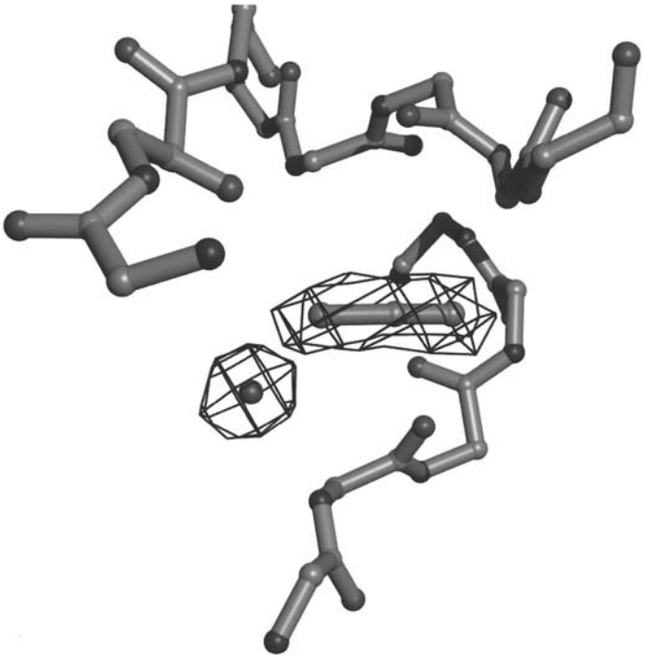

Fig. 23.4.4.3. A $2F_o - F_c$ electron-density map contoured at the 1.2σ level shows a distinct ellipsoidal density for acetonitrile 707 and a spherical density for a nearby water molecule. The protein backbone of the binding pocket is represented with nitrogen atoms shown in dark grey, oxygen atoms in medium grey and carbons in a lighter grey. *MOLSCRIPT* (Kraulis, 1991) was used in the preparation of this figure. Reprinted with permission from Allen *et al.* (1996). Copyright (1996) American Chemical Society.

23.4.4.2.1. *Elastase*

The crystal structure of porcine pancreatic elastase was solved in a variety of organic solvents, with the primary goal of mapping binding sites on the protein that could accommodate molecules representative of functional groups likely to be found in larger ligands (Ringe, 1995; Mattos & Ringe, 1996; Mattos *et al.*, 2000). Crystals of elastase cross-linked with glutaraldehyde were transferred to the following solutions: 100% acetonitrile, 95% acetone, 55% dimethylformamide, 80% ethanol, 40% trifluoroethanol, 80% isopropanol and 80% 5-hexene-1,2-diol (Allen *et al.*, 1996; Mattos & Ringe, 1996). In general, the crystals did not diffract in most neat organic solvents. However, in the acetonitrile case, where they did, the result was striking. In the structure of elastase solved in >99% acetonitrile, there were 126 water molecules visible in the electron-density maps, indicating that a good portion of the first hydration shell of the protein was still present. In contrast, only nine molecules of acetonitrile were clearly identified in the electron-density maps (Allen *et al.*, 1996). This is a powerful assertion of the evolutionary specificity of water molecules for protein surfaces. Fig. 23.4.4.3 shows the clear contrast between the elongated electron density of an acetonitrile molecule and the spherical electron density of a water molecule.

A similar result was obtained for all of the elastase structures solved in the mixtures of organic solvent and water mentioned above. A total of 11 structures were analysed, each containing 126–177 water molecules. The structures are listed in Table 23.4.4.1, together with the resolution of the data collected, the number of water molecules present and the number of organic solvent molecules observed in each case. The Cα superposition of the protein atoms in the 11 structures yielded a total of 1661 individual water molecules, occupying 426 unique water-binding sites on the elastase surface. Given that elastase has a total of 240 amino-acid residues, this represents a significant portion of the first hydration shell of the protein. This group of elastase structures served as a powerful source of information, leading to a classification of water types according to their interaction with the protein and an analysis of the specificity for water within each of the types determined (Bellamacina *et al.*, 1999).

All of the 1661 water molecules were renumbered according to the site on the protein where they were found. Any two water molecules within 1 Å of a water molecule in the cross-linked elastase structure solved in distilled water (used as the reference structure) have a common number. 39 of the 426 water-binding sites were occupied in every one of the 11 structures and were considered structurally conserved. Among these are the 16 buried water-binding sites thought to be conserved among all serine proteases (Sreenivasan & Axelsen, 1992). The 26 remaining conserved water molecules are specific to elastase and are not necessarily buried. These water molecules in general tend to have low B factors, but a few have B factors in the 30–35 Å^2 range and one conserved water molecule has a B factor of 42 Å^2.

The classification of the water sites as buried, channel, crystal contact or surface was based on the number of hydrogen-bonding interactions that a water molecule at the site could make to the protein and involved no surface-accessibility calculations (Bellamacina *et al.*, 1999). Water molecules were classified as buried if they made at least three good hydrogen-bonding interactions with protein main-chain atoms. A total of 23 buried water sites were identified in this manner, including 13 of the sites classified as buried by Sreenivasan & Axelsen (1992). One of the 16 serine protease conserved water-molecule sites is replaced by a His side chain in elastase (Sreenivasan & Axelsen, 1992). The remaining two serine protease conserved water sites were classified as channel based on the criteria used in the present study (see below). Interestingly, with the exception of these two channel water molecules, all of the buried sites found to be conserved in serine proteases are strictly conserved in all of the 11 structures in Table 23.4.4.1. The two channel water molecules are found in the aqueous structures of elastase, but are virtually absent in elastase transferred to organic solvents.

The water molecules occupying the 23 buried water sites identified in this study are tightly clustered when the protein Cα atoms are superimposed by least squares, and the interactions with the protein are conserved from structure to structure. Fig. 23.4.4.4 shows the positions of the buried water-binding sites in elastase. In general, they are found in the cleft between the two domains, in bridging elements of the secondary structure and at the base of water channels. This observation is consistent with the current

Table 23.4.4.1. *Multiple-solvent crystal structures of elastase*

Structure	Resolution (Å)	No. of water molecules	No. of organic solvent molecules
Cross-linked	1.9	165	0
Acetonitrile	2.2	126	9
Acetone	2.0	126	6
Dimethylformamide	2.0	153	6
Ethanol	2.0	135	12
Trifluoroethanol (1)	1.9	175	4
Trifluoroethanol (2)	1.85	177	3
Isopropanol	2.2	160	4
Benzene	1.9	162	4
Cyclohexane	1.95	135	7
5-Hexene-1,2-diol	2.2	147	5

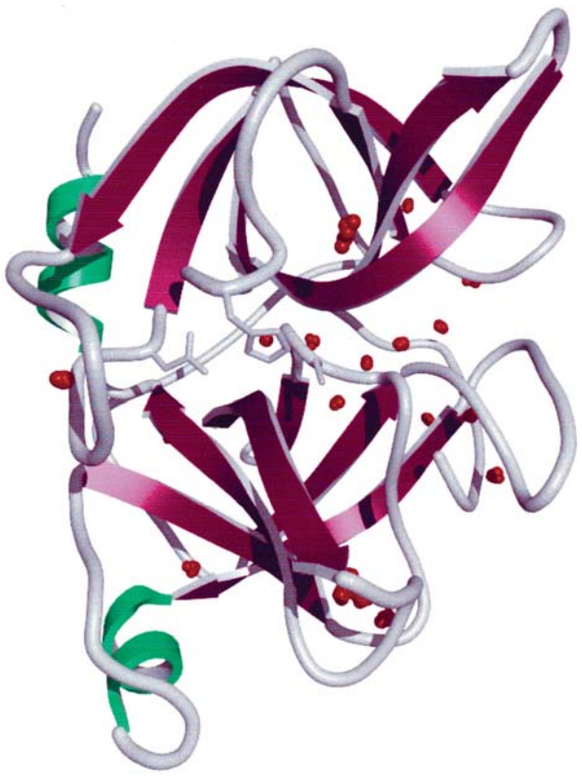

Fig. 23.4.4.4. Crystal structure of porcine pancreatic elastase represented as a ribbon diagram using *MOLSCRIPT* (Kraulis, 1991). The two α-helices are shown in green, the β-sheets are in purple and the coils are in grey. Elastase contains 240 amino-acid residues, and is composed of two β-barrel domains. The catalytic triad (Asp108, His60 and Ser203) is shown explicitly. The buried crystallographic water molecules found in 11 superimposed elastase structures solved in a variety of solvents are shown in red.

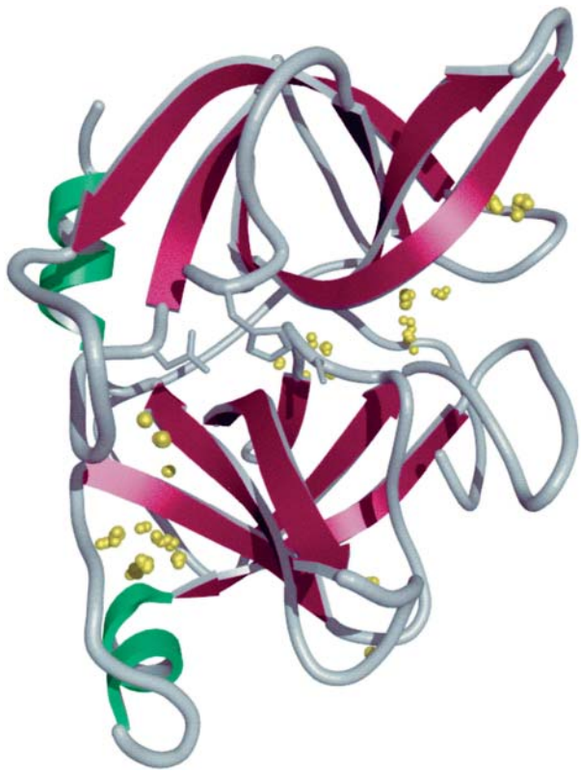

Fig. 23.4.4.5. Elastase structure represented as in Fig. 23.4.4.4. The crystallographic water molecules found in channels in 11 superimposed elastase structures solved in a variety of solvents are shown in yellow.

understanding of the functional roles played by structurally conserved water molecules as discussed above and in the following sections.

The 29 water-binding sites classified as channel contain water molecules that make hydrogen bonds with at least two other water molecules within a protein groove. The analysis of a high-resolution crystal structure of elastase (1.65 Å) revealed seven channels with a total of 32 water-binding sites (Meyer *et al.*, 1988). All of these channels were also identified in the analysis of the 11 structures in Table 23.4.4.1 (Bellamacina *et al.*, 1999). In addition, two other channels were observed. The locations of the nine elastase channels identified by the new criteria are shown in Fig. 23.4.4.5. Channels are often found in areas associated with buried water molecules, namely, at the crevice between the two domains and sandwiched between secondary-structure elements, where they lead from the surface of the protein to a buried water molecule. Fig. 23.4.4.5 also shows that the Cα superposition of the protein structures leads to a spread of water molecules within the channels. In any given structure, only two or three water molecules may be present, but the precise location and interaction with protein atoms vary so that when taken together the collection of structures gives a sense of flow inside the channels.

Of the remaining 374 water-molecule sites present within the 11 elastase structures included in this study, 56 were classified as crystal-contact sites and 318 as surface sites. Crystal-contact sites were considered to be occupied by water molecules that are within 4.0 Å of a symmetry-related protein molecule in the crystal. Fig. 23.4.4.6 shows the position of all the water molecules found to occupy these sites. The relatively large number of crystal-contact water-binding sites is a result of the somewhat broad criterion used to select them. Many of these sites are not within hydrogen-bonding distance from the nearby protein molecule, and most are not well conserved from structure to structure. Only eight of the 56 sites are occupied in the majority of the structures, and four of these make good multiple hydrogen bonds with two symmetry-related protein molecules in the crystal. These four water molecules seem to be structurally significant in the formation of the crystal contacts.

Surface water molecules were taken to be those that interact with side-chain protein atoms on the surface or make no more than two hydrogen-bonds with backbone atoms. When the 11 structures are superimposed, the surface water molecules occupying a given site are not tightly clustered. Furthermore, there is flexibility in the interactions between these water molecules and the nearby protein atoms. For example, it is often the case that all water molecules within a surface site make two or three hydrogen bonds to protein atoms, but only one of them is conserved in all of the structures where the water molecule is present at the site. Fig. 23.4.4.7 illustrates the position of all of the surface water-binding sites. Although over half of these sites are occupied in at least two of the 11 structures, a good proportion of them (178) are found in only one of the structures considered.

While crystal-contact and surface water sites were classified separately, it is important to point out that, with the exception of the four crystal-contact water-binding sites mentioned above, the crystal-contact sites exhibit very much the same traits as the surface water sites. The difference is that in the latter case, the 'surface' is provided by a single protein molecule, while in the former the interaction between two symmetry-related protein molecules constitutes the surface with which the water molecules interact.

Of the 318 surface water molecules, 21 are in the active site. The active-site water molecules were selected to be those within 4 Å of any atom belonging to either the trifluoroacetyl-Lys-Phe-*p*-

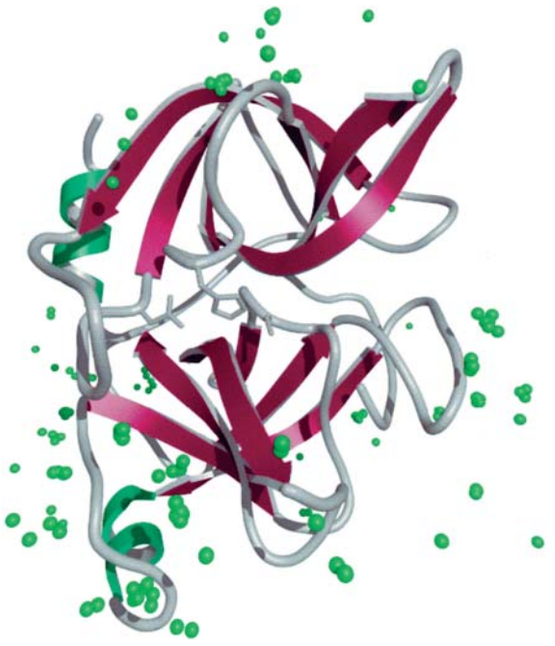

Fig. 23.4.4.6. Elastase structure represented as in Fig. 23.4.4.4. The crystallographic water molecules involved in crystal contacts in 11 superimposed elastase structures solved in a variety of solvents are shown in green.

isopropylanilide (Mattos *et al.*, 1994) or the trifluoroacetyl-Lys-Pro-*p*-trifluoromethylanilide (Mattos *et al.*, 1995) inhibitors in the structures of their complexes with elastase. These inhibitors span a large area of the active site, including an exosite not occupied by substrate analogue inhibitors (Mattos *et al.*, 1994, 1995). The water-binding sites in the active site are not very well conserved, with most sites represented in only two to four of the 11 structures. When all of the structures are superimposed, there is at least one water molecule in each of the subsites in the elastase active site. These water molecules are displaced either by inhibitors or by organic solvent molecules in the various structures. It is not surprising that in elastase, a protein with relatively broad substrate specificity, the active site in the uncomplexed native protein is populated by many displaceable surface water molecules. With the exception of a water molecule present in the oxyanion hole, these water molecules tend to make a single hydrogen bond with the protein. This hydrogen-bonding interaction is not generally conserved between different structures where a given site is occupied in multiple structures. The displacement of these water molecules upon ligand binding is entropically favourable, as they are released into bulk solvent, without too much enthalpic cost. This relatively small enthalpic cost can be compensated by the protein–ligand interactions.

Fig. 23.4.4.8 shows all of the 1661 water molecules colour-coded by the various classifications described above. Clearly, the entire surface of the protein is well hydrated. Notice how the yellow channel waters are often followed by a red buried water molecule. In addition, there is often no obvious spatial distinction between molecules categorized as crystal contacts (green) and those categorized as surface (blue).

23.4.4.2.2. *T4 lysozyme*

Over 150 mutants of T4 lysozyme have been studied to date, and, for the majority of these, the crystal structures are available. Although most of the mutant structures crystallize isomorphously to the wild type, many of them provide a view of the molecule in different crystal environments. This collection of structures leads to

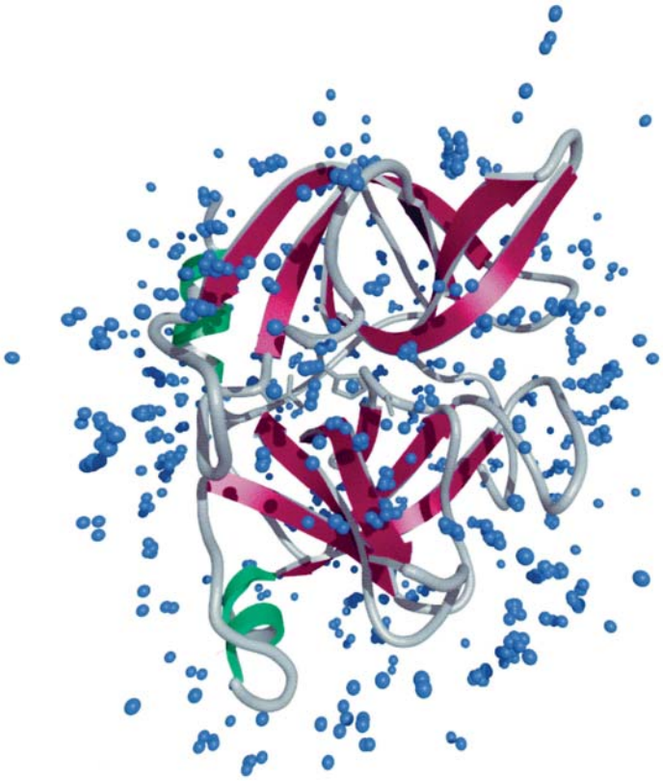

Fig. 23.4.4.7. Elastase structure represented as in Fig. 23.4.4.4. The surface crystallographic water molecules found in 11 superimposed elastase structures solved in a variety of solvents are shown in blue.

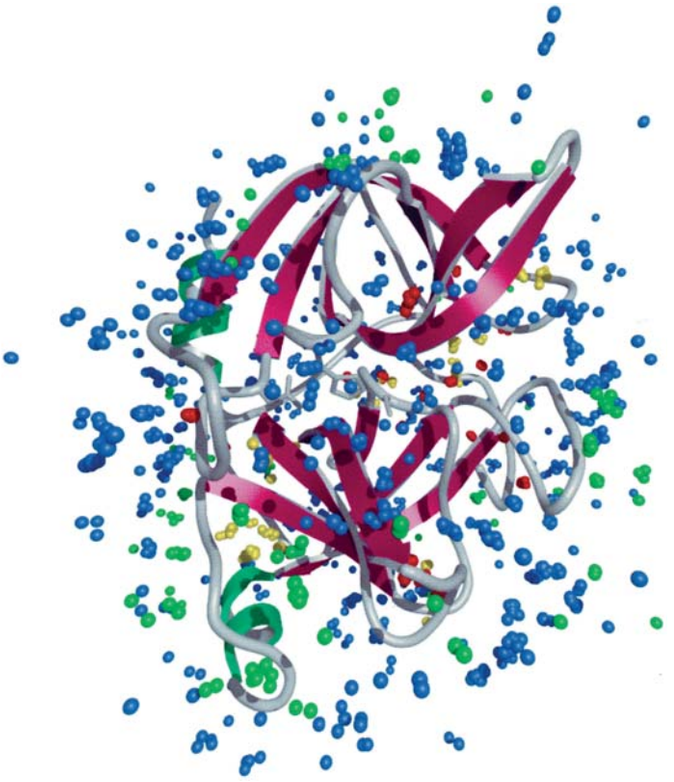

Fig. 23.4.4.8. Elastase structure represented as in Fig. 23.4.4.4. The 1661 water molecules found in 11 superimposed elastase structures of elastase are colour-coded as in Figs. 23.4.4.4–23.4.4.7.

references