the comparative analysis of the solvent positions in ten different crystal forms of T4 lysozyme, providing a clear picture of the effect of crystal contacts on the hydration sphere of a protein viewed by X-ray crystallography (Zhang & Matthews, 1994). The resolution and degree of refinement of the structures involved varied significantly, from 2.6 to 1.7 Å resolution, and the number of water molecules included per protein molecule ranged from 38 to 160. Nevertheless, this study revealed important features. A striking observation is that 95% of the solvent-exposed residues on T4 lysozyme were involved in at least one crystal contact in one or another of the crystal forms studied, showing that any part of the protein surface can be involved in crystal contacts. A corollary to this finding is that any of the surface water molecules can be displaced or involved in bridging protein–protein contacts in the crystal.

Of the 1675 individual water molecules observed in the 18 independently refined T4 lysozyme molecules included (Fig. 23.4.4.9), the ones that were within a sphere of radius 1.2 Å were considered to occupy the same site on the protein. As in the case of elastase described above, all of the water molecules observed upon superposition of the 18 T4 lysozyme structures represent a large portion of the first hydration shell. This reinforces the concept that multiple structures of a protein of interest provide a more complete picture of the protein hydration than possible with a single structure. There are four buried water sites that are occupied in at least 15 out of the 18 structures and are independent of crystal contacts. Two of these buried sites are at the hinge-bending region between the two helical domains and appear to play a functional role in the opening and closing of the active site (Weaver & Matthews, 1987). The other two play a structural role at the protein core. Other than the four buried water molecules, the most conserved water sites appear at the active-site cleft between the two domains and at the N-termini

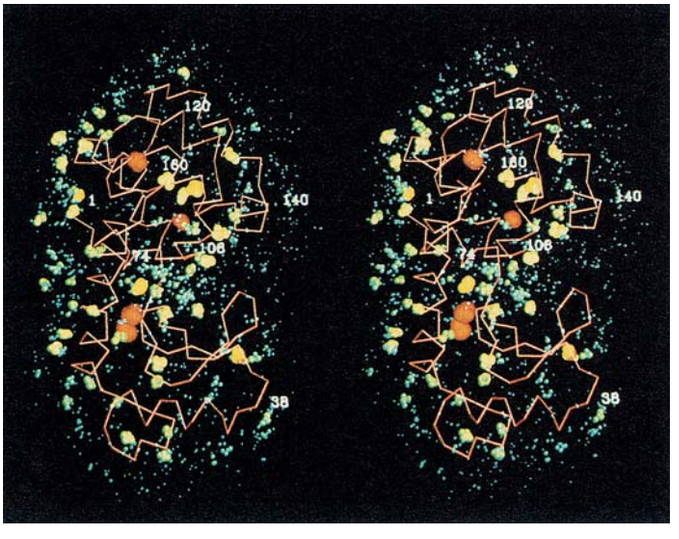

Fig. 23.4.4.9. Distribution of solvent-binding sites in 18 mutant T4 lysozymes from ten refined crystal structures. The lysozyme structures were compared to identify common sites of hydration. A total of 1675 solvent molecules were included in the comparison. Each solvent molecule is represented by a coloured sphere. The size of the sphere is proportional to the number of lysozyme structures in which solvent was observed at the same site (*i.e.* within 1.2 Å). In addition, the colour of the solvent changes from blue for the least-conserved sites to red for the most-conserved ones [*e.g.* the red spheres indicate that solvent is observed with high frequency (15–17 times) at the four internal sites]. The numbers indicate representative residue positions along the backbone of the lysozyme molecule. Reprinted with the permission of Cambridge University Press from Zhang & Matthews (1994). Copyright (1994) The Protein Society.

of α-helices. As is the case in the previous works reviewed above, the 20 most conserved water sites appear in well conserved protein environments and generally have low temperature factors. Buried or highly conserved water molecules also tend to make at least three hydrogen bonds with protein atoms or other water molecules. The less-conserved water sites appear more randomly on the protein surface and are strongly influenced by the particular crystal environment in which the structure was solved.

23.4.4.2.3. *Ribonuclease T1*

A group of four crystal structures of ribonuclease T1 in complex with guanosine, guanosine-2′-phosphate, guanylyl-2′,5′-guanosine and vanadate were used for an analysis of conserved water positions that contribute to the structural stabilization of the protein (Malin *et al.*, 1991). The four structures were obtained from isomorphous crystals and ranged in resolution from 1.7 to 1.9 Å. Conserved water molecules were considered to be those found within a sphere of 1 Å from each other in all four structures. All other water molecules were excluded from the analysis. 30 water molecules were found to be conserved. Of these, ten were observed near crystal contacts, although only one appears to be dictated by the crystal contact itself, making a single hydrogen bond with each of the symmetry-related protein molecules. Ten other water molecules form a channel that brings together an α-helix and a hairpin-like loop structure and then go on to wrap around the calcium ion, providing half of its coordination sphere. The first five of these water molecules are completely buried, holding together the two secondary-structure elements, which would otherwise collapse (Malin *et al.*, 1991). Two water molecules are found to stabilize the N and C termini, which are brought together by a disulfide bond. The remaining eight conserved water molecules hold together various elements of secondary structure or are located in the active site.

An interesting extension to this study included four additional structures: the E58A mutant in complex with guanosine-2′-monophosphate, the H92A mutant crystallized under two different conditions and wild-type RNase T1 in complex with guanosine-3′,5′-biphosphate. Two of these crystal forms were not isomorphous with the native protein crystals or with each other. Thus a total of eight structures solved in three different space groups were analysed (Pletinckx *et al.*, 1994). Although the effect of crystal packing on the three-dimensional structure of the protein is minimal, there are some significant differences in the solvent structure. In particular, there is no evidence of the calcium-binding site and its coordinating water structure in any crystal forms other than the canonical wild type. Instead, the E58A mutant has a sodium-binding site at a different position, along with three previously unobserved water molecules. It is clear that the presence of the metal ions is fortuitous and linked to the crystallization conditions.

There are 25 water molecules structurally conserved throughout the different packing arrangements studied. Ten of these are single sites, there are three clusters of two water molecules and a larger cluster originally described by Malin *et al.* (1991) to hold together the core of the protein. As was observed for the study on T4 lysozyme (Zhang & Matthews, 1994), the strictly conserved water-binding sites present in crystal structures solved across different space groups are involved in bridging protein secondary-structure elements and seem to be crucial for the integrity of the protein structure.

23.4.4.2.4. *Ribonuclease A*

Ribonuclease A is not homologous to ribonuclease T1 in either sequence or structure, but both have evolved to catalyse the same reaction with specificity for different substrates (compare Figs. 23.4.4.10 and 23.4.4.11). Ribonuclease A cleaves RNA after pyrimidines, while ribonuclease T1 cleaves specifically after

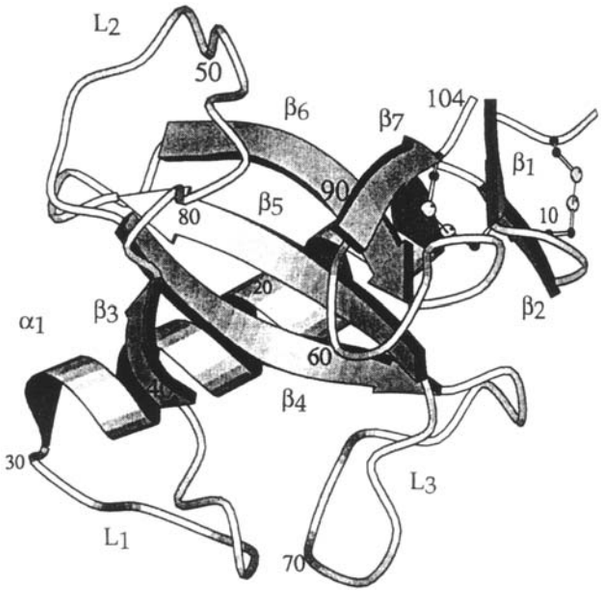

Fig. 23.4.4.10. Three-dimensional structure of RNase T1. Secondary structure is denoted as follows: α_1, α-helix; β_n, strands of β-sheet structure; L_n, loops. Drawn using *MOLSCRIPT* (Kraulis, 1991). Residue numbers indicate the beginning and end of secondary-structure elements. Reprinted with permission from Pletinckx *et al.* (1994). Copyright (1994) American Chemical Society.

guanine. Therefore, the information obtained from a study of the solvent structure in ribonuclease A is completely independent from that described above for ribonuclease T1. A collection of ten crystal structures of ribonuclease A, derived from five different crystal forms, were compared pairwise after least-squares superposition (Zegers *et al.*, 1994). 17 conserved water molecules were found to be within a sphere of 0.5 Å of each other in all of the ten structures and are shown in Fig. 23.4.4.11. These water molecules were found in small clusters of two or three or as part of a larger solvent network. Not surprisingly, they form multiple hydrogen bonds with the protein and generally have low temperature factors. Of the 17 structurally conserved sites, 13 are associated with one of the three α-helices. Most of these link the helices to one of the β-strands. Three water molecules are involved in hydrogen bonding with unpaired amido and carbonyl groups on the protein, and one is found on top of the β-pleated sheet. These interactions result in bringing together elements of secondary structure and in stabilizing distortions within these elements. Conserved water molecules are also responsible for bridging the N-terminal helix to the C-terminal β-strand, which form the two halves of the active site.

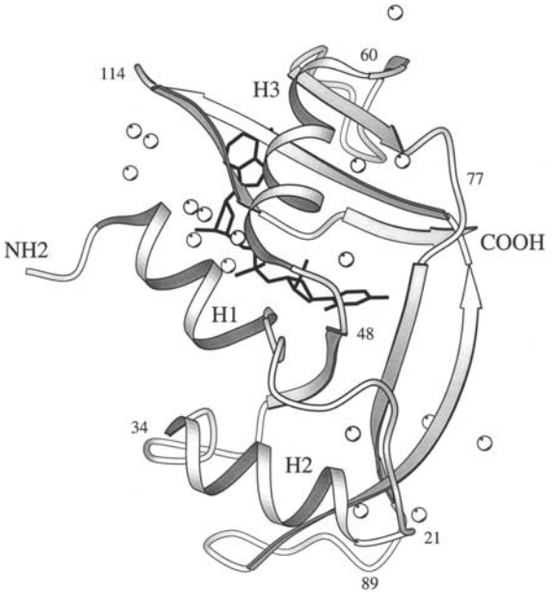

Fig. 23.4.4.11. Overall structure of RNase A. The overall structure of the d(CpA) complex of RNase A is shown as a ribbon drawing using *MOLSCRIPT* (Kraulis, 1991). The conserved water molecules are shown as white spheres and the d(CpA) inhibitor in black. The three helices are labelled H1, H2 and H3. Reprinted with the permission of Cambridge University Press from Zegers *et al.* (1994). Copyright (1994) The Protein Society.

23.4.4.2.5. *Protein kinase A*

The comparative study of water molecules in seven different protein kinase A structures in complex with different ligands focused exclusively on the active site (Shaltiel *et al.*, 1998). All of the structures were solved from isomorphous crystals to resolutions ranging from 2.0 to 2.9 Å. The more lenient cutoff of 1.5 Å for the radius of the sphere within which the conserved water molecules must be found among the different structures is consistent with the relatively low resolutions of the structures included in this study. The group of structures represents the open, the closed and an intermediate conformation of the catalytic kinase domain. There is a set of six conserved water sites in the active site, in addition to the ATP molecule and the magnesium ion. The conserved water molecules coordinate to ATP, the metal ion and a conserved Tyr residue from the carboxyl terminus of the protein. Thus, the active site consists of an extended network of interactions that weave together both domains of the core, with water molecules playing an integral role in maintaining the structural features important for catalysis. Many of these water molecules associate directly with the inhibitors. In addition, five water sites are observed in positions that would be occupied by substrates or substrate analogues. These water molecules are displaced by ligand oxygen atoms that can compensate for the water hydrogen-bonding interaction with the protein.

23.4.4.3. *Summary*

Water molecules associated with proteins can be divided between those that are conserved as a result of their functional significance and those that are partially conserved or not conserved at all. The conserved water molecules are generally classified as buried or channel (by a variety of criteria). They tend to be present in the clefts between domains, are critical components of active sites, or bridge between secondary-structure elements. The water molecules that are not conserved occupy hydration sites with favourable hydrogen-bonding characteristics, where the presence of a water molecule is not essential for the structural or functional integrity of the protein.

The displacement of water molecules by organic solvent molecules in the elastase work described above showed that most displaced waters are those classified as surface or crystal-contact waters (Mattos *et al.*, 2000). In the three cases where a buried water molecule was displaced, an alcohol hydroxyl oxygen was found to replace the protein–water hydrogen-bonding interactions. This is analogous to the active-site water molecule in the HIV aspartate protease that gets replaced by a carbonyl group of a potent cyclic urea inhibitor (Lam *et al.*, 1994). In these situations, release of a tightly bound water molecule is entropically favourable, and its enthalpic interactions with the protein are compensated by similar protein–ligand interactions.

The effect of crystal contacts on the water structure was clearly illustrated in the T4 lysozyme work (Zhang & Matthews, 1994).

references