Soft Matter

Cite this: *Soft Matter*, 2012, **8**, 324

www.rsc.org/softmatter

PAPER

Probing the structural and functional link between mutation- and pH-dependent hydration dynamics and amyloidosis of transthyretin

Xue Xu,†[ab] Xia Wang,†[ab] Zhentao Xiao,[a] Yan Li[c] and Yonghua Wang*[ab]

Received 17th August 2011, Accepted 5th September 2011
DOI: 10.1039/c1sm06569f

Protein surface hydration, providing a flexible matrix enabling the protein to respond efficiently to environmental changes, is central to its folding, structure and stability. The aberrant folding of proteins links with a huge variety of diseases, such as prion diseases, diabetes and cancer. Considering the correlation between the mechanism by which a polypeptide chain folds to a specific three-dimensional protein structure, and the role of hydration in the aggregation of misfolded proteins, we use 23 large-scale molecular dynamics simulations to study the local hydration dynamics at the surface of transthyretin (TTR), a thyroid hormone-binding protein that transports thyroxine from the bloodstream to the brain. Monitoring the effects of solvent dynamical behavior around specific mutations by density and spatial distribution entropy maps of the solvent identifies two kinds of water: (1) the hydration water surrounding the stability-bearing mutations characterized by long residential time and slow diffusion; and (2) the water adjacent to the amyloidogenic mutations in fast exchange with the bulk water. Alternative conformations of the protein induced by mutations govern the solvent dynamical behaviors, further evidenced by a map of the spatial distribution entropy of the solvent around the protein. The special behavior of the solvent around these regions is probably crucial in the folding stability and in terms of aggregation loci. These results are also proven by the global perturbations of the protein hydration shell by acidic pH that exhibits dramatic suppression of the aqueous protein motion, and meanwhile, inactivates the stability-bearing waters. This indicates that the acidic medium destroys the cluster structure of water molecules, inducing water diffusion and protein conformational transformation. The present work opens up a possibility of using the mutation and the pH value as probes for protein folding kinetics and functional dynamics measurements, and provides a clue for the treatment of amyloid diseases associated with TTR misfolding.

1 Introduction

Water is an outstanding chemical substance that has been long appreciated to be absolutely fundamental to life. It is life's ubiquitous background, and without it life cannot be sustained.[1] The interaction between waters and biological systems has received great attention since this may bring pathological concerns.[2] For a living system, one of its defining and intricate characteristics is the ability to self-assemble with precision and fidelity in the solvent medium.[3] As the most fundamental and universal example of biological self-assembly, the folding of proteins into their compact three-dimensional structures has enabled living systems to develop astonishing diversity and selectivity in their underlying chemical processes.[3] However, only correctly folded proteins have long-term stability in the water environment and are able to interact selectively with their natural partners. It is therefore not surprising that the failure of proteins to fold correctly, or to remain correctly folded, will give rise to the malfunctioning of living systems and hence to diseases.[4,5] Since proteins are the functional units of life, studies on the interactions of water and protein may provide a key to understand the basic questions in the formation of protein-folding disorders.

There is an increasing interest in the interactions of water and protein in recent years. Most of these studies so far focus on the hydration properties on the surface of proteins. For example, a recent molecular dynamics simulation has shown that due to the equilibrium processes between solubilization and aggregation, the intramolecular protection of backbone hydrogen bonds from water attack is responsible for the stability of the native structures of soluble proteins.[6] Separately, it is also shown that for the monomeric proteins with a high density of backbone

[a]*Center of Bioinformatics, Northwest A&F University, Yangling, 712100, Shaanxi, China. E-mail: yh_wang@nwsuaf.edu.cn; Tel: +86 029 87092262*
[b]*College of Life Sciences, Northwest A&F University, Yangling, 712100, Shaanxi, China*
[c]*School of Chemical Engineering, Dalian University of Technology, Dalian, 116012, Liaoning, China*
† Both authors contributed equally.

hydrogen bonds exposed to water attack, they have a great propensity to form amyloid fibrils;[7] such structural defects tightly associated with the effects of system properties (temperature, pressure, pH, and mutation), can cause local damage to the cells, thus leading to a large number of diseases, including Alzheimer's disease, Huntington's disease, Creutzfeldt-Jakob's disease and Parkinson's disease.[5,8] We note, Ferrão-Gonzales *et al.* showed experimentally that high hydrostatic pressure could induce a defective fold in the wild-type protein, converting it to an amyloidogenic conformation already present in the aggressive mutant.[9] Our previous study also found that the distribution of water molecules surrounding ligand binding regions would be more scattered with the rise in temperature, hence promoting the formation of fibrillar aggregates.[10] Despite these studies, the mechanisms, particularly, the water dynamical behaviors surrounding the mutations of protein and their changes after a mild pH treatment, which may lead to the loss of the original structure and function of proteins, are still little understood. Various experimental techniques have been developed to shed light onto this important problem;[11,12] however, the inherent difficulties involved in the dynamics of solvent can limit their applications, and the hydration dynamics may only be evaluated qualitatively but not directly and quantitatively. The computational simulations, on the other hand, might be a powerful method to study the distribution and the mobility of water around proteins.

Here, we take the most common hereditary amyloidogenic protein transthyretin (TTR) as an example to illustrate the idea. TTR serves as a secondary carrier for thyroxine (T4) and for binding retinol-binding protein (RBP) in plasma and cerebrospinal fluid.[13] It exists as a 127-residue β-sheet-rich tetramer.[14] By systematic studies of the solvation dynamics around the surface of TTR using MD simulations with femtosecond resolution, the water dynamics, their time scales and a series of dynamic correlations of water behaviors with local protein properties are explored. We find that hydration sits on the protein surface possess significant features about the stability and the potential interactions of the TTR protein. Further analysis shows that mutation and pH are crucial variables for determining the solvent motions and features of protein. We identify tightly bound water sites adjacent to the stability-bearing mutations, and the mobile behavior of water around the amyloidogenic mutations. Also, compared with the water dynamics of TTR in central mediums, we observe that the solvent molecules under acidic condition prone to have shorter residence times, and faster water exchanges.

2 Results and discussion

2.1 Water density maps and hydration behavior on the protein surface

Water is an integral part of proteins.[1] Its importance for the formation of amyloids has been deduced from a study with the amyloidogenic β-hairpin peptide of the Syrian hamster prion protein, which illustrates that solvent exposure of hydrophobic surfaces is the driving force for the folding of the peptide.[15] This raises the question: can the hydration-folding mechanism be applied to the plasma protein TTR?

Table 1 H-bond distances (Å) between water oxygen and protein atoms in the MD trajectories

Site	Protein atom	2QGB	3GRG	3A4F	3DJZ	3DO4	2QEL	2G4E	1X7S	1TSH	1G1O	1X7T	1F86	1SOK	Residence time (ns)
A	Leu58-H	2.51 ± 0.40	3.19 ± 0.70	—	—	—	—	2.26 ± 1.11	1.75 ± 0.59	2.68 ± 1.12	—	—	2.75 ± 0.60	1.68 ± 0.60	4.80 ± 3.14
	Thr59-H	1.96 ± 0.50	2.11 ± 1.00	—	—	—	—	1.90 ± 0.30	2.53 ± 0.45	2.33 ± 1.02	—	—	2.50 ± 0.58	2.39 ± 0.48	
B	Glu63-OE1	—	1.98 ± 0.32	—	—	—	—	—	—	1.78 ± 0.50	—	—	1.99 ± 0.50	2.00 ± 0.24	1.63 ± 0.48
C	His88-O	1.81 ± 0.20	1.53 ± 0.40	1.84 ± 0.20	—	—	—	2.96 ± 0.50	—	2.06 ± 0.08	3.50 ± 1.23	—	1.94 ± 1.04	1.97 ± 0.34	1.86 ± 1.07
D	Asn98-H	—	—	—	—	—	—	3.41 ± 1.05	—	—	—	3.05 ± 0.53	—	—	1.50 ± 1.41
	Gly101-O	—	—	—	—	—	—	1.60 ± 1.11	—	—	—	3.5 ± 1.24	—	—	
E	Arg104-O	1.62 ± 0.30	—	1.95 ± 0.40	3.05 ± 0.33	3.03 ± 0.38	—	2.87 ± 0.60	2.01 ± 0.40	2.74 ± 0.44	2.91 ± 0.28	—	—	2.87 ± 0.78	2.33 ± 1.22
	Thr106-H	3.16 ± 0.48	—	2.61 ± 1.00	3.10 ± 0.55	3.14 ± 0.50	—	1.68 ± 0.90	2.97 ± 0.50	2.89 ± 0.80	2.45 ± 0.45	—	—	3.04 ± 1.04	
F	Ser112-O	2.10 ± 0.10	3.40 ± 0.10	—	—	2.00 ± 0.20	—	—	2.60 ± 0.07	2.40 ± 0.90	—	—	2.80 ± 0.59	2.50 ± 0.15	6.20 ± 4.10
	Ser115-O	3.00 ± 0.30	3.50 ± 0.10	—	—	2.20 ± 0.43	—	—	2.90 ± 0.64	3.50 ± 0.21	—	—	3.50 ± 0.22	3.50 ± 0.37	2.10 ± 1.30

Table 2 H-bond distances (Å) between water oxygen and protein atoms in the X-ray structures

Site	Protein atom	2QGB	3GRG	3A4F	3DJZ	3DO4	2QEL	2G4E	1X7S	1TSH	1G1O	1X7T	1F86	1SOK
A	Leu58-H	2.91	2.60	—	—	—	—	2.57	3.11	2.42	—	—	2.92	2.82
	Thr59-H	2.15	2.08	—	—	—	—	1.82	2.19	2.07	—	—	2.07	2.22
B	Glu63-OE1	—	3.15	—	—	—	—	—		2.80	—	—	3.56	2.81
C	His88-O	2.72	3.88	2.73	—	—	—	2.68	—	3.93	2.84	—	2.75	2.82
D	Asn98-H	—	—	—	—	—	—	3.67	—	—	—	3.43	—	—
	Gly101-O	—	—	—	—	—	—	2.70	—	—	—	2.47	—	—
E	Arg104-O	3.14	—	3.19	3.34	3.03	—	3.59	3.11	3.45	3.10	—	—	3.14
	Thr106-H	3.19	—	3.16	3.10	3.64	—	3.20	3.22	3.31	3.51	—	—	3.17
F	Ser112-O	2.77	—	—	—	2.51	—	—	2.80	2.88	—	—	2.79	—
	Ser115-O	3.78	—	—	—	3.72	—	—	3.77	3.64	—	—	3.65	—

For TTR, each of its monomer is composed of eight anti-parallel β-strands (A–H) arranged as a β-barrel with a short α-helix connecting strands E and F.[16] Recent X-ray structures of wild- and mutant-types of the protein have revealed structural misfolding in its loop E-α-helix-loop F region, which indicates a possibility of this segment to participate in the generation of amyloid fibrils.[17,18] Actually, wild-type TTR proteins, in particular in the elderly, can aggregate themselves and cause non-familial cases of amyloidosis.[18] The simple and typical structure of TTR combined with its potential and sensitive role in the formation of amyloid fibrils reflect that this protein is a good choice as an appropriate model to study the relationship between hydration properties and protein misfolding. In this study, we have firstly performed 13 independent simulations to explore how the water surrounds the TTR, each with 30 ns. Additional simulations with the mutations of 1X7S and 3DO4 systems, including Leu58His and Ser112Ile respectively, are further performed to obtain another 2 trajectories.

To validate the model reliability of water behaviors in TTR, we compare the crystal water positions with the MD hydration sites. Comparison of all the water sites based on the two approaches shows that about 19% fall at the same sites, revealing significant differences between them. However, when compared to the MD hydration sites with a longer residence time (residence time (τ) >1.5 ns), the similarity shows a dramatic increase; such agreement between the X-ray crystallography and the MD simulations indicates that these are conserved waters. Tables 1 and 2 present the H-bond distances (Å) calculated between water oxygen and hydrogen and typical protein atoms that are determined in the MD trajectories and the X-ray structures. This finding indicates that for the dynamical, temporal hydration sites in the MD simulations, different solvent conditions (crystallography *vs.* MD) can alter the delicate equilibrium between the water positions and the surface backbones with a consequence of different hydration sites. However, for the more stable, longer lived hydration sites in MD, most of them (82%) have corresponding vitrified water molecules in the crystal structures, and minor changes in the solvent condition hardly perturb their positions.

From the solvent density maps, we observe a good consistency among the MD hydration sites in the simulations of the various TTR structures. For convenience, the sites are identified as local maxima in the density function, and are colored in the descending order of blue, cyan and gray depending on their maxima density value. As shown in Fig.1, the hydration shells in the 2QGB (wild type) and 3GRG (mutant type) simulations follow the shape of the protein, and also, the spatial distribution of hydration sites in them is both highly anisotropic and asymmetric. The high density regions (blue) where water molecules are tightly clustered are located at the loops of the protein, while the medium and low density regions where the waters distribute separately are mainly surrounding the β-sheets.

X-Ray crystallography has shown that the loops DE, AB and EF participate in the binding of the thyroid hormone[19] and the retinol-binding protein mediating the transport of vitamin A to the target tissues,[20] respectively, indicating that there are structural defects in these segments. The loci of these "defects" on the protein surface are associated with local destabilization and favoring of partially unfolded structures with a consequent potential for aggregation. In order to preserve the three loops, the waters are tightly bound to them, thus holding in place local elements of these folds as shown in Fig.1[21,22]

On the basis of these observations, in order to clarify the importance of hydration sites and their relationship with the amyloid fibril formation in solvent proteins, we choose two tightly bound waters with an average residence time of 3.5 ns in sites A and F to present a detailed analysis, respectively (Table 1). The water in site A interacts with the amides of Leu58 and Thr59 that are two vicinal residues at loop DE (Fig. 2 A and B). This loop partly bends as an arch where the orientations of the amides enhance dipole–dipole interactions with the water. This

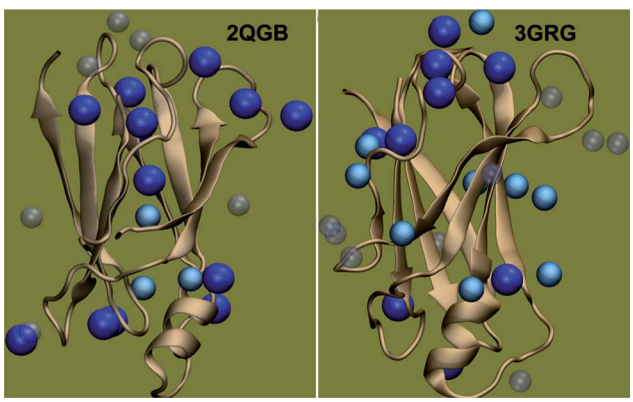

Fig. 1 Three-dimensional distribution of high-density solvent around the protein TTR (PDB codes: 2QGB and 3GRG). In the solvation map, hydration sites with low ($0.5 < \tau < 1$ ns), medium ($1 < \tau < 1.5$ ns) and long (τ>1.5 ns) residence times are colored gray, cyan and blue, respectively.

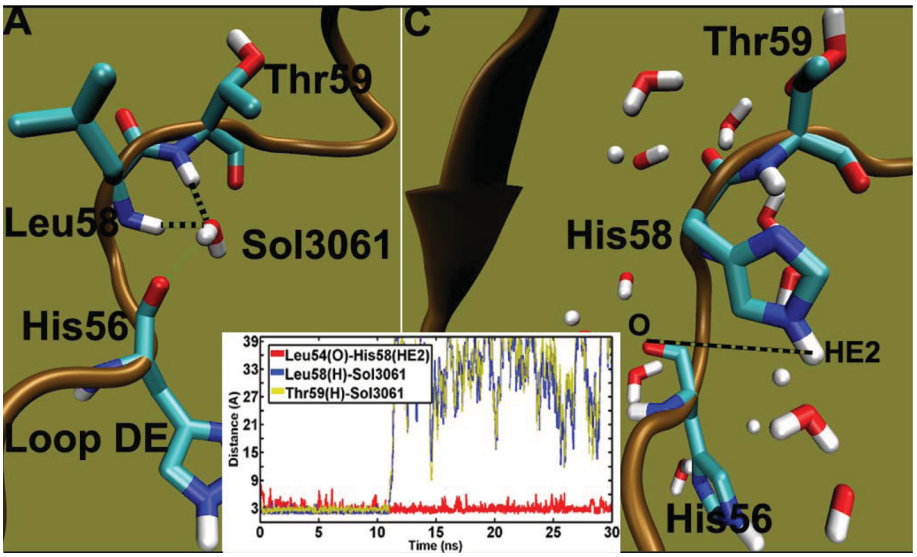

Fig. 2 The distribution model of hydration site A adjacent to the loop DE in the original and mutant 1X7S (Leu58His) systems. The loop DE is colored in brown. Residues Leu54, His56, Leu58, Leu58His and Thr59 are shown in stick presentation: cyan for carbon, white for hydrogen, blue for nitrogen and red for oxygen. (A) The close-up view of the hydration properties of site A. The red/white triangle represents the water molecule located at the hydration site. (B) Time evolutions of the distance between Leu58-H, Thr59-H and Sol 3061 in the wild 1X7S system, and the distance between Leu54-O and His58-HE2 in the mutant 1X7S system. (C) The waters in 3.5 Å radiuses centered at residues Leu58His and Thr59 in the mutant 1X7S system.

hydration site has an average residence time of 4.8 ns (Fig.2 B), and is conserved in the 2QGB, 3GRG, 2G4E, 1X7S, 1TSH, 1F86 and 1SOK systems. The calculated free energy of binding for this water molecule is -8.2 kJ mol^{-1}, providing an estimation of the local stabilization depending on the protein-water interactions.

On examination of the solvent dynamical behavior of site A, we perform an additional MD simulation by substituting Leu58 with His; and such mutation is tightly linked with amyloid diseases.[23] The mutation is carried out in Sybyl program that retains the identity of the original backbone conformations of wild and mutant residues. It is well known that histidine enables to bind a proton to the nonbonded electron pair of its imidazole ring nitrogen to become a weak acid at low pH. The pK of the acid is 6.0 so that at physiological pH histidine is about 90% in the basic form (http://class.fst.ohio-state.edu/fst605/lectures/Lect9.html). This indicates that, histidine is predominantly protonated on one imidazole nitrogen at neutral pH, and in such protonation state histidine is not charged. Thus, the system is performed with the presence of every histidine with one imidazole nitrogen protonated to guarantee normal physiological conditions. As can be seen from Fig. 2 B and C, His58 (HE2) forms a strong hydrogen bond with Leu54 (O) with distance of ~2.4 Å, causing the conformational changes involving the adjacent His56. The two histidines repel each other due to the steric effect of the imidazole ring at His58. The reason is that the bulky residue occupies a larger amount of space than Leu58 in the wild 1X7S system and when the imidazole rings of His56 and His58 are brought together, hindrance will be necessarily induced. This results in their backbone conformational changes with relative rotation angles (*N*-CA-C-O) of ~22° for His58 and ~44° for His56, as compared with those in the wild system. These conformational changes expel water molecules from site A, resulting that the original small arch is now filled by the larger iminazole ring of His58 (Fig. 2C). Moreover, the site A in the 2QGB, 3GRG, 2G4E, 1X7S, 1TSH, 1F86 and 1SOK MD systems shows that Leu58 interacts with Thr59 by the water bridge, leading to the bulging and bending of the initial segment of loop DE. However, after the replacement of Leu58 by His58, this loop region elongates and extends to an open state (~75°) and thus prevents the formation of a water-bridged H-bond between His58 and Thr59. This finally causes the rotation of Thr59 (~45°) and the displacement of site A. Clearly in this section, the conformational changes of loop DE induced by mutation govern the solvent dynamical behaviors, while taking into consideration the original backbone conformations of Leu58 and its mutant His58, and the subsequent conformational change of Leu58His. In addition, compared with the root-mean-square fluctuation (RMSF = 0.8 Å) of the two residues in the original 1X7S system, the corresponding residues in the mutant system has significantly larger RMSF of 6.0 Å, which definitely indicates that the site A facilitates to stabilize the local conformation of loop DE.

Site F has a relatively short residence time (2.1 ns) and resides in the 2QGB, 3DO4, 1X7S, 1TSH, and 1F86 systems. This site is located at the hydrophilic center formed by the strand between β-sheets H and G and enclosed by residues Ser112, Pro113, Tyr114 and Ser115 (Fig. 3 A), which undergoes significant conformational changes to assemble into aggregates.[24] During the simulations, the exposed site F causes fast solvent exchange with the bulk solution, and the loop HG that the water is located adjacent to has relatively high flexibility (RMSF = 23 Å) in all the structures. The calculations reveal a water-mediated interaction between the backbone oxygen of Ser112 and Ser115 (Fig.3 B), which implies that the loop mobility may be affected by the strength of the water-mediated H bond with a distance of 3.0–3.5 Å.

In fact, the Ser112Ile mutation of TTR (PDB code:3DO4) allows the loop to partially flatten due to the disappearance of H

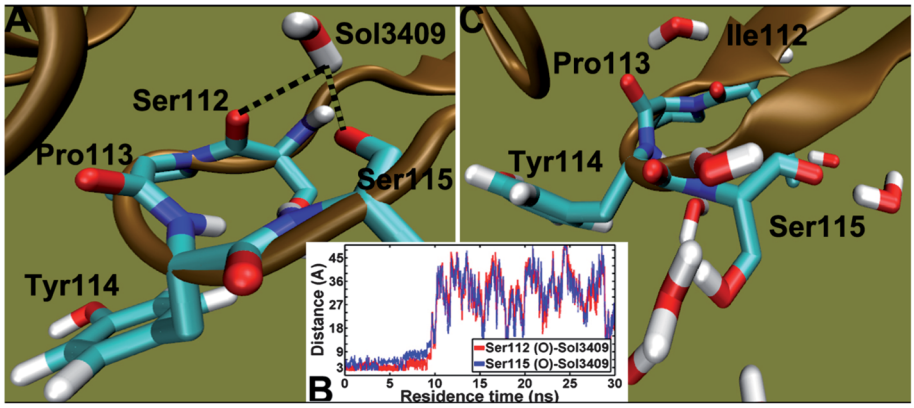

Fig. 3 The distribution model of hydration site F adjacent to loop HG in the wild and mutant 3DO4 (Ser112Ile) systems. For clarity, the loop HG is colored in brown. Residues Ser112, Ser112Ile, Pro113, Tyr114 and Ser115 are shown in stick presentation: cyan for carbon, white for hydrogen, blue for nitrogen and red for oxygen, respectively. (A) The close-up view of the hydration properties of site F. The red/white triangle represents the water molecule located at the hydration site. (B) Time evolutions of the distance between Ser112-O, Ser115-O and Sol 3409 in the wild 3DO4 system. (C) The waters in 3.5 Å radius centered at residues Ser112 and Ser115 in the mutant 3DO4 system.

bond between the site F water and Ser112, as confirmed by an increase of almost 0.5 Å in the distance between the backbones of Ile112 and Ser115. This finally results in the assembly of TTR as a spherical aggregate and the cytotoxicity to human neuroblastoma cell line.[25] Also, the conformational changes induce the spatial diffusion of water molecules adjacent to the loop HG that consequently represent a scattering state in the mutant system (Fig. 3 C). Such conformational change of loop HG and its altered solution environment suggests that the site F water is actually necessary for the local stabilization of loop HG. The calculated water-binding free energy of site F is -3.4 kJ mol^{-1}, which is much smaller than that for site A, implying that waters in site F are less stable at the surface of protein, consistent with their relatively short residence time of 2.1 ns.

2.2 Two types of mutant-dependent waters

Since in some cases the wild-type protein is present in the amyloid deposits, it is believed that there is an intrinsic propensity of the protein to form amyloid and this aspect can be enhanced by the presence of the mutations.[26] This observation brings us to the question whether local solvent dynamics are coupled to the protein mutations? How the mutant-dependent waters determine the folding and misfolding states of the protein? In this section, we find that the water molecules in the mutant systems can be divided into two groups: (1) "localizing" water; such water is located next to the mutation that can increase the local stability of the protein, thus preventing amyloid formation while preserving the native state, and (2) "delocalizing" water; such water surrounds the mutation that shifts structural equilibrium toward the mobile state, finally resulting in the transient conformations that are compatible with local aggregation and amyloid formation.

2.2.1 Localizing waters. For the various mutations of TTR identified to date, only several present non-amyloidogenic effects. Of these, a more restricted group is associated with a localizing effect against the development of the disease in heterozygotic carriers of an amyloidogenic variant.[27] Here, we have identified four "water localizing" mutations around TTR in three dynamical systems, *i.e.*, the 1SOK, 1X7T and 1F86 systems. Table 3 lists the H-bond distances calculated for the position of the "localizing" waters in the MD simulations corresponding to that of the waters in the X-ray structures. Since the discrepancies between the H-bond distances (<3.5 Å) of the two techniques are within the limits of allowed error, we conclude that our hydrations sites surrounding the "water localizing" mutations are sufficiently reliable.

To analyze the properties of the "localizing" mutation-related water behaviors, we compare this type of water observed in the three systems with that in the wild 2QGB system. In the simulation of wild type, the waters in 3.5 Å radius centered at residues Arg104, Ala108, Thr119 and Leu110 exchange fast with bulk solvent and display a partially disordered distribution (Fig. 4), which can be entropically advantageous.[1] The bulk-water dynamic exchange is strongly coupled with the local protein

Table 3 H-bond distances (Å) between water and "localizing" mutations in the MD simulations and the X-ray structures

		MD simulations		X-ray structures	
System	Mutations	Water numbers	H-bond (Å)	Water numbers	H-bond (Å)
1SOK	Ala108Tyr -O	Sol7483	3.30 ± 0.10	Sol142	3.70
	Leu110Glu -OE1	Sol7483	2.00 ± 0.15	Sol142	3.40
1X7T	Arg104His -HE2	Sol6075	2.10 ± 0.10	Sol6085	2.60
	Arg104His -ND1	Sol7475	2.00 ± 0.10	Sol6086	3.80
1F86	Thr119Met -O	Sol1499	2.10 ± 0.10	Sol463	3.70

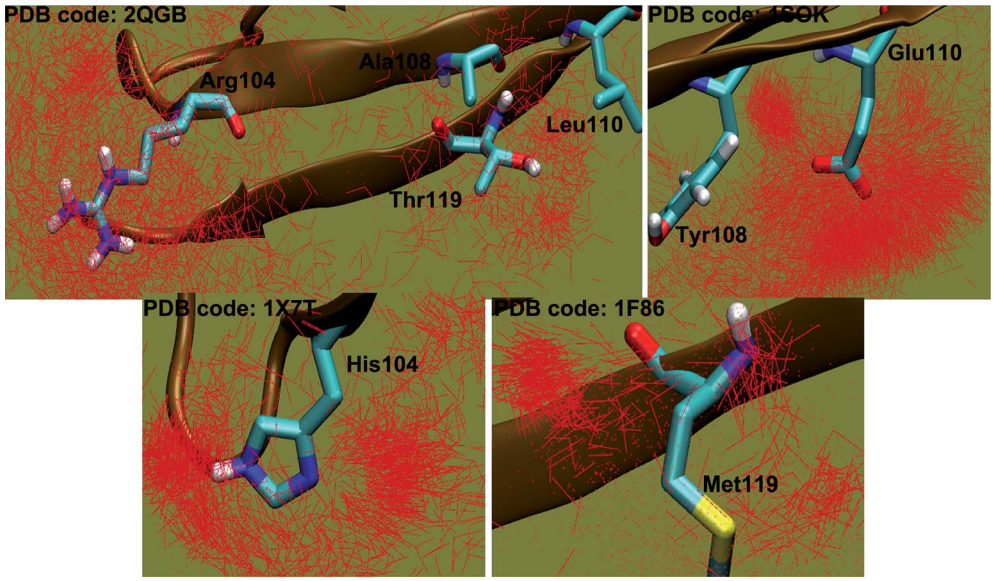

Fig. 4 Superposition of the current protein structure onto its initial structure for each snapshot in the 2QGB, 1SOK, 1X7T and 1F86 systems. For clarity, residues Arg104, Ala108, Leu110, Thr119, His104, Tyr108, Glu110 and Met119 are shown in stick presentation: cyan for carbon, white for hydrogen, blue for nitrogen and red for oxygen, respectively. The red lines represent the water molecules adjacent to the residues. For the wild 2QGB system, the water molecules diffuse spatially and represent a scattering state. However, for the remaining three mutant systems, the water molecules interact directly with the residue backbone and side chains and form clusters of water molecules.

fluctuations, with RMSF of 7.3 Å (Arg104), 4.3 Å (Ala108 and Leu110) and 5.6 Å (Thr119), respectively. On the contrary, in the 1SOK, 1X7T and 1F86 systems, waters surrounding the corresponding mutations interact directly with the protein backbone and side chains, and even form one or two clusters of water molecules around the residues (Table 3 and Fig. 4). Due to role of "localizing" water in protein stabilization, the mutations in the "water localizing" systems undergo relatively slight structural arrangement, with a significant decrease in the RMSF (0.6 Å for His104, 0.9 Å for Tyr108 and Glu110 and 1.3 Å for Met119) against that of the 2QGB system. Also, the mean residence time of the fixed water molecules in the mutant systems is much longer than that of the fast exchanged water in the wild system (2 ns for the former and ~20 ps for the latter). These results evidently indicate that these "localizing" waters, as significant contributors to the folding of TTR, give unique kinetic stability to the local conformations of protein by hydrogen-bonding with their neighboring residues. The dynamical features of the localizing waters are described in detail below.

Fig. 5 shows a series of surface maps around Tyr108 and Glu110 with the local surface topography in the 1SOK system. We find that within 3.5 Å from the H atom of water to the residues Tyr108 and Glu110 (including H), there are ~13 water molecules around them. At 7.4 ns, the water molecule (Sol) 8255 bridges the side chain of Glu110 (OE2-H distance of 2.5 Å) and the backbone oxygen atom of Tyr108 (O–H distance of 2.4 Å), by two H-bonds, and has a residence time of 100 ps. Accompanied by the movement of Tyr108 and Glu110, the water network makes small structural arrangements, and at 7.85 ns the localizing water Sol 7483 occupies this site to form the same H-bonds as Sol 8255. Subsequently, an additional H-bond between the amide of Glu110 and Sol 7483 is involved in the hydration site at 7.95 ns, revealing the local orientational and slight translational motions of the hydration water and the residues, and their intimate coupling. Ultimately, Sol 7483 dynamically exchanges with bulk water in tens of picoseconds, and the new Sol5999 immediately replaces this water at 8.2 ns. During this period, we do not observe significant conformational changes of the two mutants

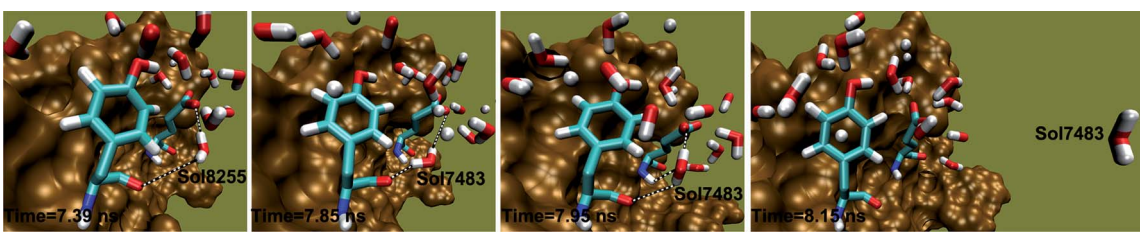

Fig. 5 Snapshots of water motions with the MD trajectory of 1SOK. Depicted here are residues Tyr108 and Glu110, and water molecules within 3.5 Å from the H atom of water to them. The residues are shown in stick presentation (cyan for backbone, red for oxygen, blue for nitrogen and white for hydrogen, respectively); the water molecules in red triangles; and the protein TTR in brown surface. At time 7.39 ns, Sol 8255 from the bulk water interacts with the side chain of Glu110 and the backbone oxygen atom of Tyr108. Subsequently at 7.85 ns, Sol 7483 occupies this position to form the same H-bonds as Sol 8255. Next, an additional H-bond between this water with the amide of Glu110 is formed at 7.95 ns. Finally at 8.15 ns, Sol 7483 dynamically exchanges with bulk water.

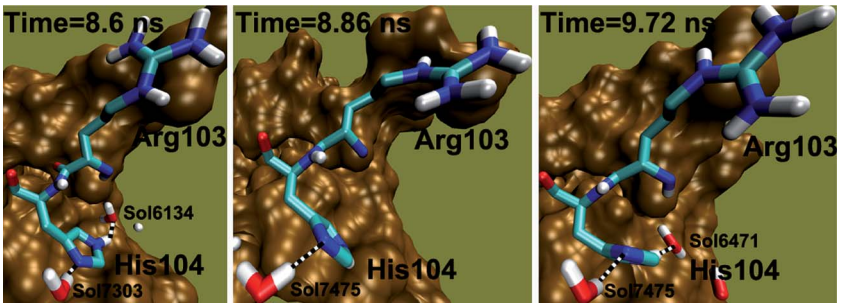

Fig. 6 Snapshots of water motions around the residue His104 with the MD trajectory of 1X7T. Arg103 and His104 are shown in stick presentation (cyan for backbone, red for oxygen, blue for nitrogen and white for hydrogen, respectively). The protein TTR is shown in brown surface. The water molecules within 3.5 Å from His104 are shown in triangle for clarity. At time 8.6 ns, Sol 7303 and Sol 6134 from the bulk water interact with the His104 ND1 and HE2 atoms, respectively. At 8.86 ns, the new Sol 7475 occupies the position to displace Sol 7303. For the water molecules interacting with the HE2 atom of His104, they are very labile, constantly in fast exchange with bulk water, and finally evolve to a new equilibrated state at 9.72 ns.

and their neighboring structure (RMSD = ~4 Å), but find that the binding affinity of Sol 7483 to the protein is ~ −17.35 kJ mol^{-1}, much stronger than those of the tightly bound waters described in section 2.1, which implies the important contribution of localizing waters on maintaining the structure of protein.

For the 1X7T system, the replacement of Arg104 with His has immediate effects of forming two hydration sites flanking the imidazole sidechain of this residue. As shown in Fig. 6, at 8.6 ns, two H-bonds are formed between the His104 imidazole nitrogen (ND1) with the Sol 7303, and the HE2 atom with the Sol 6134. The presence of His104 also causes the conformational change involving its neighboring Arg103. The two highly hydrophilic residues repel each other, successfully fixing the Arg103 side chain in the negative direction (RMSF = 0.5 Å). When Sol 7303 displaced at 8.86 ns, this segment can slightly relax, and the H-bond between the imidazole ND1 of His104 and the new Sol 7475 can be formed. This H-bond remains stable more than 1 ns, and has the effect of stabilizing the β-sheet G (RMSF = 1.7 Å). While for the water molecules interacting with the HE2 atom of His104, they are more labile, constantly in fast exchange with bulk water, and finally evolve from the initial nonequilibrium configuration to a new equilibrated state at 9.72 ns. The calculated free energy of binding for the water molecules Sol 7475 and Sol 6471 is ~ −25.25 kJ mol^{-1}. The above results confirm again the critical role of the "localizing waters" surrounding the Arg104His mutation, and reveal that the observed variations in hydration dynamics around the mutation reflect the alternation of the local landscape.

As for the 1F86 system, the observed water dynamics during 1.2 ns represent collective motions, reflecting a dynamic change of the local water molecules. Indeed, the localizing water surrounding Met119 is very labile, in fast exchange with the bulk water, and the residence time only lasts 40 ~ 100 ps. For example, at 10.5 ns, Sol 8881 interacts with the carbonyl of Met119. Subsequently at 10.6 ns, the new Sol 1844 enters the hydration site instead of Sol 8881 and stays for 40 ps (Fig. 7). Despite the fast water change around Met119, this residue still exhibits a relatively low RMSF value (1.3 Å) during the simulation, and meanwhile, the calculated free energy of binding for Sol 8881 is ~ −21.27 kJ mol^{-1}, 2-fold higher than those of tightly bound waters in section 2.1. This result indicates that the protein-water interaction enables the neighboring residue to maintain stable conformation. This also validates the importance of water localizing mutations, and their closely association with nearly hydration sites.

2.2.2 Delocalizing waters and entropy. To date, over 80 amyloidogenic TTR water mutations have been identified. Of these, V30M, L58H, T60A, I84S and V122I are responsible for the majority of familial amyloidotic polyneuropathy and familial amyloidotic cardiomiopathy cases.[23] A comparative analysis of the crystal structures of several amyloidogenic TTR has shown that only minor structural changes for the mutant forms have been observed compared to the wild-type protein, suggesting that the structure of mutant TTR is not the sole factor determining the close association between mutations with their amyloidogenic potential.[28] This raises a question: are there any well-defined and relatively stable waters around the neighborhood of these mutation regions, *i.e.*, the middle β-sheet G region in the 1SOK system (residues Tyr108-Leu111), the initial β-sheet G segment in the 1X7T system (residues Arg103-Thr106), the initial loop DE region in the 3DO4 system (residues Leu58-Glu61) and

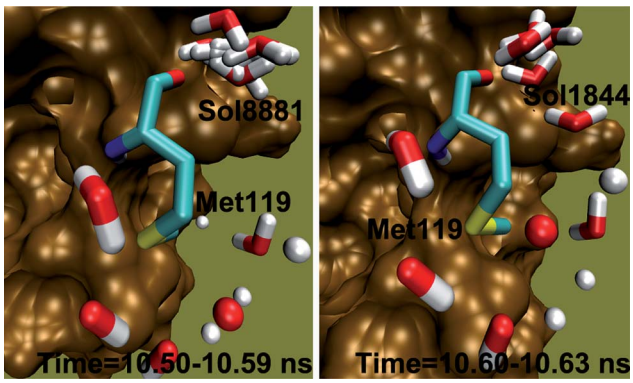

Fig. 7 Snapshots of water motions around the residue Met119 with the MD trajectory of 1F86. This residue is shown in stick presentation (cyan for backbone, red for oxygen, blue for nitrogen and white for hydrogen, respectively). The protein TTR is shown in brown surface. The water molecules within 3.5 Å from Met119 are shown in triangle for clarity. At time 10.5 ns, Sol 8881 interacts with the carbonyl of Met119. Subsequently at 10.6 ns, the new Sol 1844 enters the hydration site instead of Sol 8881.

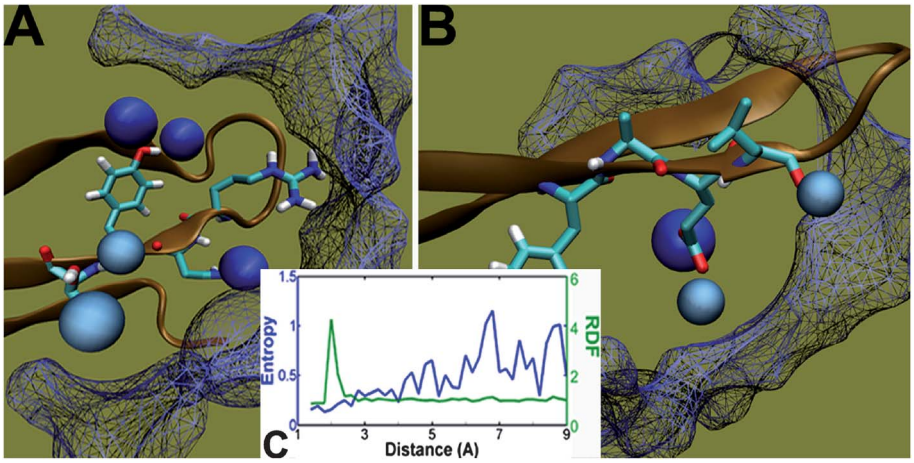

Fig. 8 The entropy-based water distribution maps of high-density regions where the "localizing" mutations (His104, Tyr108 and Glu110) are located. The blue grids represent the high entropy regions, and the blue and cyan balls are the hydration sites around the protein surface. The β-sheet G is colored in brown, and the residues are shown in stick representation: cyan for carbon, red for oxygen, white for hydrogen, and blue for nitrogen, respectively. (A) The water entropy-distribution map around residues Arg103-Thr106 in the 1X7T system (B) The water entropy-distribution map around residues Tyr108-Leu111 in the 1SOK system (C) shows the double coordinate system of entropy *vs.* distance and radial distribution function (RDF) *vs.* distance. The entropy (blue) represents several humps, but a relatively smooth and gradual increase. And the RDF (green) shows a very large hump.

the loop EF in the 2G4E system (residues Leu82-Ser85)? Unexpectedly, we did not find any particularly localized hydration site in these regions. We have therefore mapped the solvent entropy for elucidating the thermodynamic properties of "delocalizing waters", and distinguishing them from the "localizing waters".

Fig. 8 and 9 show the water-density contour plots for the four local segments of TTR. These purple contours represent entropy distribution for the water molecules. By focusing on the two high-density regions where the "water localizing" mutations (His104, Tyr108 and Glu110) are located (Fig. 8 A and B), we find that the entropy becomes weaker with decreasing distance from the protein surface, and even form an empty isolation barrier in the vicinity of the protein. While for the low-density segments bearing the water delocalizing mutations (Ala60 and Ala84) (Fig. 9 A and B), a different distribution of entropy is detected; such entropy disperses the two regions uniformly, even in close proximity to the surface of protein. In thermodynamics, entropy is an extensive state function that accounts for the effects of irreversibility in thermodynamic systems. As a measure of randomness or disorder, it is only related with the initial and final state of the system, regardless of the process by which the changes take place (http://www.science.uwaterloo.ca/~cchieh/cact/applychem/entropy.html). In this

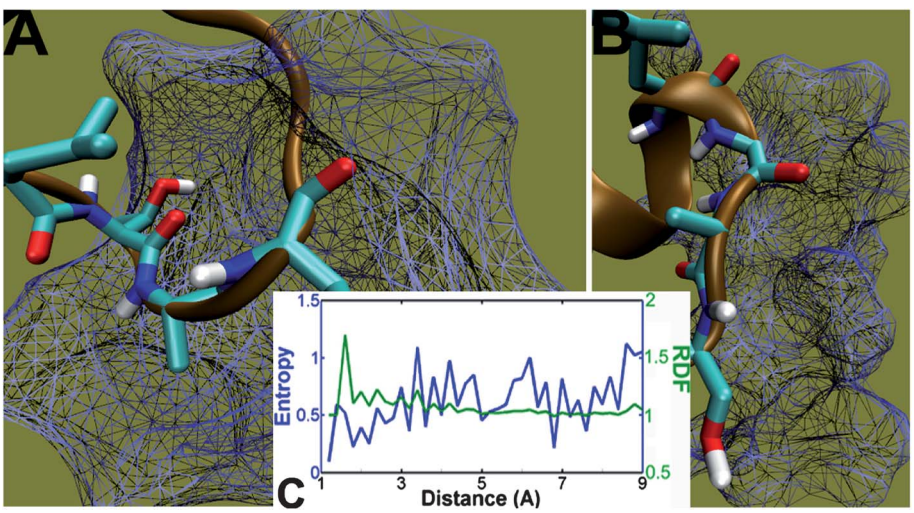

Fig. 9 The entropy-based water distribution maps of low-density segments bearing the "devastated" mutations (Ala60 and Ala84). The blue grids represent the low entropy regions, and the blue and cyan balls are the hydration sites around the protein surface. The loops DE and EF are colored in brown, and the residues are shown in stick representation: cyan for carbon, red for oxygen, white for hydrogen, and blue for nitrogen, respectively. (A) The water entropy-distribution map around residues Leu58-Glu61 in the 3DO4 system (B) The water entropy-distribution map around residues Leu82-Ser85 in the 2G4E system (C) shows the double coordinate system of entropy *vs.* distance and radial distribution function (RDF) *vs.* distance. Both the entropy and the radial distribution function curves present a steep increase at the beginning.

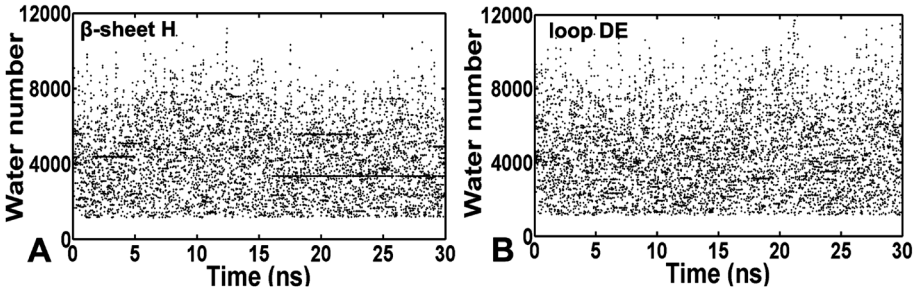

Fig. 10 The time evolutions of waters interacting with the specific H bonds of Met119 (a localizing mutation)-Val121 in β-sheet H (A) and Pro55 (a devastated mutation)-Gly57 in loop DE (B), respectively.

study, the distribution of entropy successfully reveals the stability of the tightly bound waters and the transferred energy of these waters with the bulk solvent, that is, the more stable the "localizing" waters, the weaker the entropy. Similarly, the motions of mobile water molecules are also presented by the entropy distribution; the "delocalizing" waters are in faster, more disordered exchange with the bulk solvent, the entropy shows a more progressive potentiation and a more dispersed distribution.

In order to quantitatively evaluate the entropy surrounding the protein and to evaluate its correlation with the water density, we have measured the relative variation of entropy and radial distribution function (RDF) with change in distance in the 1X7T and 3DO4 systems. As shown in Fig. 8 C, the entropy around the high density region rises upward gradually and smoothly with the increasing distance from the protein surface, and becomes approximately one for all positions at ~6 Å. This result reveals

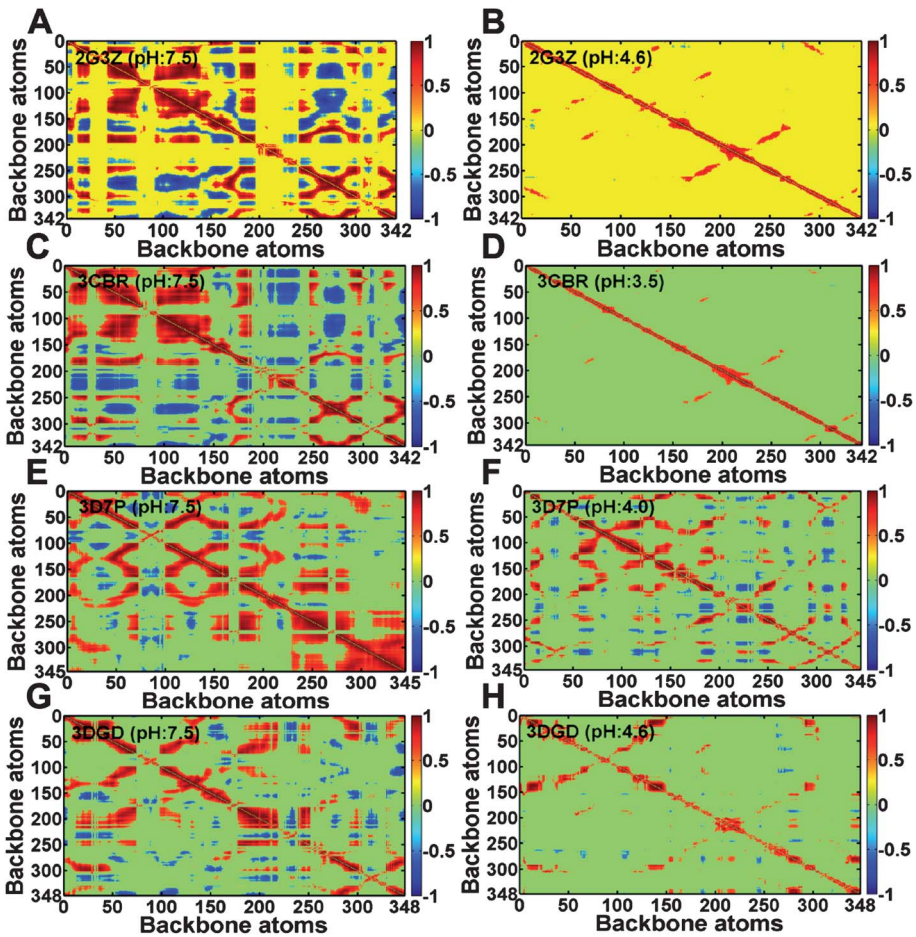

Fig. 11 Dynamical cross-correlation maps illustrating the correlation of motion between residues in (A) the 2G3Z system (at pH 7.0); (B) the 2G3Z_4.6 system (at pH 4.6); (C) the 3CBR system (at pH 7.0); (D) the 3CBR_3.5 system (at pH 3.5); (E) the 3D7P system (at pH 7.0); (F) the 3D7P_4.0 system (at pH 4.0); (G) the 3DGD system (at pH 7.0) and (H) the 3DGD_4.6 system (at pH 4.6). The color bars on the right indicate the extent of the correlation. Residue pairs with a high level of correlated motions are shown in yellow, orange, and red. Anticorrelated motions are represented by the blue and cyan regions.

the cluster and aggregation state of "localizing" water, and the faster exchange with the bulk solvent with the increase in distance. While for the low density region (Fig. 9 C), the entropy exhibits a steep increase, and immediately achieves the thermal equilibrium at ~3.5 Å, which indicates the disordered motions of "delocalizing" waters from the protein surface. As for the RDF, it is defined to be a local measure of how closely a particle distribution resembles a uniform one. For the high density region (Fig. 8 C), the RDF plot exhibits a large peak at ~2 Å, with RDF of about 4. This means that it is four times more likely that two water molecules would be positioned at this separation, revealing the greatest probability relative to a uniform distribution of waters from the protein surface, which is highly consistent with the results of our entropy-distribution map in Fig. 8 A and B. Subsequently, the RDF falls and passes through a minimum value around 3 Å, indicating that the chances of finding four water molecules with this separation are less. Conversely, the RDF in the low density regions (Fig. 9 C) shows an average value of about one with no obvious humps. The density oscillations reflect a suppressed probability relative to a uniform distribution of entropy from the protein surface, which implies the scattered state of mobile waters around the protein.

To further verify the stable state of "localizing" waters and the mobile property of "delocalizing" waters, we examine the time evolutions of waters interacting with the specific H bonds of Met119 (a localizing mutation)-Val121 in β-sheet H and Pro55 (a delocalizing mutation)-Gly57 in loop DE. Fig. 10 A shows sparsely distributed waters surrounding the "water localizing" surface of protein, which indicates relatively long residence times of the waters, and their interactions with the H-bond Met119-Val121. However, in the water delocalizing surfaces the interactions last, only for a temporal and medium time (Fig. 10 B). This reveals a fast exchange of "delocalizing" waters, which resembles the dynamical behaviors of the bulk solvent. On the basis of entropy and solvent-backbone H-bond exchange analysis for the "localizing" and "delocalizing" regions, we therefore conclude that the specific hydration profiles obtained in the calculations are a property of the protein local sequence and conformation, and the fluctuation and the structural properties of mutations are able to perturb their neighboring water behaviors.

2.3 The localizing water inactivated by acidic pH

For dynamical simulations, one of its particular advantage is the possibility of performing unnatural "experiments" that are not possible in an experimental setting. *In vitro* experiments have demonstrated that an acidic medium facilitates dissociation and conformational changes in TTR, allowing alternatively folded monomers to self-assemble into insoluble amyloid fibers.[29] However, it is still not clear what the role of acidic pH value plays in the dynamical behavior of solvent on the protein surface?

In this section, four groups of MD simulations targeting the dynamics of TTR at central and acidic pH are performed: the 2G3Z system (at pH 7.0) *vs.* the 2G3Z_4.6 system (at pH 4.6), the 3CBR system (at pH 7.0) *vs.* the 3CBR_3.5 system (at pH 3.5), the 3D7P system (at pH 7.0) *vs.* the 3D7P_4.0 system (at pH 4.0) and the 3DGD system (at pH 7.0) *vs.* the 3DGD_4.6 system (at pH 4.6). As shown in Fig. 11, dynamical cross-correlation maps (normalized covariance matrix) are constructed for the central and acidic pH trajectories to illustrate the correlation between the motions of residues. All eight maps show correlations that range from highly anticorrelated (blue) to highly correlated (red). In the cases of the pH 7.0 simulations (Fig. 11 A, C, G and E), a much greater degree of correlated motion is observed, as evidenced by the large number of red or blue patches on the surface. This is in contrast to the acidic pH maps (Fig. 11 B, D, F and H) that reveal a significant dampening of the highly correlated motions that are observed in the central pH simulations. This result indicates the significantly conformational change and enhanced stability of TTR protein in the acidic condition, which is further evidenced by its much lower RMSF values (~2.1 Å) compared with those in the central systems (~13.5 Å). Indeed, this is reasonable since the amyloid fibrils refer to elongated protein aggregates characterized by their extensively hydrogen-bonded β-sheet core structures; such conformations confer to them remarkable stability and resistance to protease cleavage.[30]

Further analysis for the time evolutions of solvent exchange with the whole protein within a distance of 3.5 Å for the 2G3Z and the 2G3Z_4.6 systems reveals that the exchange frequency of solvent with bulk water increases as the pH is lowered. As shown in Fig. 12 A, the protein TTR in the central system is surrounded by hydration sites with relatively long residence time. A large number of waters are found to directly form H-bonds with the protein. On the contrary in the acidic medium, the scattered plot shows a faster water exchange, and the interactions between the water and the protein, on average, last for short and medium time (Fig. 12 B). This indicates that although the localizing waters on the protein surface have a great effect on the conformational stability of TTR, the acidic pH is able to destroy the cluster structure of water molecules, inducing water diffusion and protein conformational transformation.[29] These ensembles thus provide a picture of how the dynamic properties of the water changes in response to pH variations.

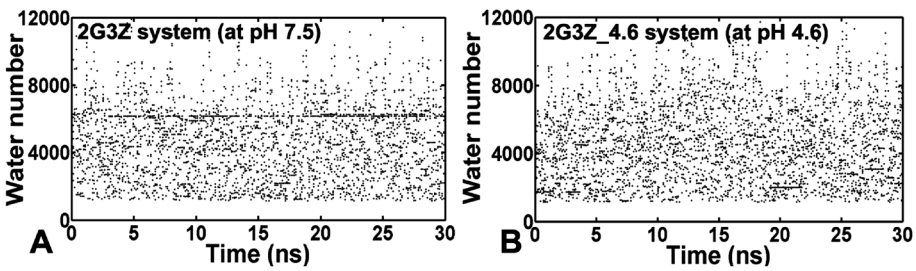

Fig. 12 The time evolutions of waters interacting with the TTR protein in the 2G3Z system (at pH 7.0) (A) and the 2G3Z_4.6 system (at pH 4.6) (B), respectively.

3 Conclusions

Water is essential for protein structure, not only with regard to defining the collapse of the hydrophobic core, but also in maintaining its stability and structure. For the amyloidogenic protein TTR, it possesses an inherent potential, albeit low, to generate amyloid fibrils and its mutations and the acidic condition can enhance the amyloidogenicity. In this study, we have studied the TTR hydration sites with MD simulations, to emphatically analyze the water behavior around the "water localizing" and the "water delocalizing" mutations of TTR, and to compare the solvent dynamics in the neutral and acidic medium, which presented enough data to address the role of the structurally conserved waters and the organization of the water at "unprotected" H-bond sites.

By substituting a typical residue with one closely associated with amyloid diseases, we observe the conformational changes of the mutation and its neighboring residues that prevent the restructuring of the hydration site adjacent to the residue; such site in turn can not provide stabilization for the structural elements in crucial regions of fold, revealing the close relationship between the solvent behavior and the structural conformation of protein. Further analysis identifies two types of water behavior with respect to the mutations resistant to highly denaturing conditions and the mutations promoting the destabilization of protein respectively, which are elucidated by comparisons of the local configurational solvent entropy and the water exchange with respect to the protein backbone H-bonds. Also, we find that acidic pH appears to be a dominant factor in modify the water motions surrounding the protein surface. Under acidic condition, the cross-correlation maps represent significantly weak coupling between the motions of the subunits of TTR, and the scattered plots with regard to the waters interacting with the backbone H-bond show relatively short residence time for the protein, both of which indicate that the water causes TTR to harden at acidic pH, likely affecting the dynamical transition and unfold of this protein that consequently leads to the amyloidogenic state. The results and concepts presented here thus provide an important basis for the water dynamics surrounding the amyloidogenic protein, and the mechanism of initiating the process of fibril formation, which provides a clue for the treatment of amyloid diseases associated with TTR misfolding.

4 Methods and materials

4.1 Simulation systems

In this study, we focused on the solvent behavior surrounding the TTR surface. The available 17 X-ray structures of the molecular monomer of the protein (PDB codes: 3GRG,[31] 3A4F,[32] 3DJZ,[33] 3DO4, 2QEL,[34] 2G4E,[17] 1X7S,[27] 1TSH,[35] 1G1O,[36] 2QGB,[37] 1X7T,[27] 1F86,[26] 1SOK,[38] 2G3Z, 3CBR,[18] 3D7P[18] and 3DGD) were taken from the protein data bank. For the 1X7S and 3DO4 systems, their leucine 58 and serine 112 were mutated to histidine and isoleucine in Sybyl 6.9 (Tripos, Inc., St. Louis, MO), respectively, in order to further evaluate the dynamical behavior of tightly bound water molecules. The structures were solvated in cubic boxes with box vectors of ~10 Å length, with an ionic concentration of ~0.15 M. The systems were energy minimized, followed by relaxation for 200 ps, with positional restraints on the protein heavy atoms by using a force constant of k = 1000 kJ mol^{-1} nm^{-2}. All simulations were carried out with the MD software packages GROMACS 4.5.1.[39] The Gromos 43A1 force field[40] was used for the protein and SPC/E for the water model.[41] All simulations were performed in the NPT ensemble. The temperature was kept constant by Berendsen coupling at $T = 310$ K, with a coupling time of $T_p = 0.1$ ps.[42] The pressure was coupled to a Berendsen barostat with $T_p = 0.5$ ps and an isotropic compressibility of 4.5×10^{-5} bar^{-1} in the x, y, and z directions.[44] All bonds were constrained with the LINCS algorithm.[43] Electrostatic interactions were calculated explicitly at a distance smaller than 14 Å; long-range electrostatic interactions were calculated by particle-mesh Ewald summation, with a grid spacing of 0.12 nm and fourth order B-spline interpolation.[44] Structures were written out every 10 ps for subsequent analysis. Consequently, a series of twenty three 30 ns MD simulations were generated for the TTR protein.

4.2 Water density analysis

To investigate the dynamic behaviors of water around the TTR protein, a water network based on isotropic harmonic and anisotropic harmonic models was reconstructed, by using occupancy weights and temperature parameters characterizing each hydration site at the protein -solvent interface.[45,46] For each frame of the MD trajectory, the solute was fitted to a reference solute structure, and immediately, the whole atom coordinates were transformed accordingly with the solute. Then, a three-dimensional grid with a mesh size of 1 Å was generated to calculated volume normalized probability distributions for the protein and the water oxygen atoms, respectively. Next, a smaller spherical volume of 0.5 Å radius centered at each grid point was applied to probe the positions of water molecules, producing an average three-dimensional number density distribution. The partial water volume occupancy ($<\rho_{water}^{1}(r)>_{time}$), the protein volume occupancy ($<\rho_{prot}^{1}(r)>_{time}$), and the total occupancy equivalent of the real space density maps ($\rho_{tot}^{1}(r)$) were:

$$\left\langle \rho_{water}^{1}(r) \right\rangle \text{time} = \frac{1}{\Delta\tau} \sum_{it=1}^{Nt} \sum_{iw=1}^{Nw} \Delta v \quad (1)$$

$$\left\langle \rho_{prot}^{1}(r) \right\rangle \text{time} = \frac{1}{\Delta\tau} \sum_{it=1}^{Nt} iw\Delta v \quad (2)$$

$$\rho_{tot}^{1}(r) = \rho_{water}^{1}(r) + \rho_{prot}^{1}(r) \quad (3)$$

where $\Delta\tau$, i_t, i_w, N_t, and N_w, are the volume element, the time increment, the water molecules/protein indices, the total number of conformations included in the average, and the total number of water molecules.

4.3 Calculation of the binding free energy

The binding free energy of water molecules buried in cavity of protein is calculated by the double decoupling method.[44] Two steps are involved: 1) calculation of the free energy for decoupling the water molecule of interest from the binding site of TTR ($\triangle G_1^\circ$); 2) calculation of the free energy cost of decoupling a water molecule from bulk solvent to gas phase ($\triangle G_2^\circ$). As shown in the following formulas, reaction [4] is the

"annihilation" of the water molecule B from the solvated TTR-water complex AB. Reaction [5] is the "annihilation" of the water from solvent. The free energy change of the two reactions, $\triangle G_1^\circ$ and $\triangle G_2^\circ$ yields the absolute free energy of binding, $\triangle G_{AB}^\circ$:

$$AB(sol) \rightarrow A(sol) + B(gas)\ \Delta G_1^\circ \quad (4)$$

$$B(sol) \rightarrow B(gas)\ \Delta G_2^\circ \quad (5)$$

$$A(sol) + B(sol) \rightarrow AB(sol)\ \Delta G_{AB}^\circ = \Delta G_2^\circ - \Delta G_1^\circ \quad (6)$$

$\triangle G_2^\circ$ is independent of the standard concentration. The calculation of $\triangle G_1^\circ$ is straightforward and can be computed by thermodynamic integration as shown in eqn (7) or by a free energy perturbation approach.

$$\Delta G_{AB}^\circ = \int_{\lambda_B}^{\lambda_A} \frac{dG}{d\lambda} d\lambda \quad (7)$$

MD simulations were performed in parallel for all steps along the decoupling pathway in this study with GROMACS 4.5.1 package using all parameters in the above section.

4.4 Measurement of water residence time

The residence time was calculated on the basis of the closest water molecules ($\leq$5) located within a shell around any particular protein atom. Here, we defined the time-autocorrelation function $P(\tau)$, which took the value 1 if the water molecule i stayed at both times t and $t + \tau$ within the first coordination shell. Otherwise, it took the value 0,[42] which was shown as follows:

$$P(\tau) = \sum_{i=1}^{N} Ci(t)Ci(t+\tau) \quad (8)$$

An exponential model is applied further to fit the resulting time-autocorrelation function.

4.5 Solvent entropy

In order to understand how the spatial distribution of the solvent around the protein affects the state of hydration sites, the solvent entropy map was calculated using the Shannon entropy. The solvent entropy map was computed in 0.1-nm-spaced circle, each boundary of which was connected to a narrower ring of 0.02-nm mesh size, which takes into count of the water distribution around the ring. The solvent entropy for a generic node of the map is calculated as follows:

$$S = -R \sum_{j,k,l} P_{j,k,l} \mathrm{In}\ P_{j,k,l} \quad (9)$$

where $P_{j,k,l}$ is the probability of finding a water in the subvolume j, k, l of the space surrounding the node, and R is the gas constant.

4.6 Cross-correlation analysis

To gain insight into the pH-induced conformational changes within the TTR protein, we investigated the correlation between the motions of residues in the 2G3Z system, the 2G3Z_4.6 system, the 3CBR system, the 3CBR_3.5 system, the 3D7P system, the 3D7P_4.0 system, the 3DGD system and the 3DGD_4.6 system, and selected the last 20 ns of the production run.

The cross-correlation coefficient C_{ij}, between atoms i and j, was a measure of the correlated nature of their atomic fluctuations and was computed as follows $C_{ij} = \langle \Delta ri \times \Delta rj \rangle / (\langle \Delta ri \times \Delta ri \rangle \langle \Delta rj \times \Delta rj \rangle)^{1/2}$ where Δr_i and Δr_j are the displacement vectors for atoms i and j, respectively. The angle brackets denote an average over the trajectory. The value of C_{ij} ranges from -1 to 1 with the correlated (positive) residue pair moving in the same direction, and the anticorrelated (negative) pair moving in the opposite direction.[47]

4.7 pH perturbation

Initial coordinates for the central and acidic systems were taken from the available X-ray crystal structures of mutant-type (PDB code: 2G3Z) and wild-type (PDB codes: 3CBR, 3D7P and 3DGD) TTR proteins. In order to represent the pH-induced conformational changes of TTR, we firstly estimated the pK_a values of all the residues within the protein for the four crystal structures respectively (http://biophysics.cs.vt.edu/H++/), and then, assigned the standard protonated states for pH 4.6 (2G3Z and 3DGD), 3.5 (3CBR) and 4.0 (3D7P) and the deprotonated states for pH 7 to their acidic side chains. All 8 simulations were subsequently carried out with GROMACS 4.5.1 using the Gromos 43A1 force field.

Acknowledgements

The research is supported by high-performance computing platform of Northwest A & F University, the Fund of Northwest A & F University and is financially supported by the National Natural Science Foundation of China (Grant No. 10801025). We are also grateful to Prof. L Yang for access of Sybyl software.

References

1 Y. Levy and J. N. Onuchic, Water Mediation in Protein Folding and Molecular recognition, *Biophysics*, 2006, **35**, 389–415.
2 J. L. Silva, T. C. R. G. Vieira, M. P. B. Gomes, A. P. A. Bom, L. M. T. R. Lima, M. S. Freitas, D. Ishimaru, Y. Cordeiro and D. Foguel, Ligand binding and hydration in protein misfolding: insights from studies of prion and p53 tumor suppressor proteins, *Acc. Chem. Res.*, 2010, **43**, 271–279.
3 C. M. Dobson, Protein folding and misfolding, *Nature*, 2003, **426**, 884–890.
4 C. M. Dobson, The structural basis of protein folding and its links with human disease, *Philos. Trans. R. Soc. London, Ser. B*, 2001, **356**, 133–145.
5 F. Chiti and C. M. Dobson, Protein misfolding, functional amyloid, and human disease, *Annu. Rev. Biochem.*, 2006, **75**, 333–366.
6 A De Simone, G. G. Dodson, C. S. Verma, A. Zagari and F. Fraternali, Prion and water: Tight and dynamical hydration sites have a key role in structural stability, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 7535–7540.
7 A. Fernández, J. Kardos, L. R. Scott, Y. Goto and R. S. Berry, Structural defects and the diagnosis of amyloidogenic propensity, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6446–6451.
8 J. R. Rodrigues, C. J. V. Simões, C. G. Silva and R. M. M. Brito, Potentially amyloidogenic conformational intermediates populate the unfolding landscape of transthyretin: insights from molecular dynamics simulations, *Protein Sci.*, 2010, **19**, 202–219.
9 A. D. Ferrão-Gonzales, S. O. Souto, S. J. L. Silva and D. Foguel, The preaggregated state of an amyloidogenic protein: hydrostatic pressure

converts native transthyretin into the amyloidogenic state, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 6445–6450.
10 X. Xu, Z. Ma, X. Wang, Z. T. Xiao, Y. Li, Z. H. Xue, S. Berry and Y. H. Wang, Water's role in protein misfolding and aggregation: insights from studies of p53 tumor suppressor protein, *Unpublished paper*, 2010.
11 B. Halle and V. P. Denisov, Magnetic relaxation dispersion studies of biomolecular solutions, *Methods Enzymol.*, 2002, **338**, 178–201.
12 L. Zhang, Y. Yang, Y. T. Kao, L. Wang and D. Zhong, Protein hydration dynamics and molecular mechanism of coupled water-protein fluctuations, *J. Am. Chem. Soc.*, 2009, **131**, 10677–10691.
13 S. E. Kolstoe, P. P. Mangione, V. Bellotti and G. W. Taylor, Trapping of palindromic ligands within native transthyretin prevents amyloid formation, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **23**, 20483–20488.
14 S. M. Johnson, H. M. Petrassi, S. K. Palaninathan, N. N. Mohamedmohaideen, H. E. Purkey, C. Nichols, K. P. Chiang, T. Walkup, J. C. Sacchettini, K. B. Sharpless and J. W. Kelly, Bisaryloxime ethers as potent inhibitors of transthyretin amyloid fibril formation, *J. Med. Chem.*, 2005, **48**, 1576–1587.
15 I. Daidone, M. B. Ulmschneider, N. A. Di, A. Amadei and J. C. Smith, Dehydration-driven solvent exposure of hydrophobic surfaces as a driving force in peptide folding, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 15230–15235.
16 J. A. Hamiltona and M. D. Benson, Transthyretin: a review from a structural perspective, *Cell. Mol. Life Sci.*, 2001, **58**, 1491–1521.
17 N. Pasquato, R. Berni, C. Folli, B. Alfieri, L. Cendron and G. Zanotti, Acidic pH-induced conformational changes in amyloidogenic mutant transthyretin, *J. Mol. Biol.*, 2007, **366**, 711–719.
18 S. K. Palaninathan, N. N. Mohamedmohaideen, W. C. Snee, J. W. Kelly and J. C. Sacchettini, Structural insight into pH-induced conformational changes within the native human transthyretin tetramer, *J. Mol. Biol.*, 2008, **382**, 1157–116.
19 A. Wojtczak, V. Cody, J. R. Luft and W. Pangborn, Structures of human transthyretin complexed with thyroxine at 2.0 A resolution and 3′,5′-dinitro-N-acetyl-L-thyronine at 2.2 Å resolution, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1996, **52**, 758–765.
20 H. M. Naylor and M. E. Newcomer, The structure of human retinol-binding protein (RBP) with its carrier protein transthyretin reveals an interaction with the carboxy terminus of RBP, *Biochemistry*, 1999, **38**, 2647–2653.
21 A. Fernandez and H. A. Scheraga, Insufficiently dehydrated hydrogen bonds as determinants of protein interactions, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 113–118.
22 A. Fernández, Keeping dry and crossing membranes, *Nat. Biotechnol.*, 2004, **22**, 1081–1084.
23 S. R. Miller, Y. Sekijima and J. W. Kelly, Native state stabilization by NSAIDs inhibits transthyretin amyloidogenesis from the most common familial disease variants, *Lab Invest*, 2004, **84**, 545–552.
24 G. Zandomeneghi, M. R. H. Krebs, M. G. McCammon and M. Fändrich, FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils, *Protein Sci.*, 2009, **13**, 3314–3321.
25 K. Matsubara, M. Mizuguchi, K. Igarashi, Y. Shinohara, M. Takeuchi, A. Matsuura, T. Saitoh, Y. Mori, H. Shinoda and K. Kawano, Dimeric Transthyretin Variant Assembles into Spherical Neurotoxins, *Biochemistry*, 2005, **44**, 3280–3288.
26 M. P. Sebastião, V. Lamzin, M. J. Saraiva and A. M. Damas, Transthyretin Stability as a Key Factor inAmyloidogenesis: X-ray Analysis at Atomic Resolution, *J. Mol. Biol.*, 2001, **306**, 733–744.
27 R. M. Neto-Silva, S. Macedo-Ribeiro, P. J. B. Pereira, M. Coll, M. J. Saraiva and A. M. Damas, X-ray crystallographic studies of two transthyretin variants: further insights into amyloidogenesis, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2005, **61**, 333–339.
28 A. Hörnberg, T. Eneqvist, A. Olofsson, E. Lundgren and A. E. Sauer-Eriksson, A comparative analysis of 23 structures of the amyloidogenic protein trans- thyretin, *J. Mol. Biol.*, 2000, **302**, 649–669.
29 N. Pasquato, R. Berni, C. Folli, B. Alfieri, L. Cendron and G. Zanotti, Acidic pH-induced conformational changes in amyloidogenic mutant transthyretin, *J. Mol. Biol.*, 2007, **366**, 711–719.
30 E. Frare, M. F. Mossuto, P. Polverino de Laureto, M. Dumoulin, C. M. Dobson and A. Fontana, Identification of the core structure of lysozyme amyloid fibrils by proteolysis, *J. Mol. Biol.*, 2006, **361**, 551–561.
31 L. C. Palmieri, L. M. T. R. Lima, J. B. B. Freire, L. Bleicher, I. Polikarpov, F. C. L. Almeida and D. Foguel, Novel Zn2 +-binding sites in human transthyretin: implications for amyloidogenesis and retinol-binding protein recognition, *J. Biol. Chem.*, 2010, **285**, 31731–31741.
32 M. Miyata, T. Sato, M. Mizuguchi, T. Nakamura, S. Ikemizu, Y. Nabeshima, S. Susuki, Y. Suwa, H. Morioka, Y. Ando, M. A. Suico, T. Shuto, T. Koga, Y. Yamagata and H. Kai, Role of the glutamic acid 54 residue in transthyretin stability and thyroxine binding, *Biochemistry*, 2009, **49**, 114–123.
33 L. Cendron, A. Trovato, F. Seno, C. Folli, B. Alfieri, G. Zanotti and R. Berni, Amyloidogenic potential of transthyretin variants: Insights from structural and computational analyses, *J. Biol. Chem.*, 2009, **284**, 25832–25841.
34 A. Karlsson and A. E. Sauer-Eriksson, Heating of proteins as a means of improving crystallization: a successful case study on a highly amyloidogenic triple mutant of human transthyretin, *Acta Cryst*, 2007, **F63**, 695–700.
35 N. Schormann, J. R. Murrell and M. D. Benson, Tertiary structures of amyloidogenic and non-amyloidogenic transthyretin variants: new model for amyloid fibril formation, *Amyloid*, 1998, **5**, 175–187.
36 T. Eneqvist, K. Andersson, A. Olofsson, E. Lundgren and A. E. Sauer-Eriksson, The β-slip: a novel concept in transthyretin amyloidosis, *Mol. Cell*, 2000, **5**, 1207–1218.
37 S. M. Johnson, S. Connelly, I. A. Wilson and J. W. Kelly, Biochemical and structural evaluation of highly selective 2-arylbenzoxazole-based transthyretin amyloidogenesis inhibitors, *J. Med. Chem.*, 2008, **51**, 260–270.
38 A. Hörnberg, A. Olofsson, T. Eneqvist, E. Lundgren and A. E. Sauer-Eriksson, The β-strand D of transthyretin trapped in two discrete conformations, *Biochim. Biophys. Acta, Proteins Proteomics*, 2008, **1700**, 93–104.
39 D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berend-sen, GROMACS: fast, flexible, and free, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
40 X. Daura, A. E. Mark and W. F. van Gunsteren, Parametrization of aliphatic CH n united atoms of GROMOS96 force field, *J. Comput. Chem.*, 1998, **19**, 535–547.
41 H. J. C. Berendsen, J. R. Grigerra and T. P. Straatsma, The missing term in effective pair potentials, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
42 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola and J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
43 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. Fraaije, A linear constraint solver for molecular simulations, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
44 T. Darden, D. York and L. Pedersen, Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
45 V. Lounnas and B. M. Pettitt, Distribution function implied dynamics *versus* residence times and correlations: solvation shells of myoglobin, *Proteins: Struct., Funct., Genet.*, 1994, **18**, 148–160.
46 V. Lounnas and B. M. A. Pettitt, Connected-cluster of hydration around myoglobin: Correlation between molecular dynamics simulations and experiment, *Proteins: Struct., Funct., Genet.*, 1994, **18**, 133–147.
47 T. Ichiye and M. Karplus, Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations, *Proteins: Struct., Funct., Genet.*, 1991, **11**, 205–217.

