BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

The putative role of some conserved water molecules in the structure and function of human transthyretin

Avik Banerjee,[a] Subrata Dasgupta,[a] Bishnu P. Mukhopadhyay[a]* and Kanagaraj Sekar[b]

[a]Department of Chemistry, National Institute of Technology–Durgapur, Durgapur 713 209, India, and [b]Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560 012, India. *Correspondence e-mail: bpmk2@yahoo.com

Received 9 February 2015
Accepted 26 August 2015

Edited by P. Langan, Oak Ridge National Laboratory, USA

Keywords: human transthyretin; X-ray structure analysis; conserved water; hydrogen bonding; molecular-dynamics simulation.

Supporting information: this article has supporting information at journals.iucr.org/d

Human transthyretin (hTTR) is a multifunctional protein that is involved in several neurodegenerative diseases. Besides the transportation of thyroxin and vitamin A, it is also involved in the proteolysis of apolipoprotein A1 and Aβ peptide. Extensive analyses of 32 high-resolution X-ray and neutron diffraction structures of hTTR followed by molecular-dynamics simulation studies using a set of 15 selected structures affirmed the presence of 44 conserved water molecules in its dimeric structure. They are found to play several important roles in the structure and function of the protein. Eight water molecules stabilize the dimeric structure through an extensive hydrogen-bonding network. The absence of some of these water molecules in highly acidic conditions (pH $\leq$ 4.0) severely affects the interfacial hydrogen-bond network, which may destabilize the native tetrameric structure, leading to its dissociation. Three pairs of conserved water molecules contribute to maintaining the geometry of the ligand-binding cavities. Some other water molecules control the orientation and dynamics of different structural elements of hTTR. This systematic study of the location, absence, networking and interactions of the conserved water molecules may shed some light on various structural and functional aspects of the protein. The present study may also provide some rational clues about the conserved water-mediated architecture and stability of hTTR.

1. Introduction

Human transthyretin (hTTR), a 55 kDa homotetrameric protein with 127 residues per monomer, acts as a transporter of the hormone thyroxin (T4) in cerebrospinal fluid (CSF) and blood plasma (Bartalena & Robbins, 1993). In association with retinol-binding protein (RBP), it also plays an important role in retinol (vitamin A) transportation (Monaco *et al.*, 1995; Naylor & Newcomer, 1999; Zanotti *et al.*, 2008). Recent studies also indicate its involvement in regulating the pathogenesis of Alzheimer's disease by scavenging Aβ peptide (Liu & Murphy, 2006). Normally, hTTR circulates as a soluble protein, but in some cases it polymerizes to form an insoluble toxic amyloid fibril which is deposited in the heart, peripheral/autosomal nerves and other organs, causing several diseases such as autonomic and peripheral neuropathy, carpal tunnel syndrome, vitreous deposition, cardiomyopathy *etc.* (Buxbaum & Tagoe, 2000; Gambetti & Russo, 1998; Sekijima *et al.*, 2005).

Water molecules are frequently found to play various roles in the stability, function and dynamics of protein molecules (Chaplin, 2006; Kanaujia & Sekar, 2009; Smolin & Winter, 2008). Hydration forces are significant in packing and

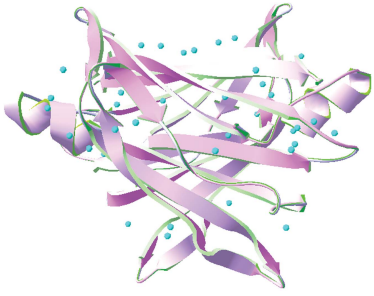

Table 1
Preliminary X-ray and neutron diffraction structural data for human transthyretin (hTTR).

Temp, template structure. Those structures which were considered for MD simulation are indicated with an asterisk.

PDB code	Resolution (Å)	*R* value (obs.)	Molecules in asymmetric unit	No. of water molecules	Crystallization pH	Mutation/ligand	Reference
1f41* (temp)	1.30	0.188	2 (*A*, *B*)	186	7.0	—	Hörnberg *et al.* (2000)
2qgb*	1.40	0.167	2 (*A*, *B*)	166	5.5	—	Johnson *et al.* (2008*a*)
3cfm*	1.6	0.192	2 (*A*, *B*)	282	7.5	—	Lima *et al.* (2010)
1tta*	1.7	0.168	2 (*A*, *B*)	186	—	—	Hamilton *et al.* (1993)
3u2i	1.7	0.197	2 (*A*, *B*)	84	7.4	—	Yokoyama *et al.* (2012)
3cbr*	1.7	0.225	2 (*A*, *B*)	254	3.5	—	Palaninathan *et al.* (2008)
3d7p	1.72	0.223	2 (*A*, *B*)	157	4.0	—	Palaninathan *et al.* (2008)
2g4g	1.85	0.212	2 (*A*, *B*)	150	4.6	—	Pasquato *et al.* (2007)
3i9p	1.9	0.201	2 (*A*, *B*)	174	7.5	—	L. Lima & D. Foguel (unpublished work)
3a4d	2.0	0.171	2 (*A*, *B*)	121	5.3	—	Miyata *et al.* (2010)
4pvm*	2.0	0.158	2 (*A*, *B*)	89	6.4	—	Haupt *et al.* (2014)
3u2j	2.0	0.180	2 (*A*, *B*)	55	7.4	—	Yokoyama *et al.* (2012)
1x7s*	1.55	0.195	2 (*A*, *B*)	196	4.8	Y78F	Neto-Silva *et al.* (2005)
3bt0	1.59	0.210	2 (*A*, *B*)	140	7.5	V20S	Zanotti *et al.* (2008)
1x7t*	1.6	0.201	2 (*A*, *B*)	187	5.1	R104H	Neto-Silva *et al.* (2005)
1sok	1.6	0.225	2 (*A*, *B*)	148	6.0	A108Y/L110E	Hörnberg *et al.* (2004)
1ttb*	1.70	0.157	2 (*A*, *B*)	178	—	A109T	Hamilton *et al.* (1993)
1ttc	1.7	0.179	2 (*A*, *B*)	319	—	V30M	Hamilton *et al.* (1993)
1fhn*	1.75	0.153	2 (*A*, *B*)	128	4.9	T119M	Sebastião *et al.* (2001)
1f86*	1.10	0.140	2 (*A*, *B*)	251	—	T119M/3,5,3′,5′-tetraiodo-L-thyronine	Sebastião *et al.* (2001)
1etb	1.7	0.163	2 (*A*, *B*)	172	—	A109T/3,5,3′,5′-tetraiodo-L-thyronine	Hamilton *et al.* (1993)
2b16	1.75	0.191	2 (*A*, *B*)	154	—	Y78F/2,4-dinitrophenol	Morais-de-Sá *et al.* (2006)
2qgc*	1.30	0.420	2 (*A*, *B*)	196	5.5	2-(3,5-Dimethyl-4-hydroxyphenyl)benzoxazole	Johnson *et al.* (2008*a*)
3esp*	1.31	0.178	2 (*A*, *B*)	209	5.5	*N*-(3,5-Dibromo- 4-hydroxyphenyl)-4-hydroxy-3,5-dimethylbenzamide	Johnson *et al.* (2009)
3eso	1.31	0.179	2 (*A*, *B*)	208	5.5	2,5-Dichloro-*N*-(3,5-dibromo-4-hydroxyphenyl)-benzamide	Johnson *et al.* (2009)
3esn*	1.35	0.175	2 (*A*, *B*)	215	5.5	*N*-(3,5-Dibromo-4-hydroxyphenyl)-2,6-dimethyl-benzamide	Johnson *et al.* (2009)
3cn4	1.40	0.175	2 (*A*, *B*)	126	5.5	*N*-(3,5-Dibromo-4-hydroxyphenyl)benzamide	Johnson *et al.* (2008*b*)
2qge*	1.45	0.166	2 (*A*, *B*)	148	5.5	2-(3,5-Dimethylphenyl)benzoxazole	Johnson *et al.* (2008*a*)
2fbr	1.46	0.211	2 (*A*, *B*)	181	7.0	4′-(4-{4-[(2-Carboxyphenyl)amino]phenoxy}butoxy)-1,1′-biphenyl-4-carboxylic acid	Green *et al.* (2003)
2qgd	1.50	0.162	2 (*A*, *B*)	142	5.5	2-(3,5-Dibromo-4-hydroxyphenyl)benzoxazole	Johnson *et al.* (2008*a*)
3cn0	1.52	0.162	2 (*A*, *B*)	126	5.5	3,5-Dimethyl-4-hydroxystilbene	Johnson *et al.* (2008*b*)
3cn1	1.52	0.162	2 (*A*, *B*)	132	5.5	3,5-Dibromo-4-hydroxystilbene	Johnson *et al.* (2008*b*)

maintaining the integrity of the three-dimensional architecture (Raschke, 2006). Protein–water interactions (hydrophilic and hydrophobic) play key roles in many biological functions. Some of these water molecules are indispensable for the activity, recognition and self-assembly of proteins, and hence water molecules are considered as an integral part. Buried water molecules form several strong hydrogen bonds to the polar atoms of the interior amino-acid residues, and maintain the structure and conformation of proteins. Many of these water molecules remain highly ordered in protein structures, especially at their interfaces, metal-binding and catalytic sites, and play pivotal roles in function and intra/intermolecular association (Jiang *et al.*, 2005; Banerjee & Mukhopadhyay, 2015). Water molecules also play crucial roles in drug/inhibitor–protein complexation (Poornima & Dean, 1995). Again, water molecules sometimes control the ligand-binding propensity by their presence in and subsequent displacement from the protein active site (Banerjee *et al.*, 2013). Several studies have already been reported on conserved water molecules in various biomolecular systems (Ogata & Wodak, 2002; Nandi *et al.*, 2012). Here, a systematic study has been undertaken to determine the conserved or other water molecules in high-resolution X-ray and neutron diffraction structures (crystallized under different conditions, containing different numbers of water molecules and with or without minor mutations) and their dynamic propensities in order to better understand the multifunctional aspects of human transthyretin.

In this investigation, the location and hydrogen-bonding interactions of water molecules, especially the conserved water centres, have been compared in several X-ray, neutron diffraction and MD-simulated structures of hTTR to obtain an overall idea of the role of these water molecules. A study on the mobility or residential frequency of these conserved or semi-conserved water molecules may provide some plausible rationale for the water-mediated recognition of intramolecular strands and intermolecular subunits. The hydrophilic susceptibility of these conserved water molecules and their subsequent hydrogen-bonding potentialities to the functionally/structurally important residues may provide an insight into the stabilization of hTTR. Again, knowledge of the geometry, orientation and dynamics of these water molecules

in the ligand-binding cavity may be useful for *de novo* ligand/drug design.

2. Materials and methods

A total of 30 high-resolution (1.3–2.0 Å) X-ray and two high-resolution (2.0 Å) neutron diffraction crystal structures of human transthyretin (taken from the PDB) were analyzed to determine the conserved water molecules using the 3*DSS* web server (Sumathi *et al.*, 2006) and *Swiss-PdbViewer* (Guex & Peitsch, 1997). Some of the essential crystallographic statistics and structural information for the structures used in the present study are given in Table 1. The structure with PDB code 1f41 (Hörnberg *et al.*, 2000) was taken as reference molecule and the remaining structures were treated as mobile molecules during structural superimposition (Sumathi *et al.*, 2006). The cutoff distance between pairs of superposed water molecules was taken as 1.8 Å, and only those water molecules which form at least one hydrogen bond to the protein molecule were considered (Balamurugan *et al.*, 2007). However, in certain instances water molecules were considered as equivalent if a similar type of hydrogen-bonding pattern was encountered even if the pairwise distance criterion was not satisfied owing to varying side-chain conformations (Kanaujia & Sekar, 2009).

2.1. Normalized *B* factor, hydrogen-bonding analysis and accessible surface-area calculation

Neutron protein crystallography is more useful as it can locate the position of H atoms, in contrast to the X-ray crystallographic method. Thus, hydrogen-bonding studies would be more informative (especially about water molecules) if we include neutron diffraction results together with X-ray structures. Hence, two available neutron diffraction structures of hTTR (PDB entries 4pvm and 3u2j) have been included in the present study. Although some of the H atoms of the water molecules have not been located in these structures, their inclusion may provide some rationale for the hydrogen-bonding character of the water molecules. For decades, hydrogen bonds were identified using the criteria that the donor–acceptor distance must be less than 3.5 Å and the donor–hydrogen–acceptor angle must be greater than 90° (Baker & Hubbard, 1984), but recently it has been suggested that a donor–acceptor distance of greater than 3 Å does not necessarily imply a hydrogen-bond interaction, especially when the donor is pointing away from the acceptor (Chen *et al.*, 2012). Since the water molecules are mobile in molecular-dynamics (MD) simulations and can easily switch the H atoms involved in hydrogen bonding by rotation, so the hydrogen bonds of water molecules are described by the distance between the polar atoms of the protein and the O atom of the water molecule (Mladenovic *et al.*, 2008). In this situation, we have considered a donor–acceptor distance of $\leq$3.0 Å as a hydrogen-bond interaction; however, we have included interactions of up to 3.5 Å for convenience. The normalized *B* factor (B_i') for all conserved water molecules was calculated using the formula $B_i' = (B_i - \langle B_i \rangle)/\sigma(B)$, where B_i is the *B* factor for each atom, $\langle B_i \rangle$ is the mean *B* factor of the protein molecule and $\sigma(B)$ is the standard deviation (Smith *et al.*, 2003). The accessible surface area of the conserved water molecules was computed using the *POPS* server (Fraternali & Cavallo, 2002; Cavallo *et al.*, 2003) with a probe radius of 1.4 Å. Water molecules with an accessible surface area of less than or equal to 2.5 Å^2 were considered to be buried (Kanaujia & Sekar, 2009).

2.2. Free-energy calculation

To evaluate the contribution of water molecules towards the stability of the protein, the stability free energy (ΔG) was calculated (Banerjee *et al.*, 2010; Nandi *et al.*, 2012) using *FoldX* (Schymkowitz, Borg *et al.*, 2005) with and without these water molecules (Schymkowitz, Rousseau *et al.*, 2005). The difference in the stability free energy [$\Delta\Delta G = \Delta G$(with water) $- \Delta G$(without water)] was considered to be the contribution of these water molecules to protein stability.

2.3. Molecular-dynamics (MD) simulations

A total of 15 high-resolution structures (six apo, four mutant, four ligand-bound and one ligand-bound mutant structures, chosen on the basis of their crystallographic parameters) were subjected to molecular-dynamics simulations using *NAMD* v.2.9 (Kalé *et al.*, 1999) with the CHARMM27 force field (Brooks *et al.*, 1983; MacKerell *et al.*, 1998). The necessary topology and parameter files for the ligand molecules were generated by the *SwissParam* program (Zoete *et al.*, 2011). Each structure was then converted to a protein structure file (PSF) by the Automatic PSF Generation Plug-in within *VMD* v.1.9.1 (Humphrey *et al.*, 1996) and solvated explicitly with the TIP3P water model, retaining the crystallographic water molecules (Jorgensen *et al.*, 1983). Cuboid water boxes were constructed around the protein with 5 Å boundaries and 6–8 Å padding (containing approximately 5000 water molecules) using the *Solvate* program of *VMD*. Subsequently, energy minimization was performed for 100 ps by the conjugate-gradient procedure and all-atom MD simulations (time step 1 fs) were carried out with these energy-minimized structures at constant temperature (300 K) and constant pressure (101.325 kPa) by means of Langevin dynamics (Ramachandran & Schlick, 1995) using periodic boundary conditions. Initially, 1 ns of molecular dynamics was carried out for the equilibration of each system, followed by 3 ns MD simulations with the equilibrated protein structures to investigate the interaction and dynamics of the conserved water molecules identified in X-ray and neutron diffraction structures. To evaluate the importance of some conserved water molecules, a few MD-simulation studies were carried out using the template structure (PDB entry 1f41) after removing the intended water molecules.

2.4. Normalized r.m.s.f. (root-mean-square fluctuation) and volume calculations

The r.m.s.f. or the averaged standard deviation of atomic positions in the trajectory was calculated for C^{α} atoms using the equation

$$\text{r.m.s.f.} = \left[\frac{1}{T}\sum_{t=1}^{T}(r_t - r_{\text{m}})^2\right]^{1/2},$$

where

$$r_{\text{m}} = \left(\frac{1}{T}\sum_{t=1}^{T} r_t\right)^{1/2},$$

T is the total number of frames (or time steps) during the MD simulation, r_t is the averaged standard deviation of atomic positions in the trajectory of the C^{α} atom of a particular residue at a time step t and r_{m} is the mean r.m.s.f. value. The normalized r.m.s.f. was calculated using the formula $R_{\text{NRD}} = (R_i - \langle R_i\rangle)/\sigma(R)$, where R_{NRD} is the normalized r.m.s.f., R_i is the r.m.s.f. for each residue, $\langle R_i\rangle$ is the average r.m.s.f. of the protein molecule and $\sigma(R)$ is the standard deviation of the r.m.s.f. The volume of the ligand-binding cavities in the MD-simulated structures was calculated using the *CASTp* (*Computed Atlas of Surface Topography of proteins*) server (Dundas *et al.*, 2006).

2.5. Mutational study

To study the effect of mutation on the conserved water-molecule cluster or its hydrogen-bonding network, His88 (in both chains) of the template structure (PDB entry 1f41) was mutated (Banerjee *et al.*, 2010; Dasgupta *et al.*, 2003) to alanine using *Swiss-PdbViewer*. The program assesses the possible rotamer combinations using the GROMOS96 force field (Scott *et al.*, 1999) to reach the minimum energy conformation of the mutated residues. Some favoured conformations are generated by its scoring scheme and are ranked accordingly. The lowest energy rotamers of the mutated residues, which do not make steric clashes with other residues, are included in the study.

3. Results and discussion

The investigation of 32 high-resolution X-ray and neutron diffraction crystal structures of hTTR revealed the presence of 44 conserved or invariant water molecules in the protein molecule (Fig. 1). If a water molecule is observed in most of the structures (more than 80%) and shows higher average residential frequency in the MD-simulated average structures, it is considered to be a conserved water molecule. These can be grouped into 19 pairs (*i.e.* if one water molecule forms a hydrogen bond to the Z atom of a residue in chain A, its counter pair involves a similar interaction with the Z atom of the same residue in chain B), and the remaining (six) water

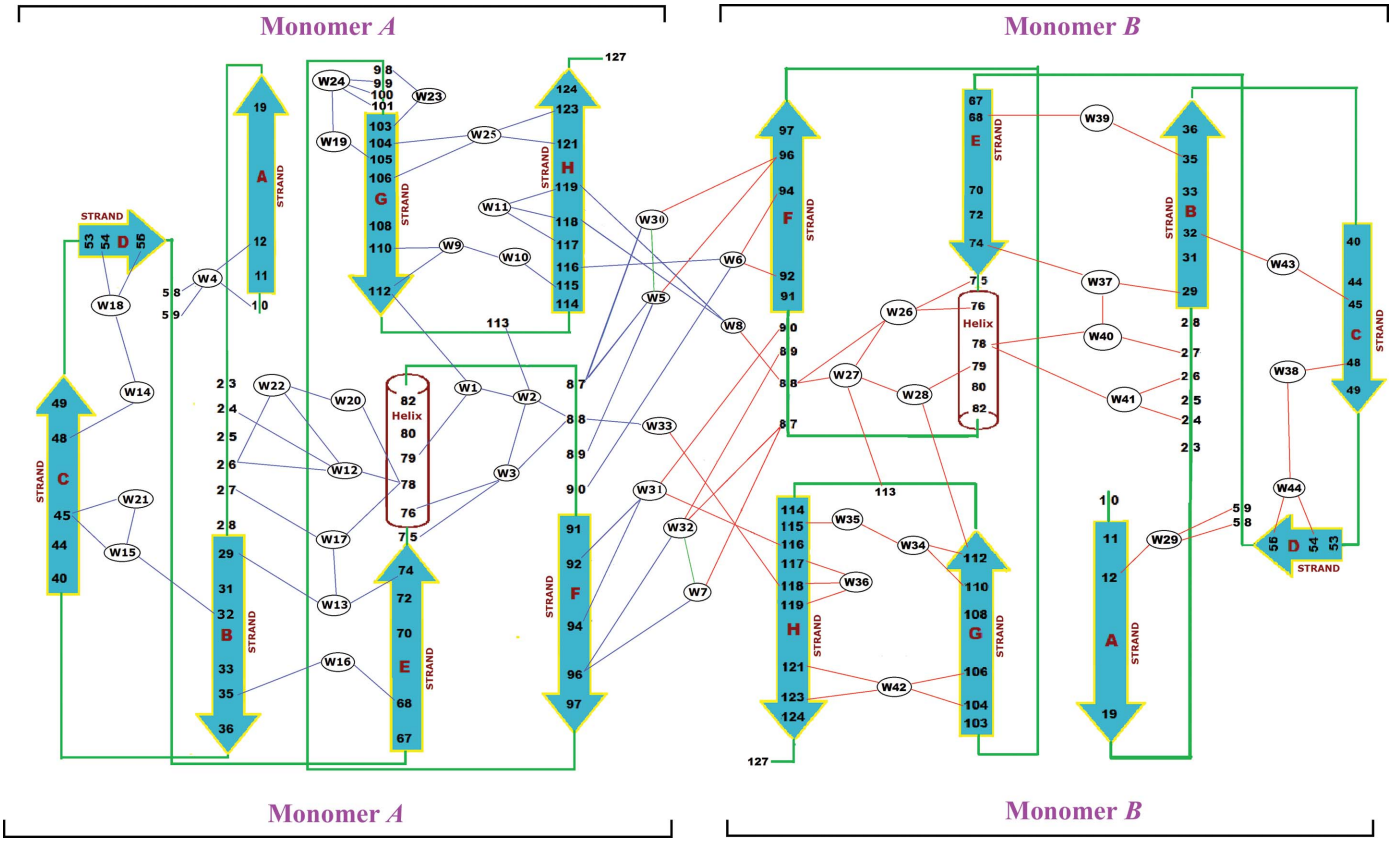

Figure 1
Schematic representation of the hydrogen-bond interactions provided by the conserved water molecules in hTTR (dimer).

Table 2
Interaction of the conserved water molecules and protein residues in the X-ray and neutron diffraction structures of hTTR.

The numbers in parentheses indicate the corresponding distances; hydrogen-bond interactions (distance $\leq$ 3 Å) are indicated in bold. The ID numbers of the nonconserved water molecules are as were found in the template structure; an asterisk indicates the absence of a water molecule and a dash denotes that the data are not available.

(*a*) Conserved waters.

Conserved water	Protein atoms		Interaction with other conserved waters
	Main chain	Side chain	
W1	*A* Ser112 O^{β} (**2.85**)	*A* Trp79 N^{ε} (**2.92**)	W2 (**3.23**)
W2	*A* His88 N^{β} (3.13), *A* Pro113 O^{β} (**2.78**)	*A* His88 N^{δ} (**2.74**)	W1 (3.23), W3 (**2.78**)
W3	*A* His88 O^{β} (**2.74**), *A* Lys76 N^{β} (3.31)	*A* Thr75 O^{γ} (**2.65**)	W2 (**2.78**)
W4	*A* Leu12 O^{β} (**2.81**), *A* Leu12 N^{β} (3.27), *A* Leu58 N^{β} (3.48), *A* Thr59 O^{β} (3.12), *A* Thr59 N^{β} (3.07)	*A* Cys10 S^{γ} (3.39)	—
W9	*A* Leu110 O^{β} (3.06), *A* Ser112 N^{β} (3.30)	*A* Ser112 O^{γ} (**2.70**)	W10 (2.96)
W10	—	*A* Ser115 O^{γ} (**2.68**)	W9 (2.96)
W11	*A* Ser117 O^{β} (3.02), *A* Thr118 N^{β} (3.39), *A* Thr119 N^{β} (3.35)	*A* Ser117 O^{γ} (**2.74**), *A* Thr119 O^{γ} (**2.74**)	—
W12	*A* Pro24 O^{β} (3.18), *A* Ile26 N^{β} (**3.00**)	*A* Tyr78 OH (**2.83**)	W22 (3.50)
W13	*A* Ala29 N^{β} (**2.90**), *A* Asp74 O^{β} (**2.72**)	—	W17 (**2.75**)
W14	*A* Lys48 O^{β} (**2.71**)	—	W18 (**2.70**)
W15	*A* Val32 O^{β} (**2.85**), *A* Ala45 O^{β} (3.18)	—	W21 (**2.66**)
W16	*A* Lys35 O^{β} (**2.69**), *A* Ile68 N^{β} (**2.87**)	—	—
W17	*A* Asn27 O^{β} (**2.95**), *A* Tyr78 N^{β} (3.49)	—	W13 (**2.75**)
W18	*A* His56 N^{β} (**2.95**)	—	W14 (**2.70**)
W19	—	*A* Tyr105 OH (**2.56**)	W24 (3.41)
W20	*A* Tyr78 O^{β} (**2.95**)	—	W22 (**2.74**)
W21	*A* Ala45 N^{β} (3.21)	—	—
W22	*A* Ile26 O^{β} (**2.70**)	—	W20 (**2.74**)
W23	*A* Arg103 O^{β} (**2.82**)	*A* Asn98 O^{δ} (3.36)	—
W24	*A* Asp99 N^{β} (3.34), *A* Ser100 N^{β} (**2.99**), *A* Gly101 O^{β} (3.50)	—	W19 (3.41)
W25	*A* Arg104 O^{β} (3.22), *A* Val121 O^{β} (**2.90**)	*A* Thr106 O^{γ} (3.35), *A* Thr123 O^{γ} (3.34)	—
W26	*B* His88 O^{β} (**2.75**) *B* Lys76 N^{β} (3.20)	*B* Thr75 O^{γ} (**2.63**)	W27 (**2.76**)
W27	*B* His88 N^{β} (3.17), *B* Pro113 O^{β} (**2.71**)	*B* His88 N^{δ} (**2.71**)	W28 (3.14), W26 (**2.76**)
W28	*B* Ser112 O^{β} (**2.83**)	*B* Trp79 N^{ε} (**2.99**)	W27 (3.14)
W29	*B* Leu12 O^{β} (**2.67**) *B* Leu12 N^{β} (3.24), *B* Leu58 N^{β} (3.39), *B* Thr59 O^{β} (3.17), *B* Thr59 N^{β} (3.06)	—	—
W34	*B* Leu110 O^{β} (3.22)	*B* Ser112 O^{γ} (**2.82**)	W35 (**2.67**)
W35	—	*B* Ser115 O^{γ} (**2.77**)	W34 (**2.67**)
W36	*B* Ser117 O^{β} (3.47), *B* Thr118 N^{β} (3.43), *B* Thr119 N^{β} (3.33)	*B* Ser117 O^{γ} (**2.73**), *B* Thr119 O^{γ} (**2.53**)	—
W37	*B* Ala29 N^{β} (**2.93**), *B* Asp74 O^{β} (**2.70**)	—	W40 (**2.79**)
W38	*B* Lys48 O^{β} (**2.84**)	—	W44 (**2.68**)
W39	*B* Lys35 O^{β} (**2.73**), *B* Ile68 N^{β} (**2.97**)	—	—
W40	*B* Asn27 O^{β} (**2.81**)	—	W37 (**2.79**)
W41	*B* Pro24 O^{β} (**2.96**), *B* Ile26 N^{β} (**2.96**)	*B* Tyr78 OH (**2.85**)	—
W42	*B* Arg104 O^{β} (3.30), *B* Val121 O^{β} (**2.78**)	*B* Thr106 O^{γ} (3.11), *B* Thr123 O^{γ} (3.01)	—
W43	*B* Val32 O^{β} (**2.85**), *B* Ala45 O^{β} (3.18)	—	—
W44	*B* His54 O^{β} (3.40), *B* His56 N^{β} (**2.95**)	—	W38 (**2.68**)

(*b*) Interfacial conserved waters.

Interfacial conserved waters	Protein atoms (*A* chain)		Protein atoms (*B* chain)		Interaction with other conserved waters
	Main chain	Side chain	Main chain	Side chain	
W5	Phe87 O^{β} (**2.80**)	Glu89 O^{ε} (**2.39**)	Thr96 N^{β} (**2.82**)	Thr96 O^{γ} (3.25)	W30 (3.19)
W6	His90 O^{β} (**2.63**)	Tyr116 OH (**2.84**)	Glu92 O^{β} (3.37), Val94 N^{β} (**2.87**)	—	—
W7	Thr96 O^{β} (**2.71**)	—	Phe87 O^{β} (**2.88**)	—	W32 (3.34)
W8	Thr119 O^{β} (3.10)	Thr118 O^{γ} (**2.71**)	—	His88 N^{ε} (**2.96**)	—
W30	Phe87 O^{β} (**2.84**)	—	Thr96 O^{β} (**2.73**)	—	W5 (3.19)
W31	Glu92 O^{β} (3.28), Val94 N^{β} (**2.88**)	—	His90 O^{β} (**2.66**)	Tyr116 OH (3.05)	—
W32	Thr96 N^{β} (**2.86**)	Thr96 O^{γ} (3.44)	Phe87 O^{β} (**2.71**)	Glu89 O^{ε} (**2.62**)	W7 (3.34)
W33	—	His88 N^{ε} (**2.84**)	Thr119 O^{β} (3.41)	Thr118 O^{γ} (**2.68**)	—

molecules have been located only in chain *A*. The details of the polar interaction(s) of these water molecules are given in Table 2. Six conserved water molecules (W1, W2 and W3 in chain *A* and W28, W27 and W26 in chain *B*) are located near the helix and the loop before strand *F*. These internal water molecules are interconnected (in each monomer), forming a cluster (Fig. 2). Among them, water molecule W1 (W28 in chain *B*) forms hydrogen bonds to Trp79 N^{ε}, Ser112 O^{β} and

Table 3
Normalized *B* factor (Norm. *B*) and solvent-accessible surface area (SASA; Å^2) of the conserved water molecules interacting with the catalytic and RBP-binding residues in the X-ray and neutron diffraction crystal structures of hTTR.

The numbers in the first row indicate the IDs of the water molecules in the respective crystals; normalized *B* factors of the O atoms of these water molecules are given in the second row and their SASAs are listed in the third row. RF, residential frequency. The absence of a water molecule in a structure is indicated by an asterisk and a dash denotes that the data are not available. 'o' denotes that the water position is occupied by some other atom of the protein in the crystal.

		W1	W2	W3	W4	W26	W27	W28	W29
1	1f41	1	2	3	10	5	6	8	12
	Norm. *B*	−0.95	−0.75	−0.92	0.11	−0.95	−0.71	−0.93	0.57
	SASA	1.65	1.40	1.40	1.78	1.28	1.46	1.44	2.51
	RF	100	100	100	100	100	100	100	100
2	2qgb	130 *A*	145 *A*	161 *A*	131 *A*	136 *B*	135 *B*	142 *B*	149 *B*
	Norm. *B*	−0.89	−0.88	−0.98	0.24	−0.84	−0.65	−0.87	0.67
	SASA	2.02	1.71	1.52	2.79	1.43	1.54	1.81	2.65
	RF	98	100	100	100	100	100	100	100
3	3cfm	129 *A*	139 *A*	126 *A*	159 *A*	126 *B*	128 *B*	131 *B*	151 *B*
	Norm. *B*	−0.96	−0.77	−0.80	−0.15	−0.81	−0.88	−0.95	0.78
	SASA	1.67	1.47	1.41	1.45	1.42	1.53	1.88	1.43
	RF	100	100	100	100	100	100	100	65
4	**3cbr**	**131 *A***	**199 *A***	**130 *A***	**202 *A***	**176 *A***	**140 *B***	**192 *B***	**165 *B***
	Norm. *B*	**−1.03**	**−0.57**	**−0.98**	**−0.13**	**0.54**	**0.20**	**0.11**	**1.47**
	SASA	**1.81**	**1.55**	**1.34**	**1.65**	**2.15**	**4.30**	**2.23**	**2.60**
	RF	**100**	**100**	**100**	**92**	**100**	**74**	**100**	**70**
5	1tta	201 *A*	210 *A*	202 *A*	215 *A*	402 *B*	410 *B*	401 *B*	415 *B*
	Norm. *B*	−1.75	−1.60	−1.62	−0.82	−1.08	−1.20	−1.55	−0.12
	SASA	2.36	1.86	1.93	1.35	1.66	1.71	2.15	1.48
	RF	100	100	100	100	100	100	100	100
6	3d7p	130 *A*	137 *A*	128 *A*	151 *A*	180 *B*	*	*	*
	Norm. *B*	−1.04	−1.06	−1.04	−0.22	1.38	–	–	–
	SASA	1.92	1.53	1.33	1.36	5.06	–	–	–
7	3a4d	138 *A*	192 *A*	195 *A*	153	131 *B*	179 *B*	136 *B*	140 *B*
	Norm. *B*	−1.40	−1.39	−1.23	−0.69	−1.09	−1.00	−1.42	−0.38
	SASA	1.88	1.64	1.54	1.83	1.56	1.89	2.09	1.54
8	2g4g	130 *A*	129 *A*	128 *A*	180 *A*	128 *B*	131 *B*	129 *B*	158 *B*
	Norm. *B*	−1.31	−0.67	−0.80	0.49	−0.61	−0.69	−0.91	1.74
	SASA	2.13	1.81	1.53	1.61	1.66	2.01	2.30	1.38
9	3i9p	132 *A*	159 *A*	4 *A*	141 *A*	157 *B*	1 *B*	127 *B*	135 *B*
	Norm. *B*	−1.25	−1.17	−1.35	0.50	−1.52	−0.81	−1.57	0.39
	SASA	1.35	1.11	1.16	1.33	1.27	1.29	1.69	1.57
10	3u2i	10 *A*	2 *A*	5 *A*	141 *A*	9 *B*	6 *B*	4 *B*	148 *B*
	Norm. *B*	−0.90	−0.80	−0.90	0.30	−0.80	−0.80	−0.80	0.70
	SASA	2.20	1.90	1.89	3.72	1.87	2.01	2.36	3.08
11	**3u2j**	**2 *A***	**135 *A***	**8*A***	**132 *A***	**150 *B***	**1 *B***	**151 *B***	**141 *B***
	Norm. *B*	**−0.32**	**−0.22**	**−0.31**	**−0.06**	**−0.12**	**−0.12**	**−0.31**	**−0.05**
	SASA	**2.19**	**1.93**	**2.09**	**3.37**	**1.92**	**2.06**	**2.45**	**3.72**
12	**4pvm**	**208 *A***	**206 *A***	**201 *A***	*	**202 *B***	**201 *B***	**205 *B***	**230 *B***
	Norm. *B*	**−0.57**	**−0.46**	**−0.42**	–	**−0.34**	**−0.41**	**−0.49**	**0.15**
	SASA	**2.35**	**2.03**	**1.89**	–	**1.85**	**1.89**	**2.45**	**1.82**
	RF	**95**	**91**	**88**	**86**	**95**	**91**	**100**	**89**
11	1f86	429 *A*	432 *A*	430 *A*	449 *A*	530 *B*	533 *B*	532 *B*	552 *B*
	Norm. *B*	−0.91	−0.84	−0.87	−0.08	−0.53	−0.51	−0.63	0.18
	SASA	1.74	1.43	1.31	1.66	1.46	1.58	1.94	1.44
	RF	100	100	100	100	100	100	100	100
12	1x7s	6019	6020	6021	6005	6069	6070	6028	6026
	Norm. *B*	−0.78	−0.70	−0.77	−0.07	−0.44	−0.18	−0.52	0.30
	SASA	1.71	1.47	1.47	1.67	1.49	1.62	1.88	2.28
13	3bt0	129 *A*	130 *A*	133 *A*	144 *A*	130 *B*	131 *B*	129 *B*	140 *B*
	Norm. *B*	−0.98	−0.83	−0.35	−0.07	−0.60	−0.48	−1.03	0.82
	SASA	2.08	1.64	1.46	1.33	1.44	1.56	1.84	1.47
14	1x7t	6003	6016	6056	6082	6029	6031	6030	6103
	Norm. *B*	−0.87	−0.64	−0.67	−0.57	−0.75	−0.40	−0.96	0.69
	SASA	1.83	1.49	1.42	1.33	1.54	1.67	2.12	2.19
	RF	100	100	100	100	100	100	100	90
15	1sok	128 *A*	129 *A*	130 *A*	132 *A*	128 *B*	129 *B*	131 *B*	132 *B*
	Norm. *B*	−0.44	−0.42	−0.54	0.86	−0.58	−0.79	−0.74	1.29
	SASA	1.68	1.41	1.40	1.48	1.38	1.49	1.81	1.52
16	1ttb	303 *A*	319 *A*	304 *A*	408 *A*	404 *B*	419 *B*	403 *B*	409 *B*
	Norm. *B*	−1.18	−0.58	−0.77	0.19	−0.53	−0.69	−0.87	0.64
	SASA	2.33	1.91	1.86	1.30	1.92	2.00	2.50	1.36
	RF	100	100	100	85	100	100	100	70
17	1ttc	202 *A*	216 *A*	203 *A*	209 *A*	403 *B*	416 *B*	402 *B*	409 *B*
	Norm. *B*	−1.45	−1.54	−1.17	0.35	−1.03	−1.15	−1.49	0.49
	SASA	2.04	1.88	1.66	1.33	1.91	1.85	2.24	1.20

Table 3 (continued)

		W1	W2	W3	W4	W26	W27	W28	W29
18	1etb	132 *A*	140 *A*	133 *A*	147 *A*	140 *B*	147 *B*	139 *B*	142 *B*
	Norm. *B*	−1.66	−0.76	−0.31	0.39	−2.05	−0.89	−1.10	0.92
	SASA	2.29	1.97	1.86	1.88	1.71	1.93	2.59	1.64
19	1fhn	128 *A*	131 *A*	129 *A*	137 *A*	131 *B*	130 *B*	129 *B*	143 *B*
	Norm. *B*	−0.74	−0.56	−0.54	−0.14	−0.60	−0.51	−0.60	0.34
	SASA	2.00	1.61	1.61	1.53	1.68	1.65	2.10	1.69
	RF	100	100	100	100	100	100	100	155
20	2b16	2041	2043	2040	2028	2016	2017	2015	2001
	Norm. *B*	−2.25	−1.98	−1.97	−1.22	−2.05	−1.61	−2.01	−0.93
	SASA	2.03	1.60	1.51	1.69	1.56	1.79	2.18	1.74
21	2qgc	129 *A*	131 *A*	164 *A*	134 *A*	162 *B*	161 *B*	129 *B*	131 *B*
	Norm. *B*	−0.73	−0.76	−0.86	0.15	−0.76	−0.61	−0.80	0.19
	SASA	2.09	1.67	1.57	3.12	1.59	1.71	2.17	4.00
	RF	100	100	100	100	100	100	100	100
22	3esp	230 *A*	130 *A*	229 *A*	195 *A*	232 *B*	*/o	*	*
	Norm. *B*	−0.69	−0.76	−0.71	0.09	−0.40	—	—	—
	SASA	2.11	1.65	1.54	3.98	1.88	—	—	—
	RF	100	100	100	100	100	100	85	100
23	3eso	127 *A*	129 *A*	128 *A*	227 *A*	126 *B*	137 *B*	128 *B*	*
	Norm. *B*	−0.72	−0.77	−0.62	−0.04	−0.54	−0.41	−0.71	—
	SASA	2.08	1.69	1.54	4.07	1.52	1.59	1.88	—
24	3esn	230 *A*	229 *A*	126 *A*	*	128 *B*	130 *B*	*	164 *B*
	Norm. *B*	−0.79	−0.77	−0.67	—	−0.47	−0.23	—	−0.53
	SASA	2.09	1.65	1.49	—	1.59	1.70	—	4.49
	RF	100	100	100	100	100	100	85	100
25	3cn4	138 *A*	145 *A*	135 *A*	137 *A*	130 *B*	139 *B*	132 *B*	129 *B*
	Norm. *B*	−0.73	−0.80	−0.69	0.01	−0.44	−0.34	−0.64	−0.37
	SASA	2.09	1.69	1.58	4.44	1.46	1.59	1.90	4.24
26	2fbr	7	10	11	73	6	17	5	57
	Norm. *B*	−0.89	−0.99	−0.89	0.16	−0.96	−0.61	−0.97	0.29
	SASA	1.92	1.54	1.60	1.75	1.62	1.64	2.03	2.18
27	2qgd	201 *A*	206 *A*	204 *A*	221 *A*	203 *B*	209 *B*	204 *B*	210 *B*
	Norm. *B*	−0.74	−0.86	−0.78	−0.04	−0.70	−0.56	−0.86	0.51
	SASA	2.20	1.76	1.65	2.95	1.70	1.81	2.26	3.28
28	2qge	203 *A*	207 *A*	204 *A*	214 *A*	202 *B*	203 *B*	201 *B*	206 *B*
	Norm. *B*	−0.76	−0.81	−0.75	0.32	−0.78	−0.60	−0.86	0.37
	SASA	2.10	1.72	1.57	2.73	1.63	1.84	2.19	3.74
	RF	100	100	100	100	100	100	100	100
29	3cn0	132 *A*	136 *A*	129 *A*	142 *A*	130 *B*	131 *B*	129 *B*	*
	Norm. *B*	−0.90	−0.85	−0.84	0.15	−0.62	−0.52	−0.74	—
	SASA	2.20	1.71	1.56	3.99	1.48	1.56	1.96	—
30	3cn1	134 *A*	136 *A*	133 *A*	*	131 *B*	144 *B*	132 *B*	*/o
	Norm. *B*	−0.78	−0.77	−0.64	—	−0.43	−0.24	−0.49	—
	SASA	2.15	1.67	1.54	—	1.52	1.56	1.89	—
Average Norm. *B*		**−0.98**	**−0.85**	**−0.84**	**0.00**	**−0.61**	**−0.61**	**−0.89**	**0.41**
Average SASA		**2.01**	**1.66**	**1.57**	**2.22**	**1.73**	**1.79**	**2.08**	**2.31**
Average RF		**99**	**99**	**99**	**97**	**97**	**98**	**98**	**96**

W2 (W27 in chain *B*). The water molecule W2 also interacts with Pro113 O^{β}, His88 $N^{\beta}/N^{\delta1}$ and W3 (W26 in chain *B*). The conserved water molecule W3 (W26) interacts with His88 O^{β}, Lys76 N^{β} and Thr75 O^{γ}. The water molecule within parentheses is the corresponding counter pair in the other chain, and this notation is used throughout the manuscript. The normalized *B* factors and solvent-accessible surface areas (SASAs) indicate the stability of these water molecules in the buried core of the protein (Table 3). A recent study has revealed their involvement in the recognition of the catalytic core and RBP-binding residues of hTTR (Banerjee & Mukhopadhyay, 2015). Moreover, these water molecules also play an important role in the side-chain dynamics of His88, which is considered to be crucial for the proteolytic activity of the protein (Liz *et al.*, 2012). The MD-simulation studies carried out to review the dynamic character of these water molecules revealed their presence at the respective positions with very high residential frequency (RF; Table 3), maintaining a hydrogen-bonding network, which indicates their important role in the structural integrity of the protein. The average residential frequency of the water molecules W1, W2, W3, W26 and W27 is 100%, whereas it is 98% for W28. Another water molecule W4 (W29) is found to form a hydrogen bond to Leu58 N^{β}, Thr59 O^{β}/N^{β} of the D–E loop and Leu12 O^{β}/N^{β} of the A strand. The water molecule W4 also stabilizes Cys10 S^{γ}, which may be involved in the formation of intramolecular disulfide bonds and facilitate amyloidosis (Thylén *et al.*, 1993). Most of the MD-simulation trajectories show the presence of water molecules with 100% residential frequency (Table 3), but in some cases (W4 in PDB entries 1x7t and 1ttb; W29 in PDB entries 3cfm, 1fhn and 1ttb), it decreases slightly when the hydrophilic position is occupied by Cys10 S^{γ}. All of the water molecules (W1, W2 and W3 in chain *A* and W28, W27 and W26 in chain *B*) have been found at the

respective hydrophilic positions in both of the neutron diffraction structures (PDB entries 4pvm and 3u2j), except for W4, which is absent in 4pvm. Moreover, all of the water molecules show a high residential frequency (Table 3) during MD simulations. However, during dynamics one water molecule comes from the bulk-solvent zone to occupy the hydrophilic centre (W4) within 100 ps and shows a high residential frequency.

The effect of amino-acid replacement may be related to direct changes associated with the binding and interaction of the residue as well as indirect changes owing to water-molecule reorganization near the mutation site. By combining the site-directed mutagenesis and MD-simulation studies, further insight could be gained into the roles of these conserved water molecules in the structure and function of the protein. To verify the importance of the conserved water-molecule cluster, His88 was mutated to alanine. A 10 ns MD-simulation study of the mutated His88Ala structure reveals that Ala88 O^{β} forms a hydrogen bond to W3, and W2 forms a hydrogen bond to Pro113 O^{β} and Ala88 N^{β}. Probably owing to the absence of a long polar side chain, Ala88 could not interact with W1, and the water molecule (W1) shifts towards the W8 hydrophilic centre, which stabilizes it through hydrogen bonding. In this mutant form, Ser115 forms a hydrogen bond to W2, whereas Trp79 is stabilized by W3. Therefore, it is evident that the conserved water molecules in the mutant structures are also trying to maintain a similar type of stereochemical topology as prevails in the X-ray and MD-simulated structures of native hTTR.

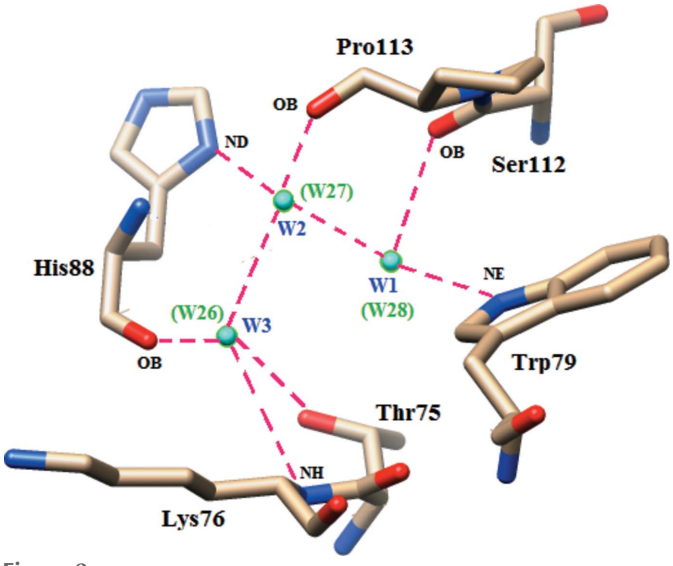

Figure 2
Conserved water cluster interacting with the catalytic and RBP-binding residues of hTTR.

The study of the X-ray and neutron diffraction structures and their MD-simulation trajectories has also revealed the presence of four pairs of conserved water molecules (W9, W10, W11 and W25 in chain *A* and W34, W35, W36 and W42 in chain *B*) in the vicinity of the G and H strands (Fig. 3). These strands mostly participate in the formation of the ligand-binding cavity of the protein (Klabunde *et al.*, 2000). Moreover, along with the F strand and the loop following it, the H strand also participates in formation of the monomer–monomer interface (Yang *et al.*, 2003). The water molecules (W9, W10, W11, W25, W34, W35, W36 and W42) are exposed to the bulk solvent, although they have low normalized *B* factors, with the exception of W10 (Table 4). Among them W11 (W36), found in all apo crystal structures, forms hydrogen bonds to Ser117 O^{β}/O^{γ}, Thr118 N^{β} and Thr119 N^{β}/O^{γ}. In ligand-bound structures, these residues interact with the ligands and stabilize them (Klabunde *et al.*, 2000; Tomar *et al.*, 2012); however, these water molecules (W11 and W36) are absent in all of the holo structures. Analyses of the MD-simulation trajectories of all apo and mutant structures reveal high residential frequency for both W11 and W36, except in PDB entries 1ttb and 1fhn (Table 4), which may be owing to the mutational effect. A recent study has also revealed the role of W11 and W36 in protein–ligand complexation (Banerjee *et al.*, 2013). Water molecule W9 interacts with Ser112 N^{β}/O^{γ} and Leu110 O^{β} of the G strand and

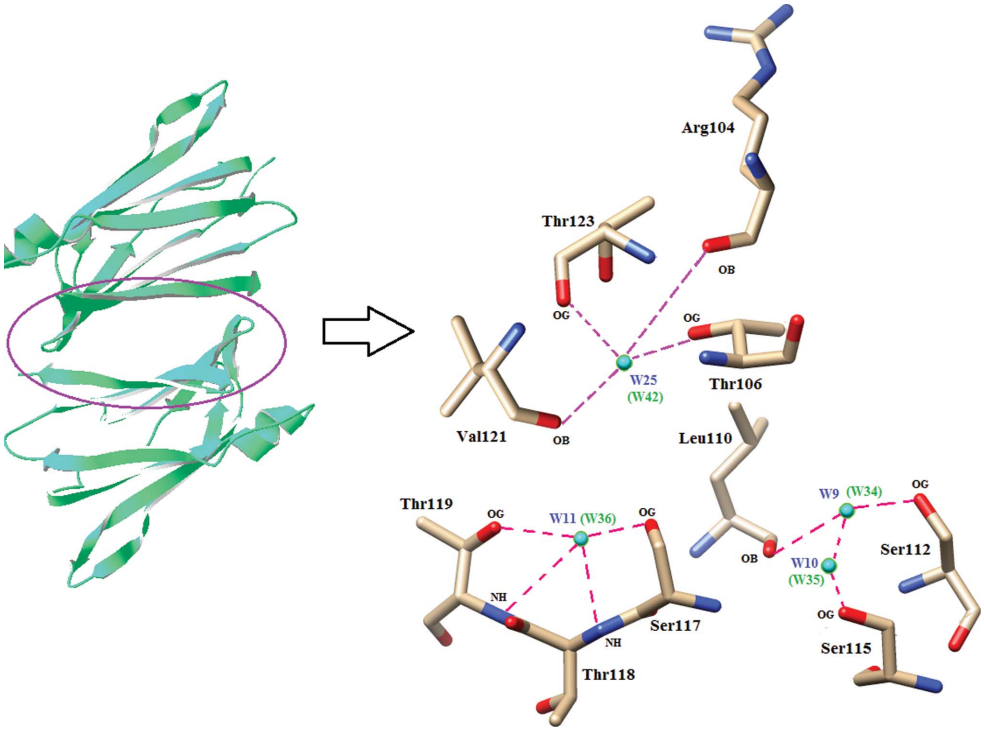

Figure 3
Hydrogen-bond interactions of the conserved water molecules in the ligand-binding cavity of hTTR.

Table 4
Normalized *B* factor (Norm. *B*) and solvent-accessible surface area (SASA; Å^2) of the conserved water molecules in the vicinity of the ligand-binding cavity in the X-ray and neutron-diffraction crystal structures of hTTR.

The numbers in the first row indicate the IDs of the water molecules in the respective crystals; normalized *B* factors of the O atoms of these water molecules are given in the second row and their SASAs are listed in the third row. RF, residential frequency. The absence of a water molecule in a structure is indicated by an asterisk and a dash denotes that the data are not available. 'o' denotes that the water position is occupied by some other atom of the protein in the crystal.

	W9	W10	W11	W25	W34	W35	W36	W42
1f41	31	32	36	171	29	30	33	59
Norm. *B*	0.14	1.09	0.67	0.67	0.35	0.52	1.08	0.76
SASA	9.21	6.44	8.01	21.96	13.98	9.50	15.01	24.71
RF	87	100	100	100	90	80	92	100
2qgb	157 *A*	178 *A*	166 *A*	153 *A*	147 *B*	165 *B*	168 *B*	148 *B*
Norm. *B*	0.52	1.16	0.74	0.87	0.25	1.04	1.06	1.03
SASA	3.77	7.01	4.29	4.09	4.01	6.79	5.73	3.77
RF	68	100	100	100	83	100	100	100
3cfm	143 *A*	167 *A*	135 *A*	151 *A*	142 *B*	145 *B*	127 *B*	152 *B*
Norm. *B*	−0.12	1.18	−0.34	0.06	0.05	0.48	−0.26	0.70
SASA	3.23	9.54	3.25	3.96	4.54	5.73	3.38	3.24
RF	50	100	100	100	90	100	100	100
3cbr	**197 *A***	**152 *A***	**132 *A***	**154 *A***	**170 *B***	**129 *B***	**195 *B***	**143 *B***
Norm. *B*	**0.24**	**0.37**	**−0.65**	**0.16**	**−0.65**	**−0.58**	**0.08**	**0.90**
SASA	**8.36**	**7.29**	**3.42**	**4.05**	**2.67**	**9.40**	**2.88**	**4.31**
RF	**20**	**100**	**100**	**100**	**60**	**100**	**87**	**72**
1tta	218 *A*	219 *A*	726 *A*	211 *A*	418 *B*	419 *B*	507 *B*	411 *B*
Norm. *B*	−0.47	−0.29	−0.83	−0.52	−0.20	0.28	0.42	−0.12
SASA	4.35	8.31	4.55	4.79	4.24	8.04	5.31	1.75
RF	88	100	100	100	90	100	95	90
3d7p	161 *A*	*	136 *A*	162 *A*	129 *B*	132 *B*	141 *B*	154 *B*
Norm. *B*	0.36	–	−0.65	0.33	−0.91	−0.68	1.01	1.07
SASA	10.06	–	3.85	4.06	2.61	8.70	4.83	4.33
3a4d	194 *A*	193 *A*	196 *A*	133 *A*	132 *B*	133 *B*	172 *B*	173 *B*
Norm. *B*	−0.06	−0.33	−0.42	0.46	0.50	−0.02	−0.86	0.30
SASA	3.82	7.49	4.24	3.44	6.75	5.80	2.08	3.70
2g4g	166 *A*	158 *A*	195 *A*	146 *A*	159 *B*	179 *B*	155 *B*	133 *B*
Norm. *B*	0.53	1.21	1.05	0.69	1.25	0.72	0.43	0.67
SASA	3.78	7.72	4.89	4.70	4.19	7.54	4.93	3.15
3i9p	7 *A*	197 *A*	157 *A*	156 *A*	131 *B*	129 *B*	130 *B*	156 *B*
Norm. *B*	−0.15	0.86	−0.39	0.05	−0.09	0.24	−0.32	0.81
SASA	2.58	10.07	2.91	3.63	4.02	4.34	3.32	3.49
3u2i	131 *A*	165 *A*	130 *A*	137 *A*	132 *B*	140 *B*	130 *B*	142 *B*
Norm. *B*	−0.30	0.20	−0.40	−0.10	−0.10	0.20	−0.40	0.20
SASA	3.71	10.08	4.23	3.34	4.43	7.47	5.20	5.33
3u2j	**147 *A***	**148 *A***	*	**129 *A***	**147 *B***	**148 *B***	**128 *B***	**10 *B***
Norm. *B*	**−0.34**	**−0.26**	–	**−0.14**	**−0.48**	**−0.50**	**−0.37**	**−0.06**
SASA	**3.48**	**15.93**	–	**3.77**	**4.56**	**11.93**	**4.14**	**3.54**
4pvm	**203 *A***	**222 *A***	**245 *A***	**224 *A***	**207 *B***	**210 *B***	**216 *B***	**209 *B***
Norm. *B*	**−0.18**	**−0.02**	**−0.17**	**0.03**	**−0.05**	**0.18**	**−0.04**	**0.10**
SASA	**4.32**	**7.67**	**4.68**	**4.11**	**3.84**	**8.07**	**5.13**	**3.99**
RF	**78**	**92**	**100**	**100**	**86**	**89**	**100**	**100**
1f86	441 *A*	462 *A*	LIG	452 *A*	542 *B*	544 *B*	LIG	535 *B*
Norm. *B*	0.06	0.94	–	0.44	−0.22	0.19	–	0.68
SASA	4.00	7.32	–	3.19	3.64	9.50	–	1.69
RF	20	100	–	45	70	85	–	100
1x7s	6049	6050	6024	6023	6100	6098	6111	6039
Norm. *B*	0.32	0.41	0.15	0.52	1.13	0.42	0.56	0.33
SASA	3.67	6.09	3.79	3.96	3.62	6.79	4.13	4.16
RF	60	100	100	100	50	100	93	100
3bt0	168 *A*	*	171 *A*	146 *A*	163 *B*	*	165 *B*	150 *B*
Norm. *B*	1.09	–	0.51	0.53	0.99	–	1.26	0.62
SASA	5.11	–	4.17	3.94	6.92	–	3.82	3.84
1x7t	6071	6094	6125	6075	6080	6033	6123	6036
Norm. *B*	0.49	0.89	−0.02	0.55	0.51	0.80	1.20	0.99
SASA	3.73	6.63	3.47	3.53	3.30	5.36	6.13	2.17
RF	50	100	100	100	70	100	100	
1sok	140 *A*	141 *A*	142 *A*	200 *A*	137 *B*	138 *B*	139 *B*	148 *B*
Norm. *B*	0.66	2.32	2.32	2.58	0.50	2.28	2.03	1.44
SASA	5.30	14.21	2.52	3.19	4.96	6.88	2.64	3.44
1ttb	425 *A*	426 *A*	*	320 *A*	465 *B*	442 *B*	451 *B*	420 *B*
Norm. *B*	−0.17	0.44	–	−0.32	0.04	0.60	0.17	0.80
SASA	3.42	9.49	–	3.57	3.50	9.23	5.27	2.61
RF	100	100	45	100	100	100	55	100

Table 4 (continued)

	W9	W10	W11	W25	W34	W35	W36	W42
1ttc	237	238	241	217	437	438	*	417
Norm. *B*	−0.11	0.89	1.28	0.94	0.62	1.10	—	0.78
SASA	3.78	6.92	4.30	4.48	3.60	11.2	—	2.39
1etb	209 *A*	208 *A*	215 *A*	141 *A*	197 *B*	196 *B*	211 *B*	148 *B*
Norm. *B*	0.52	0.91	0.41	0.56	−0.31	1.37	1.38	0.82
SASA	4.33	7.72	7.28	4.33	4.24	9.80	6.36	2.43
1fhn	138 *A*	151 *A*	Mut	135 *A*	134 *B*	137 *B*	Mut	139 *B*
Norm. *B*	0.02	0.10	—	−0.18	−0.25	0.25	—	0.32
SASA	5.55	10.31	—	3.12	5.07	9.94	—	2.77
RF	100	100	—	100	45	100	—	
2b16	2076	2103	LIG	2075	2061	2091	LIG	2026
Norm. *B*	−0.69	0.24	—	−0.81	−0.26	0.38	—	−0.92
SASA	3.97	7.65	—	5.14	4.24	11.05	—	4.89
2qgc	165 *A*	223 *A*	LIG	130 *A*	173 *B*	192 *B*	LIG	168 *B*
Norm. *B*	0.08	0.90	—	0.75	0.08	0.78	—	0.95
SASA	4.44	9.24	—	4.09	3.74	6.12	—	3.99
RF	75	100	—	100	85	100	—	100
3esp	132 *A*	178 *A*	LIG	*	143 *B*	159	LIG	*/o
Norm. *B*	−0.01	0.96	—	—	0.18	0.63	—	—
SASA	4.42	8.11	—	—	4.37	8.27	—	2.49
RF	65	100	—	100	80	100	—	70
3eso	135 *A*	185 *A*	LIG	134 *A*	132 *B*	*	LIG	143 *B*
Norm. *B*	0.21	1.14	—	0.82	0.34	—	—	−0.49
SASA	4.13	7.13	—	3.26	5.08	—	—	4.02
3esn	128 *A*	191 *A*	LIG	171 *A*	132 *B*	*	LIG	169 *B*
Norm. *B*	0.40	0.54	—	0.55	0.59	—	—	1.16
SASA	3.82	7.63	—	3.28	4.37	—	—	3.78
3cn4	151 *A*	189 *A*	*	133 *A*	169 *B*	*	LIG	138 *B*
Norm. *B*	0.79	1.78	—	0.91	0.77	—	—	1.15
SASA	4.03	8.84	—	4.22	4.59	—	—	4.15
RF	50	100	—	100	75	75	—	100
2fbr	139	135	LIG	43	97	77	LIG	154
Norm. *B*	1.41	0.85	—	0.53	0.73	0.39	—	1.33
SASA	7.79	3.28	—	3.27	8.86	2.50	—	3.79
2qgd	214 *A*	249 *A*	LIG	210 *A*	217 *B*	228 *B*	LIG	213 *B*
Norm. *B*	0.28	0.96	—	0.69	0.33	0.85	—	1.07
SASA	3.58	9.61	—	3.61	3.71	7.98	—	3.51
2qge	227 *A*	245 *A*	LIG	211 *A*	224 *B*	226 *B*	LIG	222 *B*
Norm. *B*	0.57	0.64	—	0.80	0.45	0.53	—	0.70
SASA	4.10	6.98	—	4.53	3.67	5.48	—	3.78
RF	65	100	—	100	85	100	—	100
3cn0	152 *A*	179 *A*	LIG	135 *A*	146 *B*	162 *B*	LIG	140 *B*
Norm. *B*	0.77	1.86	—	0.72	0.82	1.53	—	1.03
SASA	3.46	6.22	—	4.18	3.65	7.56	—	4.28
3cn1	148 *A*	174 *A*	LIG	130 *A*	142 *B*	163 *B*	LIG	136 *B*
Norm. *B*	0.51	1.22	—	0.73	0.54	1.73	—	1.29
SASA	3.83	8.03	—	4.16	3.78	9.22	—	4.17
Average Norm. *B*	**0.23**	**0.77**	**0.19**	**0.45**	**0.23**	**0.56**	**0.47**	**0.66**
Average SASA	**4.60**	**8.30**	**4.34**	**4.48**	**4.65**	**7.86**	**5.02**	**4.18**
Average RF	**65**	**99**	**94**	**96**	**78**	**95**	**91**	**95**

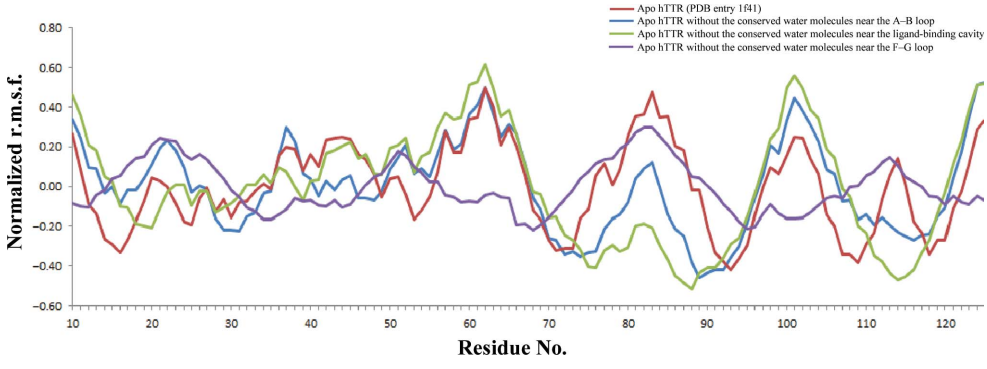

Figure 4
Normalized root-mean-square fluctuation (r.m.s.f.) of the C^{α} atoms of hTTR plotted against residue number, showing the changes in the backbone flexibility of the protein monomer in the presence and the absence of different conserved water molecules during dynamics.

W10. The water molecule W10 also forms a hydrogen bond to Ser115 O^{γ} of the H strand, whereas water molecule W25 forms hydrogen bonds to Arg104 O^{β}, Thr106 O^{γ} of the G strand and Val121 O^{β} and Thr123 O^{γ} of the H strand. During MD simulations, conservation of the hydrogen-bond interaction has been observed with high residential frequency (∼90–100%) for the water molecules W10, W11, W25, W35 and W36, although it seems low for

Table 5
Normalized *B* factor (Norm. *B*) and solvent-accessible surface area (SASA; Å^2) of the conserved water molecules in the interface of the two monomers in the X-ray and neutron diffraction crystal structures of hTTR.

The numbers in the first row indicate the IDs of the water molecules in the respective crystals; normalized *B* factors of the O atoms of these water molecules are given in the second row and their SASAs are listed in the third row. RF, residential frequency. The absence of a water molecule in a structure is indicated by an asterisk and a dash denotes that the data are not available.

	W5	W6	W7	W8	W30	W31	W32	W33
1f41	13	15	18	20	14	16	17	19
Norm. *B*	0.29	−0.43	−0.36	−0.06	1.05	−0.62	−0.46	−0.30
SASA	19.91	3.35	22.68	39.61	35.85	6.79	25.95	3.33
RF	100	100	85	100	100	100	100	100
2qgb	139 *A*	146 *A*	148 *A*	149 *A*	175 *A*	144 *B*	129 *B*	143 *B*
Norm. *B*	0.72	−0.39	−0.17	−0.14	0.93	−0.51	−0.33	−0.65
SASA	1.72	1.36	4.61	2.09	3.93	1.54	2.60	2.01
RF	100	100	100	100	100	100	100	100
3cfm	162 *A*	131 *A*	141 *A*	132 *A*	170 *B*	135 *B*	129 *B*	133 *B*
Norm. *B*	0.18	−0.50	−0.58	−0.24	1.05	−0.56	−0.60	−0.01
SASA	1.64	1.14	2.89	1.93	4.18	1.15	2.03	1.85
RF	83	100	100	100	85	100	100	100
3cbr	**228 *A***	**138 *A***	**167 *B***	*	**130 *B***	**His90 *B***	**133 *A***	**128 *B***
Norm. *B*	**0.45**	**0.09**	**1.30**	–	**0.13**	–	**−0.25**	**−0.93**
SASA	**1.87**	**1.23**	**2.07**	–	**4.56**	–	**4.16**	**1.92**
RF	**100**	**100**	**25**	**45**	**85**	**0**	**0**	**100**
1tta	213 *A*	204 *A*	220 *A*	400 *A*	420 *B*	404 *B*	413 *B*	200 *B*
Norm. *B*	−1.62	−1.52	−1.01	−0.46	−1.12	−1.02	−0.86	−0.49
SASA	2.63	1.09	4.10	2.27	4.56	1.66	2.55	2.29
RF	100	100	100	100	83	100	100	100
3d7p	140 *A*	143 *A*	176 *A*	*	131 *B*	His90 *B*	143 *B*	128 *B*
Norm. *B*	−0.34	−0.37	1.14	–	0.05	–	−0.03	−1.04
SASA	2.32	1.18	3.88	–	4.56	–	3.53	1.93
3a4d	181 *A*	159 *A*	132 *A*	158 *A*	139 *A*	174 *B*	137 *B*	143 *B*
Norm. *B*	0.08	−1.23	−0.77	−0.97	2.33	−1.00	−1.11	−0.41
SASA	2.27	1.49	3.50	1.93	19.12	1.63	2.51	1.92
2g4g	156 *A*	134 *A*	140 *A*	157 *A*	150 *B*	144 *B*	132 *B*	136 *B*
Norm. *B*	0.77	0.06	−0.61	0.14	0.13	−0.35	−0.77	−0.54
SASA	2.41	1.50	3.49	2.07	5.35	1.40	2.08	1.95
3i9p	198 *A*	134 *A*	6 *A*	146 *A*	217 *A*	126 *B*	143 *B*	145 *B*
Norm. *B*	0.84	−0.88	−0.89	0.02	1.15	−0.86	0.12	0.44
SASA	1.42	0.96	2.35	1.40	2.51	1.19	1.81	1.19
3u2i	142 *A*	3 *A*	134 *A*	129 *A*	149 *B*	8 *B*	131 *B*	129 *B*
Norm. *B*	0.10	−0.70	−0.30	−0.50	0.10	−0.80	−0.50	−0.40
SASA	2.42	1.76	5.10	2.10	6.85	1.78	2.57	2.03
3u2j	*	**3 *A***	**133 *A***	**4 *A***	*	**140 *B***	*	**134 *B***
Norm. *B*	–	**0.12**	**−0.65**	**−0.17**	–	**0.09**	–	**−0.06**
SASA	–	**1.89**	**5.39**	**2.19**	–	**2.03**	–	**1.95**
4pvm	**240 *A***	**207 *A***	**216 *A***	**204 *A***	**219 *B***	**203 *B***	**204 *B***	**206 *B***
Norm. *B*	**0.27**	**−0.38**	**−0.05**	**−0.33**	**0.14**	**−0.40**	**−0.26**	**−0.20**
SASA	**2.62**	**1.92**	**3.95**	**2.22**	**4.87**	**1.93**	**2.73**	**2.23**
RF	**77**	**100**	**89**	**100**	**86**	**100**	**82**	**100**
1f86	460 *A*	445 *A*	438 *A*	463 *A*	536 *B*	627 *B*	534 *B*	531 *B*
Norm. *B*	−0.10	−0.44	−0.39	−0.16	0.08	−0.42	−0.54	−0.58
SASA	1.68	0.86	3.21	2.18	3.41	1.45	2.16	2.00
RF	100	100	75	100	100	100	100	100
1x7s	6002	6003	6036	6022	6038	6016	6004	6037
Norm. *B*	0.53	−0.20	0.44	−0.23	0.37	0.17	−0.19	0.14
SASA	1.81	1.21	3.30	2.03	3.82	1.46	2.33	1.91
RF	100	100	93	100	100	100	100	100
3bt0	*	132 *A*	137 *A*	184 *A*	*	128 *B*	137 *B*	144 *B*
Norm. *B*	–	−0.31	−0.10	1.26	–	−0.35	0.15	−0.19
SASA	–	1.36	4.71	2.18	–	1.20	2.86	1.96
1x7t	6055	6017	6022	6026	6115	6020	6021	6018
Norm. *B*	0.28	−0.57	−0.43	0.35	0.84	−0.44	−0.64	0.00
SASA	1.66	1.48	2.63	2.01	4.40	1.54	2.04	1.80
RF	100	100	75	100	100	100	100	100
1sok	133 *A*	135 *A*	136 *A*	137 *A*	134 *A*	133 *B*	134 *B*	135 *B*
Norm. *B*	1.29	−0.22	−0.15	0.22	2.34	−0.03	−0.42	0.28
SASA	2.39	1.58	3.01	1.78	5.35	1.55	2.33	1.68
1ttb	324 *A*	407 *A*	474 *A*	424 *A*	437 *B*	307 *B*	327 *B*	301 *B*
Norm. *B*	1.53	−0.63	0.09	−0.66	−0.51	−0.78	−0.48	−0.73
SASA	2.28	1.63	3.97	2.25	5.33	1.63	2.58	2.25
RF	100	100	75	100	60	100	100	100

W9 and W34 (Table 4). The decrease is probably owing to rotation of the side-chain atoms of Ser112. These two water molecules (W9 and W34) shift ∼3 Å (from the positions observed in the crystal structures) to maintain the hydrogen bond to Ser112 (O^{γ}). Again, free-energy calculation reveals that these water clusters (W9, W10, W11 and W25 in chain *A* and W34, W35, W36 and W42 in chain *B*) contribute ∼250 cal mol^{-1} towards the structural stability of the protein. Further, a 10 ns MD-simulation study (with PDB entry 1f41 as the template structure), after the removal of those eight water molecules, reveals the appearance of water molecules at the W10, W25 and W34 positions. During the simulation, significant change in the normalized r.m.s.f. is noticed for the ligand-binding residues (Fig. 4). Comparison of the ligand-binding cavities of the apo hTTR (PDB entry 1f41) structure with the structure obtained after removing these conserved water molecules (after 10 ns MD simulations of both structures) reveals a cumulative contraction in volume of ∼150 Å^3 for the two pockets (Fig. 5). A careful observation reveals that the decrease in the volume of the pockets is owing to collapse of the side chains of some ligand-binding residues which were hydrogen-bonded to those conserved water molecules. Thus, these water molecules may play some role in maintaining the geometry of the ligand-binding cavity of the protein.

In multimeric proteins, generally the interfacial water molecules form an extensive hydrogen-bonding network for interlinking the monomers. Such water-mediated interfacial stabilization has been observed in many biological systems (Janin & Chothia, 1990; Miller & Krause, 1996; Shankar *et al.*, 2007). In the crystal structure of hTTR, eight conserved water molecules (W5, W6, W7 and W8 and their counter pairs W32, W31, W30 and W33) have been identified at the interface of the two monomers. These water molecules are stable at their positions (with low normalized *B* factors) and are completely buried (Table 5), except for W7 and W30, which are semi-exposed to the bulk solvent with SASA values of

Table 5 (continued)

	W5	W6	W7	W8	W30	W31	W32	W33
1ttc	222	406	440	200	240	206	422	400
Norm. *B*	0.04	−1.19	0.62	−0.40	−0.49	−0.80	0.01	−0.20
SASA	2.71	1.74	4.17	2.38	4.19	1.65	2.33	2.42
1etb	143 *A*	146 *A*	161 *A*	150 *A*	150 *B*	131 *B*	135 *B*	130 *B*
Norm. *B*	2.27	−0.76	0.50	−0.84	1.23	−0.76	−0.41	−0.90
SASA	2.27	1.60	4.26	2.29	5.35	1.53	2.26	2.22
1fhn	161 *A*	130 *A*	136 *A*	140 *A*	169 *A*	128 *B*	133 *B*	132 *B*
Norm. *B*	0.45	−0.52	−0.21	−0.03	0.72	−0.54	−0.31	−0.37
SASA	2.19	1.79	4.77	2.34	5.92	1.44	2.80	1.96
RF	100	100	100	100	100	100	100	100
2b16	2059	2024	2019	2058	2025	2023	2022	2042
Norm. *B*	−0.89	−1.75	−1.16	−1.41	−1.04	−1.65	−1.40	−1.67
SASA	2.36	1.41	4.12	2.14	3.70	1.65	2.52	1.93
2qgc	190 *A*	135 *A*	171 *A*	173 *A*	195 *B*	130 *B*	163 *B*	166 *B*
Norm. *B*	0.51	−0.59	−0.27	−0.47	0.58	−0.58	−0.46	−0.55
SASA	1.90	1.38	3.54	2.20	4.16	1.62	2.46	2.13
RF	100	100	100	100	100	100	100	100
3esp	166 *A*	131 *A*	143 *A*	*	154 *B*	136 *B*	139 *B*	129 *B*
Norm. *B*	0.72	−0.29	−0.10	—	0.23	−0.38	−0.12	−0.67
SASA	2.17	1.32	4.38	—	4.62	1.31	2.43	2.18
RF	100	100	100	0	80	100	83	100
3eso	189 *A*	138 *A*	152 *A*	*	175 *B*	134 *B*	139 *B*	229 *B*
Norm. *B*	1.31	−0.21	−0.04	—	0.56	−0.46	0.01	−0.55
SASA	1.63	1.34	3.72	—	2.97	1.57	2.12	2.16
3esn	*	*	153 *A*	*	213 *B*	170 *B*	177 *B*	127 *B*
Norm. *B*	—	His90 *B*	0.40	—	0.74	−0.40	0.10	−0.59
SASA	—	2.47	4.88	—	3.00	1.65	2.73	2.14
3cn4	*	144 *A*	149 *A*	161 *B*	158 *B*	150 *B*	141 *B*	134 *B*
Norm. *B*	—	−0.12	−0.06	0.61	0.60	−0.20	0.15	−0.60
SASA	—	1.47	4.96	2.16	5.57	1.43	2.66	2.12
RF	100	100	88	100	100	100	100	100
2fbr	148	90	21	35	98	30	29	70
Norm. *B*	1.27	0.36	−0.30	−0.19	0.76	−0.34	−0.34	−0.37
SASA	2.18	1.37	3.90	1.88	4.13	1.34	2.85	1.88
2qgd	*	203 *A*	228 *A*	218 *A*	225 *B*	202 *B*	211 *B*	206 *B*
Norm. *B*	—	−0.36	−0.17	−0.18	0.80	−0.45	−0.30	−0.42
SASA	—	1.51	5.08	2.22	5.47	1.66	2.81	2.13
2qge	236 *A*	205 *A*	213 *A*	212 *A*	231 *B*	204 *B*	207 *B*	209 *B*
Norm. *B*	0.31	−0.57	−0.27	−0.30	0.64	−0.43	−0.36	−0.52
SASA	2.13	1.41	3.54	2.06	4.77	1.34	2.25	1.99
RF	100	100	100	100	100	100	100	100
3cn0	*	130 *A*	143 *A*	147 *A*	153 *B*	132 *B*	136 *B*	133 *B*
Norm. *B*	—	−0.28	0.01	0.51	1.06	−0.37	−0.05	−0.46
SASA	—	1.35	4.45	2.25	5.80	1.57	2.72	2.12
3cn1	*	138 *A*	150 *A*	161 *A*	168 *B*	135 *B*	145 *B*	133 *B*
Norm. *B*	—	−0.22	0.07	0.80	0.69	−0.43	0.27	−0.50
SASA	—	1.40	4.98	2.15	4.93	1.68	2.87	2.17
Average Norm. *B*	**0.45**	**−0.48**	**−0.14**	**−0.14**	**0.54**	**−0.52**	**−0.33**	**−0.44**
Average SASA	**2.82**	**1.49**	**4.55**	**2.49**	**6.11**	**1.77**	**2.47**	**2.05**
Average RF	**97**	**100**	**87**	**88**	**92**	**93**	**96**	**100**

3.0–5.0 Å^2. Among them, W5 (W32) interacts with Phe87 O$^{\beta}$ and the catalytic Glu89 (O$^{\varepsilon 2}$) of chain *A* and Thr96 (N$^{\beta}$/O$^{\gamma}$) of chain *B*, whereas the water molecule W30 (W7) forms hydrogen bonds to Phe87 O$^{\beta}$ of chain *A* and Thr96 O$^{\beta}$ of chain *B* (Fig. 6). These water molecules (W5 and W30), interconnected through a hydrogen bond, hold the F strands of the monomers together. During dynamics, both the water molecules W5 and W32 occupy the hydrophilic sites with very high residential frequency (>95%) and maintain the hydrogen-bonding network, which signifies their importance in the structure. The values are slightly lower for W7 and W30 (87 and 92%), which may be owing to the displacement (∼3–3.5 Å) of these water molecules from their positions by hydrophobic interaction of the ring atoms of Phe87 in some structures. Water molecule W8 (W33), which bridges the thyroxin-binding residues Thr118 (O$^{\gamma}$) and Thr119 (O$^{\beta}$) of chain *A* with the catalytic His88 (N$^{\varepsilon}$) of chain *B*, may play some role in the intermonomeric recognition between the catalytic and thyroxin-binding sites (Banerjee & Mukhopadhyay, 2015). During dynamics, W33 is found in all of the snapshots. The water molecule W8 also shows 100% residential frequency in all the MD-simulation trajectories except for the 3esp and 3cbr structures. Possibly, the rotation of the His88 side chain destabilizes the water molecule at this position. Again, W6 (W31), which forms hydrogen bonds to Val94 N$^{\beta}$ of chain *A* and Tyr116 OH of chain *B*, also interacts with His90 O$^{\beta}$ of chain *A* and Glu92 O$^{\beta}$ of chain *B*. These residues (His90, Glu92 and Val94) play important roles in the proteolytic activity of hTTR (Liz *et al.*, 2012). Both W6 and W31 maintain the hydrogen-bond architecture with high residential frequency in all of the MD-simulation trajectories. Free-energy calculation shows that the interfacial water molecules (W5, W6, W7, W8, W32, W31, W30 and W33) contribute ∼780 cal mol^{-1} towards the stabilization of the protein, which is ∼22.81% of the total contribution of all of the water molecules present in the crystal (Supplementary Table S1).

It is well known that *in vitro* low pH can promote destabilization of the tetrameric hTTR structure, resulting in its dissociation into monomers. To study the effect of pH on the distribution of conserved water molecules and their hydrogen-bonding network at the monomer–monomer interface, we have compared the conserved water positions in apo hTTR X-ray and neutron diffraction structures obtained at pH ≥ 5.0 (PDB entries 1f41, 3cfm, 3i9p, 3u2j, 4pvm, 3a4d, 1tta and 2qgb) with those crystallized below pH 5.0 [PDB entries 3cbr (pH 3.5), 3d7p (pH 4.0) and 2g4g (pH 4.6)]. Careful observation of the structures reveals that the water molecule W7 is absent in the structure crystallized at pH ≤ 4.0 (PDB entry 3d7p), whereas W31 and W32 (along with W7) are absent in the 3cbr crystal (pH 3.5). MD simulations of 3cbr also support this observation and no water molecule is found to occupy these positions during the simulation period (Table 5). In the crystal structure 3d7p no water molecule is found at the W8 position; however, one water molecule is observed very close to this position (∼1.8 Å away) in 3cbr, although it is involved in a different hydrogen-bonding network and bridges the O$^{\beta}$ and O$^{\gamma}$ atoms

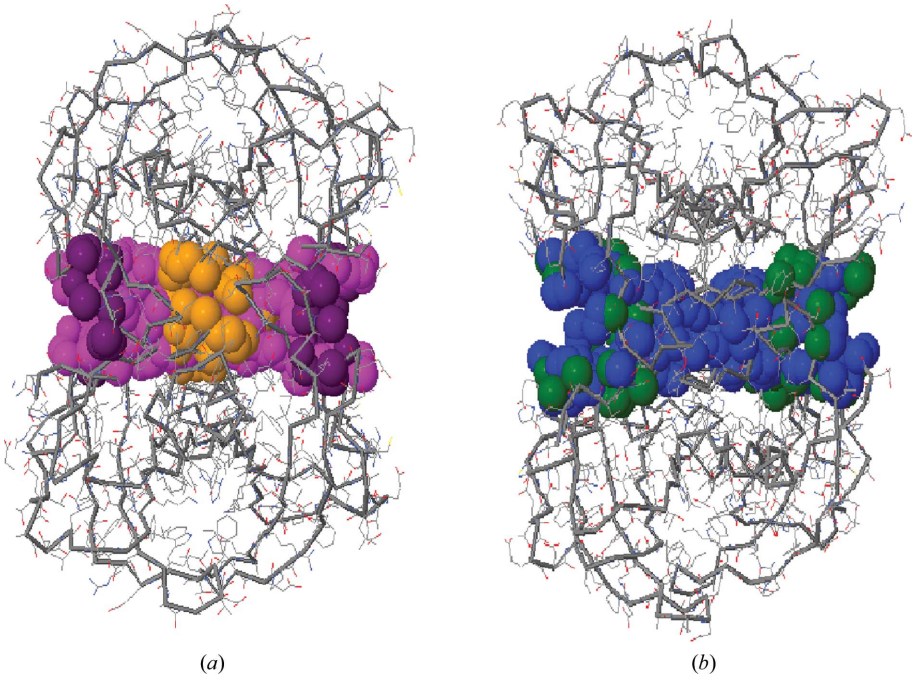

Figure 5
Ligand-binding pockets after 10 ns MD simulation of hTTR in (*a*) the presence and (*b*) the absence of the conserved water molecules.

of Thr118 (chain *A*) with the $O^{\varepsilon 1}$ and $O^{\varepsilon 2}$ atoms of Glu89 (chain *B*). Structural analysis reveals significant changes in the side-chain conformations of His88 and Glu89 in the absence of the water molecule W8 in these structures crystallized at low pH. During MD simulations of 3cbr, no water is found at these (W7, W31 and W32) hydrophilic sites. The water molecule at W30 and the water molecule found near the W8 position also show a sharp decrease in their residential frequencies (Table 5). Some other water molecules near the ligand-binding site (W9 and W10), the A–B loop (W39) and the dimer–dimer interface (W20 and W22) also show a significant decrease in their residential frequencies. Thus, it appears that a decrease in pH results in the displacement of some conserved water molecules at the monomer–monomer interface and causes the reorientation of some residues which could affect the

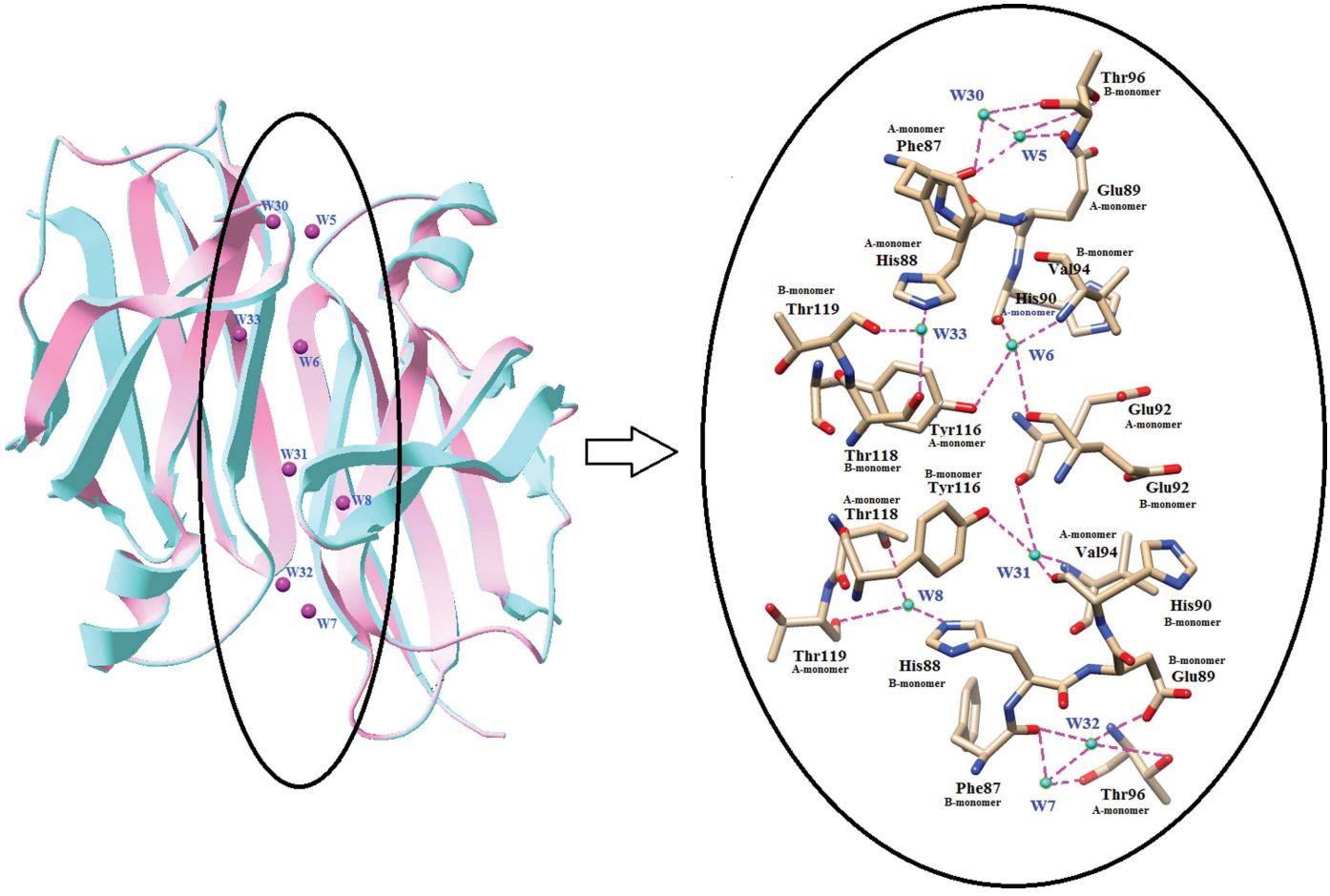

Figure 6
The interfacial conserved water molecules stabilizing the dimeric structure of hTTR. The left panel shows the location of the water molecules and the right panel shows their interaction.

Table 6
Normalized *B* factor (Norm. *B*) and solvent-accessible surface area (SASA; Å^2) of the conserved water molecules in the region encompassed by the A–B loop, B and E strands and the α-helix in the X-ray and neutron diffraction crystal structures of hTTR.

The numbers in the first row indicate the IDs of the water molecules in the respective crystals; normalized *B* factors of the O atoms of these water molecules are given in the second row and their SASAs are listed in the third row. RF, residential frequency. The absence of a water molecule in a structure is indicated by an asterisk and a dash denotes that the data are not available.

	W12	W13	W15	W16	W17	W20	W21	W22	W37	W39	W40	W41	W43
1f41	39	40	43	44	49	63	64	65	37	51	52	54	60
Norm. *B*	0.04	−0.63	0.34	−0.08	0.19	1.04	1.09	0.63	−0.34	0.52	0.35	0.28	0.36
SASA	9.46	22.36	12.00	13.93	10.19	15.03	24.20	28.78	27.11	5.16	15.79	23.88	9.95
RF	100	100	100	100	100	100	100	100	100	72	100	100	90
2qgb	134 *A*	144 *A*	147 *A*	129 *A*	171 *A*	164 *A*	168 *A*	170 *A*	138 *B*	141 *B*	153 *B*	140 *B*	*
Norm. *B*	0.23	−0.67	−0.11	0.17	0.32	0.12	0.96	0.32	0.26	1.05	1.15	0.32	—
SASA	3.68	2.63	1.62	3.51	7.56	3.00	8.44	5.24	2.38	4.02	4.84	4.00	—
RF	62	100	100	100	88	100	100	50	100	100	100	100	80
3cfm	160 *A*	138 *A*	137 *A*	130 *A*	180 *A*	146 *A*	156 *A*	161 *A*	132 *B*	130 *B*	149 *B*	143 *B*	*
Norm. *B*	0.07	−0.58	−0.36	−0.36	0.61	0.42	−0.34	−0.07	−0.27	0.80	0.67	0.29	—
SASA	3.16	3.69	1.98	2.57	5.66	4.34	9.11	6.94	2.50	2.91	3.60	4.05	—
RF	83	100	60	100	100	100	60	100	100	100	83	100	90
3cbr	**134 *A***	**225 *A***	**223 *A***	**233 *A***	**141 *A***	**155 *A***	**137 *A***	**212 *A***	**189 *B***	**139 *B***	**159 *B***	**172 *B***	**174 *B***
Norm. *B*	**−0.31**	**−0.71**	**−0.36**	**−0.38**	**0.19**	**0.39**	**0.07**	**1.10**	**1.37**	**0.73**	**2.43**	**1.23**	**0.74**
SASA	**4.59**	**3.90**	**1.71**	**3.33**	**5.61**	**5.56**	**8.81**	**7.69**	**6.36**	**4.82**	**13.83**	**11.00**	**1.26**
RF	**62**	**87**	**59**	**81**	**82**	**40**	**85**	**60**	**100**	**89**	**100**	**100**	**100**
1tta	203 *A*	207 *A*	780 *A*	212 *A*	217 *A*	216 *A*	781 *A*	302 *A*	407 *B*	412 *B*	417 *B*	403 *B*	405 *B*
Norm. *B*	−0.43	−1.35	−0.31	0.39	−0.70	−0.03	0.45	0.39	0.10	−0.11	−0.53	−0.51	−0.53
SASA	5.31	2.99	2.23	4.15	6.15	5.78	8.72	13.24	2.49	5.54	5.88	4.91	2.44
RF	100	100	40	100	100	100	92	100	100	100	100	100	90
3d7p	141 *A*	131 *A*	134 *A*	129 *A*	171 *A*	146 *A*	149 *A*	172 *A*	*	136 *B*	Tyr78 *B*	*	148 *B*
Norm. *B*	−0.39	−0.74	−0.60	−0.72	0.78	0.20	−0.41	2.54	—	1.68	—	—	0.57
SASA	3.59	4.19	1.64	2.89	6.02	5.49	7.91	7.45	—	4.18	—	—	1.66
3a4d	198 *A*	130 *A*	150 *A*	157 *A*	189 *A*	170 *A*	136 *A*	183 *A*	178 *B*	138 *B*	145 *B*	177 *B*	142 *B*
Norm. *B*	−0.05	−1.22	−0.50	−0.30	0.86	0.26	0.16	0.60	−1.24	0.08	−0.06	−0.71	−0.19
SASA	3.38	3.28	1.83	4.16	5.46	5.36	9.09	8.60	3.40	6.59	4.33	3.00	2.03
2g4g	159 *A*	131 *A*	135 *A*	150 *A*	160 *A*	151 *A*	168 *A*	148 *A*	130 *B*	140 *B*	177 *B*	146 *B*	139 *B*
Norm. *B*	−0.16	−0.66	−0.07	−0.12	0.47	0.38	0.89	−0.30	0.16	0.54	2.68	0.26	0.85
SASA	4.31	2.50	1.82	4.39	7.74	3.33	7.89	5.33	2.59	4.32	7.75	4.92	2.41
3i9p	176 *A*	144 *A*	2 *A*	165 *A*	174 *A*	183 *A*	182 *A*	190 *A*	133 *B*	9 *B*	159 *B*	140 *B*	176 *B*
Norm. *B*	−0.36	−0.99	−0.58	−0.25	0.17	0.28	0.85	2.87	−0.84	0.29	0.43	−0.23	1.50
SASA	3.47	1.86	1.49	2.72	2.77	3.52	8.32	7.79	2.16	3.17	5.59	3.16	2.46
3u2i	143 *A*	1 *A*	132 *A*	139 *A*	140 *A*	138 *A*	148 *A*	161 *A*	128 *B*	141 *B*	150 *B*	135 *B*	157 *B*
Norm. *B*	0.10	−0.70	0.00	0.00	0.00	0.20	0.40	0.50	0.00	0.30	0.60	0.00	0.20
SASA	5.24	3.10	1.69	4.42	7.08	6.76	8.01	14.51	3.48	4.96	7.03	5.29	3.38
3u2j	**139 *A***	**6 *A***	**7 *A***	**130 *A***	*	**138 *A***	*	*	**9 *B***	**129 *B***	*	**130 *B***	**135 *B***
Norm. *B*	**−0.02**	**−0.19**	**0.35**	**−0.38**	—	**−0.26**	—	—	**−0.11**	**0.18**	—	**0.04**	**0.02**
SASA	**3.99**	**2.57**	**4.10**	**5.27**	—	**5.25**	—	—	**2.82**	**4.65**	—	**4.00**	**1.62**
4pvm	**241 *A***	**202 *A***	**210 *A***	**211 *A***	**215 *A***	**205 *A***	**226 *A***	*	**208 *B***	**213 *B***	**224 *B***	**226 *B***	**221 *B***
Norm. *B*	**0.10**	**−0.31**	**0.00**	**0.02**	**0.30**	**0.10**	**0.25**	—	**−0.01**	**0.26**	**0.29**	**0.06**	**0.10**
SASA	**4.81**	**3.22**	**1.82**	**3.67**	**9.10**	**7.05**	**9.31**	—	**3.34**	**5.79**	**6.32**	**4.64**	**2.74**
RF	**87**	**92**	**100**	**100**	**89**	**67**	**100**	**71**	**100**	**100**	**100**	**92**	**75**
1f86	440 *A*	431 *A*	436 *A*	435 *A*	455 *A*	458 *A*	442 *A*	439 *A*	529 *B*	543 *B*	549 *B*	537 *B*	539 *B*
Norm. *B*	−0.43	−0.72	−0.16	−0.22	−0.13	−0.06	0.22	−0.12	−0.32	0.47	0.43	−0.22	−0.01
SASA	3.42	2.61	1.31	2.75	6.69	2.74	5.89	4.53	2.66	3.30	2.92	2.61	1.88
RF	85	100	70	100	100	100	100	100	100	100	100	100	100
1x7s	6009	6040	6010	6001	6047	6042	*	6046	6032	6080	6062	6035	6033
Norm. *B*	−0.29	−0.13	0.22	−0.39	1.18	0.40	—	1.08	−0.13	0.48	1.30	−0.08	0.37
SASA	5.18	3.40	2.16	3.40	5.77	3.18	—	5.28	3.77	3.46	7.05	4.10	2.20
RF	90	100	95	100	10	100	65	85	100	100	100	70	100
3bt0	138 *A*	131 *A*	141 *B*	139 *A*	173 *A*	147 *A*	194 *A*	160 *A*	134 *B*	146 *B*	161 *B*	132 *B*	*
Norm. *B*	−0.10	−0.57	−0.90	0.27	1.00	0.26	2.05	0.87	−0.17	0.62	1.62	0.00	—
SASA	2.74	2.25	1.83	2.61	6.24	4.50	10.97	6.17	2.41	4.38	6.10	3.23	—
1x7t	6087	6006	6007	6010	6100	6040	6037	6005	6028	6054	6096	6106	6047
Norm. *B*	0.37	−0.61	−0.52	−0.73	1.64	0.50	0.01	0.22	−0.09	0.29	0.57	0.12	0.51
SASA	4.13	2.49	1.66	3.76	5.27	4.23	7.82	7.73	2.78	4.63	6.09	4.93	2.00
RF	100	100	85	100	100	100	93	100	100	100	100	100	100
1sok	143 *A*	144 *A*	146 *A*	147 *A*	149 *A*	156 *A*	157 *A*	158 *A*	140 *B*	145 *B*	146 *B*	147 *B*	149 *B*
Norm. *B*	1.30	−0.16	0.60	0.93	1.31	2.45	0.72	1.04	0.41	0.80	0.78	0.98	0.80
SASA	4.14	3.08	1.45	3.99	4.83	7.00	8.81	10.77	1.99	3.82	4.62	4.20	2.18
1ttb	305 *A*	316 *A*	605 *A*	323 *A*	421 *A*	434 *A*	606 *A*	306 *A*	416 *B*	428 *B*	491 *B*	405 *B*	414 *B*
Norm. *B*	0.80	−0.58	−0.42	0.11	1.61	1.12	−0.69	0.61	0.34	0.24	1.92	0.88	0.52
SASA	3.67	2.43	2.16	3.62	5.51	5.60	7.81	12.47	2.54	6.25	5.06	4.01	1.93
RF	75	100	50	100	100	35	55	100	100	100	100	100	25

Table 6 (continued)

	W12	W13	W15	W16	W17	W20	W21	W22	W37	W39	W40	W41	W43
1ttc	204	214	212	221	235	246	*	228	414	421	435	404	412
Norm. *B*	−0.18	−0.95	−0.36	−0.61	0.23	0.70	—	0.54	−0.43	1.12	0.55	−0.09	0.24
SASA	3.98	2.67	2.35	3.33	5.55	3.78	—	6.86	2.55	5.27	5.72	3.42	2.05
1etb	134 *A*	137 *A*	199 *A*	142 *A*	149 *A*	155 *A*	200 *A*	135 *A*	144 *B*	149 *B*	163 *B*	141 *B*	143 *B*
Norm. *B*	1.39	−0.28	−0.43	0.94	1.88	1.34	−2.33	1.23	1.17	0.11	2.98	2.11	−0.35
SASA	4.52	2.70	2.26	4.59	5.87	5.77	9.83	12.15	2.41	7.23	4.51	2.96	1.93
1fhn	145 *A*	132 *A*	134 *A*	141 *A*	152 *A*	146 *A*	153 *A*	166 *A*	136 *B*	140 *B*	148 *B*	141 *B*	146 *B*
Norm. *B*	0.00	−0.51	−0.31	−0.32	0.20	0.11	0.23	0.77	0.00	0.03	0.20	0.04	0.87
SASA	4.38	4.04	1.70	3.40	7.80	5.41	6.79	9.74	3.29	3.67	5.10	3.98	2.96
RF	100	100	70	100	100	77	100	100	100	100	100	100	65
2b16	2064	2031	2034	2035	2032	2071	2066	2065	2008	2013	2010	2005	*
Norm. *B*	−1.28	−1.52	−1.29	−1.39	−0.63	−0.72	−0.37	−0.20	−1.57	−1.13	−0.66	−1.35	—
SASA	7.58	3.06	1.84	3.84	7.01	6.47	10.56	11.39	2.39	4.58	5.67	5.66	—
2qgc	187 *A*	167 *A*	132 *A*	170 *A*	*	172 *A*	181 *A*	191 *A*	164 *B*	172 *B*	181 *B*	177 *B*	178 *B*
Norm. *B*	0.08	−0.61	0.13	0.20	—	0.68	0.71	0.29	−0.04	0.64	0.51	0.22	0.31
SASA	3.68	3.34	2.17	2.86	—	7.24	10.99	5.54	1.96	3.52	3.40	3.02	2.00
RF	100	100	100	100	85	100	100	100	100	100	100	90	75
3esp	141 *A*	*	193 *A*	135 *A*	*	139 *A*	147 *A*	142 *A*	*	148 *B*	*	134 *B*	145 *B*
Norm. *B*	0.14	—	0.28	−0.05	—	−0.04	0.80	0.05	—	0.66	—	−0.01	0.46
SASA	3.14	—	1.60	3.09	—	4.12	9.97	4.86	—	2.68	—	2.77	2.14
RF	100	80	100	100	95	78	100	100	85	100	100	85	65
3eso	161 *A*	131 *A*	139 *A*	136 *A*	146 *A*	147 *A*	157 *A*	145 *A*	133 *B*	149 *B*	160 *B*	142 *B*	*
Norm. *B*	0.25	−0.52	0.14	0.05	0.23	0.10	1.03	0.27	0.24	0.73	1.03	0.10	—
SASA	4.10	2.58	2.18	2.83	7.30	3.12	11.34	5.59	2.05	4.00	3.72	2.53	—
3esn	129 *A*	164 *A*	159 *A*	163 *A*	146 *A*	167 *A*	180 *A*	147 *A*	229 *B*	134 *B*	178 *B*	129 *B*	167 *B*
Norm. *B*	0.18	−0.61	−0.03	0.08	0.07	−0.09	1.05	0.09	0.04	0.89	1.11	0.19	0.49
SASA	4.30	2.64	1.72	2.86	6.16	2.73	10.85	5.29	1.88	3.22	3.52	2.52	2.05
3cn4	154 *A*	132 *A*	134 *A*	130 *A*	146 *A*	159 *A*	163 *A*	150 *A*	135 *B*	160 *B*	168 *B*	157 *B*	136 *B*
Norm. *B*	0.31	−0.55	0.36	0.39	0.65	0.05	1.24	0.24	0.26	0.95	1.52	0.29	0.51
SASA	4.35	2.22	1.71	3.49	6.13	21.15	8.93	6.34	2.27	5.23	4.89	3.88	2.43
RF	100	100	100	100	100	100	100	100	100	100	100	100	85
2fbr	55	9	15	26	19	20	40	45	33	87	149	64	179
Norm. *B*	0.14	−0.93	−0.43	0.01	−0.33	0.10	0.20	−0.15	−0.01	0.70	1.56	0.57	1.05
SASA	3.71	2.58	1.56	3.07	8.10	3.31	9.03	4.63	2.63	3.41	5.34	3.94	2.24
2qgd	227 *A*	205 *A*	212 *A*	207 *A*	231 *B*	211 *A*	237 *A*	241 *A*	207 *B*	218 *B*	224 *B*	212 *B*	215 *B*
Norm. *B*	0.16	−0.64	−0.11	−0.80	1.25	0.73	0.66	0.44	0.25	0.75	0.64	0.33	0.52
SASA	4.63	3.01	1.63	3.01	8.13	4.71	9.52	6.99	2.33	4.45	4.41	4.69	2.42
2qge	233 *A*	202 *A*	208 *A*	209 *A*	222 *A*	225 *A*	234 *A*	232 *A*	208 *B*	212 *B*	230 *B*	220 *B*	219 *B*
Norm. *B*	0.18	−0.58	−0.07	0.07	0.97	0.60	0.55	0.21	0.08	0.68	0.57	0.30	0.26
SASA	4.53	2.66	1.70	3.27	8.91	5.87	9.93	7.01	1.96	4.13	3.93	4.60	1.98
RF	100	83	95	100	100	75	100	85	100	100	100	100	75
3cn0	151 *A*	133 *A*	139 *A*	137 *A*	165 *A*	166 *A*	154 *A*	160 *A*	141 *B*	137 *B*	155 *B*	143 *B*	142 *B*
Norm. *B*	0.19	−0.58	0.22	0.16	1.49	0.29	1.08	0.36	0.47	0.75	1.30	0.51	0.65
SASA	3.58	2.83	1.75	3.66	7.40	3.96	10.50	6.13	2.51	4.99	5.38	3.69	2.36
3cn1	146 *A*	132 *A*	135 *A*	139 *A*	153 *A*	166 *A*	168 *A*	151 *A*	137 *B*	147 *B*	148 *B*	143 *B*	140 *B*
Norm. *B*	0.26	−0.66	0.08	−0.06	0.79	0.09	1.40	0.24	0.43	0.80	0.64	0.27	0.42
SASA	3.99	2.38	1.64	3.41	7.03	3.93	9.95	6.71	2.40	4.59	5.33	3.83	2.19
Average Norm. *B*	**0.07**	**−0.66**	**−0.16**	**−0.11**	**0.57**	**0.37**	**0.45**	**0.56**	**0.00**	**0.53**	**0.92**	**0.20**	**0.42**
Average SASA	**4.34**	**3.52**	**2.20**	**3.81**	**6.66**	**5.60**	**9.63**	**8.39**	**2.51**	**4.47**	**5.78**	**4.76**	**2.48**
Average RF	**90**	**96**	**82**	**99**	**90**	**85**	**90**	**91**	**99**	**97**	**99**	**96**	**81**

intermonomeric hydrogen-bonding network and possibly destabilize the hTTR association. This may be owing to changes in the protonation states of some acidic or basic residues (His88, Glu89, Glu92 *etc.*) present at the interface. The interfacial hydrogen-bonding network of MD-simulated structures of 1f41 (pH 7.0) and 3cbr (pH 3.5) is shown in Fig. 7.

Analyses of the X-ray and neutron diffraction structures and their corresponding MD-simulated structures have revealed the presence of six conserved water molecules (W12, W13, W16, W17, W20 and W22) in the region encompassed by the A–B loop, the B and E strands and the α-helix of monomer *A* (Fig. 1) and all have counter pairs (W41, W37, W39 and W40) in monomer *B* except for W20 and W22. The semi-accessible (3.0 < ASA < 5.0 Å^2) water molecule W12 (W41) interacts with Pro24 O$^{\beta}$ and Ile26 N$^{\beta}$ of the A–B loop situated at the dimeric interface during the formation and stabilization of the tetramer. It also forms hydrogen bonds to Tyr78 OH and W22. These two water molecules (W12 and W22) together with W20 form a triad cluster. A 10 ns MD-simulation study of the apo hTTR (1f41) structure without the water cluster (W12, W20 and W22) shows an increase in the normalized r.m.s.f. of the residues belonging to the A–B loop (Fig. 4), which signifies the characteristic change in the dynamics of the loop. Again, free-energy calculation shows that these four conserved water molecules (W12, W20, W22 and W41) contribute ∼350 cal mol^{-1} (∼10% of the total contribution of all the water molecules in the crystal) towards the stabilization of the protein. Two water molecules (W20 and W22) have been located at the dimer–dimer interface of the native tetramer, as was observed by Hörnberg *et al.* (2000). During dynamics,

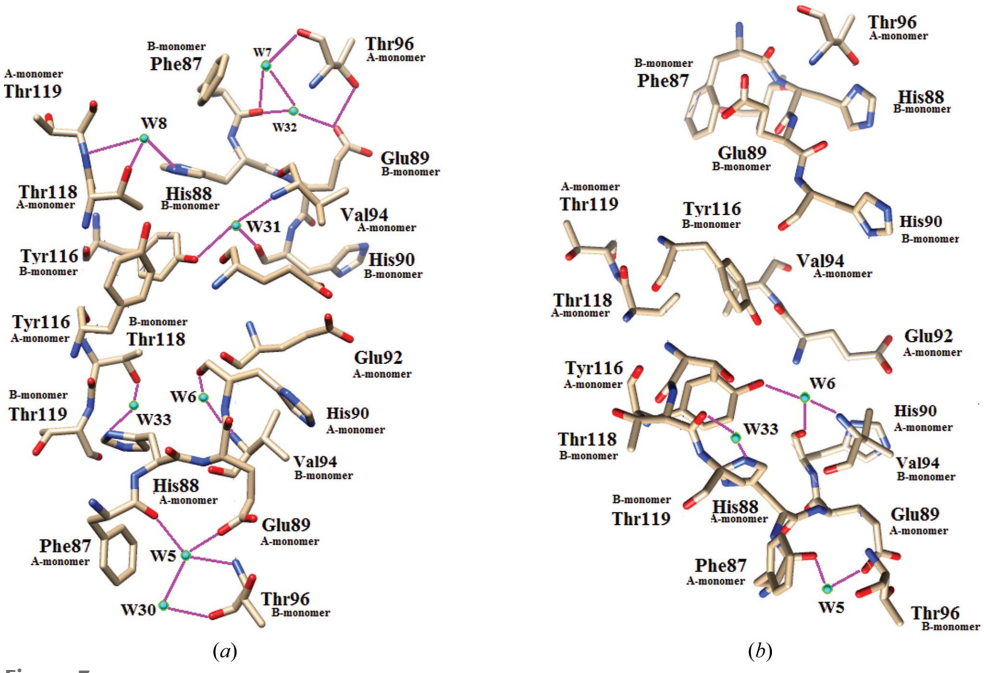

Figure 7
The interfacial hydrogen-bonding network in MD-simulated (*a*) 1f41 (crystallized at pH 7.0) and (*b*) 3cbr (crystallized at pH 3.5) structures.

W41 shows a residential frequency of more than 95% at its position; but its counter pair (W12) shows a lower value (90%), which may be owing to the interdependency of the members of the triad. During simulations, owing to a large shift of the Arg21 side chain, the water molecule W22 forms a new hydrogen bond to Arg21 (NH1) and is displaced (∼4 Å) from its position (in the crystal structures), resulting in the decrease in its residential frequency (91%). During the shift, W22 also displaces the water molecules W12 and W20, thus reducing their residential frequencies (Table 6). In the X-ray and neutron diffraction structures, water molecules W13 (W37) and W16 (W39) are sandwiched between the B and E strands of the inner sheet, where water molecule W13 seems to form a hydrogen bond to Asp74 O^{β}, Ala29 N^{β} and W17 (W40). Free-energy calculation shows that these water molecules (W13, W16, W17, W37, W39 and W40) contribute ∼460 cal mol^{-1} towards the stability of the protein, which is ∼13.5% of the total contribution of all of the water molecules in the crystal (Supplementary Table S1). Again, together with W5 and W6, water molecule W13 brings stability to the negatively charged canyon (Hörnberg *et al.*, 2000) through hydrogen bonding to Glu72, Asp74, Glu89 and Glu92 (Fig. 8). Water molecule W17, which remains hydrogen-bonded to Asn27 O^{β} of the A–B loop and Tyr78 N^{β} of the α-helix, forms an extended water network with W12, W13, W17, W20 and W22 to stabilize a series of important residues of hTTR (Fig. 1). These interactions are also observed when analyzing the MD snapshots. Water molecule W16 (W39) forms a hydrogen bond to Ile68 N^{β} of the E strand and Lys35 O^{β} of the B strand, whereas water molecule W15 (W43) interacts with Val32 O^{β} of the B strand and Ala45 O^{β} of the C strand. Again, Ala45 N^{β} also interacts with W21, which is hydrogen-bonded to W15. All of these conserved water molecules (W13, W15, W16 and W21) are stable at their positions, forming hydrogen bonds to the backbone polar atoms of the β-strands. They are also involved in similar types of interaction during MD simulations, with high residential frequency (>90%) at their respective positions.

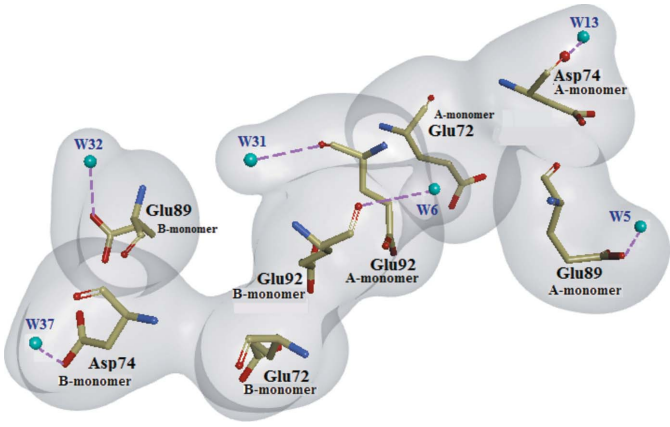

Figure 8
Stabilization of the negatively charged canyon (shown as a molecular surface) in hTTR by the conserved water molecules.

Careful observation of the high-resolution X-ray and neutron diffraction structures shows that water molecules W14 (W38) and W18 (W44) stabilize Lys48 O^{β} of the C strand and His56 N^{β} of the D strand. In addition, these water molecules (residential frequencies of >90%) also interact with each other, leading to the formation of a hydrogen-bond bridge (Fig. 1) between the inner (CBEF) and outer (DAGH) sheets of hTTR. This water-mediated bridge is found to be conserved in the MD-simulation trajectories. Thus, these water molecules may play an important role in stabilizing β-barrel formation, thereby helping to maintain the overall tertiary structure of hTTR. The conserved water molecules W19, W23 and W24 are located near the F–G loop of monomer *A*; however, they are not conserved in monomer *B*. The F–G loops (of all the monomers) which share the dimer–dimer interface contribute to the formation of the native tetramer (Hörnberg *et al.*, 2000) and mediate the interaction with the retinol-binding protein (Monaco *et al.*, 1995; Naylor & Newcomer, 1999). The crystal structure analyses and the solvent-accessible surface-area calculations reveal that water molecules W23 and W24 are buried within the F–G loop;

Table 7
Normalized *B* factor (Norm. *B*) and solvent-accessible surface area (SASA; Å^2) of the conserved water molecules between the C and D strand and near the F–G loop in the X-ray and neutron diffraction crystal structures of hTTR.

The numbers in the first row indicate the IDs of the water molecules in the respective crystals; normalized *B* factors of the O atoms of these water molecules are given in the second row and their SASAs are listed in the third row. RF, residential frequency. The absence of a water molecule in a structure is indicated by an asterisk and a dash denotes that the data are not available.

	W14	W18	W19	W23	W24	W38	W44
1f41	41	57	58	67	82	48	74
Norm. *B*	0.65	0.73	0.12	1.30	1.15	0.77	1.08
SASA	18.75	19.30	13.02	21.11	33.85	13.04	16.56
RF	100	100	100	100	85	100	100
2qgb	154 *A*	155 *A*	158 *A*	176 *A*	128 *A*	139 *B*	146 *B*
Norm. *B*	0.59	0.59	0.31	1.38	0.30	0.69	1.07
SASA	12.95	9.11	9.94	4.07	1.14	6.83	6.08
RF	100	100	100	100	100	100	100
3cfm	157 *A*	134 *A*	144 *A*	152 *A*	127 *A*	148 *B*	159 *B*
Norm. *B*	0.07	−0.15	−0.61	−0.18	−0.18	0.47	0.97
SASA	8.20	5.59	6.03	2.40	1.36	6.86	6.88
RF	100	100	100	90	100	100	85
3cbr	**195 *A***	**136 *A***	**203 *A***	**140 *A***	**165 *A***	**156 *B***	**183 *B***
Norm. *B*	**6.47**	**−0.06**	**0.24**	**0.17**	**0.46**	**1.48**	**2.05**
SASA	**−0.17**	**6.50**	**6.04**	**3.06**	**1.52**	**6.69**	**5.32**
RF	**100**	**84**	**100**	**78**	**55**	**100**	**82**
1tta	206 *A*	*	208 *A*	322 *A*	312 *A*	*	779 *B*
Norm. *B*	0.10	—	0.42	0.14	−0.14	—	−0.50
SASA	11.21	—	7.72	3.62	1.81	—	1.60
RF	100	100	100	100	30	85	100
3d7p	173 *A*	156 *A*	181 *A*	163 *A*	177 *A*	135 *B*	140 *B*
Norm. *B*	0.38	0.60	0.80	0.51	0.52	0.52	0.73
SASA	8.19	5.17	9.09	3.24	1.61	9.59	8.15
3a4d	197 *A*	134 *A*	151 *A*	154 *A*	147 *A*	166 *B*	141 *B*
Norm. *B*	−0.61	−0.70	0.61	0.05	1.18	0.13	0.54
SASA	11.74	9.85	5.87	2.32	1.30	10.97	8.36
2g4g	142 *A*	139 *A*	149 *A*	136 *A*	144 *A*	143 *B*	151 *B*
Norm. *B*	0.03	−0.08	−0.08	0.26	0.06	0.03	0.32
SASA	13.31	8.57	6.57	2.88	1.21	11.48	8.73
3i9p	160 *A*	158 *A*	161 *A*	170 *A*	171 *A*	179 *B*	149 *B*
Norm. *B*	−0.26	−0.52	−0.48	0.64	0.01	0.87	0.25
SASA	8.83	6.90	4.16	1.43	1.11	8.62	5.52
3u2i	136 *A*	133 *A*	153 *A*	*	147 *A*	152 *B*	138 *B*
Norm. *B*	0.00	0.10	0.50	—	0.50	0.80	0.40
SASA	13.57	10.30	12.99	—	1.87	13.51	10.60
3u2j	**131 *A***	**137 *A***	**146 *A***	*	*	**136 *B***	*
Norm. *B*	**−0.18**	**−0.07**	**−0.21**	—	—	**0.14**	—
SASA	**12.04**	**12.88**	**11.76**	—	—	**11.51**	—
4pvm	**217 *A***	**212 *A***	**225 *A***	**230 *A***	**233 *A***	**220 *B***	**222 *B***
Norm. *B*	**0.10**	**0.13**	**0.29**	**0.24**	**0.12**	**0.37**	**0.21**
SASA	**12.87**	**9.80**	**10.97**	**3.95**	**1.94**	**12.64**	**8.97**
RF	**100**	**89**	**100**	**100**	**73**	**100**	**100**
1f86	444 *A*	457 *A*	450 *A*	454 *A*	461 *A*	538 *B*	556 *B*
Norm. *B*	0.42	0.31	−0.13	0.07	0.61	0.32	0.54
SASA	9.92	8.61	6.40	3.20	1.50	9.40	6.77
RF	100	100	100	95	90	100	100
1x7s	6189	6014	6056	6043	6058	6031	6066
Norm. *B*	2.49	0.79	0.82	0.74	0.67	0.48	1.26
SASA	10.44	6.23	7.39	4.31	1.73	8.53	6.06
RF	100	100	100	100	100	100	100
3bt0	140 *A*	158 *A*	149 *A*	163 *A*	143 *A*	143 *B*	156 *B*
Norm. *B*	0.34	0.77	0.49	0.55	0.28	0.64	1.18
SASA	11.26	9.59	9.11	3.27	2.05	9.02	7.60
1x7t	6001	6070	6023	6013	6025	6107	*
Norm. *B*	0.20	0.98	−0.64	−0.46	−0.19	0.65	—
SASA	10.69	7.74	6.87	2.66	1.44	10.26	—
RF	100	90	85	100	100	100	100
1sok	145 *A*	153 *A*	154 *A*	160 *A*	167 *A*	143 *B*	153 *B*
Norm. *B*	1.41	2.62	0.36	1.06	0.75	1.81	2.65
SASA	10.27	7.05	6.76	2.56	1.41	8.74	6.35
1ttb	485 *A*	488 *A*	317 *A*	522 *A*	427 *A*	608 *B*	341 *B*

Table 7 (continued)

	W14	W18	W19	W23	W24	W38	W44
Norm. *B*	−0.66	1.62	0.97	−0.36	0.58	0.78	0.20
SASA	11.83	8.62	6.76	4.93	1.83	3.19	2.99
RF	100	100	100	100	100	100	45
1ttc	213	248	215	482	567	413	448
Norm. *B*	0.54	−0.75	−0.01	1.27	−0.12	0.60	0.61
SASA	6.66	7.87	6.43	2.73	1.19	11.72	6.72
1etb	163 *A*	165 *A*	138 *A*	210 *A*	151 *A*	199 *B*	138 *B*
Norm. *B*	1.20	3.39	1.93	0.84	1.54	1.23	0.78
SASA	11.73	9.62	8.95	3.09	1.87	12.05	10.17
1fhn	150 *A*	156 *A*	157 *A*	148 *A*	139 *A*	158 *B*	160 *B*
Norm. *B*	0.13	0.58	0.13	0.22	−0.15	1.83	1.36
SASA	12.35	8.02	9.49	2.68	1.71	18.16	9.34
RF	100	100	100	100	90	100	100
2b16	2037	2069	2109	2029	2074	2052	2053
Norm. *B*	−0.69	−0.36	−0.51	0.35	−0.47	−1.01	−0.75
SASA	11.95	8.17	8.62	3.10	1.36	9.43	7.22
2qgc	184 *A*	168 *A*	176 *A*	178 *A*	133 *A*	191 *B*	176 *B*
Norm. *B*	0.62	0.56	0.07	0.71	0.19	0.55	0.82
SASA	10.62	8.03	6.73	2.31	1.23	7.16	6.20
RF	100	100	100	100	83	100	100
3esp	*	133 *A*	150 *A*	149 *A*	148 *A*	131 *B*	132 *B*
Norm. *B*	—	0.37	0.39	0.58	0.46	0.52	0.56
SASA	—	9.66	6.88	2.71	1.20	7.68	7.04
RF	100	100	100	100	65	100	100
3eso	*	140 *A*	159 *A*	178 *A*	148 *A*	136 *B*	131 *B*
Norm. *B*	—	0.66	0.46	0.87	0.72	0.49	0.55
SASA	—	7.58	7.64	2.75	1.34	6.78	6.41
3esn	169 *A*	*	174 *A*	183 *A*	187 *A*	226 *B*	171 *B*
Norm. *B*	0.40	—	0.20	0.82	0.68	0.45	0.31
SASA	15.56	—	9.04	2.57	1.25	7.54	6.28
3cn4	139 *A*	136 *A*	169 *A*	152 *A*	143 *A*	149 *B*	140 *B*
Norm. *B*	0.61	0.66	0.76	1.26	1.04	0.57	0.62
SASA	13.15	8.66	9.73	2.95	1.45	9.80	7.03
RF	100	100	100	80	100	100	100
2fbr	25	22	49	53	44	72	58
Norm. *B*	0.47	−0.14	−0.19	0.39	0.17	0.51	0.85
SASA	11.32	7.64	8.41	4.15	1.49	7.79	6.82
2qgd	224 *A*	209 *A*	223 *A*	235 *A*	215 *A*	227 *B*	216 *B*
Norm. *B*	0.58	0.54	0.21	1.01	0.20	0.70	0.87
SASA	12.59	8.70	10.29	3.40	1.43	9.40	7.85
2qge	228 *A*	215 *A*	230 *A*	239 *A*	216 *A*	227 *B*	218 *B*
Norm. *B*	0.60	0.28	0.15	0.75	0.21	0.52	0.86
SASA	9.49	7.73	7.04	2.63	1.20	9.07	6.93
RF	100	100	100	100	50	100	100
3cn0	146 *A*	134 *A*	161 *A*	155 *A*	145 *A*	147 *B*	139 *B*
Norm. *B*	0.53	0.65	0.84	1.22	0.67	0.75	1.08
SASA	10.82	7.44	8.32	2.37	1.44	9.45	6.86
3cn1	142 *A*	129 *A*	181 *A*	157 *A*	167 *A*	141 *B*	134 *B*
Norm. *B*	0.59	0.44	0.66	1.18	0.80	0.61	0.71
SASA	11.94	6.54	9.73	2.87	1.46	7.77	6.37
Average Norm. *B*	**0.57**	**0.48**	**0.28**	**0.59**	**0.41**	**0.62**	**0.74**
Average SASA	**11.14**	**8.58**	**8.27**	**3.64**	**2.51**	**9.51**	**7.26**
Average RF	**100**	**98**	**99**	**97**	**81**	**99**	**95**

however, W19 remains highly accessible to the solvent molecules (Table 7). Water molecule W24 forms hydrogen bonds to Gly101 O$^{\beta}$, Asp99 N$^{\beta}$ and Ser100 N$^{\beta}$, whereas Asn98 O$^{\delta}$ and Arg103 O$^{\beta}$ are stabilized by a water molecule (W23) through hydrogen bonds (Fig. 9). The water molecule W19, which is hydrogen-bonded to W24, also interacts with the hydroxyl group of Tyr105. The temperature factors of these water molecules found in the X-ray structures (Table 7) indicate their stability at their positions and the MD simulations also reveal high residential frequency (more than 95%) for these water molecules, except for W23 (81%). Free-energy calculations show that these three conserved water molecules

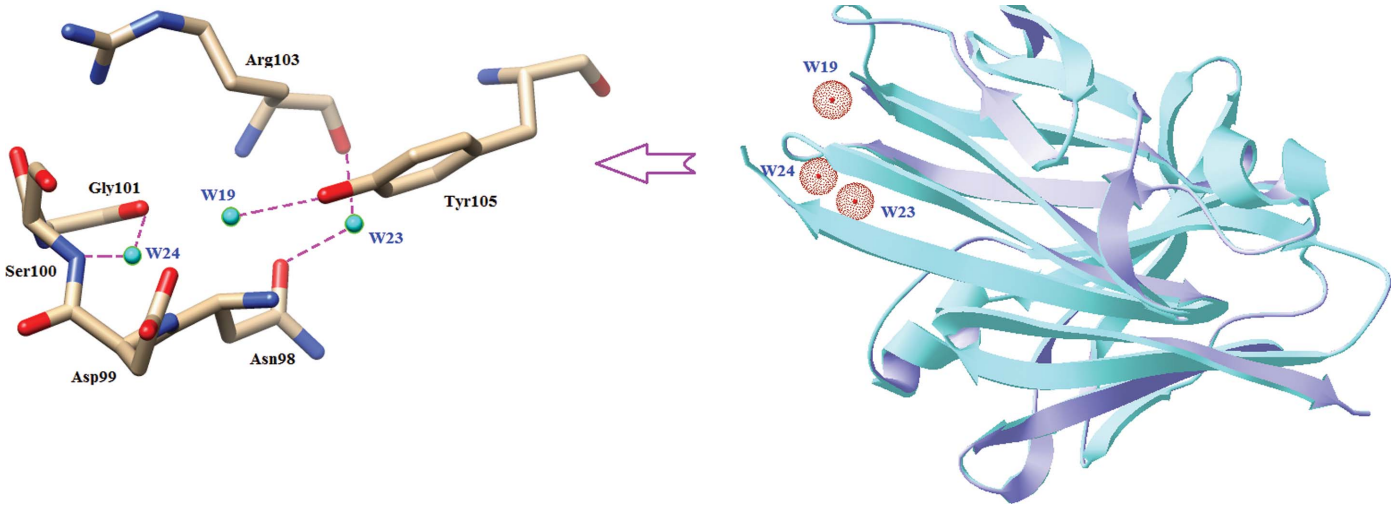

Figure 9
Stabilization of the F–G loop in hTTR by the conserved water molecules. The right panel shows the location of the water molecules and the left panel shows their interaction.

(W19, W23 and W24) contribute ∼140 cal mol^{-1} towards the stability of the protein (Supplementary Table S1). Again, the normalized r.m.s.f. values obtained from 10 ns MD simulations of the apo hTTR (PDB entry 1f41) structure(s) show that their removal significantly affects the dynamics, especially in the F–G loop region (Fig. 5).

4. Conclusion

Extensive analyses using a set of 32 high-resolution X-ray and neutron diffraction structures of human transthyretin followed by MD-simulation studies of a selected set of 15 structures revealed the presence of 44 conserved or invariant water molecules. Among them, four pairs of water molecules (W9, W10, W11 and W25 in chain *A* and their counter pairs W34, W35, W36 and W42 in chain *B*) were located near the G and H strands, which may be involved in maintaining the geometry of the ligand-binding cavities. Four pairs of water molecules (W12, W13, W16, W17, W37, W39, W40 and W41) are important for keeping the A–B loop and the α-helix in a proper orientation. Eight water molecules (W5, W6, W7, W8, W30, W31, W32 and W33) stabilize the monomer–monomer interface through an extensive hydrogen-bonding network, whereas water molecules W20 and W22 were located at the dimer–dimer interface of the native tetramer. These interfacial water molecules play a significant role in the stability of the protein. The absence of some of these water molecules (W7, W8, W31 and W32) in highly acidic conditions (pH $\leq$ 4.0) severely affects the interfacial hydrogen-bond network, which may destabilize the native tetrameric structure, leading to its dissociation. Water molecules W19, W23 and W24 may help in positioning the F–G loop to provide different intramolecular and intermolecular interactions. The present study provides some clues about the water-mediated architecture and stability of hTTR.

Acknowledgements

AB, SD and BPM acknowledge the National Institute of Technology (Government of India)–Durgapur for providing research facilities at the Department of Chemistry.

References

Baker, E. N. & Hubbard, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.
Balamurugan, B. *et al.* (2007). *J. Appl. Cryst.* **40**, 773–777.
Banerjee, A., Bairagya, H. R., Mukhopadhyay, B. P., Nandi, T. K. & Bera, A. K. (2010). *Indian J. Biochem. Biophys.* **47**, 197–202.
Banerjee, A., Bairagya, H. R., Mukhopadhyay, B. P., Nandi, T. K. & Mishra, D. K. (2013). *J. Mol. Graph. Model.* **44**, 70–80.
Banerjee, A. & Mukhopadhyay, B. P. (2015). *J. Biomol. Struct. Dyn.* **33**, 1973–1988.
Bartalena, L. & Robbins, J. (1993). *Clin. Lab. Med.* **13**, 583–598.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). *J. Comput. Chem.* **4**, 187–217.
Buxbaum, J. N. & Tagoe, C. E. (2000). *Annu. Rev. Med.* **51**, 543–569.
Cavallo, L., Kleinjung, J. & Fraternali, F. (2003). *Nucleic Acids Res.* **31**, 3364–3366.
Chaplin, M. (2006). *Nature Rev. Mol. Cell Biol.* **7**, 861–866.
Chen, J. C.-H., Hanson, B. L., Fisher, S. Z., Langan, P. & Kovalevsky, A. Y. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 15301–15306.
Dasgupta, J., Sen, U. & Dattagupta, J. K. (2003). *Protein Eng.* **16**, 489–496.
Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Res.* **34**, W116–W118.
Fraternali, F. & Cavallo, L. (2002). *Nucleic Acids Res.* **30**, 2950–2960.
Gambetti, P. & Russo, C. (1998). *Nephrol. Dial. Transplant.* **13**, 33–40.
Gowri Shankar, B. A., Sarani, R., Michael, D., Mridula, P., Vasuki Ranjani, C., Sowmiya, G., Vasundhar, B., Sudha, P., Jeyakanthan, J., Velmurugan, D. K. & Seka, K. (2007). *J. Biosci.* **32**, 693–704.
Green, N. S., Palaninathan, S. K., Sacchettini, J. C. & Kelly, J. W. (2003). *J. Am. Chem. Soc.* **125**, 13404–13414.
Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.
Hamilton, J. A., Steinrauf, L. K., Braden, B. C., Liepnieks, J., Benson, M. D., Holmgren, G., Sandgren, O. & Steen, L. (1993). *J. Biol. Chem.* **268**, 2416–2424.
Haupt, M., Blakeley, M. P., Fisher, S. J., Mason, S. A., Cooper, J. B., Mitchell, E. P. & Forsyth, V. T. (2014). *IUCrJ*, **1**, 429–438.

Hörnberg, A., Eneqvist, T., Olofsson, A., Lundgren, E. & Sauer-Eriksson, A. E. (2000). *J. Mol. Biol.* **302**, 649–669.
Hörnberg, A., Olofsson, A., Eneqvist, T., Lundgren, E. & Sauer-Eriksson, A. E. (2004). *Biochim. Biophys. Acta*, **1700**, 93–104.
Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.
Janin, J. & Chothia, C. (1990). *J. Biol. Chem.* **265**, 16027–16030.
Jiang, L., Kuhlman, B., Kortemme, T. & Baker, D. (2005). *Proteins*, **58**, 893–904.
Johnson, S. M., Connelly, S., Wilson, I. A. & Kelly, J. W. (2008*a*). *J. Med. Chem.* **51**, 260–270.
Johnson, S. M., Connelly, S., Wilson, I. A. & Kelly, J. W. (2008*b*). *J. Med. Chem.* **51**, 6348–6358.
Johnson, S. M., Connelly, S., Wilson, I. A. & Kelly, J. W. (2009). *J. Med. Chem.* **52**, 1115–1125.
Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W. & Klein, M. L. (1983). *J. Chem. Phys.* **79**, 926–935.
Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. & Schulten, K. (1999). *J. Comput. Phys.* **151**, 283–312.
Kanaujia, S. P. & Sekar, K. (2009). *Acta Cryst.* D**65**, 74–84.
Klabunde, T., Petrassi, H. M., Oza, V. B., Raman, P., Kelly, J. W. & Sacchettini, J. C. (2000). *Nature Struct. Biol.* **7**, 312–321.
Lima, L. M., Silva, V. de A., Palmieri, L. de C., Oliveira, M. C., Foguel, D. & Polikarpov, I. (2010). *Bioorg. Med. Chem.* **18**, 100–110.
Liu, L. & Murphy, R. M. (2006). *Biochemistry*, **45**, 15702–15709.
Liz, M. A., Leite, S. C., Juliano, L., Saraiva, M. J., Damas, A. M., Bur, D. & Sousa, M. M. (2012). *Biochem. J.* **443**, 769–778.
MacKerell, A. D. Jr *et al.* (1998). *J. Phys. Chem. B*, **102**, 3586–3616.
Miller, M. D. & Krause, K. L. (1996). *Protein Sci.* **5**, 24–33.
Miyata, M., Sato, T., Mizuguchi, M., Nakamura, T., Ikemizu, S., Nabeshima, Y., Susuki, S., Suwa, Y., Morioka, H., Ando, Y., Suico, M. A., Shuto, T., Koga, T., Yamagata, Y. & Kai, H. (2010). *Biochemistry*, **49**, 114–123.
Mladenovic, M., Fink, R. F., Thiel, W., Schirmeister, T. & Engels, B. (2008). *J. Am. Chem. Soc.* **130**, 8696–8705.
Monaco, H. L., Rizzi, M. & Coda, A. (1995). *Science*, **268**, 1039–1041.
Morais-de-Sá, E., Neto-Silva, R. M., Pereira, P. J. B., Saraiva, M. J. & Damas, A. M. (2006). *Acta Cryst.* D**62**, 512–519.
Nandi, T. K., Bairagya, H. R., Mukhopadhyay, B. P., Mallik, P., Sukul, D. & Bera, A. K. (2012). *J. Mol. Model.* **18**, 2633–2644.
Naylor, H. M. & Newcomer, M. E. (1999). *Biochemistry*, **38**, 2647–2653.
Neto-Silva, R. M., Macedo-Ribeiro, S., Pereira, P. J. B., Coll, M., Saraiva, M. J. & Damas, A. M. (2005). *Acta Cryst.* D**61**, 333–339.
Ogata, K. & Wodak, S. J. (2002). *Protein Eng.* **15**, 697–705.
Palaninathan, S. K., Mohamedmohaideen, N. N., Snee, W. C., Kelly, J. W. & Sacchettini, J. C. (2008). *J. Mol. Biol.* **382**, 1157–1167.
Pasquato, N., Berni, R., Folli, C., Alfieri, B., Cendron, L. & Zanotti, G. (2007). *J. Mol. Biol.* **366**, 711–719.
Poornima, C. S. & Dean, P. M. (1995). *J. Comput. Aided Mol. Des.* **9**, 500–512.
Ramachandran, G. & Schlick, T. (1995). *Phys. Rev. E*, **51**, 6188–6203.
Raschke, T. M. (2006). *Curr. Opin. Struct. Biol.* **16**, 152–159.
Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F. & Serrano, L. (2005). *Nucleic Acids Res.* **33**, W382–W388.
Schymkowitz, J. W. H., Rousseau, F., Martins, I. C., Ferkinghoff-Borg, J., Stricher, F. & Serrano, L. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 10147–10152.
Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Krüger, P. & van Gunsteren, W. F. (1999). *J. Phys. Chem. A*, **103**, 3596–3607.
Sebastião, M. P., Lamzin, V., Saraiva, M. J. & Damas, A. M. (2001). *J. Mol. Biol.* **306**, 733–744.
Sekijima, Y., Wiseman, R. L., Matteson, J., Hammarström, P., Miller, S. R., Sawkar, A. R., Balch, W. E. & Kelly, J. W. (2005). *Cell*, **121**, 73–85.
Smith, D. K., Radivojac, P., Obradovic, Z., Dunker, A. K. & Zhu, G. (2003). *Protein Sci.* **12**, 1060–1072.
Smolin, N. & Winter, R. (2008). *J. Phys. Chem. B*, **112**, 997–1006.
Sumathi, K., Ananthalakshmi, P., Roshan, M. M. N. A. & Sekar, K. (2006). *Nucleic Acids Res.* **34**, W128–W132.
Thylén, C., Wahlqvist, J., Haettner, E., Sandgren, O., Holmgren, G. & Lundgren, E. (1993). *EMBO J.* **12**, 743–748.
Tomar, D., Khan, T., Singh, R. R., Mishra, S., Gupta, S., Surolia, A. & Salunke, D. M. (2012). *PLoS One*, **7**, e43522.
Yang, M., Lei, M. & Huo, S. (2003). *Protein Sci.* **12**, 1222–1231.
Yokoyama, T., Mizuguchi, M., Nabeshima, Y., Kusaka, K., Yamada, T., Hosoya, T., Ohhara, T., Kurihara, K., Tomoyori, K., Tanaka, I. & Niimura, N. (2012). *J. Struct. Biol.* **177**, 283–290.
Zanotti, G., Folli, C., Cendron, L., Alfieri, B., Nishida, S. K., Gliubich, F., Pasquato, N., Negro, A. & Berni, R. (2008). *FEBS J.* **275**, 5841–5854.
Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. (2011). *J. Comput. Chem.* **32**, 2359–2368.