Lys and (b) Leu substituted into the end-capped AXA tripeptide motif. Areas of high water oxygen density are shown with points and vectors. The density color scale is ascending ROYGB (red, orange, yellow, green, blue). Vector length indicates orientation persistence for an area of high density. For reference, the final simulation structure of the central residue is displayed with ball and stick.

3.3. Chymotrypsin inhibitor 2

A native simulation of CI2 was chosen to illustrate protein–water interactions. CI2 is very stable in simulations and has been extensively studied in vitro and in silico [30]. Fig. 4 is the 15-ns snapshot of CI2. It is apparent from the explicit residues in this figure that side chains interact with one another on the protein surface and that not all 'surface' residues are totally exposed to solvent. This partial solvation is the case for both hydrophobic and polar residues, as well as structured and unstructured parts of the protein. To illustrate this behavior, we focus on two examples: Asp 52 and Lys 53, which are the first two residues in a solvent-exposed loop that ends at Asn 56; and Tyr 42, Ile 44 and Val 63, which participate in a solvent-exposed hydrophobic cluster.

Such side chain–side chain interactions can interrupt some protein–water interactions and perturb others. Fig. 5a portrays the 15-ns snapshot focusing on Asp 52 and Lys 53. As with the Lys tripeptide, the first-solvation-shell waters have molecular planes parallel to the side-chain hydrophobic stem. Another view of the Asp 52 and Lys 53 neighbors is shown in Fig. 5b. Note that the amide hydrogen of Asp 52 and the carbonyl of Asn 56 are hydrogen-bonded at the entrance of the loop. The core loop region, formed by Lys 53, Leu 54, Asp 55 and Asn 56, is 'crowned' by the water molecule in the center of the snapshot. The core residues of the loop are forming hydrogen bonds to the crown water with their main-chain amide hydrogens. Leu 54, in particular, has a short hydrogen bond with the crown water. The significance of the crown water is demonstrated by the high average residence time of 12.84 ± 2.84 ps. Residues 52–55 have their main-chain carbonyl vectors oriented away from the loop itself. A second water (center top of the image) coordinates with the Asp 52 side-chain carboxyl group, the crown water and another main-chain amide group to further stabilize this loop.

Hydration-shell water statistics for these residues are provided in Table 3. The carboxyl groups and main-chain carbonyl group of Asp 52 have much higher water oxygen residence times and water density than the tripeptide. The Lys 53 side-

Table 2
Properties of the hydration shell and side chains of all 20 L-amino acids in the tripeptide Ala–X–Ala

Central residue	Group	Shell density[a] ($\times 10^2$ Å^{-3})	Residence time[b] (ps)
Non-polar, aliphatic			
Gly	NH	2.48	1.02
	Cα (CH_3)	2.56	0.73
	CO	2.79	1.13
Ala	NH	2.72	1.40
	Cα (CH)	2.72	0.68
	Cβ (CH_3)	2.22	0.73
	CO	2.84	1.26
Pro	N	0.87	0.24
$O(\chi_1) = 0.99$	Cα (CH)	3.28	0.97
$O(\chi_2) = 0.99$	Cβ (CH_2)	2.37	0.65
	Cγ (CH_2)	2.25	0.61
	Cδ (CH_2)	2.40	0.65
	CO	2.94	1.61
Val	NH	3.44	1.86
$O(\chi_1) = 0.87$	Cα (CH)	2.38	0.71
	Cβ (CH_2)	2.27	0.86
	Cγ^1 (CH_3)	1.93	0.44
	Cγ^2 (CH_3)	1.90	0.45
	CO	3.20	1.64
Leu	NH	3.24	1.84
$O(\chi_1) = 0.63$	Cα (CH)	2.42	0.61
$O(\chi_2) = 0.53$	Cβ (CH_2)	2.94	0.86
	Cγ (CH)	2.27	0.50
	Cδ^1 (CH_3)	2.05	0.60
	Cδ^2 (CH_3)	2.02	0.61
	CO	2.87	1.33
Ile	NH	3.40	2.31
$O(\chi_1) = 0.93$	Cα (CH)	2.76	0.65
$O(\chi_2) = 0.60$	Cβ (CH_2)	3.46	0.90
	Cγ^2 (CH_3)	1.99	0.66
	Cγ^1 (CH_2)	2.14	0.55
	Cδ (CH_3)	2.01	0.59
	CO	3.10	1.38
Aromatic			
Phe	NH	3.29	1.72
$O(\chi_1) = 0.60$	Cα (CH)	2.85	0.70
$O(\chi_2) = 0.19$	Cβ (CH_2)	2.60	0.79
	Cγ (C)	1.72	0.36
	Cδ^1 (CH)	2.26	0.50
	Cε^1 (CH)	2.18	0.49
	Cζ (CH)	2.17	0.51
	Cε^2 (CH)	2.21	0.49
	Cδ^2 (CH)	2.34	0.51
	CO	2.91	1.38
Tyr	NH	2.57	1.35
$O(\chi_1) = 0.54$	Cα (CH)	2.89	0.75
$O(\chi_2) = 0.36$	Cβ (CH_2)	2.52	0.77
	Cγ (C)	1.48	0.34
	Cδ^1 (CH)	2.28	0.54

Table 2 *(Continued)*

Central residue	Group	Shell density[a] ($\times 10^2$ Å^{-3})	Residence time[b] (ps)
	Cε^1 (CH)	2.45	0.56
	Cζ (C)	3.41	0.46
	Oη (OH)	2.18	0.92
	Cε^2 (CH)	2.39	0.59
	Cδ^2 (CH)	2.21	0.54
	CO	3.32	1.38
Trp	NH	2.91	1.68
O(χ$_1$) = 0.57	Cα (CH)	2.61	0.66
O(χ$_2$) = 0.18	Cβ (CH$_2$)	2.57	0.80
	Cγ (C)	1.62	0.36
	Cδ^1 (CH)	2.12	0.55
	Nε^1 (NH)	2.12	0.74
	Cε^2 (C)	2.17	0.43
	Cζ^1 (CH)	2.20	0.51
	Cη (CH)	2.10	0.50
	Cζ^2 (CH)	2.10	0.50
	Cε^3 (CH)	2.19	0.51
	Cδ^2 (C)	2.67	0.43
	CO	2.93	1.31
Polar, uncharged			
Asn	NH	2.84	1.64
O(χ$_1$) = 0.76	Cα (CH)	2.63	0.67
O(χ$_2$) = 0.54	Cβ (CH$_2$)	2.81	0.89
	Cγ (CO)	2.32	1.01
	Nδ (NH$_2$)	1.99	0.75
	CO	3.06	1.25
Gln	NH	3.30	1.45
O(χ$_1$) = 0.60	Cα (CH)	3.02	0.77
O(χ$_2$) = 0.59	Cβ (CH$_2$)	2.73	0.78
	Cγ (CO)	2.39	0.63
	Cδ (CO)	2.20	0.96
	Nε (NH$_2$)	1.94	0.73
	CO	3.06	1.40
Ser	NH	2.53	1.17
	Cα (CH)	3.14	0.85
	Cβ (CH$_2$)	2.59	0.72
	Oγ (OH)	2.42	1.06
	CO	3.34	1.37
Thr	NH	2.48	1.34
	Cα (CH)	3.15	0.86
	Cβ (CH)	3.00	0.70
	Oγ^1 (OH)	2.53	1.12
	Cγ^2 (CH$_3$)	2.25	0.72
	CO	3.52	1.45
Met	NH	3.12	1.54
O(χ$_1$) = 0.61	Cα (CH)	2.68	0.69
	Cβ (CH$_2$)	2.53	0.77
	Cγ (CH$_2$)	2.04	0.57
	Sδ (S)	1.55	0.27
	Cε (CH$_3$)	1.92	0.56
	CO	2.97	1.32
Cys	NH	2.70	1.72
	Cα (CH)	2.34	0.63
	Cβ (CH$_2$)	2.37	0.74
	Sγ (SH)	1.75	0.27
	CO	2.80	1.27
Polar, charged			
Asp	NH	2.65	1.34
O(χ$_1$) = 0.88	Cα (CH)	3.60	0.93
O(χ$_2$) = 0.35	Cβ (CH$_2$)	3.32	0.97
	Cγ (COO)	3.23	5.40
	CO	4.76	1.31
Glu	NH	3.19	1.75
O(χ$_1$) = 0.62	Cα (CH)	2.52	0.64
O(χ$_2$) = 0.88	Cβ (CH$_2$)	3.73	1.00
	Cγ (CH$_2$)	3.12	0.78
	Cδ (COO)	2.99	5.38
	CO	2.75	1.22
Arg	NH	2.91	1.54
O(χ$_1$) = 0.47	Cα (CH)	2.64	0.67
O(χ$_2$) = 0.81	Cβ (CH$_2$)	2.82	0.83
	Cγ (CH$_2$)	2.40	0.67
	Cδ (CH$_2$)	2.49	0.74
	Nε (NH)	2.15	0.93
	Cζ (C)	2.44	0.42
	Nη^1 (NH)	1.94	1.04
	Nη^2 (NH)	2.03	1.11
	CO	2.96	1.33
Lys	NH	2.79	1.69
O(χ$_1$) = 0.62	Cα (CH)	2.48	0.64
O(χ$_2$) = 0.64	Cβ (CH$_2$)	2.70	0.83
	Cγ (CH$_2$)	2.33	0.63
	Cδ (CH$_2$)	2.56	0.68
	Cε (CH$_2$)	2.45	0.61
	Nζ (NH$_3$)	1.58	1.63
	CO	2.85	1.32
His	NH	2.84	1.69
O(χ$_1$) = 0.52	Cα (CH)	2.68	0.70
	Cβ (CH$_2$)	2.70	0.82
	Cγ (C)	0.98	0.30
	Nδ^1 (NH)	2.17	0.70
	Nδ^2 (N,2CH)	2.07	0.75
	CO	2.88	1.25

Statistics are calculated from the last 14 ns with a sampling frequency of 0.2 ps.

[a] Water oxygen density in the first solvent-accessible shell (0.0–3.5 Å). The range of standard deviations for density values in this table is 0.12–0.23$\times 10^2$ Å^{-3}.

[b] Water oxygen residence time in the first solvent-accessible shell (0.0–3.5 Å). The range of standard deviations for residence time values in this table is 0.01–0.11 ps.

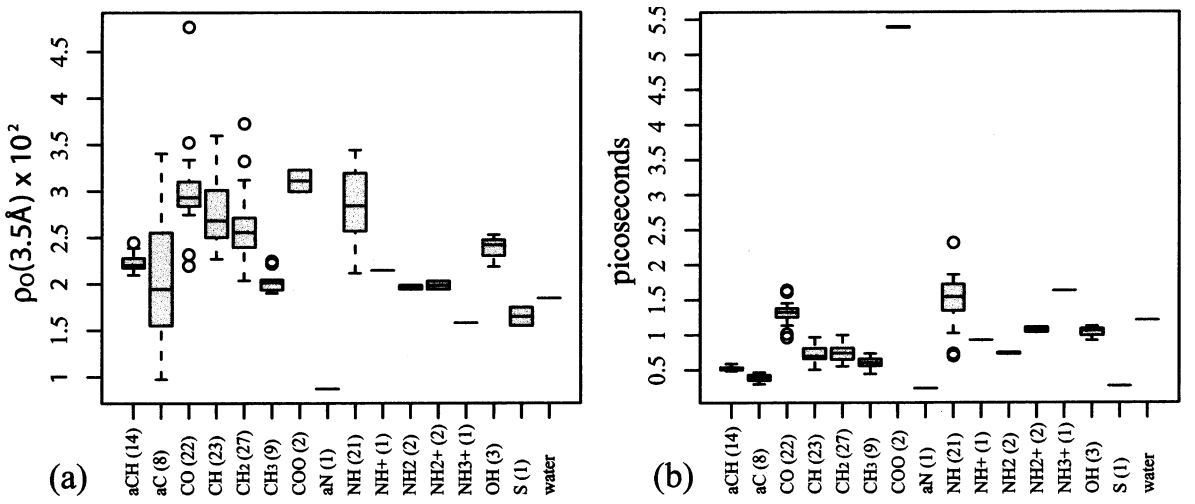

Fig. 2. Properties of water oxygens in the first solvent-accessible shell of tripeptide charge groups: (a) water oxygen density; and (b) water-oxygen residence time in the solvent-accessible volume extending 3.5 Å from a given group. In parentheses is the number of samples that contributed to the statistics from the 20 end-capped AXA tripeptide simulations. Aromatic charge groups are denoted with an 'a' before their name. Water–water simulation data are shown for reference.

chain amide group and main-chain carbonyl group have higher residences time when compared with the tripeptide data.

The time-averaged water density and molecular orientation image for these residues are shown in Fig. 6. Areas of low solvent water density illustrate the volume excluded to water by side chain interactions and neighboring residues. Also present are two hyper-ordered water bridges between the main-chain polar groups shown by molecular dipole orientation vectors of the blue (high water density) fading to white (water hydrogen).

The SDF data for Asp 52 and Lys 53 are plotted in Fig. 3e,f, respectively. The Asp amide does not have a first solvation site peak, as is present in the tripeptide plot (a). Lys 53 has a wider first site peak than that of the tripeptide Lys (b); this group also has a pronounced second peak at 3.6 Å. In addition, most of the side-chain aliphatic hydrocarbon groups have right-shifted peaks with respect to the tripeptide.

A hydrophobic cluster is formed by Tyr 42, Ile 44 and Val 63 on the surface of CI2. Over the last 14 ns of the simulation, Tyr 42 and Ile 44 have a mean side-chain distance of 4.65 ± 0.69 Å. Similarly, Tyr 42 and Val 63 have a mean side-chain distance of 6.27 ± 0.91 Å, and Ile 44 and Val 63, 3.86 ± 0.76 Å. The accessible first-solvation-shell water oxygen properties for these residues are reported in Table 3. The Tyr statistics for both properties appear very similar to those from the tripeptides, and indeed this residue is exposed to solvent in much the same way as in the peptide simulations. The exception is the higher residence time of 2.12 ps for the carbonyl group as compared with 1.38 ps in the peptides. Some water is excluded from Ile 44 by the bulky Tyr 42, which may account for its much lower Cα solvent density and lower main-chain residence times. Val 63 is oriented such that the N-terminal portion of its backbone and most of its side chain are not exposed to solvent. The $C\gamma^1$ lies closest to solvation waters, which is reflected in a higher solvent density than most of its other groups, but lower still than the tripeptide data. The Val 63 carbonyl group has an extremely high density.

Finally, in Fig. 3g,h, SDF plots for Tyr 42 and Ile 44, respectively, are shown. The Tyr 42 hydroxyl group has an exaggerated first peak over the tripeptide (c). Also different is the main-chain amide group, which has a wider, reduced-density first site peak. Ile 44 is the most strikingly different

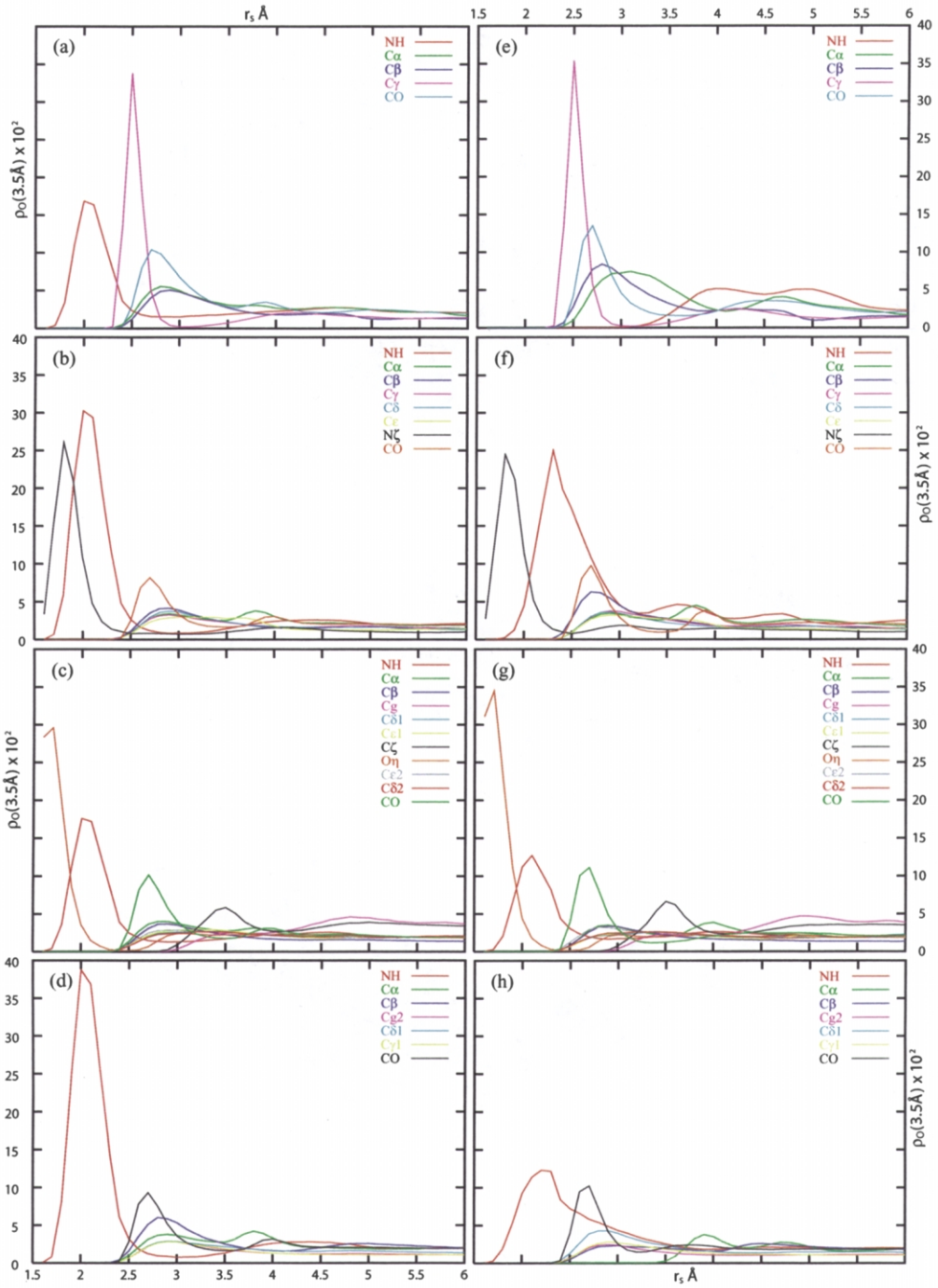

Fig. 3. Solvent density function plots of four residues in varying solvent environments. Data presented are from simulations of (a–d) end-capped AXA tripeptides and (e–h) CI2. Residues are: (a,e) Asp; (b,f) Lys; (c,g) Tyr; and (d,h) Ile. Asp 52 (e) and Lys 53 (f) are directly interacting at side chain termini and the backbone is initiating a crowned loop. Tyr 42 (g) and Ile 44 (h) are involved in a solvent-exposed hydrophobic patch.

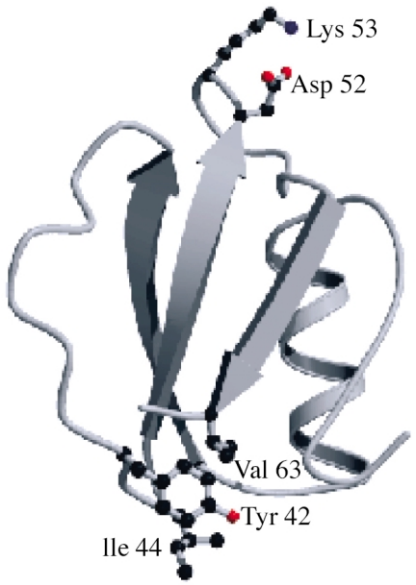

Fig. 4. 15-ns snapshot of CI2 from molecular dynamics simulation at 298 K. Two sets of locally interacting surface residues have been explicitly displayed: the local polar interaction of Asp 52 and Lys 53; and the hydrophobic cluster formed by Tyr 42, Ile 44 and Val 63.

when compared to the tripeptide SDF (d)—the amide peak is dramatically attenuated and widened, while the Cα first site peak is completely absent.

4. Discussion

Both the macroscopic and microscopic structural and dynamic properties of liquid water at 298 K are well reproduced by the F3C water model [27]. Dynamic properties, such as the self-diffusion coefficient, are important indicators of the accuracy of solvent model behavior, as well as the simulation protocols. The experimental value for self-diffusion is easily and accurately obtained. Structural quantities, such as the radial distribution function, solvent density function and solvation-shell water oxygen statistics, give well-defined descriptions of average water structure.

The choice of 3.5 Å as the distance cut-off for the first hydration shell is appropriate, as it captures only those pairs in the first peak, i.e. it most precisely maps the waters of hydration to the charge groups under investigation. Other studies have used 4.5 Å as the end of the first shell—this cut-off counts many pairs that are truly in the second shell. Further evidence for the use of a

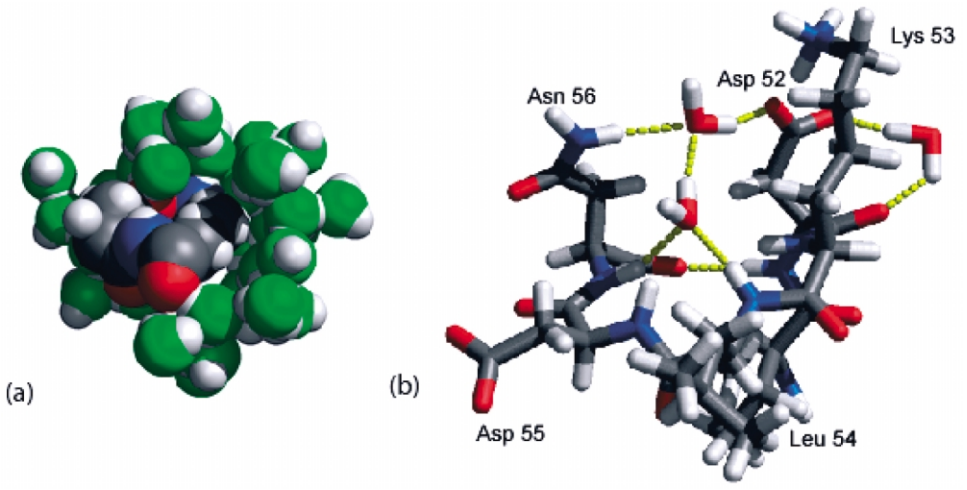

Fig. 5. Snapshots from CI2 simulation focusing on Asp 52 and Lys 53: (a) the side chains (Asp 52 left, Lys 53 right) with first shell waters (H_2O oxygens in green); and (b) the five-residue loop, starting at Asp 52 and its crowning water and associated hydrogen bonding network (*dashed lines*).

Table 3
Properties of the hydration shell and side chains of CI2

Residue	Group	Shell density[a] ($\times 10^2$ Å^{-3})	Residence time[b] (ps)
Polar, charged			
ASP 52	NH	0.56	0.22
$O(\chi_1)=0.98$	Cα (CH)	5.64	0.71
$O(\chi_2)=0.93$	Cβ (CH_2)	4.77	1.20
	Cγ (COO)	3.45	6.49
	CO	3.97	2.16
Lys 53	NH	6.90	1.54
$O(\chi_1)=0.80$	Cα (CH)	2.47	0.75
$O(\chi_2)=0.85$	Cβ (CH_2)	3.46	0.88
	Cγ (CH_2)	2.71	0.69
	Cδ (CH_2)	2.46	0.74
	Cε (CH_2)	2.53	0.71
	Nζ (NH_3)	2.09	1.76
	CO	2.56	2.07
Aromatic			
Tyr 42	NH	2.63	1.16
$O(\chi_1)=0.98$	Cα (CH)	2.33	0.68
$O(\chi_2)=0.54$	Cβ (CH_2)	2.35	0.78
	Cγ (C)	1.43	0.37
	$C\delta^1$ (CH)	2.05	0.54
	$C\varepsilon^1$ (CH)	2.14	0.55
	Cζ (C)	3.33	0.43
	Oη (OH)	2.17	1.15
	$C\varepsilon^2$ (CH)	2.42	0.66
	$C\delta^2$ (CH)	2.19	0.54
	CO	2.99	2.12
Non-polar, aliphatic			
Ile 44	NH	4.68	1.59
$O(\chi_1)=0.36$	Cα (CH)	0.10	0.38
$O(\chi_2)=0.87$	Cβ (CH_2)	1.78	0.61
	$C\gamma^2$ (CH_3)	1.66	0.74
	$C\gamma^1$ (CH_2)	2.76	0.77
	Cδ (CH_3)	1.97	0.63
	CO	3.03	1.37
Val 63	NH	1.34	0.35
$O(\chi_1)=0.98$	Cα (CH)	0.27	0.34
	Cβ (CH_2)	0.06	0.47
	$C\gamma^1$ (CH_3)	0.06	0.39
	$C\gamma^2$ (CH_3)	1.46	0.62
	CO	7.39	2.20

Statistics are calculated from the last 14 ns with a sampling frequency of 0.2 ps.

[a] Water oxygen density in the first solvent accessible shell (0.0–3.5 Å). The range of standard deviations for densities in this table is 0.20–0.34$\times 10^2$ Å^{-3}.

[b] Water oxygen residence time in the first solvent accessible shell (0.0–3.5 Å). The range of standard deviations for residence time values in this table is 0.02–0.25 ps.

short cut-off distance for the hydration shell can be found in a recent study describing very good agreement between simulation and experiment [41], in which the authors concluded that the first hydration layer is approximately 3 Å thick. In addition, Perutz et al. [42] also asserted that the hydration layer is 3 Å, based on Perutz's findings in 1946 that these waters are unavailable as a solvent for diffusible electrolytes. Finally, recent studies by Zewail and co-workers aimed at mapping protein hydration on the femtosecond timescale support the residence times described in this work [43]. They found that the primary component of dynamic solvation times is 1.1 ps—a value in good agreement with the average residence times calculated using the methods employed here.

The hydration shell properties are useful as references for bulk solvent behavior. Water that is perturbed by protein atoms will have dramatically

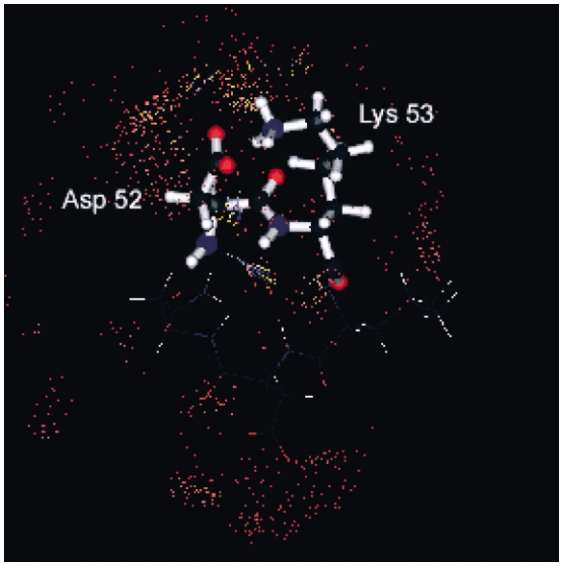

Fig. 6. Time-averaged water oxygen density and orientation persistence data for the CI2 simulation. Only density and atoms within a 10-Å radius of Asp 52 and Lys 53 are displayed. Areas of high water oxygen density are shown with points and vectors. The density color scale is ascending ROYGB. Vector length indicates orientation persistence for an area of high density. For reference, the final simulation structure for the residues used in frame alignment is displayed with ball and stick. Neighboring residues are displayed with lines between atomic centers.

different structural and dynamic characteristics. From extension of the SDF for terminal methyl groups of the Ile in a tripeptide to 8 Å, it is apparent that waters beyond 7 Å have bulk properties—this is in agreement with other theoretical studies [43].

When considering the tripeptide simulations, it is important to note how free the central residue is to sample local solvation configurations. The residue is relatively unconstrained compared with a surface residue in a protein system such as CI2. In this tripeptide motif, there are very few steric restrictions from the peptide chain and no bulky neighboring residues. As a result, these systems provide excellent descriptions of possible solvation sites and the expected behavior in fully solvent-exposed situations.

Despite the lack of a strict correlation between the polarity of a charge group and its first-shell solvent density in the boxplot of Fig. 2a, several interesting relationships are observed. Positive charge groups exhibit lower first-shell solvent density and higher residence times than most groups. From Figs. 1 and 6, it can be observed that the waters in this shell are well ordered. These groups will hydrogen bond to a water oxygen—this orientation (oxygen in and hydrogens out) permits fewer degrees of freedom for the hydrogen bond, and by association reduces the freedom of other potential hydrogen-bonding partners. As a result, fewer waters can occupy this shell. The residence times for positively charged polar groups increase with polarity and available hydrogen bonding sites (i.e. number of charge group hydrogens). The water oxygen is electronegative, and therefore has a deep interaction-potential well with these groups. The depth of this well contributes to the residence times.

Negative charge groups have higher first-shell density, in addition to longer residence times, when compared with most other groups. The residence times of Asp and Glu carboxyl groups are more than two-fold as long as any other group. These negatively charged groups effectively orient the molecular bisector of solvent-accessible shell waters (Figs. 1 and 6).

Every polar side-chain group organizes one or more water solvation sites. Examples are shown clearly by the ‘tufts’ of water in Figs. 1 and 6. In all cases, the main-chain amide and carbonyl groups were able to hyper-orient waters. It is not uncommon to find tufts from two oppositely charged groups that are clearly interacting with one another indirectly through a water, or directly with water solvating the interaction.

Aliphatic hydrocarbon groups exhibit decreasing first-shell solvent density with increasing number of hydrogens, in agreement with previous theoretical studies of the solvation of hydrocarbon chains [29]. The outliers and distribution skews in the statistics for these groups are a result of highly polar neighboring groups, such as COO and NH_3^+. Not one of the solvent density/orientation persistence images shows these hydrocarbon groups ordering first-shell waters. The residence time for all groups of this type is sub-picosecond.

Aromatic groups have low residence times, but a very concise density range. This is a result of nearly identical environments for most of these groups—they are found clustered together on the rings, and thus experience very similar solvent conditions. The aromatic single-atom groups (carbons and the N in Pro) and the neutral sulfur have very low residence times, reflecting their lack of charge. In these cases, their local environments are much more a function of neighboring charge groups. Like the aliphatic hydrocarbon groups, there are no examples of directly ordered waters around these groups in the solvent density/orientation persistence images.

There is a relationship between a group’s ability to orient local solvent molecules and the residence time of local waters. What is not clear is the nature of this relationship. Is it the case that, by being able to hold an individual water for a long period of time, a charge group can orient the water? Or, is it the opposite: by orienting a water to optimize its interaction, a group can hold it for a long period of time? The latter is more likely—a water entering into a group’s solvation shell in a favorable orientation will have a longer residence time than one entering in an unfavorable orientation. In conjunction, highly polar groups have the ability to recruit and pre-orient waters at longer distances, thereby increasing the local water residence time.

An interesting feature of the SDF plots in Fig. 3, which is reflected in the statistics in Tables 2 and 3, is the prominent increase in water oxygen density around the main-chain groups, the amide group in particular, of residues with hydrophobic side chains. This main-chain effect occurs even in the cases of Arg and Lys, where the side-chain polar groups are relatively distant from the main chain. This finding suggests that these main-chain groups enjoy more water interactions when there are fewer favorable interaction partners for water, i.e. more hydrocarbon and neutral groups in the local environment (<6.0 Å). Furthermore, hydrophobic groups sample fewer waters with shorter residence times.

In protein systems, using CI2 as an example, residue movement is rather restricted by contacts with neighboring residues and main chain orientation. As a result, surface residues sample solvent less freely than when contained in a peptide. The extent of side chain mobility can be quantified through χ_1 and χ_2 order parameters (Tables 2 and 3). It is notable that the selected surface residues of CI2 exhibit order parameters closer to 1.0 than those from the tripeptide simulations. This lack of freedom is primarily responsible for the deviations from the tripeptide data. In general, however, the characteristic tufts of persistently oriented waters around solvent-exposed polar groups in the tripeptide simulations are also observed for the surface residues of CI2.

The SDF profiles for Asp 52 and Lys 53 have significant deviations from their tripeptide equivalents. Asp 52 does not have an amide first-solvation-site peak, due to its stable and tight hydrogen bond with the carbonyl of the loop exit residue, Asn 56, and the exclusion of water by the hydrophobic side chains of Val 51 and Ile 57. These hydrophobic caps serve to protect the interaction from interference and competition from water, thereby increasing the integrity of the loop. All of the other Asp 52 groups have higher, and in some cases wider, peaks. The differences in these peaks is an indication of the desire of water to solicit alternative interactions with neighboring groups when polar groups are unavailable for partnering; this is an interesting corollary to the high solvent density of polar groups when in hydrophobic contexts, as previously mentioned.

The Lys 53 amide peak is much wider than the tripeptide peak for this amino acid. This perturbation may be the result of competition between the NH groups of four residues for the crown water. The Nζ has only a slightly lowered peak, suggesting that its interaction with the Asp 52 carbonyl group only partially excludes water; indeed, from Fig. 6 it is possible to observe that this interaction is mediated by a water with a high average residence time of 22.63 ± 4.2 ps, and may necessitate its presence for formation. The pronounced second NH peak at 3.6 Å arises from the crown-water hydrogen bonding partner(s) in the first shell of the side chain, possibly even the Lys 53 Nζ–Asp 52 Cγ mediating water.

In both the tripeptides and CI2, main-chain and side-chain polar groups can indirectly interact through first- and second-solvation-shell bridging waters. This effect is reflected in the short orientation vectors in Fig. 1 and the crown water of the CI2 Asp 52–Lys 53 loop region in Fig. 5. The crown water bridges a set of main chain–main chain interactions, in addition to hydrogen bonding with another water in the second shell of the main chain. This second water is in the first shell of the Asp 52 carboxyl group. A similar configuration is sampled by the Lys head-group first-solvation-shell water and the main-chain crown water. In this manner, main-chain and side-chain solvation shells are inextricably linked by water–water interactions between them.

This crown water bridging may seem startling, but a cursory search of the Protein Databank (PDB) turned up over 1000 structures with a water oxygen within 3.5 Å of four consecutive main-chain amide nitrogens. Of these structures, 41 have resolution at or below 1.3 Å. Inspection of randomly chosen structures indicates that this solute/solvent configuration occurs even in static, low-solvent-content crystal structures. Most of these randomly chosen structures exhibit this configuration in loop regions (often with higher B-factors) and in two distinct types: one crown water with the main-chain amides of the loop residues oriented towards it; and two crown waters on opposite sides of the main-chain loop. In this latter

motif, two amides interact with one of the crown waters, while the other two interact with the second crown water. Indirect main chain–side chain interactions of this type could have important implications for flexible portions of proteins that become immobilized in response to exclusion from solvent or the introduction of neighboring polar groups.

For the hydrophobic cluster on the surface of CI2 (Tyr 42, Ile 44 and Val 63), much is revealed in the density and residence time values in Table 3, as well as the SDF profiles in Fig. 3. For Tyr 42, the widened SDF peaks and elevated density demonstrate preferential solvation of polar groups (in this case, the hydroxyl), allowing neighboring hydrocarbon groups to participate in hydrophobic interactions. Tyr, with its mix of aliphatic and aromatic hydrocarbons and well-dispersed polar groups, is well suited to cap the cluster and act as a solvent interface. There are also solvent interfaces for this cluster at the Ile 44 amide group and the Val 63 main-chain polar groups, which accept a much elevated solvent density in exchange for burial of their hydrophobic groups. For Val 63, this is extreme, with the carbonyl exposed to two orders of magnitude more solvent $Å^{-3}$ than some of the hydrophobic groups of the side chain.

These data, in conjunction with the elevated density of polar groups in hydrophobic contexts from the tripeptide simulations, demonstrate that polar groups can aid in the formation and stabilization of hydrophobic clusters. By accepting more interactions with local waters, polar groups enrich their local water density from the hydration shell of neighboring hydrocarbon groups, thereby allowing them to pack in the relative absence of water. Furthermore, when an environment is flooded with hydrocarbons, such as the cluster of Tyr 42, Ile 44 and Val 63, these groups can exhibit higher local solvent density—similar evidence has been experimentally observed in binary methanol/water mixtures [48].

5. Conclusions

There is a clear correlation between charge group polarity, solvent-accessible shell-water residence times and local water orientation. Groups that have deep interaction-potential wells have higher residence times and more ordered first-shell waters. A more mixed relationship exists for density and polarity: negatively charged groups have higher first-shell density than positively charged groups. Non-polar groups fall into the middle of the range of hydration density values, with totally neutral groups at the low end.

Polar groups can be preferentially solvated over their less polar neighbors, especially when in the context of hydrophobic groups. Hydrophobic groups are unable to orient local waters and are often observed with their solvation water planes parallel to the hydrocarbon chain. The low water density and residence time around these groups leaves them free to sample more solvent configurations and form interactions with other neighboring groups. The combination of freedom and polar-group overloading permits (partial) burial of hydrophobic groups to form patches and clusters on a protein's surface.

Finally, the solvation shells of the main chain and side chain groups, for all residues, are entangled by the water–water interactions between them. Water directly mediates a variety of interactions, including the crown water motif and main chain–side chain bridges. Orientation of water around salt bridges is stabilizing. In some circumstances, water must be excluded in order to preserve the integrity of polar interactions and reduce competition for interaction sites.

Acknowledgments

This work was supported by the National Institutes of Health (GM 50789 to V.D.) and the Office of Naval Research (N00014-95-1-0484 to V.D.). Simulations were computed on hardware donated by Intel. UCSF MIDASPLUS [49] was used to prepare Fig. 5, and Figs. 1, 4 and 6 were rendered using a locally modified version of MOLSCRIPT [38].

References

[1] J.T. Edsall, H.A. McKenzie, Water and proteins. I. The significance and structure of water; its interaction with electrolytes and non-electrolytes, Adv. Biophys. 10 (1978) 137–207.

[2] A.K. Soper, The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 440 MPa, Chem. Phys. 258 (2000) 121–137.
[3] J.L. Finney, A.K. Soper, Solvent structure and perturbations in solutions of chemical and biological importance, Chem. Soc. Rev. 23 (1994) 1–10.
[4] A.K. Soper, A. Luzar, Orientation of water molecules around small polar and non-polar groups in solution: a neutron diffraction and computer simulation study, J. Phys. Chem. 100 (1996) 1357–1367.
[5] A.K. Soper, Probing the structure of water around biological molecules: concepts, constructs and consequences, Physica B 276-278 (2000) 12–16.
[6] A. Pertsemlidis, A.M. Saxena, A.K. Soper, T. Head-Gordon, R.M. Glaesar, Direct evidence for modified solvent structure within the hydration shell of a hydrophobic amino acid, Proc. Natl. Acad. Sci. USA 93 (1996) 10769–10774.
[7] K.E. Laidig, V. Daggett, Testing the modified hydration-shell hydrogen-bond model of hydrophobic effects using molecular dynamics simulation, J. Phys. Chem. 100 (1996) 5616–5619.
[8] T. Head-Gordon, J.M. Sorenson, A. Pertsemlidis, R.M. Glaeser, Differences in hydration structure near hydrophobic and hydrophilic amino acids, Biophys. J. 73 (1997) 2106–2115.
[9] R.P. Kesselring, The hydrophobic effect. 3. A key ingredient in predicting *n*-octanol–water partition coefficients, J. Pharm. Sci. 87 (1998) 1015–1024.
[10] A. Bakk, J.S. Hoye, A. Hansen, Apolar and polar solvation thermodynamics related to the protein unfolding process, Biophys. J. 82 (2002) 713–719.
[11] T. Lazaridis, M. Karplus, Effective energy function for proteins in solution, Proteins 35 (1999) 133–152.
[12] F.T. Burling, W.I. Weis, K.M. Flaherty, A.T. Brunger, Direct observation of protein solvation and discrete disorder with experimental crystallographic phases, Science 271 (1996) 72–77.
[13] M. Feig, B.M. Pettitt, Crystallographic water sites from a theoretical prospective, Structure 6 (1998) 1351–1354.
[14] V.A. Makarov, M. Feigh, K. Andrews, B.M. Pettitt, Diffusion of solvent around biomolecular solutes: a molecular dynamics simulation study, Biophys. J. 75 (1998) 150–158.
[15] V.A. Makarov, B.K. Andrews, P.E. Smith, B.M. Pettitt, Residence times of water molecules in the hydration sites of myoglobin, Biophys. J 79 (2000) 2966–2974.
[16] G. Otting, NMR studies of water bound to biological molecules, Prog. NMR Spectrosc. 31 (1997) 259–285.
[17] V.P. Denisov, B.H. Jonsson, B. Halle, Hydration of denatured and molten globule proteins, Nat. Struct. Biol. 6 (1999) 253–260.
[18] A.R. Dinner, T. Lazaridis, M. Karplus, Understanding β-hairpin formation, Proc. Natl. Acad. Sci. USA 96 (1999) 9068–9073.
[19] R. Zhou, B. Berne, R. Germain, The free energy landscape for β-hairpin folding in explicit water, Proc. Natl. Acad. Sci. USA 98 (2001) 14931–14936.
[20] V. Daggett, A.R. Fersht, Transition States in Protein Folding. In: Mechanisms of Protein Folding, 2nd edition, Frontiers in Molecular Biology Series, R.H. Pain (Ed.), Oxford University Press, Oxford, UK, Chapter 7, 2000, pp. 175–211.
[21] D.O.V. Alonso, V. Daggett, Molecular dynamics simulations of the hydrophobic collapse of ubiquitin, Protein Sci. 7 (1998) 860–874.
[22] D.O.V. Alonso, V. Daggett, Molecular dynamics simulations of protein unfolding and limited refolding—characterization of partially unfolded states of ubiquitin in 60% methanol and in water, J. Mol. Biol. 247 (1995) 501–520.
[23] Y. Duan, L. Wang, P.A. Kollman, The early stage of folding of villain headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation, Proc. Natl. Acad. Sci. USA 95 (1998) 9897–9902.
[24] B. Li, D.O.V. Alonso, B.J. Bennion, V. Daggett, Hydrophobic hydration is an important source of elasticity in elastin-based biopolymers, J. Am. Chem. Soc. 123 (2001) 11991–11998.
[25] A. Purliss, S. Skoulakis, J.M. Goodfellow, The protein–solvent interface: a big splash, Philos. Trans. R. Soc. Lond. A 359 (2001) 1515–1527.
[26] V. Makarov, B.M. Pettitt, M. Feig, Solvation and hydration of proteins and nucleic acids: a theoretical view of simulation and experiment, Acc. Chem. Res. 35 (2002) 376–384.
[27] M. Levitt, M. Hirshberg, R. Sharon, K.E. Laiding, V. Daggett, Calibration and testing of a water model for simulation of the molecular dynamics of proteins and nucleic acids in solution, J. Phys. Chem. B 101 (1997) 5051–5061.
[28] M. Levitt, R. Sharon, Accurate simulation of protein dynamics in solution, Proc. Natl. Acad. Sci. USA 85 (1988) 7557–7561.
[29] K.E. Laidig, J.L. Gainer, V. Daggett, Altering diffusivity in biological solutions through modification of solution structure and dynamics, J. Am. Chem. Soc. 120 (1998) 9394–9395.
[30] R. Day, B.J. Bennion, S. Ham, V. Daggett, Increasing temperature accelerates protein unfolding without changing the pathway of unfolding, J. Mol. Biol. 322 (2002) 189–203.
[31] K.E. Laidig, B.J. Bennion, D.A.C. Beck, D.O.V. Alonso, V. Daggett, Macromolecular simulations in mixed solvents: development of solvent systems, properties of the solvent systems and the modulation of the hydrophobic effect, in preparation.
[32] M. Levitt, ENCAD—Energy Calculations and Dynamics, Yeda, Rehovat, Israel, 1990.
[33] M. Levitt, M. Hirshberg, R. Sharon, V. Daggett, Potential energy function and parameters for simulations of

the molecular dynamics of proteins and nucleic acids in solution, Comp. Phys. Commun. 91 (1995) 215–231.
[34] G.S. Kell, Precise representation of volume properties of water at 1 atmosphere, J. Chem. Eng. Data 12 (1967) 66.
[35] J.M. Haile, Molecular Dynamics Simulation, Wiley, New York, 1992.
[36] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Oxford, New York, 1987.
[37] S.K. Kearsley, On the orthogonal transformation used for structural comparisons, Acta Cryst. A 5 (1989) 208–210.
[38] P.J. Kaulis, MOLSCRIPT—a program to produce both detailed and schematic plots of protein structures, J. Appl. Cryst. 24 (1991) 946–950.
[39] S.G. Hyberts, M.S. Goldberg, T.F. Havel, G. Wagner, The solution structure of Eglin-C based on measurements of many NOEs and coupling constants and its comparison with X-ray structures, Protein Sci. 1 (1992) 736–751.
[40] R.W. Impey, P.A. Madden, I.R. McDonald, Hydration and mobility of ions in solution, J. Phys. Chem. US 87 (1983) 5071–5083.
[41] F. Merzel, J.C. Smith, Is the first hydration shell of lysozyme of higher density than bulk water?, Proc. Natl. Acad. Sci. USA 99 (2002) 5378–5383.
[42] M.F. Perutz, J.T. Finch, J. Berriman, A. Lesk, Amyloid fibers are water-filled nanotubes, Proc. Natl. Acad. Sci. USA 99 (2002) 5591–5595.
[43] S.K. Pal, J. Peon, A. Zewail, Biological water at the protein surface: dynamics solvation probed directly with femtosecond resolution, Proc. Natl. Acad. Sci. USA 99 (2002) 1763–1768.
[44] K. Krynicki, C.D. Green, D.W. Sawyer, Pressure and temperature dependence of self-diffusion in water, Discuss. Faraday Soc. 66 (1978) 199–208.
[45] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A 105 (2001) 9954–9960.
[46] R.H. Tromp, P. Postorino, G.W. Neilson, M.A. Ricci, A.K. Soper, Neutron diffraction studies of H_2O/D_2O at supercritical temperatures—a direct determination of Ghh(r), Goh(r), and Goo(r), J. Chem. Phys. 101 (1994) 6210–6215.
[47] A.H. Narten, H.A. Levy, Observed diffraction pattern and proposed models of liquid water, Science 165 (1969) 447–454.
[48] S. Dixit, J. Crain, W.C.K. Poon, J.L. Finney, A.K. Soper, Molecular segregation observed in a concentrated alcohol–water solution, Nature 416 (2002) 829–832.
[49] T.E. Ferrin, C.C. Huang, L.E. Jarvis, R. Langridge, The MIDAS display system, J. Mol. Graphics 6 (1988) 13–27.