ELSEVIER

Biophysical Chemistry 105 (2003) 195–209

Biophysical Chemistry

www.elsevier.com/locate/bpc

Analysis of thermal hysteresis protein hydration using the random network model

Kelly Ryan Gallagher, Kim A. Sharp*

E.R. Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, 3700 Hamilton Walk, Philadelphia, PA 19104-6059, USA

Received 24 June 2002; received in revised form 22 August 2002; accepted 22 August 2002

Abstract

The hydration of polar and apolar groups can be explained quantitatively, via the random network model of water, in terms of differential distortions in first hydration shell water–water hydrogen bonding angle. This method of analyzing solute induced structural distortions of water is applied to study the ice-binding type III thermal hysteresis protein. The analysis reveals subtle but significant differences in solvent structuring of the ice-binding surface, compared to non-ice binding protein surface. The major differences in hydration in the ice-binding region are (i) polar groups have a very apolar-like hydration. (ii) there is more uniform hydration structure. Overall, this surface strongly enhances the tetrahedral, or ice-like, hydration within the primary hydration shell. It is concluded that these two specific features of the hydration structure are important for this surface to recognize, and preferentially interact with nascent ice crystals forming in liquid water.

Keywords: Hydration; Thermal hysteresis protein; Liquid water; Random network mode

1. Introduction

It is well established that the unique properties of water dictate much of a protein's structure and function. One central theme that has emerged is that hydrophobic, or solvophobic, interactions drive or influence many molecular 'transactions' [1–6]. In the past, descriptions of this effect have focused on the entropic penalty for ordering water around an apolar solute [7]. Yet the hydration of polar solutes is also usually accompanied by a decrease in entropy. In contrast, the hydration heat capacities of these two types of hydration have different signs [8–12] and thus are a more revealing indicator of the difference between apolar and polar hydration. We have previously directed our efforts towards developing a deeper understanding of heat capacity changes using as a starting point the random network model (rnm) of Henn and Kauzmann [13]. A combination of explicit water simulations and this rnm yields quantitative agreement between measured and calculated hydration heat capacities [2,14,15], as well as a revealing description of the change in water structure induced by apolar and polar solutes [16]. The main

*Corresponding author. Tel.: +1-215-573-3506; fax: +1-215-898-4217.

E-mail address: sharpk@mail.med.upenn.edu (K.A. Sharp).

0301-4622/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0301-4622(03)00087-5

features of this picture are: The bulk water–water angular distribution function ($P(\theta)$) is bimodal, with a low angle population and a high angle population, centered at approximately 12° and approximately 54°, respectively. Solutes perturb the structure of water principally in their first hydration shell by changing the ratio of these two populations, apolar solutes increasing the low angle population, polar solutes increasing the high angle population. There is a direct link between these angular structural changes in the water hydrogen bond network in the first hydration shell of the solute and heat capacity changes. Changes in water–water distance in the first hydration shell are highly correlated with these angle changes, but are much smaller, explaining why previous experimental and theoretical analyses of water using the usual analyses of radial distribution functions [17,18] show apparently little change in water structure due to apolar solvation [14].

The sensitivity of $P(\theta)$ as a reporter of solute induced structure changes suggests that it could be used more generally to enhance our understanding of protein and nucleic acid hydration. A particularly intriguing potential application is that of antifreeze or thermal hysteresis proteins (THPs). These proteins are produced by a variety of organisms, including fish, spiders, insects and bacteria, in order to facilitate their survival at sub-freezing temperatures. Although the precise mechanism is unclear, these proteins act to depress the freezing point of a solution below the melting point in a non-colligative manner by accumulating at the surface and inhibiting the growth of ice crystals [19]. The ice-binding sites on several THPs have been identified and are distinguished by the relative planarity and nonpolar character of the surface (67% type I THP, 59% type III THP) [20,21]. The type III THP is a 66-residue protein that, unlike the α-helical type I THP, consists mainly of β-strands connected by large loops. Although type I THPs present regular arrays of hydrogen bonding groups on their surface, the geometric arrangement of potential hydrogen bonding groups cannot fully explain their ice-binding properties [25]. Recent studies have also identified an important role for hydrophobic interactions in the type I ice-binding mechanism [44]. An ice-binding face of the type III THP molecule has been proposed based on several studies [22–24], and it is notable for its flatness [21] and hydrophobicity. Thus, many questions remain regarding the mechanism of specificity and affinity in these proteins, the principal one being how such a hydrophobic surface is able to manifest a strong enough interaction with the face of an ice crystal in order to preferentially bind to ice nuclei even in the presence of a large excess (55 M) of liquid water.

The goals of this paper were: First, to refine our previous analysis of the effect of solutes on the water–water angular distribution function, and extend it to the study of protein hydration. Second, to apply the analysis to a type III THP protein in order to better understand its ice binding properties. Two specific questions include (i) whether there are significant differences in the hydration structure of the highly hydrophobic THP III ice-binding surface compared to either its non-binding surface, or an equivalently hydrophobic binding surface of a non-THP protein. (ii) Whether this analysis can shed light on the interaction between the THP III binding site and the ice surface.

2. Methods

2.1. Selection of proteins

In addition to the analysis of a type III THP, it is also necessary to apply the analysis to a non-THP of similar general physical properties as a control. The selection of both the THP and control was guided by several criteria: (i) Availability of a high-resolution crystal structure. (ii) Moderate size, in order to be able to include a sufficient number of explicit water molecules to form at least two complete hydration shells around the protein, yet run the simulations to convergence in a reasonable time. (iii) No post-translational modifications or heteroatoms. This both simplifies the simulation procedure and contributes to the generality of our results. (iv) Monomeric, so that dimer interfaces did not complicate our analysis of surface hydration. It is also desirable that the two proteins be of similar size and constructed of mainly the same type of secondary structure. Two proteins that fulfilled these requirements were the

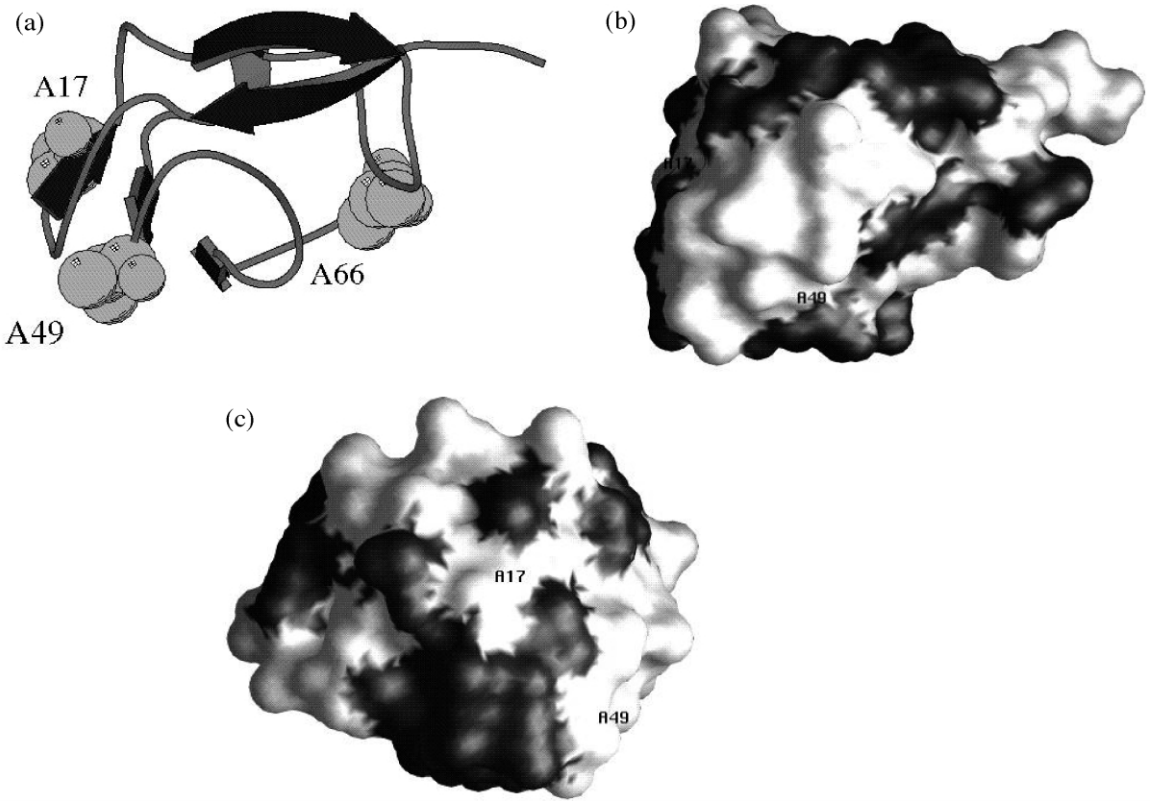

Fig. 1. THP. (a) Secondary structure. The ice binding face is on the left. Ala 17 in the center of the ice-binding site, Ala 49 is on the edge, and Ala 66 is on the opposite side, shown in CPK representation. Molecular surface in the same orientation (b) and rotated 90° to show binding face (c). Apolar residue surface is white, polar residue surface is dark grey.

66-residue type III THP (pdb entry 1MSI, [26]), and the 61-residue protein fasciculin-1 (pdb entry 1FAS, [27]). Both proteins are mainly comprised of β-sheets (Fig. 1a and Fig. 2a). Fasciculin-1 is isolated from mamba venom, and belongs to a family of three-fingered toxins that include α-neurotoxins and cardiotoxins. It is an extremely potent and highly selective inhibitor of most synaptic acetylcholinesterases [28,29]. The structure of fasciculin is composed of four anti-parallel β-sheets that connect three loops. Four disulfide bonds stabilize the globular core of the molecule (Fig. 2a). Despite the predominance of basic amino acids on its surface, it has been shown that hydrophobic interactions involving residues on the tips of loops I (residues 4–16) and II (residues 23–28) play a vital role in binding (Fig. 2a) by making van der Waals contacts with residues on the acetylcholinesterase surface [29].

2.2. Analysis of protein hydration

The overall approach to the analysis of hydration is as follows: we first run a molecular dynamics simulation of the protein in periodic box of explicit water molecules. For each saved frame in the trajectory, waters in the first hydration shell of the protein are identified, and the angular distribution function of all pairs of waters within a certain distance cut-off is calculated. The shape of the bimodal angular distribution function is quantified by obtaining the areas beneath the two populations by numerical integration. The ratio of these areas is characteristic of the hydration of different types

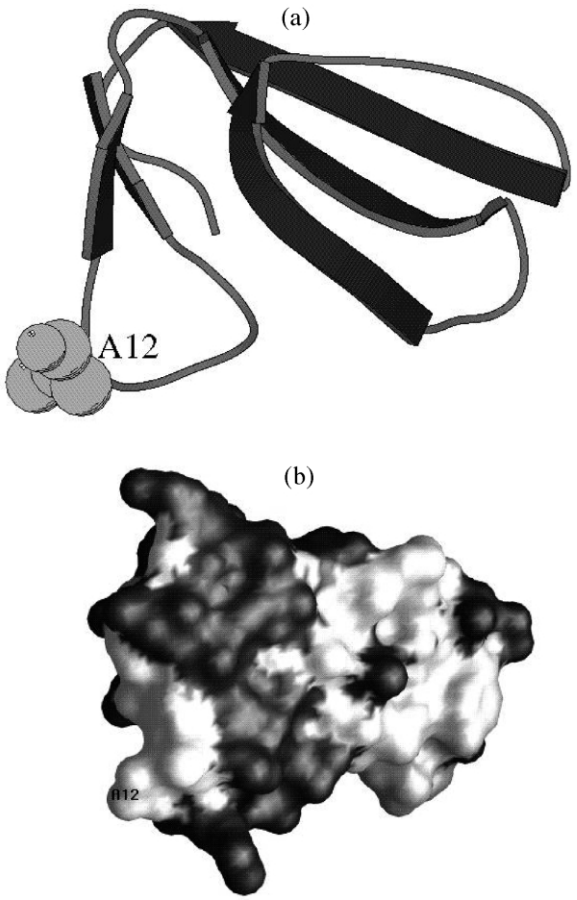

Fig. 2. Fasciculin-1. (a) The secondary structure. The single surface alanine on the biding loop is shown in CPK representation. (b) Molecular surface in the same orientation. Apolar residue surface is white, polar residue surface is dark grey.

of groups or atoms (i.e. polar, nonpolar, mixed). Finally, the area ratios are compared to the ratios derived from small solutes of varying polarity, and between different protein surfaces or surface subsets.

2.3. Molecular dynamics simulations

Molecular dynamics simulations were carried out using the CHARMM22 parameter set to model the solutes and the c22b1 version of the program CHARMM to generate the trajectories [30–32]. The solute–water systems were constructed in the same manner as those used in previous Monte Carlo simulations [33]. However, a very similar, but slightly simpler water potential, the TIP3P potential [34] was used to model the water because CHARMM has been highly optimized for this potential, and because the four-site TIP4P model could not be accommodated by the available c22b1 CHARMM code. Comparison of TIP3P and TIP4P for small solute hydration is possible using the Monte Carlo program BOSS [35], and we found very similar hydration structures [36]. A 63 cube of TIP3P water was built using InsightII (Accelerys, San Diego) and equilibrated at 298 K. The solvent cube was superimposed on the protein structure and overlapping atoms were eliminated. This resulted in a system containing approximately 25 000 atoms, of which 998 (THP III) and 921 (fasciculin) were protein atoms, providing at least 5 layers of water. A time step of 1 fs was used. The simulations were done at constant pressure and constant temperature was maintained ($T=300$ K) using the Nose-Hoover algorithm. Minimum image periodic boundary conditions were used. Electrostatic and van der Waals interactions were truncated using a shifting function between cut on and cut off distances of 11.5 and 13.0, respectively. For each solute system, a 50 ps equilibration preceded a 1 ns simulation. Configurations were saved at 0.1 ps intervals (every 100 steps). For both the Monte Carlo and the molecular dynamics simulations, convergence was determined by monitoring the batch averages of the energy and the hydration structural parameters. All protein atoms were allowed to move during the simulation,

Table 1
Position of solute atom–water $g(r)$ minima used to define the first hydration shell

Atom type	Hydration radius
C	5.60
O	3.40
N	3.60
S	5.30
H	0.00

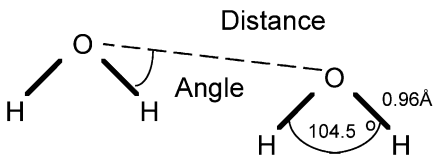

Fig. 3. Water structure. The geometry of the TIP3P water is shown. The water–water distance is taken as the distance between O atoms. The water–water angle is defined as the minimum of the four HO–O angles. Note that given the TIP3P HOH angle of 104°, the maximum water–water angle possible with this definition is 128°.

whereas the internal geometry of the solvent molecules was maintained by means of the SHAKE algorithm.

2.4. Analysis of water–water angular distribution functions

The first hydration shell waters were identified by finding all waters that lay within the first hydration shell cutoff of any protein or solute atom. The first shell hydration cutoff for each atom type is defined as the first minimum in that atom–water radial distribution function, $g(r)$ [37]. The $g(r)$ for each atom type was obtained from preliminary simulations on the model compounds alanine dipeptide and cysteine which contain all the required atom types (Table 1). The position of the first minimum in each function was identified using the peak identification function of Origin 6.0 [38]. Every pair of waters within a water–water distance cutoff was then analyzed, the distance and angle being defined as in Fig. 3. Previous work used a cutoff of 3.4, because this is the position of the first minimum in the bulk water–water $g(r)$ [37]. However, in order to increase the sensitivity of our water structure analysis, we examined the effect of increasing the water–water cutoff distances up to 5.2, in order to include more water–water pairs. Previously we referred to any water pair within the cutoff as

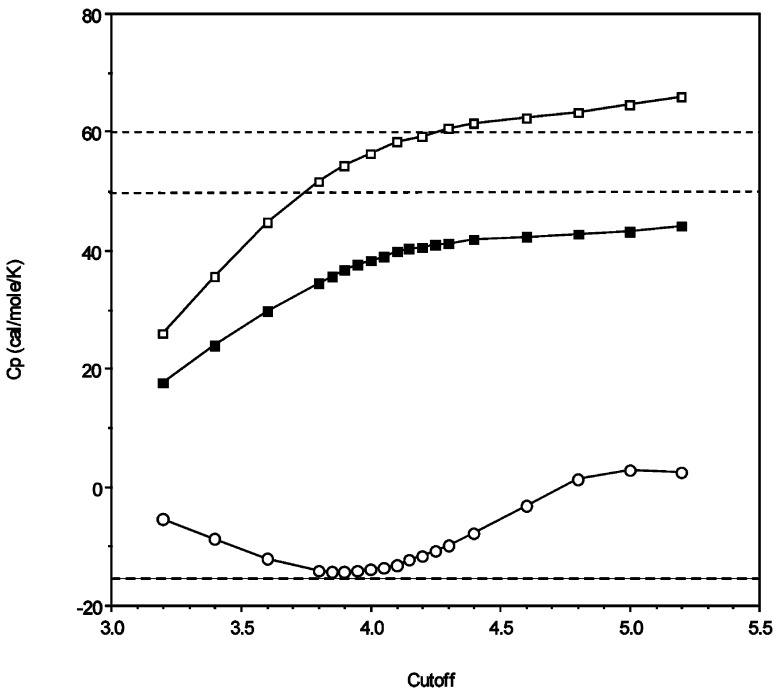

Fig. 4. Effect of cutoff on discrimination between polar and apolar hydration. Hydration heat capacity is calculated using the water–water angles via the rnm. (n) Methane, (q) ethane, (o) K^+. Experimental values are shown as horizontal lines.

'hydrogen bonded' regardless of angle. In view of our use of an increased distance cutoff, and the confusing lack of consensus on the geometric definition of a hydrogen bond, we use here the more general terms water–water angular distribution function $P(\theta)$ and reserve the term hydrogen bond for waters within 3.4. Fig. 4 demonstrates the difference in sensitivity as a function of the water–water distance. The plot shows the hydration heat capacity of small reference solute calculated via the rnm using the root mean square water–water angle, θ accumulated over the course of each simulation [37]. For our final analysis of THP III and FAS hydration we employed a cutoff of 4.0, which, judging from the difference in hydration heat capacities of the reference solutes in Fig. 4 provides the most discrimination in water structure perturbation. (It should be pointed out that this cutoff is only used for the structure analysis. The water–water non-bonded interaction cutoff used in the simulations to generate the protein/water configurations is much greater.) Standard deviations in other quantities were obtained from the variation in batch means between different Monte Carlo runs, or from the variation in batch means from successive segments of the molecular dynamics simulation.

For the protein hydration, the water–water angle distribution analysis was carried out using the program PRAM (available from the authors), which uses the CHARMM trajectory files as input. All atoms in the protein are first classified as being apolar, polar, or weakly polar, based on the partial charge assigned by the molecular dynamics force field. Those atoms possessing a charge magnitude less than 0.35 are considered nonpolar, charge magnitude 0.35–0.45 are classified as weakly polar, and atoms with a partial charge magnitude greater than 0.45 are considered polar. Solute atoms of interest can also be 'tagged' before analyzing the trajectories, allowing the hydration structure of subsets of the protein surface to be examined. Water–water angular distributions are also accumulated separately for the three possible types of solute atom polarity class: polar–polar, apolar–apolar, and polar–apolar (mixed). For example, if the first water oxygen was closest to an apolar solute atom and the second water oxygen was closest to a polar solute atom, that water–water angle would be added to the mixed angle frequency histogram.

2.5. Analysis of angle probability distributions

In order to obtain a quantitative comparison of $P(\theta)$ curves, the area under each peak, corresponding to the populations of low angle and high angle water–water geometry, is obtained by numerical integration using the trapezoidal rule as implemented in Origin 6.0 [38,39]. The boundary between the peaks is determined from the position of the minimum in the pure TIP3P water $P(\theta)$ curve (Fig. 5). The ratio of these two areas is characteristic of the type of solute (nonpolar, polar, etc.). Reference ratios were calculated from simulations of several small molecules of differing polarity studied previously by Madan and Sharp [2] using the rnm-explicit water method. Thus, we may interpret the distribution of water–water angles surrounding the protein surface in relation to small test solutes.

2.6. Surface subsets and controls

The hydration structures of the entire THP III and fasciculin surfaces were evaluated according to the procedures outlined above. We also examined several subsets of the protein surfaces. These included the ice-binding site of THP III, which is composed of residues 10, 16–19 and 45 (Q, TALT and Q, respectively—see Fig. 1c) and the acetylcholinesterase-binding loops of fasciculin, residues 4–16 and 23–28 (YSHTTTSRAILTN and YRKSRR, respectively—see Fig. 2b). We also analyzed separately the hydration structure surrounding the alanine 17 residue of THP III, which defines the center of the ice-binding site. The hydration structures of two other alanines on the THP III surface were studied, alanine 49, which is located adjacent to the binding site, and alanine 66, which is far removed from the ice-binding site (Fig. 1a). We also examined the hydration structure surrounding alanine 12 of fasciculin, which is in a very similar secondary structure environment as A17 of THP III (Fig. 2a). As another control, we included a subset of water hydrating a roughly

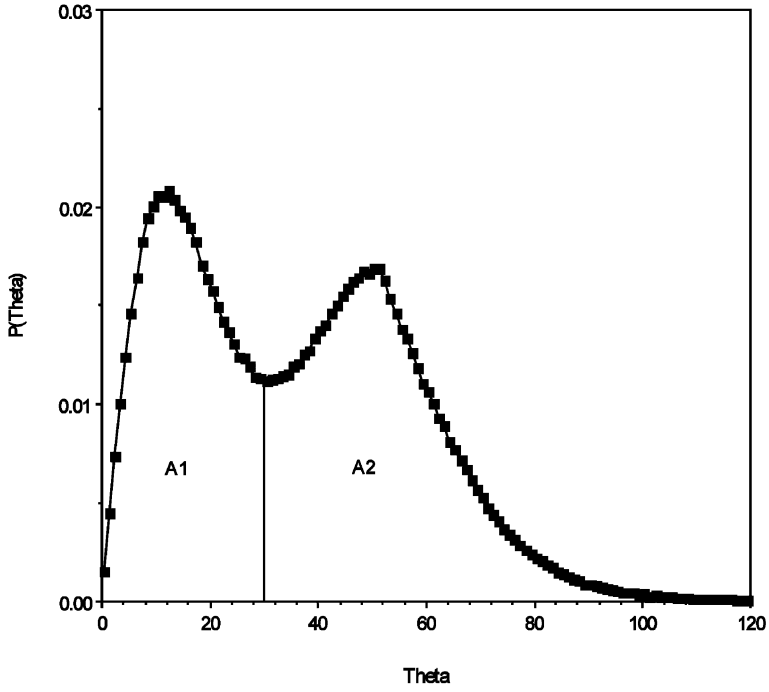

Fig. 5. Integration of water–water angular distribution populations. The division between the low angle and high angle populations is 30°, based on the minimum in the pure water curve shown for the TIP3P potential.

10-diameter patch of surface on the face of THP III that is directly opposite the binding site. This site is comprised of residues 3–7, 26–32 and 55–56 (QAS, EVVTPVG and TT, respectively).

3. Results

3.1. Protein surface hydrophobicity

The hydration structure of groups is primarily determined by their polarity. We first made a comparison of the overall hydrophobicity of Type III AFP with other proteins of similar size. We found in the PDB data base nine other proteins of similar size (61–70 residues) and high resolution (<2.1), including fasciculin (1FAS, 1A80, 1FSC, 1IGD, 1NXB, 1PIX, 1QKE, 1SNB, 2SNB) [40]. On average the surface of these proteins is $51 \pm 3\%$ apolar, which is typical for globular proteins of all sizes. This should be compared to 62% apolar surface for Type III AFP. This is very high, and comparable to one of the most hydrophobic soluble globular proteins known, Crambin, which has a 64% apolar surface. The binding face is even more hydrophobic, but does contain some polar regions (Fig. 1c).

3.2. Simulation results

After the 50-ps equilibration period, the energy and temperature values for all simulations were stable and all solute structures remained intact. In order to determine if the simulation length was sufficient, we examined the convergence of the parameter of interest, the water–water angle distribution population ratio (A1/A2). The angle distributions converged to their final profiles rather quickly. After only half of the configurations were analyzed, the population ratios had reached their final values (Fig. 6). This figure also shows that there is little variation and that after as few as 500 configurations, all classes are within 4% of their final values. Fig. 6 depicts the population ratios calculated for the entire first hydration shell of

THP III. Slower convergence of the polar class, which contains the fewest number of first hydration shell water–water pairs, is seen. All the surface subsets converge more slowly given the smaller still number of water pairs, but we looked at the convergence around the polar atoms of an alanine residue (the worst case scenario, in terms of statistics), and it converged to within 4% of the final value within half of the run length. Thus 10 000 configurations are sufficient to establish precise population ratios for all classes and subsets.

3.3. Population ratios for small molecules

Table 2 lists the calculated low angle (A1) and high angle (A2) populations in the hydration structure for pure water, methane, ethane, ethanol and the potassium ion. The ratio of A1 to A2 is distinct for different types of hydration. This ratio is close to 1.00 for nonpolar groups and the more polar the solute, the lower the A1/A2 ratio. For example, the polar atoms of ethanol decrease the ratio slightly (to 0.81), whereas the highly polar potassium ion reduces this ratio to 0.07. Pure water falls between these two extremes at 0.60. It is interesting to note that the ratio for the mixed class in ethanol is lower than the ratio for the polar atoms (0.72 and 0.81, respectively). As we have shown previously by detailed analysis of specific water structures corresponding to these different $P(\theta)$ distributions [16], the higher the low angle

Table 2
Water–water low angle and high angle populations for small solutes

Solute	A1	A2	A1/A2
TIP3P water	0.37	0.63	0.60
Methane	0.49	0.51	0.96
Ethane	0.49	0.51	0.96
Ethanol-nonpolar	0.51	0.49	1.02
Ethanol-polar	0.45	0.55	0.81
Ethanol-mixed	0.42	0.58	0.72
Potassium ion	0.06	0.94	0.07

Populations defined as the areas A1, A2 in Fig. 5.

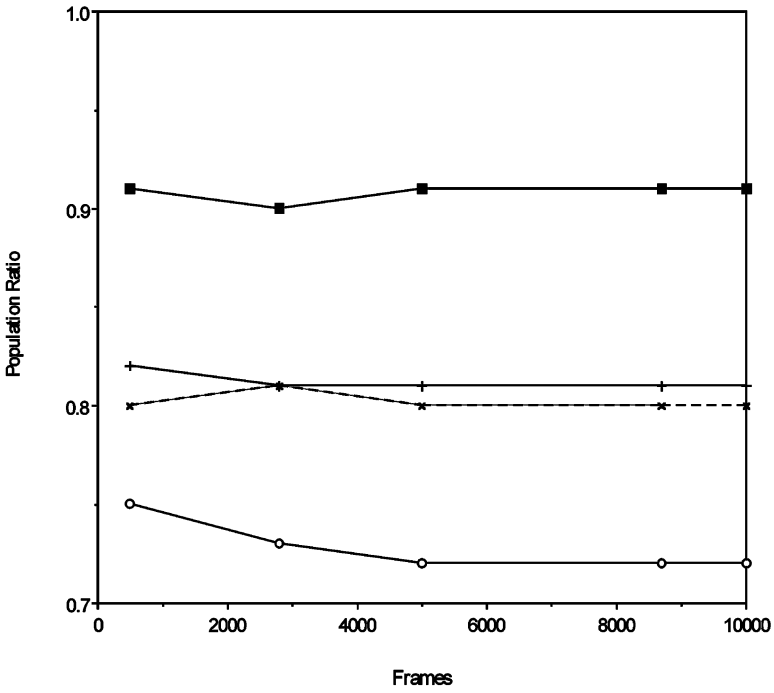

Fig. 6. Convergence of low angle/high angle population ratios. The ratios for the entire surface of THP III are shown as a function of the number of configurations analyzed. Hydrating water–water class: (n) nonpolar, (o) polar, (x) mixed, (-) all classes.

population, the more tetrahedral-like, or ice-like the hydrating water structure is.

3.4. Overall $P(\theta)$ distributions for THP III and fasciculin

The water–water angle distributions for the entire first hydration shell of THP III and fasciculin are shown in Fig. 7a and b, and are typical of the profiles seen for the different classes of solvating water pairs seen previously, and in Table 2. The apolar–apolar pairs have a higher low angle population, the polar–polar pairs a lower angle population, and the mixed pairs are intermediate. Population ratios for the nonpolar and polar classes are summarized in Fig. 9. The small molecule ratios are included as points of reference and are indicated by the horizontal lines on the histograms. In general, the nonpolar character of the THP III surface is reflected in higher A1/A2 ratios than fasciculin. The water structuring surrounding nonpolar and polar atoms differs little between the two molecules. The most significant differences are found in the mixed class (ratio data for the mixed class not shown in Fig. 9). In this class, the THP III is significantly (8%) more nonpolar than fasciculin.

3.5. The THP III binding site—internal comparison

In order to examine the differences between ice-binding and non ice-binding surfaces, the water structure surrounding the proposed binding site of THP III was compared with two controls: a patch of surface on the opposite face, and the hydrophobic binding loops of fasciculin. The $P(\theta)$ distributions are shown in Fig. 8a–c. Population ratios for all three surface subsets are summarized in Fig. 9. Comparison of these profiles indicates significant differences between the ice binding surface and the two non-ice binding surfaces. For the ice-binding surface there is now almost no distinction between the apolar, polar and mixed water-pair angle distributions, in contrast to the other two surfaces. This difference comes primarily from a shift in the angular distributions of polar and mixed pairs towards the apolar-type distribution around the ice binding site, indicating two things: (i) a less distorted, more tetrahedral or hydrophobic-like water structuring around the ice binding site. (ii) A homogenous, or regular hydration structure around the ice-binding site. These changes are apparent in the population ratios in Fig. 9, where the most significant change for the ice-binding site is in the A1/A2 ratio for polar groups. At 0.89 it is approximately 16% greater than for polar groups on the opposite face (0.77). Comparing this to the ratio for polar groups on the entire THP III surface (0.72), the increase in the ice-binding site is even greater, at 24%. This same trend is also observed in the mixed hydrating water class (ratio data not shown in Fig. 9), although the differences between binding site and the opposite face or the entire surface are smaller (10 and 14%, respectively). Although the surface of fasciculin is on average significantly more polar than that of THP III, it also contains two loops that form important nonpolar contacts upon binding. While the hydration water structure surrounding these loops is more tetrahedral than the average of the entire fasciculin surface, it is still significantly more distorted than the network hydrating the THP III binding site. The nonpolar, polar, mixed and combined (i.e. 'All Types') classes are 6, 17, 20 and 17% less tetrahedral, respectively.

3.6. Water structure surrounding alanine residues

In order to isolate more specifically the structural origin of the different hydration structure seen in the ice-binding site, the hydration around the key Ala 17 residue in the center of the ice-binding site, Ala 49 on the edge of this site, and two alanines not in ice-binding regions (Ala 69 of Type III AFP and Ala 12 of Fasciculin) was analyzed. The population ratios for all classes of water pairs hydrating these four alanines are summarized in Fig. 9. The most striking feature is that the amount of the low angle population is significantly enhanced for all classes of water around the alanines in the ice-binding site, Ala 17 and Ala 49. The largest effect is seen for polar groups neighboring Ala 17 in the center of the site, but the effect is noticeable for Ala 49 at the edge of the site. Even the water hydrating polar groups around these two alanines has a very apolar-like

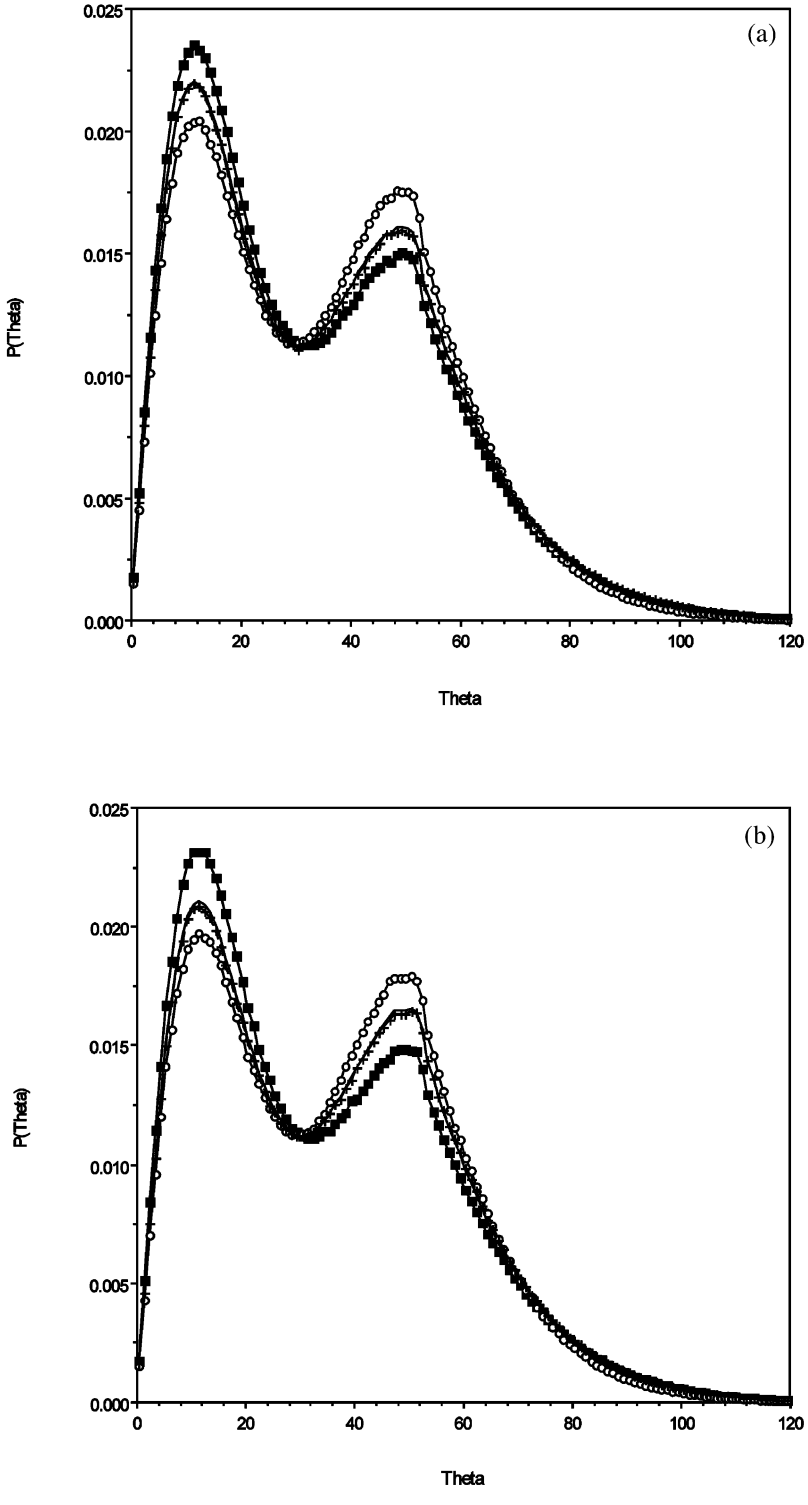

Fig. 7. Overall water structure around THP III and Fasciculin. The water–water angle probability distributions for the entire first hydration shell of THP III (a) and Fasciculin (b). Hydrating water–water class: (n) nonpolar, (o) polar, (x) mixed, (-) all classes.

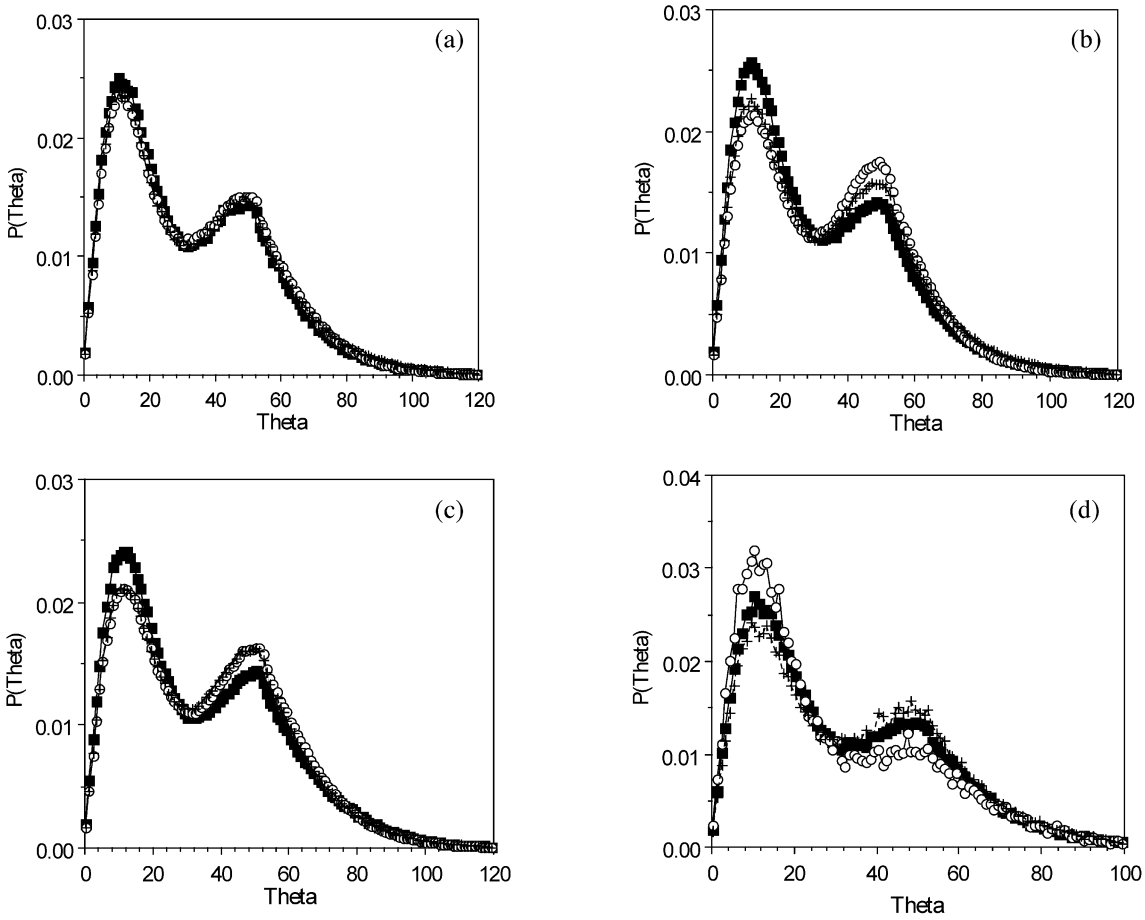

Fig. 8. Water structure around surface subsets. The water–water angle probability distributions for the hydration shell of the THP III ice-binding site (a); a 10 patch opposite the THP III ice-binding site (b); the binding site of fasciculin (c); around Ala 19 in the center of the THP III ice-binding site (d). Hydrating water–water class: (n) nonpolar, (o) polar, (x) mixed.

population ratio. In the case of Ala 17 this exceeds that of the apolar class, because of the large shift to low angles in the $P(\theta)$ distribution for the polar hydrating water class (Fig. 8d), indicating a significant increase in more tetrahedral-like water structure. The important polar groups for the Ala 17 and Ala 49 hydration are the carbonyl oxygens of the same residues, which in both cases are on the outer edge of beta sheet, i.e. not making protein–protein H-bonds. An important neighbor for Ala 17 is the amide side-chain of Asn 45, which also is not involved in protein–protein H-bonds, and is near enough to affect the hydration structure of Ala 17.

4. Discussion

This work was motivated by the ability of the rnm analysis of water hydration to reveal subtle but important changes in water structure induced by polar and apolar solutes, and by the desire to better understand the mode of action of THPs. To this end we have extended our analysis protocol so that we can easily and accurately analyze water structure around the entire surface of a protein or any specified subset(s). We have performed extended molecular dynamics simulations of a Type III THP and a suitable control protein, Fasciculin, in water, using the CHARMM/TIP3P

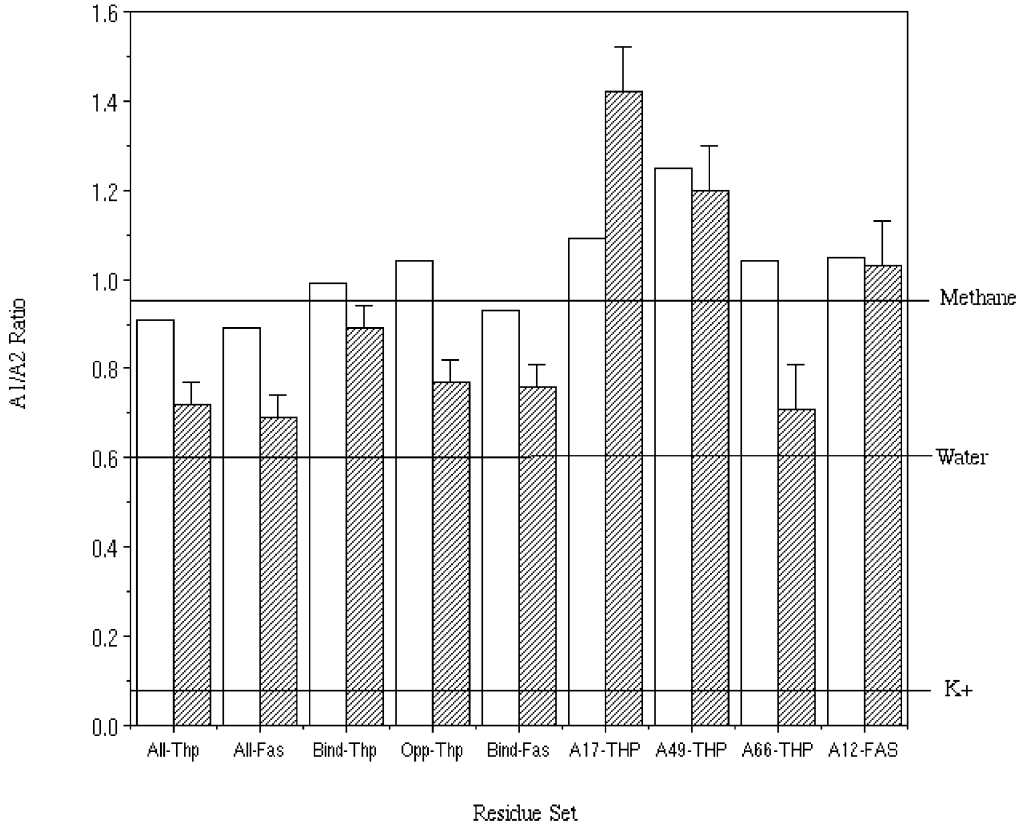

Fig. 9. Summary of $P(\theta)$ population ratios. The ratio of low angle to high angle populations in the water–water angle probability distributions for the entire protein surface (All), the binding site (Bnd), the opposite face of THP (Opp), and the alanines (A17, A49, A66 of THP and A12 of FAS). Unshaded bars, shaded bars are the nonpolar and polar hydrating water–water classes, respectively. Ratios for pure water, hydration of methane and potassium are shown as horizontal lines for reference.

potential functions and analyzed the water structure. A newly developed software program, PRAM, can then be used to calculate the distribution of water–water angles, $P(\theta)$, for macromolecular, as well as simple solute, surfaces. We have also developed a better quantitative method to analyze changes in $P(\theta)$: numerical integration of the two peaks in the $P(\theta)$ distribution and calculation of the low to high angle population ratio (A1/A2). This value can be easily computed for the entire protein surface or for any part of it, and the results interpreted, using values obtained from small test solutes as a reference. In previous work, numerical comparisons of $P(\theta)$ curves were performed by comparing θ_{rms} (i.e. the width of the $P(\theta)$ distribution) [2,41]. While θ_{rms} and the A1/A2 population ratios follow the same rank order for apolar, mixed, polar classes of hydrating water, the population analysis can incorporate subtleties in $P(\theta)$ that the θ_{rms} does not because the latter model contains less information on both the shape of the distribution and the relative amounts of the two populations. Furthermore, the A1/A2 ratio defines the populations in relation to pure water and the reference values for hydration of apolar/polar small molecule provides us with a quantitative measure of changes in hydrating water structure around different protein groups.

Applying this analysis to the type III THP and the control protein Fasciculin (Figs. 7–9), we can draw several conclusions. First, we note that the relative population ratios for apolar, mixed and

polar hydrating water pairs follow the expected trend based on polarity, and are similar for the two proteins. However, the ratios for every class (and thus the average value for all types of hydrating water) are higher for THP III compared to FAS. Thus, these ratios may be interpreted as a reflection of the overall hydrophobicity of the protein surface. Second, the differences we see are much smaller than the range seen in small molecule hydration. For example, compared to our small molecule polar standard, K^+, the polar protein atoms' $P(\theta)$ distributions are not as distorted, indicating that either water has a relatively weak interaction with these groups or the effect of neighboring polar and apolar groups on the protein surface tend to cancel. When we examine subsets of the surface, however, two significant differences between the ice binding and non-ice binding surfaces emerges. First, the hydration of the polar atoms in the functional, or ice-binding site, of THP III contains a greater amount of tetrahedral structure or ice-like water structure than either the nonfunctional portion of the molecule or the fasciculin control surface. Second, the hydration structure is more uniform over the ice-binding site, i.e. there is less variation between the polar and apolar hydration even in the absence of a regular structure of protein surface hydrogen bonding groups like those in Type I THPs. The flatness of the ice-bonding surface no doubt contributes to its more uniform hydration structure compared to the other surfaces.

If we further narrow our focus to study the structuring of water about individual residues, we observe even greater differences between the ice-binding surface and other surfaces, primarily with regards to polar atom hydration. The water solvating the polar groups near the ice-binding site alanines A49, and especially A17, has structure very uncharacteristic of polar hydration, and it is in fact as, or more, apolar-like as water hydrating the apolar side-chains of Ala 17 and Ala 49. These polar groups include the carbonyl oxygens of the two alanines. The unique hydration structure of these polar groups must be a direct consequence of the protein structure in this region, and the neighboring polar/apolar environment, including the Asn 45 side chain. However, even detailed comparative analysis of this part the protein surface structure and that of non ice-binding surfaces is not capable, a priori, of explaining the differences in water hydration we see. These clearly depend on aspects of the structure that are not obvious to the eye, but can be revealed by explicit solvent simulations averaging over many water configurations, and using sensitive structural measures (e.g. angular distortions) in the analysis. This analysis effectively tells us that the effect these polar groups have on water structure can be more easily modulated (compared to polar groups on protein surfaces in general) by the opposing tendency of surrounding apolar groups to enhance the tetrahedral structure of water. This happens to such an extent for these polar groups surrounding the ice-binding alanines that the apolar neighbor group effect actually dominates.

5. Conclusions

Although many studies have attempted to elucidate the binding mechanism of THP III, their main successes have been in identifying and confirming which residues in the protein play an important role in binding. In addition, it has been difficult to unambiguously establish which plane of ice comprises the protein-binding site. Many modeling studies have presumed, as initial evidence suggested, that THP III binds to the primary prism $(1\,0\,\bar{1}\,0)$ plane, yet recent data suggests that the protein may, in fact, interact with the pyramidal plane $(2\,0\,\bar{2}\,1)$ and other faces, as well [23,42]. Some have proposed that the growing ice surface is semi-ordered [23]. Thus, there has also been debate about whether THP III recognizes an existing site on a nascent ice crystal, whether the protein itself influences the formation of the binding site, or can bind at several sites.

Regardless of the precise mechanism of THP action, general consideration of the problem of ice binding suggests that the main results of our analysis: (i) more uniform hydration structure in the ice-binding region primarily due to (ii) polar groups with very apolar-like hydration, are significant for the following reasons. The major problem for a THP is how to recognize, and preferentially

bind nascent ice crystals forming in liquid water, with the two types of surface groups it has available: polar and apolar. If it uses polar groups, then ordinarily, such groups, being polar will have a high affinity for liquid water which is present at large excess (55 M), and net binding of ice, present initially at low concentrations, will be negligible. On the other hand, if it uses apolar groups, these presumably prefer more ice-like water, but because of their hydrophobicity, their affinity for any kind of water—liquid or ice—will be low. These unsatisfactory situations may be summarized as either: high affinity but low specificity, or high specificity but low affinity. We suggest that the solution for THPs is to provide some polar surface (for affinity) that yet prefers more ice-like water (specificity). If there is such a polar, pycrophilic surface, one would expect to recognize it, when surrounded by bulk water, by its tendency push its hydrating water towards a more ice-like structure rather than away like a typical polar surface. This is exactly what we observe in the ice-binding region of Type III THP. This also has the effect of making the hydration structure of the ice-binding site more uniform across polar and apolar regions, which also would favor binding to the more structured ice form of water over the liquid form. The highly conserved flatness of the ice-binding face [42] would also seem to play a role in enhancing the uniformity of the hydration structure.

Explicit water simulations of the binding of small ice-nuclei in the presence of bulk water would in principle be able to test some of these conclusions, but simultaneous protein/water/ice simulations at thermodynamic equilibrium are currently beyond the ability of potential function/simulation technology. Other future directions would include a study of the series of mutants at the A17 position (located at the center of the ice-binding site) made by DeLuca et al. [43]. These mutations were specifically designed to disrupt the topology of the binding surface. The crystal structures of several of these mutants have been determined (A17M, A17H, A17T, A17C and A17R). An analysis of $P(\theta)$ for these mutants, and also mutations affecting the neighboring polar groups would both increase our database of protein A1/A2 values and could improve our understanding of THP action.

Acknowledgments

We thank Dr Frank Sönnichsen for helpful discussions. Financial support is gratefully acknowledged from NIH (GM54105).

References

[1] R.J. Doyle, Contribution of the hydrophobic effect to microbial infection, Microbes Infect. 2 (2000) 391–400.
[2] K.A. Sharp, B. Madan, The hydrophobic effect, water structure and heat capacity changes, J. Phys. Chem. 101 (1997) 4343–4348.
[3] W.E. Stites, Protein–protein interactions: interface structure, binding thermodynamics, and mutational analysis, Chem. Rev. 97 (1997) 1233–1250.
[4] C.H. Tanford, The Hydrophobic Effect: Formation of Micelles and Biological Membranes, Wiley, New York, 1980.
[5] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133.
[6] R.S. Spolar, M.T. Record, Coupling of local folding to site-specific binding of proteins to DNA, Science 263 (1994) 777–784.
[7] W. Kauzmann, Some factors in the interpretation of protein denaturation, Adv. Prot. Chem. 14 (1959) 1–63.
[8] A. Ben-Naim, Y. Marcus, Solvation thermodynamics of nonionic solutes, J. Chem. Phys. 81 (1984) 2016–2027.
[9] K.P. Murphy, P.L. Privalov, S.J. Gill, Common features of protein unfolding and dissolution of hydrophobic compounds, Science 247 (1990) 559–561.
[10] R. Baldwin, Temperature dependence of the hydrophobic effect in protein folding, Proc. Natl. Acad. Sci. 83 (1986) 8069–8072.
[11] P. Privalov, G. Makhatadze, Contribution of hydration to protein folding thermodynamics. II The entropy and Gibbs free energy of hydration, J. Mol. Biol. 232 (1993) 660–679.
[12] P. Privalov, G. Makhatadze, Contribution of hydration to protein folding thermodynamics. I The enthalpy of hydration, J. Mol. Biol. 232 (1993) 639–659.
[13] A.R. Henn, W. Kauzmann, Equation of state of a random network, continuum model of liquid water, J. Phys. Chem. 93 (1989) 3770–3783.
[14] B. Madan, K. Sharp, Changes in water structure induced by a hydrophobic solute probed by simulation of the water hydrogen bond angle and radial distribution functions, Biophys. Chem. 78 (1999) 33–41.
[15] F. Vanzi, B. Madan, K. Sharp, Effect of the protein denaturants urea and guanidinium on water structure: a structural and thermodynamic study, JACS 120 (1998) 10748–10753.

[16] K. Sharp, B. Madan, E. Manas, J. Vanderkooi, Water structure changes induced by hydrophobic and polar solutes revealed by simulations and IR spectroscopy, J. Chem. Phys. 114 (2001) 1791–1796.
[17] J. Turner, A. Soper, The effect of apolar solutes on water structure: alcohols and tetraalkylammonium ions, J. Chem. Phys. 101 (1994) 6116–6125.
[18] J. Turner, A. Soper, J. Finney, Ionic versus apolar behavior of the teramethylammonium ion in water, J. Chem. Phys. 102 (1995) 5438–5443.
[19] Y. Yeh, R.E. Feeney, Antifreeze proteins: structures and mechanisms of function, Chem. Rev. 96 (1996) 601–617.
[20] P.L. Davies, B.D. Sykes, Antifreeze proteins, Curr. Opin. Struct. Biol. 7 (1997) 828–834.
[21] D.S.C. Yang, W.C. Hen, S. Bubanko, et al., Identification of the ice-binding surface on a type III AFP with a 'flatness function' algorithm, Biophys. J. 74 (1998) 2142–2151.
[22] H. Chao, F.D. Sönnichsen, C.I. DeLuca, P.L. Davies, B.D Sykes, Structure-function relationship in the globular type III antifreeze protein: identification of a cluster of surface residues required for binding to ice, Protein Sci. 3 (1994) 1760–1769.
[23] G. Chen, Z. Jia, Ice-binding surface of fish type III antifreeze, Biophys. J. 77 (1999) 1602–1608.
[24] F.D. Sönnichsen, C.I. DeLuca, P.L. Davies, B.D. Sykes, Refined solution structure of type III antifreeze protein: hydrophobic groups may be involved in the energetics of the protein-ice interaction, Structure 4 (1996) 1325–1337.
[25] H.M. Chao, M.E. Houston, R.S. Hodges, et al., A diminished role for hydrogen bonds in antifreeze protein binding to ice, Biochemistry 36 (1997) 14652–14660.
[26] Z. Jia, C.I. DeLuca, H. Chao, P.L. Davies, Structural basis for the binding of a globular antifreeze protein to ice, Nature 384 (1996) 285–288.
[27] M.H. le Du, P. Marchot, P.E. Bougis, J.C. Fontecilla-Camps, 1.9-Resolution structure of fasciculin 1, an anti-acetylcholinesterase toxin from green mamba snake venom, J. Biol. Chem. 267 (1992) 22122.
[28] C. Cervenansky, F. Dajas, A.L. Harvey, Snake Toxins, Pergamon Press, New York, 1991.
[29] Z. Radic, R. Duran, D.C. Vellom, Y. Li, C. Cervernasky, P. Taylor, Site of fasciculin interaction with acetylcholinesterase, J. Biol. Chem. 269 (1994) 11233–11239.
[30] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization and dynamics calculations, J. Comp. Chem. 4 (1983) 187.
[31] MacKerell Jr., et al., Self-consistent parameterization of biomolecules for molecular modeling and condensed phase simulations, FASEB J. 6 (1992) A143.
[32] A.D.J. MacKerell, B. Brooks, I.C.L. Brooks, et al., in: P.v.R. Schleyer (Ed.), The Encyclopedia of Computational Chemistry, vol. 1, Wiley, Chichester, 1998, pp. 271–277.
[33] B. Madan, K.A. Sharp, Heat capacity changes accompanying hydrophobic and ionic solvation: a Monte-Carlo and random network model study, J. Phys. Chem. 100 (1996) 7713–7721.
[34] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926.
[35] W.L. Jorgensen, BOSS, Version 3.3, Copyright Yale University, New Haven, CT, 1992.
[36] K.R. Gallagher, Electrostatic Contributions to Heat Capacity Changes and an Analysis of Thermal Hysteresis Protein Hydration Using the Random Network Model, University of Pennsylvania, Philadelphia, 2002.
[37] B. Madan, K.A. Sharp, Heat capacity changes accompanying hydrophobic and ionic solvation: a Monte-Carlo and random network model study, J. Phys. Chem. 104 (2000) 12047–12047.
[38] Origin6.0, Microcal Software, Inc. Northampton, MA, 1999.
[39] W. Press, B. Flannery, S. Teukolsky, W. Vetterling, Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, New York, 1985.
[40] H.M. Berman, J. Westbrook, Z. Feng, et al., The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[41] B. Madan, K.A. Sharp, Molecular origin of hydration heat capacity changes of hydrophobic solutes: perturbation of water structure around alkanes, J. Phys. Chem. 101 (1997) 11237–11242.
[42] A.A. Antson, D.J. Smith, D.I. Roper, et al., Understanding the mechanism of ice binding by type III antifreeze proteins, J. Mol. Biol. 305 (2001) 875–889.
[43] C.I. DeLuca, P.L. Davies, Q. Ye, Z. Jia, The effects of steric mutations on the structure of type III antifreeze protein and its interaction with ice, J. Mol. Biol. 275 (1998) 515.
[44] M.M. Harding, L.G. Ward, A.D.J. Haymet, Type I antifreeze proteins: structure-activity studies and mechanism of ice growth inhibition, Eur. J. Biochem. 264 (1999) 653–665.