Molecular mimicry of substrate oxygen atoms by water molecules in the β -amylase active site

GERARD PUJADAS AND JAUME PALAU

Unitat de Biotecnologia Computacional, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Tarragona, Catalonia, Spain

(RECEIVED March 1, 2001; FINAL REVISION May 23, 2001; ACCEPTED May 23, 2001)

Abstract

Soybean β -amylase (EC 3.2.1.2) has been crystallized both free and complexed with a variety of ligands. Four water molecules in the free-enzyme catalytic cleft form a multihydrogen-bond network with eight strategic residues involved in enzyme-ligand hydrogen bonds. We show here that the positions of these four water molecules are coincident with the positions of four potential oxygen atoms of the ligands within the complex. Some of these waters are displaced from the active site when the ligands bind to the enzyme. How many are displaced depends on the shape of the ligand. This means that when one of the four positions is not occupied by a ligand oxygen atom, the corresponding water remains. We studied the functional/structural role of these four waters and conclude that their presence means that the conformation of the eight side chains is fixed in all situations (free or complexed enzyme) and preserved from unwanted or forbidden conformational changes that could hamper the catalytic mechanism. The water structure at the active pocket of β -amylase is therefore essential for providing the ligand recognition process with plasticity. It does not affect the protein active-site geometry and preserves the overall hydrogen-bonding network, irrespective of which ligand is bound to the enzyme. We also investigated whether other enzymes showed a similar role for water. Finally, we discuss the potential use of these results for predicting whether water molecules can mimic ligand atoms in the active center.

Keywords: TIM barrel; β -amylase; molecular mimicry; water molecules; active site; energy of interaction

Supplemental material: Kinemage. See www.proteinscience.org.

β -Amylase (BAMY; EC 3.2.1.2) is a glycoside hydrolase that hydrolyzes the (1,4)- α -D-glucosidic linkages of a variety of substrates such as starch, glycogen, and maltooligosaccharides from the nonreducing end of their polysaccharide chains. The enzyme follows an α -inverting mechanism (McCarter and Withers 1994), and the final products of the enzymatic action are β -maltose and a β -limit dextrin. The general acid residue involved in the hydrolytic reaction is

Glu186, whereas Glu380 acts as the general base (henceforth, soybean β -amylase (SBA) numbering) (Mikami et al. 1994). BAMY also catalyzes other reactions such as the condensation of two β -maltose units to maltotetraose (Hehre et al. 1969) and the slow irreversible hydration of maltal to form β -2-deoxymaltose (Kitahata et al. 1991). In maltal hydration and β -maltose condensation, the role of the two catalytic residues is interchanged (i.e., Glu380 acts as the general acid and Glu186 as the general base) (Mikami et al. 1994).

X-ray structures for SBA are known; they are either free or complexed with ligands. The Protein Data Bank (PDB) code is 1BYA for the free enzyme and 1BYB, 1BYC, 1BYD, 1BTC, and 1BFN for the complexes. From a structural point of view, BAMY has a core with a TIM barrel fold (Pujadas and Palau 1999) plus a smaller globular region formed by long loops (L3, L4, and L5) extending from β -strands β 3, β 4, and β 5. A region of the L3 (residues

Reprint requests to: Gerard Pujadas, Unitat de Biotecnologia Computacional, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Plaça de la Imperial Tàrraco, 1. Tarragona 43005, Catalonia, Spain; e-mail: pujadas@quimica.urv.es; fax: 34 (977) 55 95 97.

Abbreviations: BAMY, β -amylase; DHODA, dihydroorotate dehydrogenase from gene *pyrDa*; EDM, electron density map; HIV-PR, human immunodeficiency virus protease; LLM, lowest local minima; LSCWs, ligand-specific conserved waters; PDB, Protein Data Bank; RSFit, real space fit; SBA, soybean β -amylase.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1101/ps.8201>.

96–103) (Mikami et al. 1994) and the $^{341}\text{FTC}^{343}$ segment of $\beta 6$ (Pujadas and Palau 1997) undergo conformational changes that are associated with substrate binding. These conformational changes allow the interaction of Asp101 and Thr342 with the substrate. Recent docking studies have suggested that residues 96–103 contribute not only to position the glycosidic bond next to the catalytic residues but also to distort the substrate conformation so that its shape approaches the transition-state structure (Laederach et al. 1999). The active-site topology for BAMY is of the pocket or crater type, which is optimal recognizing the nonreducing extremity of saccharides (Davies and Henrissat 1995).

A set of SBA residues (Asp53, His93, Asp101, Glu186, Arg188, Tyr192, Lys295, Ser297, Gly298, His300, Thr342, Glu380 and Arg420) was identified as proton donors or acceptors in hydrogen bonds with maltotetraose (1BYB and 1BYC; since both structures fully coincide, hereafter we will deal with only 1BYB) and β -2-deoxymaltose (1BYD) oxygen atoms (Mikami et al. 1994). Moreover, Ala382 makes an additional hydrogen bond with the former ligand and Asn381 makes an additional one with the latter. The full set of these ligand-binding residues, except Ser297, is conserved for 28 BAMY complete sequences from 16 different sources, and all (except His93) belong to some of the eight highly preserved sequence motifs that define the BAMY fingerprint (Pujadas et al. 1996). Obviously, the enzyme/ligand hydrogen bonds are not possible for the free SBA structure (1BYA). Alternatively, the binding pocket (5.0 Å around the volume that is to be occupied by any atom of maltotetraose or β -2-deoxymaltose) is filled with 26 water molecules. Five of these water molecules (Wat602_{1BYA}, Wat784_{1BYA}, Wat785_{1BYA}, Wat786_{1BYA}, and Wat789_{1BYA}) are hydrogen bonded with a substantial number of the protein atoms involved in the enzyme/ligand interactions.

In this study, we used the structure of SBA to investigate the functional/structural role of water molecules in the BAMY active site. SBA has been crystallized in different situations (free and complexed with a variety of ligands) and therefore provides an excellent framework for studying this role. We also compare our results from the analysis of the SBA active pocket with those for other enzymes that have a similar role for water molecules in their active sites.

Results and Discussion

Molecular mimicry of the ligand shape by a chain of water molecules

We show, by structural superposition of 1BYA with 1BYB and 1BYD, the existence of a general molecular mimicry of the ligand shape by a chain of 13 water molecules located at distances ≤ 1.5 Å around ligand atoms (see waters in cyan in Fig. 1A and B and Kinemage 1). This number is high, since it represents 50% of the water molecules found in the free-enzyme binding pocket (see above). In addition, when we compare these 13 water oxygen coordinates with those of the nearest-neighbor oxygen atoms of the ligand, we find rmsd values that are significantly low (1.3 and 1.2 Å for maltotetraose and β -2-deoxymaltose, respectively). We also found that, when the substrate binds, these 13 water molecules are expelled (and also Wat684_{1BYA} for maltotetraose and Wat779_{1BYA} and Wat827_{1BYA} for β -2-deoxymaltose; see green water molecules in Fig. 1A and B).

Four water molecules present in 1BYA are coincident (< 0.9 Å) with oxygen atoms of the ligands (Wat602 with O4 Glc496_{1BYB}/Dom496_{1BYD}; Wat784 with O1' Dom496_{1BYD}; Wat785 with O6 Glc497_{1BYB}/O6' Dom496_{1BYD}; Wat789 with O6 Glc498_{1BYB}/Dom498_{1BYD}). The behavior of Wat784_{1BYA} is different for maltotetraose and β -2-deoxymaltose because of the local structure of these ligands around this water molecule (i.e., in the same place as Wat784_{1BYA} there is no atom for maltotetraose but there is O1' Dom496_{1BYD} for β -2-deoxymaltose; see Fig. 1A and B and Kinemage 1).

How reliable are these crystallographic waters?

There is some doubt about the placement of at least some crystallographic waters in X-ray structures (certainly, many are not as reliable as well-defined protein atoms). We therefore need to know how reliable the water molecules shown in Figure 1 are before further analyzing their possible functional/structural role.

The best way of analyzing the correctness of water molecules in crystallographic models is to justify their presence from a biochemical point of view. Wat784_{1BYA} has been suggested as the reactant water molecule needed for the

Fig. 1. Stereo pairs. Water molecules from 1BYA located at ≤ 1.5 Å around the area that will be occupied by any ligand atom in the complexed structures. The relative positions of these water molecules with respect to maltotetraose (A), β -2-deoxymaltose (B), α -cyclodextrin (C), or β -cyclodextrin (D) are shown. Water molecules in cyan or green are absent for the corresponding complexes, whereas yellow waters remain. Cyan water molecules superpose well with maltotetraose and β -2-deoxymaltose oxygen atoms (rmsd = 1.3 and 1.2 Å for maltotetraose and β -2-deoxymaltose, respectively). Water molecules with square labels are involved in multihydrogen bonds with catalytically important residues (Table 2). Three of these multihydrogen-bonded water molecules are exactly replaced by ligand oxygen atoms in the case of maltotetraose (Wat602, Wat785, and Wat789) and all four in the case of β -2-deoxymaltose. Note that the multihydrogen-bonded Wat785 is expelled by β -cyclodextrin although no inhibitor oxygen atom or equivalent water molecule replaces it. The lack of Wat785 causes a sidechain rearrangement for Lys295, as described in Fig. 2. See Table 1 for equivalences between active site waters in the different SBA structures.

Fig. 1.

hydrolytic reaction of BAMY (Mikami et al. 1994). Wat866_{1BYB}, its equivalent in 1BYB, is: (a) hydrogen bonded to the proposed general base (i.e., Glu380); (b) located at 2.76 Å from the C1 atom of the second glucose unit in maltotetraose (the carbon atom that is expected to be water-attacked in the hydrolytic reaction); (c) on the opposite side of the glycosidic oxygen in the (1,4)- α -D-glucosidic linkage to be hydrolyzed (and is therefore ready to invert the configuration and obtain β -anomeric products); (d) substituted by the O1' Dom496 atom of β -2-deoxymaltose in the hydration of maltal (1BYD); (e) present prior to the loop 3 closure that shields the reaction center from the solvent (i.e., Wat784_{1BYA}); and, finally, (f) present in all structures that have the potential to hydrolyze carbohydrates (i.e., Wat784 in 1BYA; Wat866 in 1BYB; Wat778 in 1BTC, and Wat648 in 1BFN). No comparable ordered water is present near Glu186. Therefore, we can conclude that the correctness of Wat784_{1BYA} — and its equivalents in other SBA structures — is fully justified from the biochemical point of view.

No biochemical role has yet been described for the other water molecules. We therefore studied their reliability: (a) by directly comparing the PDB file with its corresponding electron density map (EDM); (b) by using temperature factors found in the PDB files; and, finally, (c) by analyzing the energy of these waters and comparing their positions in the crystallographic models with the positions of local energy minima for water around them. Results for the reactant water (i.e., Wat784_{1BYA} and counterparts in the other SBA

structures) are also included for comparison with data from the other waters whose reliability has been studied (see Table 1 for equivalences between active site waters in the different SBA structures).

1BFN (Adachi et al. 1998) is the only SBA structure whose structure factors have been deposited in the PDB. The “sphericity” of the electron density surrounding Wat648_{1BFN}, Wat660_{1BFN}, Wat678_{1BFN}, Wat749_{1BFN}, and Wat752_{1BFN} on a σ_A -weighted 2Fo-Fc map displayed at a contour level of 1 σ (contour level 1 and scaled on an absolute level) shows a real space fit (RSFit) of 0.079, 0.058, 0.073, 0.093, and 0.338, respectively. The lower the RSFit is, the better the water molecule is defined at the point at which it has been modeled. We can therefore conclude that four of five water molecules in the SBA active pocket of 1BFN (Wat648, Wat660, Wat678, and Wat749) have been correctly modeled and are coherent with its EDM. As we will describe in detail in the next section, these four water molecules are hydrogen bonded with a substantial number of the protein atoms involved in the enzyme/ligand interactions. Wat752_{1BFN} fits rather poorly to the density, judging from the RSFit values. Moreover, this water molecule is not hydrogen bonded to any enzyme atom (result not shown), and any functional/structural role it has in SBA must therefore be very limited.

There is no EDM available for the other SBA structures. We have therefore evaluated the reliability of their active-pocket water molecules by comparing their atomic displacement parameters (i.e., temperature or B-factors) with those

Table 1. Temperature factors for the water molecules in the SBA active pocket

	1BYA			1BYB			1BYD			1BTC			1BFN		
	Label	B-factor	B'-factor	Label	B-factor	B'-factor	Label	B-factor	B'-factor	Label	B-factor	B'-factor	Label	B-factor	B'-factor
a	602	23.5	−0.1	—	—	—	—	—	—	596	33.8	2.1	749	41.4	2.4
	684	70.7	4.5	—	—	—	667	78.4	4.9	—	—	—	—	—	—
	710	32.0	0.7	—	—	—	—	—	—	—	—	—	—	—	—
	767	55.5	3.0	—	—	—	—	—	—	—	—	—	—	—	—
	774	33.3	0.9	—	—	—	—	—	—	—	—	—	—	—	—
	779	58.1	3.3	500	24.2	0.3	—	—	—	—	—	—	—	—	—
	784	43.8	1.9	866	19.2	−0.2	—	—	—	778	26.4	0.7	648	39.2	2.1
a	785	30.9	0.6	—	—	—	—	—	—	779	26.1	0.6	—	—	—
b	786	25.9	0.1	—	—	—	—	—	—	780	33.4	2.1	678	37.6	1.9
a	789	31.1	0.6	—	—	—	—	—	—	783	26.1	0.6	660	38.9	2.1
	809	29.2	0.4	—	—	—	—	—	—	—	—	—	—	—	—
	810	37.4	1.3	—	—	—	—	—	—	801	37.4	2.9	752	54.9	3.9
	824	39.2	1.4	—	—	—	—	—	—	—	—	—	—	—	—
	825	45.5	2.1	—	—	—	—	—	—	—	—	—	—	—	—
	826	49.0	2.4	—	—	—	—	—	—	—	—	—	—	—	—
	827	44.7	2.0	—	—	—	—	—	—	—	—	—	—	—	—

Data for equivalent water molecules but different SBA structures are shown in the same row (the first column for each SBA structure shows the label used to name water molecules in the corresponding PDB file). The B-factor (in Å²) and B'-factor columns show the temperature factor and the normalized temperature factor, respectively, for each crystallographic water.

^a Water molecules that form a multihydrogen-bond network with a substantial number of the protein atoms involved in the enzyme/ligand interactions (see Table 2). These water molecules are coincident with oxygen atoms of the ligands (see Figures 1A, B).

^b This water molecule is hydrogen-bonded to OD2 Asp53.

from equivalent waters in 1BFN (Adachi et al. 1998). According to Matthews (1993), atoms with B-factors in the range 0 to 40 Å² are usually well defined. When the B-factor exceeds about 60 Å², the atom is so mobile or disordered that it can no longer be seen reliably in EDMs. Nevertheless, B-factor values in protein crystal structures vary not only because of genuine physical differences but also because of refinement strategies (Frauenfelder and Petsko 1980; Ringe and Petsko 1986; Tronrud 1996), which makes Matthews' rules difficult to apply when comparing independent structures [there are usually large variations in the B-factors even when X-ray structures obtained at high resolution are compared (Parthasarathy and Murthy 1997)]. We therefore used normalized B values (see B'-factors in Table 1) to compare the reliability of equivalent water molecules in the full set of SBA structures (Parthasarathy and Murthy 1997; Carugo 1999). Table 1 shows that B'-factors for three of four water molecules in 1BYA that are coinci-

dent with oxygen atoms of the ligands (Wat602, Wat784, and Wat789) are lower than those for equivalent waters in 1BFN (Wat749, Wat648, and Wat660, respectively) and that their reliability therefore compares favorably with that of equivalent waters in 1BFN. This is also true for equivalent waters in 1BYB (Wat866) and 1BTC (Wat596, Wat778, and Wat783). The fourth water molecule coincident with oxygen atoms of the ligands (Wat785_{1BYA}) has no equivalent in 1BFN (see Table 1). However, three factors support the reliability of this water molecule and its equivalent in 1BTC (Wat779): first, it has a relatively low B'-factor (0.6; also for Wat779_{1BTC}); second, it forms two hydrogen bonds with the enzyme (NZ Lys295 and OE1 Glu380; see Table 2), and third, the absence of a hydrogen bond partner (either from a water molecule or from the ligand) in the position occupied by Wat785 in 1BYA causes Lys295 in 1BFN to change its side-chain conformation (this will be seen later in greater detail; see Fig. 2). Another water

Table 2. Hydrogen bonds between ligand atoms or water molecules mimicking oxygen ligand atoms and SBA

		1BYA	1BYB	1BYD	1BTC	1BFN
OD1	Asp 53	Wat602	O4 Glc496	O4 Dom496	Wat596	Wat749
OD2	Asp 53	—	O6 Glc496	O6 Dom496	—	—
NE2	His 93	—	O3 Glc496	O3 Dom496	—	—
NE2	His 93	Wat602	O4 Glc496	O4 Dom496	Wat596	Wat749
OD1	Asp101	—	O2 Glc496	O2 Dom496	—	—
OE1	Glu186	Wat789	O6 Glc498	O6 Dom498	Wat783	Wat660
OE2	Glu186	—	O4 Glc498	O4 Dom498	—	—
NH2	Arg188	Wat789	O6 Glc498	O6 Dom498	Wat783	Wat660
OH	Tyr192	—	O5 Glc498	O5 Dom498	—	—
OH	Tyr192	Wat789	O6 Glc498	O6 Dom498	Wat783	Wat660
NZ	Lys295	Wat785	O6 Glc497	O6' Dom496	Wat779	—
OG	Ser297	—	O3 Glc498	O3 Dom498	—	—
N	Gly298	—	O3 Glc498	O3 Dom498	—	—
O	Gly298	—	O2 Glc498	O2 Dom498	—	—
NE2	His300	—	O3 Glc499	O3' Dom498	—	—
NE2	His300	—	—	—	O3 Glc004	O3 Glc004
O	Thr342	—	O3 Glc498	O3 Dom498	—	—
O	Thr342	^a	Wat866	O1' Dom496	Wat778	^a
OG1	Thr342	—	O3 Glc498	O3 Dom498	—	—
OG1	Thr342	—	O4 Glc498	O4 Dom498	—	—
OE1	Gln351	—	—	—	—	O2 Glc004
OE1	Gln351	—	—	—	—	O3 Glc005
OE1	Glu380	Wat784	Wat866	O1' Dom496	Wat778	Wat648
OE1	Glu380	Wat785	O6 Glc497	O6' Dom496	Wat779	—
O	Asn381	Wat784	Wat866	O1' Dom496	Wat778	Wat648
O	Ala382	—	O2 Glc497	—	—	—
O	Ala382	—	—	—	O2 Glc002	O2 Glc002
O	Ala382	—	—	—	O3 Glc003	O3 Glc003
NH2	Arg420	—	O5 Glc496	O5 Dom496	—	—
NH2	Arg420	—	O6 Glc496	O6 Dom496	—	—

Ligand atoms and water molecules that make spatially equivalent hydrogen bonds with the same enzyme atom in different SBA structures are shown in the same row. Atom labels in italics indicate hydrogen bonds not detected by LIGPLOT/HBPLUS (McDonald and Thornton 1994; Wallace et al. 1995). However, visual inspection shows that these hydrogen bonds are only just outside the detection limits of the programs.

^a The Thr342 conformation in 1BYA and 1BFN is different from the one in 1BYB, 1BYD and 1BTC (Pujadas and Palau 1997). Consequently, the main chain oxygen atom of Thr342_{1BYA/1BFN} cannot interact with Wat784_{1BYA} or Wat648_{1BFN}, which mimic O1' Dom496.

Fig. 2. Stereo pairs. The absence of a hydrogen-bond acceptor in 1BFN (*D*) located at an equivalent position to Wat785 in 1BYA (*A*) causes a strong conformational change in Lys295. In 1BFN, the Lys295 sidechain occupies part of the virtual space assigned to maltotetraose or β -2-deoxymaltose; compare the position for the ligand in 1BYB (*B*) and the Lys295 sidechain in 1BFN (*D*). This indicates that the presence of a hydrogen-bond acceptor equivalent to the free enzyme water molecule Wat785 is essential for fixing an interacting metastable sidechain conformation for Lys295. Moreover, in 1BFN Lys295 plays the same role as the two water molecules hydrogen bonded to OE1 Glu380 in 1BYA and 1BTC (*C*) and is enough to protect the Glu380 sidechain from conformational changes.

molecule (Wat786_{1BYA}) is also hydrogen bonded to one protein atom (OD2 Asp53) involved in enzyme/ligand interactions. Wat786_{1BYA} is equivalent to Wat678_{1BFN}, one of the four water molecules in the SBA active pocket of 1BFN that have been correctly modeled (see above). Table 1 shows that the B'-factor for Wat786_{1BYA} (0.1) is significantly lower than the B'-factor for Wat678_{1BFN} (1.9). In our opinion, this is proof of the reliability of Wat786_{1BYA}. The equivalent water in 1BTC (Wat780) has a slightly higher B'-factor (2.1) than Wat678_{1BFN} (1.9).

We also performed an analysis with GRID (Goodford 1985) to evaluate the energy of interaction between the enzyme and each of the crystallographic water molecules Wat602, Wat784, Wat785, Wat786, and Wat789 in 1BYA (and their equivalents in the rest of the SBA structures) and to compare the location of these water molecules with local minima of interaction energy between SBA and water. For this second objective, we searched for all of the local minima present in a cubic "cage" 5.0 Å long. Each cage is approximately centered at the position of the 1BYA crystallographic water that is equivalent to the one being studied. A common cage definition for equivalent water molecules makes it easier to compare the position in the cage of the lowest local minima (LLM) for the full set of SBA structures. Results of this energetic analysis are summarized in Table 3 and Kinemage 2. We can see that: (a) 14 of 15 crystallographic waters favorably interact with their protein environment (attractive energies range from -6.6 to -12.0 kcal/mol); (b) Wat784_{1BYA} has a positive energy (+3.6 kcal/mol) and is therefore not energetically favored at its present location (although it is only 0.7 Å from the LLM in its cage; -10.4 kcal/mol); (c) 11 of 15 crystallographic waters are less than 1.2 Å from their LLM (a distance of 1.2 Å between oxygen atoms in two waters results in considerable overlap and provides a conservative criterion for defining

equivalent water sites in superimposed structures (Zhang and Matthews 1994)); and (d) two of the four waters that are more than 1.2 Å from their LLM (Wat783_{1BTC} and Wat660_{1BFN}) are near other important local minima (Table 3). We also observed that LLM for waters equivalent to Wat785_{1BYA} do not cluster (this is also true for the LLM of waters equivalent to Wat789_{1BYA}). The LLMs of Wat785_{1BYA} and Wat779_{1BTC} have different positions. This is because the protein environments of these two equivalent waters are different; the conformation of Thr342 in 1BYA is different from that in 1BTC (Pujadas and Palau 1997). The immediate protein environment around waters equivalent to Wat789_{1BYA} is equivalent in 1BYA, 1BTC, and 1BFN. We therefore suggest that the different positions of the LLMs are due to medium- to long-range effects. In addition, the positive value for the energy of Wat784_{1BYA} and its location just beside the LLM suggest that this water would be better modeled if it were slightly displaced from its present location in 1BYA. Nevertheless, this does not affect the fact that Wat784_{1BYA} is fully reliable (see discussion above about its role in the catalytic mechanism).

There are 11 more water molecules in the active pocket of 1BYA (Wat684, Wat710, Wat767, Wat774, Wat779, Wat809, Wat810, Wat824, Wat825, Wat826, and Wat827; see Fig. 1). None of these waters is directly hydrogen bonded to enzyme atoms, and they therefore are not part of the first hydration shell of SBA. Their B'-factors are generally larger than those of the water molecules that form hydrogen bonds in the 1BYA active pocket (Table 1). Only three of these 11 water molecules (Wat684, Wat779, and Wat810) have an equivalent in the other SBA structures (although the counterpart of Wat810 in 1BFN, i.e., Wat752, fits to the EDM rather poorly, judging from the RSFit value; see above).

Table 3. Comparison of the energy of interaction of SBA with (i) selected crystallographic waters in its active site and (ii) probe water molecules at the LLM around these crystallographic waters

1BYA				1BYB				1BTC				1BFN			
Crystallographic water		Water probe at the local minima		Crystallographic water		Water probe at the local minima		Crystallographic water		Water probe at the local minima		Crystallographic water		Water probe at the local minima	
Label ^a	Energy	Distance ^b	Energy	Label ^a	Energy	Distance ^b	Energy	Label ^a	Energy	Distance ^b	Energy	Label ^a	Energy	Distance ^b	Energy
602	-8.1	0.6	-11.0	—	—	—	—	596	-9.8	0.6	-11.0	749	-10.7	0.2	-11.3
784	+3.6	0.7	-10.4	866	-9.9	0.6	-13.1	778	-12.0	0.2	-12.2	648	-10.9	1.0	-14.9
785	-7.1	2.0	-10.9	—	—	—	—	779	-8.4	1.1	-11.5	—	—	—	—
786	-6.6	1.6	-12.8	—	—	—	—	780	-10.7	0.8	-12.4	678	-11.3	0.2	-11.8
789	-9.3	0.9	-11.5	—	—	—	—	783	-10.0	1.3 ^c	-12.0	660	-8.0	2.9 ^d	-11.0

Energies of interaction are given in kcal/mol and distances in Å. Data for equivalent waters in the different SBA structures are in the same row.

^a Labels of the crystallographic waters in the corresponding SBA structures.

^b Distance between the crystallographic water and the LLM in a 5.0 Å cubic "cage" and approximately centered in the corresponding crystallographic water of 1BYA.

^c Wat783_{1BTC} is at 0.7 Å of the second LLM (-10.7 kcal/mol) in the cage defined around Wat789_{1BYA}.

^d Wat660_{1BFN} is at 0.7 Å of the third LLM (-10.1 kcal/mol) in the cage defined around Wat789_{1BYA}.

We may conclude, therefore, that there are objective reasons to support the reliability of the hydration sites defined by the five water molecules that are hydrogen bonded to active pocket residues in 1BYA (Wat602, Wat784, Wat785, Wat786, and Wat789) and their equivalents in the other SBA structures. Four of these five waters are the ones that are coincident with oxygen atoms of the ligands (Wat602, Wat784, Wat785, and Wat789), and Wat784 may be identified as the reactant water. The reliability of the other 11 water molecules that are also inside the active pocket of 1BYA is not well supported by this analysis. However, their lack of direct interaction with the active center means that any role they have in the catalytic mechanism of BAMY must be very marginal. The study of the functional/structural role of Wat602_{1BYA}, Wat784_{1BYA}, Wat785_{1BYA}, Wat786_{1BYA}, Wat789_{1BYA}, and their counterparts in the other SBA structures now becomes fully justified.

Role of key oxygen atoms (from ligand or water) in the enzyme catalytic pocket

The ligand oxygen atoms that are coincident with Wat602_{1BYA}, Wat784_{1BYA}, Wat785_{1BYA}, and Wat789_{1BYA} (O4 Glc496/Dom496, O1' Dom496, O6 Glc497/O6' Dom496, and O6 Glc498/Dom498, respectively) are actively involved in enzyme/ligand hydrogen bonding (see Table 2 and Kinemage 1). Similarly, these four water molecules keep the same hydrogen-bond partners in the free enzyme as those of the ligand oxygen atoms that they mimic [the only exception is Thr342, whose conformation is known to change when the substrate is bound (Pujadas and Palau 1997); see Table 2 and Kinemage 1]. Consequently, in the free enzyme, Wat602, Wat784, and Wat785 form two hydrogen bonds and Wat789 forms three. Apparently, the water-mediated mimicry avoids conformational changes for Asp53, His93, Glu186, Arg188, Tyr192, Glu380, and Asn381, either when the ligand is absent or when the corresponding residue atom is not used in enzyme/ligand hydrogen bonds (i.e., the interaction of Wat866_{1BYB} with OE1 Glu380 and O Asn381; Table 2). With Lys295, the position of the NZ atom (responsible for Lys sidechain hydrogen bonds) is equivalent in 1BYA, 1BYB, and 1BYD despite the fact that χ_2 for 1BYA is 120.1° and $-172.4 \pm 5.5^\circ$ for the enzyme/ligand structures (Fig. 2A and B). Therefore, these four key oxygen atoms (from ligand or water) fix the conformation of the eight side chains in all situations (free or complexed enzyme) and preserve them from unwanted or forbidden conformational changes that could hamper the catalytic mechanism.

We used the Consolv algorithm (Raymer et al. 1997) to predict whether Wat602_{1BYA}, Wat784_{1BYA}, Wat785_{1BYA}, and Wat789_{1BYA} could be displaced by the ligands. Only Wat785_{1BYA} was not correctly predicted as displaceable. Wat786_{1BYA}, which does not mimic any oxygen ligand

atom but is hydrogen bonded to the active site residue Asp53, was correctly predicted to be displaceable by the ligand.

Not all residues that make enzyme/ligand hydrogen bonds have water that mimics the ligand donor/acceptor atom in the free enzyme

In six of the 15 residues involved in hydrogen bonds with maltotetraose or β -2-deoxymaltose (Asp101, Ser297, Gly298, His300, Ala382, and Arg420), no water mimics the ligand donor/acceptor in the free-enzyme structure (Table 2). The search for their hydrogen-bond partners in 1BYA shows that: (a) Gly298 and Arg420 are not hydrogen bonded; (b) Asp101 and Ala382 have only one hydrogen bond that is also found for the corresponding complexed structures (OD2 Asp101 with N Asn98 and N Ala382 with OE1 Glu345); and (c) Ser297 and His300 have the same residue-residue interactions as the complexed structures (N Ser297 with OH Tyr316, OG Ser297 with OH Tyr192; O His300 with N Tyr303 and N Gly358, ND1 His300 with OG Ser357). Therefore, the number of hydrogen bonds where these six residues are involved is increased by the binding of the ligand.

Thr342 has a water molecule that mimics the ligand donor/acceptor in 1BYA (Wat784), but Thr342_{1BYA} has a metastable conformation that is different from the one in 1BYB and 1BYD (Pujadas and Palau 1997). Consequently, the main chain oxygen atom of Thr342_{1BYA} is not able to interact with Wat784 (Table 2). Moreover, Asp101 belongs to the hinged-lid motif (residues 96–103) that moves to cover the reaction center when the enzyme/substrate complex forms (Mikami et al. 1993; Mikami et al. 1994). In addition, χ_2 for Arg420 in 1BYA shows an unusual value (1.6°) where C $^\alpha$ and C $^\delta$ are almost eclipsed and so have the maximum steric hindrance for χ_2 (Dunbrack and Karplus 1994). On the other hand, χ_2 for Arg420 in the complexes changes to its most stable conformation: trans ($143.3 \pm 15.6^\circ$) (Dunbrack and Karplus 1994).

All the above observations suggest that Asp101, Ser297, Gly298, His300, Thr342, Ala382, and Arg420 are more stable in the complexes than in the free enzyme. Once again, a local surplus of macromolecule instability in the free enzyme structure appears to maximize the functional role of the active site (Shoichet et al. 1995; Pujadas and Palau 1997). Therefore, this supports the “stability-function” hypothesis for proteins that recognize a ligand; this states that enzyme residues involved in catalysis or ligand binding are not optimized for stability (Shoichet et al. 1995). Our previous finding that there is a negative enthalpy change (-27 kcal/mol) for the $^{341}\text{FTC}^{343}$ segment of $\beta 6$ as a result of conformational transitions in the ligand-binding process also supports this hypothesis for SBA (Pujadas and Palau 1997).

Analysis of SBA-inhibitor complexes provides proof that water molecules have a conformational stabilizing role on hydrogen bonding with β -amylase residues

SBA has also been crystallized with two different inhibitors, α - and β -cyclodextrin (1BTC and 1BFN structures, respectively) (Mikami et al. 1993; Adachi et al. 1998). The binding site for α - and β -cyclodextrin is not so deep as for maltotetraose and β -2-deoxymaltose (Mikami et al. 1994; Adachi et al. 1998). This means that some of the water molecules in 1BYA that were expelled by maltotetraose and β -2-deoxymaltose remain at the original site of the free enzyme (yellow water molecules in Fig. 1C and D; see also Table 2), protecting the sidechains of the “nonused” enzyme ligand-binding atoms from conformational changes. The sequence for 1BTC coincides with that for 1BYA, but 1BFN has three differences (Phe76Leu, Arg202Gly, and Lys399Arg). These three residues are located far from the binding contact surface of the inhibitor (Adachi et al. 1998), and such changes are therefore not expected to influence the contacts very much.

No water molecule equivalent to Wat785 in 1BYA is present in 1BFN (see Fig. 1D, Table 1, and Kinemage 3), unlike with 1BTC (Wat779). This absence causes a strong conformational change in Lys295_{1BFN} (the whole conformation changes from *g+ttt* to *tttt*, in which χ_1 moves from -61.3° for 1BTC to -169.8° for 1BFN; Fig. 2C and D). In 1BFN, the Lys295 sidechain occupies part of the virtual space assigned to maltotetraose or β -2-deoxymaltose (the virtual distances between NZ Lys295_{1BFN} and C6 Glc497/C6' Dom496 are about 1.3–1.5 Å). Therefore, in the absence of maltotetraose or β -2-deoxymaltose, a water molecule equivalent to Wat785_{1BYA} appears to be essential for fixing an interacting metastable sidechain conformation for Lys295. This provides additional evidence for the conformational stabilizing role of water molecules upon hydrogen bonding with β -amylase residues and supports results of studies that show how water molecules in the interior of a protein contribute favorably to its conformational stability (Funahashi et al. 1996; Takano et al. 1997). The OE1 atom for Glu380 sidechain is also hydrogen bonded to Wat785 in 1BYA and to its equivalent water molecule in 1BTC (Wat779) (Fig. 2A and C). Nevertheless, the Glu380 sidechain conformation is the same for 1BYA, 1BTC, and 1BFN. One explanation for the different behavior of Glu380 and Lys295 in 1BFN is that there is a strong interaction between OE1 Glu380_{1BFN} and NZ Lys295_{1BFN} (2.8 Å) that would be impossible without the conformational change of the latter residue (Fig. 2D). Lys295 therefore plays the same role in 1BFN as the two water molecules hydrogen bonded to OE1 Glu380 in 1BYA (Wat784 and Wat785) and 1BTC (Wat778 and Wat779) and is enough to protect the Glu380_{1BFN} sidechain from conformational changes.

Other examples of molecular mimicry of polar ligand atoms by water molecules in the enzyme active site

We looked for other examples of molecular mimicry of polar ligand atoms by water molecules in the enzyme active site and found few, although significant, examples of this role in the literature (Lam et al. 1994; Rowland et al. 1997, 1998; Shaltiel et al. 1998; Niefind et al. 1999). We will compare these with our own results on SBA.

Lam and coworkers made a rational design of a series of nonpeptide cyclic ureas that inhibit the human immunodeficiency virus protease (HIV-PR). The fundamental feature of these inhibitors is that the carbonyl oxygen of the cyclic urea mimics the hydrogen-bonding features of a key structural water molecule present in previously published HIV-PR/inhibitor complexes (Erickson et al. 1990). This water molecule accepts two hydrogen bonds from backbone amide hydrogens of residue Ile50 in chains A and B (strictly preserved by evolution in all known HIV-PR sequences) and donates two hydrogen bonds to two carbonyl oxygens of the inhibitor. Incorporating a mimic for this water within the inhibitor was a key feature of the inhibitor design that ensures specificity for the HIV protease against other aspartic acid proteases (the water molecule is only found in retroviral proteases). The Consolv algorithm (Raymer et al. 1997) correctly predicts that Wat308_{1HSG} is displaceable by the inhibitors. The B'-factor for Wat308_{1HSG} is -1.1 .

Using the free (1DOR) and the product-complexed (2DOR) structures of the enzyme dihydroorotate dehydrogenase from gene *pyrDa* of *Lactococcus lactis* (DHODA) determined to 2.0 Å resolution, Rowland and coworkers showed that the substrate-binding cavity of DHODA is filled with three water molecules that leave when the substrate binds. The three water molecules form hydrogen bonds with seven residues (Wat1047_{1DOR} with ND2 Asn67, ND2 Asn132, and OG Ser194; Wat1108_{1DOR} with SG Cys130 and OD1 Asn193; and Wat1246_{1DOR} with NZ Lys43 and N Leu71) that have been preserved throughout evolution in all dihydroorotate dehydrogenases from family 1A [formed by sequences from anaerobic yeasts, some protozoa and milk-fermenting bacteria (Rowland et al. 2000)]. These water molecules indicate where the atoms of the substrate capable of forming hydrogen bonds must be and, therefore, help to orient the substrate in the active site (Wat1047, Wat1108, and Wat1246 in subunit A of 1DOR are replaced by the hydrophilic O2, N3, and O71 orotate atoms, respectively, in 2DOR). The same results are found for subunit B of DHODA. The Consolv algorithm (Raymer et al. 1997) correctly predicted that Wat1108_{1DOR} and Wat1246_{1DOR} are displaceable by the ligand, whereas Wat1047_{1DOR} was incorrectly predicted as being conserved in enzyme-ligand complexes. The B'-factors for Wat1047_{1DOR}, Wat1108_{1DOR}, and Wat1246_{1DOR} are 0.4, 2.3, and 0.0, respectively.

Shaltiel and coworkers studied the conserved water molecules that contribute to the extensive network of interactions at the active site of protein kinase A (Shaltiel et al. 1998). Their study was done on a set of seven catalytic domain subunits of protein kinase A obtained at a resolution range of 2.0–2.9 Å. On the basis of their survey, they coined the concept of ligand-specific conserved waters (LSCWs). LSCWs are those water molecules that are displaced when ligands bind and that are found in the active site only with “incomplete” complexes (those formed with ligands that do not bind to the full set of residues that interact with ATP). They found six LSCWs (labeled from **g** to **m**) that are less than 1.5 Å from six ATP atoms: (a) the N1 and N6 nitrogen atoms of the adenine ring (waters **g** and **h**, respectively); (b) the 2' and the 3' OH of the ribose (waters **i** and **j**, respectively); and (c) the β and γ -phosphates in the triphosphate chain (waters **m** and **k**, respectively). The level of matching between these six water molecules in the 1APM structure (Knighton et al. 1993) and the corresponding ATP atoms in the 1ATP structure (Zheng et al. 1993) shows that only three of the six LSCWs are involved in the mimicry of oxygen or nitrogen atoms of ATP (Wat452_{1APM} or **g**, Wat507_{1APM} or **h** and Wat547_{1APM} or **m**). Only backbone atoms from the enzyme make hydrogen bonds with these three water molecules (Wat452_{1APM} with N and O from Val123; Wat507_{1APM} with O Glu121; and Wat547_{1APM} with N Ser53), but these residues have been qualitatively preserved throughout evolution (Ser53_{1APM}, Glu121_{1APM}, and Val123_{1APM} are replaced in some sequences by Thr, Asp, and Ile, respectively). Consolv (Raymer et al. 1997) correctly predicted that Wat507_{1APM} and Wat547_{1APM} are replaceable by the ligand, whereas Wat452_{1APM} was incorrectly predicted as conserved in enzyme-ligand complexes. Most of the protein kinase inhibitors bind to the active site with the same hydrogen bond interactions as Wat452_{1APM}/N1 ATP and Wat507_{1APM}/N6 ATP (Taylor and Radzio-Andzelm 1997). At this point, we should note the importance of a water molecule equivalent to Wat507_{1APM} (Wat350_{1DAY}) in the dual-cosubstrate specificity of protein kinase CK2 by ATP or GTP (Niefind et al. 1999). Wat350_{1DAY} mimics the N6 atom of ATP in protein kinase CK2 and switches the active site from an ATP- to a GTP-compatible state without affecting the protein active site geometry but preserving the overall hydrogen-bonding network irrespective of whether ATP or GTP is bound. The B'-factors for Wat452_{1APM}, Wat507_{1APM}, Wat547_{1APM}, and Wat350_{1DAY} are 0.1, 0.7, 1.2, and -0.5, respectively.

We would like to stress, therefore, that to the best of our knowledge, BAMY shows the highest number of water molecules involved in the molecular mimicry of polar ligand atoms (four) known to date (one for HIV-PR and protein kinase CK2 and three for DHODA and protein kinase A).

Conclusions

Integration of findings in a BAMY catalytic mechanism

The molecular mimicry of substrate oxygen atoms by water molecules in the SBA active site reported in this study is an important mechanistic element for this enzyme activity. Our previous findings [i.e., the high level of conservation throughout evolution of the eight residues that form the two C-terminal CA-layers of the barrel (Pujadas et al. 1996) and the implication of the ³⁴¹FTC³⁴³ segment in the catalytic mechanism (Pujadas and Palau 1997)], together with the results of the present study draw a consistent model for BAMY catalysis, where the enzyme operates either as a “cutter device” (exo-acting carbohydrolase giving β -maltose as a final product, which is the most common process) or as a “hydration device” (β -2-deoxymaltose as a final product).

In this model, four isolated water molecules are hydrogen bonded to strategic active-site residues and avoid unwanted conformational changes either for all of them in the free enzyme or for those which are not used for binding the ligand in the complexed structures. A “trap trigger,” formed by residues 341–343 (Phe-Thr-Cys), waits for the nonreducing end of the polysaccharide chain to appear in the pocket (Pujadas and Palau 1997). How the substrate adapts itself to the pocket may be related to cohesive forces near the active center of the enzyme (Adachi et al. 1998). We therefore believe that the last two CA-layers of the barrel, which show a high level of conservation throughout evolution, may play an important role in the substrate-binding process (Pujadas et al. 1996). Therefore, the BAMY catalytic mechanism implies that: (a) the strategic water molecules are replaced by oxygen atoms in the substrate (as shown in this report); (b) the hinged lid is closed to form part of the active center (Mikami et al. 1994); (c) the trap trigger is discharged by a conformational transition of β -strand 6 (Pujadas and Palau 1997); (d) the reaction process is activated; and (e) the hinged lid is opened to allow the product to depart (Mikami et al. 1994; Adachi et al. 1998). After this operation, the cycle would therefore be reactivated by “rearming” the trap trigger and reintroducing the water molecules into the pocket in positions that were previously used by substrate oxygen atoms involved in hydrogen bonds with the enzyme and that are vacated when the product is expelled.

We can conclude that the water structure at the active pocket of BAMY is essential for providing the ligand recognition with plasticity. It does not affect the protein active-site geometry and preserves the overall hydrogen-bonding network irrespective of which substrate is bound to the enzyme. Moreover, the water structure is expected to help orient the ligand in the active site and indicates where the atoms of the ligand that are capable of forming hydrogen

bonds must be. These findings highlight water's important role in protein-ligand docking.

Water as a guide to finding positions that polar ligand atoms might occupy

Predicting how and which ligands will bind in protein active sites is currently of great interest for protein-ligand docking studies. Our results indicate that some enzymes are able to use a preexisting pattern of water molecules in their active sites to mimic where polar ligand atoms should be. Unlike water's other roles in biomolecular interactions, for example in the protein-ligand interface to water-mediated hydrogen bonds between the protein and the ligand (Ladbury 1996), its role in molecular mimicry is still rather unknown. It would therefore be interesting, for designing drugs or for finding new active-site compatible ligands, to predict when waters in free-enzyme structures mimic potential polar ligand atoms.

Obviously, to know how to discriminate between waters that "can" or "cannot" mimic polar ligand atoms, we must analyze the characteristics common to the known mimicking water molecules. Our results on BAMY together with those from HIV-PR (Lam et al. 1994), DHODA (Rowland et al. 1997, 1998) and protein kinases A/CK2 (Shaltiel et al. 1998; Niefind et al. 1999) show that these kinds of waters: (a) are involved in hydrogen bonds — usually, multihydrogen bonds — with active site residues that have been well preserved by evolution; (b) are predicted to be easily displaceable by the ligand when the Consolv algorithm (Raymer et al. 1997) is applied (our results on SBA, HIV-PR, DHODA and protein kinases A/CK2 indicate that the displacement status of 75.0% of these waters is correctly predicted by Consolv); and (c) have low B'-factors (all are below 2.4, which is the highest B'-factor that corresponds to Wat749_{1BFN}, a water that has been correctly modeled according to its EDM; see Table 1). Unfortunately, these common characteristics were derived from few data, so their ability to discriminate is limited. Nevertheless, they may be useful for discounting some waters when we are trying to predict the positions that might be occupied by polar ligand atoms (e.g., active site waters with high B'-factors that are not part of the first hydration shell of the enzyme).

Therefore, the main problem is how to recognize which active site waters that fulfill the mentioned requirements (low B'-factors, easily displaceable by the ligand, and hydrogen bonded to evolutionary preserved active-site residues) are potential candidates for mimicking polar ligand atoms. This is not easy to solve, but a computational study of all PDB structures that have been crystallized free and complexed with ligands may help. Such a huge computational study should be able to identify undescribed examples of the molecular mimicry of polar ligand atoms by water molecules in protein active sites. It would characterize mim-

icking waters better, and this in turn could be used to fine-tune predictive methods. However, the lack of reliability of many water placements (as exhaustively shown in this paper) seriously limits massive — and often "blind" — computational studies. We hope that the increasing number of high-resolution structures delivered to the PDB, together with the deposition of structure factors for these structures and an "intimate" knowledge of the corresponding catalytic mechanism, will soon help to overcome these limitations and enhance our knowledge of the mimicking role of active-site water molecules.

Materials and methods

Protein structures were imported from the last on-line release of PDB (<http://www.rcsb.org/pdb/>). Swiss-PdbViewer v3.7 (Guex and Peitsch 1997) was used for superposing and posterior merging the structures in the same file. Visual analysis of the superposed structures contained in this merged file allows conformational changes and water movements between structures to be easily identified on the same computer screen. This visual analysis was done with Rasmol (Sayle and Milner-White 1995). We checked the whole set of hydrogen bonds, as well as distances between donors and acceptors, with the LIGPLOT/HBPLUS programs (McDonald and Thornton 1994; Wallace et al. 1995). Default values for parameters that define the interactions were used with both programs.

The evolutionary conservation of residues that are hydrogen bonded to waters that mimic polar ligand atoms was analyzed by: (a) a FASTA search [<http://www.ebi.ac.uk/fasta3/>; (Pearson and Lipman 1988)] in the Swiss-All database [formed by SWISS-PROT, TrEMBL and updates (Bairoch and Apweiler 2000)] for sequences similar to those from 1BYA (BAMY), 1HSG (HIV-PR), 1DOR (DHODA), 1APM (protein kinase A), and 1DAY (protein kinase CK2); (b) retrieving similar sequences from Swiss-All (<http://www.expasy.ch/sprot/sprot-retrieve-list.html>); and (c) carrying out the multialignment between each "seed" sequence and those provided by the FASTA similarity searches. The multialignments were carried out with the CLUSTALV algorithm (Higgins and Sharp 1989) and the commercial program MEGALIGN v3.16 from the Lasergene software package (1997, DNASTAR, Inc.). Default values for parameters that define sequence similarity were used. Visual analysis of the multialignments provides rapid information about the evolutionary conservation of the key residues.

The EDM for 1BFN was obtained from the Uppsala Electron Density Server (<http://xray.bmc.uu.se/eds/>). RSFit values for Wat648_{1BFN}, Wat660_{1BFN}, Wat678_{1BFN}, Wat749_{1BFN}, and Wat752_{1BFN} were kindly provided by Dr. Tom Taylor of the Department of Molecular Biology at Uppsala (Sweden). The RSFit parameter is calculated using the CSN program (Brünger et al. 1998). Its value is 1 minus the correlation coefficient between each point in (a) the map calculated using the data and (b) a map calculated using the model alone. In practice, for a water molecule, the RSFit value corresponds to how spherical the density surrounding that water is. Therefore, the lower the RSFit is, the better that water molecule is defined at the point at which it has been modeled.

B'-factors were calculated from the following equation (Parthasarathy and Murthy 1997; Carugo 1999):

$$B' - \text{factor} = \frac{((B - \text{factor}) - \langle B - \text{factor} \rangle_{\text{prot}})}{s(B - \text{factor})_{\text{prot}}}$$

where (B-factor) is the temperature factor indicated in the PDB file for the corresponding water molecule; $\langle \text{B-factor} \rangle_{\text{prot}}$ and $s(\text{B-factor})_{\text{prot}}$ are the mean and standard deviation, respectively, for the temperature factors in the PDB file corresponding to protein atoms (not waters, heteroatoms, etc.). Only protein atoms are used to calculate $\langle \text{B-factor} \rangle_{\text{prot}}$ and $s(\text{B-factor})_{\text{prot}}$ in order to calibrate the temperature factors of the waters and compare them with the well-positioned part of the structure.

The energy of interaction between SBA and each of the crystallographic water molecules of interest was calculated using GRID v18.0 [(Goodford 1985); Molecular Discovery Ltd. 2000]. GRID was also used to search the location of all the local minima that are present in a cubic "cage" of 5.0 Å on each side and approximately centered at the position of the 1BYA crystallographic water equivalent to the one being studied. The search for these local minima was done by setting a low separation between grid points (0.1 Å) and further interpolating using the program MINIM (Molecular Discovery Ltd. 2000). It was easy to compare the location of local minima from the coordinates of the corresponding crystallographic water molecule by merging coordinates for minima and water in the same file (see Kinemage 2) and further visualizing them with Rasmol (Sayle and Milner-White 1995). In all calculations with GRID, the "target" is defined as the set formed by the protein atoms plus the sulfate ion, the "probe" is water and has been treated as an extended oxygen atom, and default values were taken for the parameters that define the "probe" (Van der Waals radius, effective number of electrons, polarizability, electrostatic charge, optimal hydrogen-bond energy, hydrogen-bonding radius, number of hydrogen bonds donated, number of hydrogen bonds accepted, and hydrogen-bonding type).

The displacement status of water molecules upon ligand binding was studied with the Consolv algorithm (Raymer et al. 1997). This program was applied to ligand-free protein structures except 1HSG (Chen et al. 1994). This structure is complexed with an inhibitor of the HIV proteases that does not displace the structural water of interest (Wat308). In all cases, the Consolv option that is tuned for displacement-prediction in active-site water molecules was chosen.

MEGALIGN was run in a Power Macintosh. The other programs were run on a Silicon Graphics Indigo2 XZ workstation.

Acknowledgments

We thank Kevin Costello of the Language Service of our University for his help in writing the manuscript; Drs. Oliviero Carugo, Ignasi Fita and Tom Taylor for their advice on defining the quality criteria for studying the reliability of crystallographic water molecules, and Prof. Peter J. Goodford for his help with Molecular Discovery software. This work has not been awarded grants by any research-supporting institution.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Adachi, M., Mikami, B., Katsube, T., and Utsumi, S. 1998. Crystal structure of recombinant soybean β -amylase complexed with β -cyclodextrin. *J. Biol. Chem.* **273**: 19859–19865.
- Bairoch, A. and Apweiler, R. 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **28**: 45–48.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, et al. 1998. Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**: 905–921.
- Carugo, O. 1999. Correlation between occupancy and B factor of water molecules in protein crystal structures. *Protein Eng.* **12**: 1021–1024.
- Chen, Z., Li, Y., Chen, E., Hall, D.L., Darke, P.L., Culbertson, C., Shafer, J.A., and Kuo, L.C. 1994. Crystal structure at 1.9 Å resolution of human immunodeficiency virus (HIV) II protease complexed with L-735,524, an orally bioavailable inhibitor of the HIV proteases. *J. Biol. Chem.* **269**: 26344–26348.
- Davies, G. and Henrissat, B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**: 853–859.
- Dunbrack, R.L. and Karplus, M. 1994. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nat. Struct. Biol.* **1**: 334–340.
- Erickson, J., Neidhart, D.J., VanDrie, J., Kempf, D.J., Wang, X.C., Norbeck, D.W., Plattner, J.J., Rittenhouse, J.W., Turon, M., and Wideburg, N. 1990. Design, activity, and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. *Science* **249**: 527–533.
- Frauenfelder, H. and Petsko, G.A. 1980. Structural dynamics of liganded myoglobin. *Biophys. J.* **32**: 465–483.
- Funahashi, J., Takano, K., Ogasahara, K., Yamagata, Y., and Yutani, K. 1996. The structure, stability, and folding process of amyloidogenic mutant human lysozyme. *J. Biochem.* **120**: 1216–1223.
- Goodford, P.J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **28**: 849–857.
- Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.
- Hehre, E.J., Okada, G., and Genghof, D.S. 1969. Configurational specificity: Unappreciated key to understanding enzymic reversions and de novo glycosidic bond synthesis. I. Reversal of hydrolysis by α -, β - and glucoamylases with donors of correct anomeric form. *Arch. Biochem. Biophys.* **135**: 75–89.
- Higgins, D.G. and Sharp, P.M. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.* **5**: 151–153.
- Kitahata, S., Chiba, S., Brewer, C.F., and Hehre, E.J. 1991. Mechanism of maltal hydration catalyzed by β -amylase: Role of protein structure in controlling the steric outcome of reactions catalyzed by a glycosylase. *Biochemistry* **30**: 6769–6775.
- Knighton, D.R., Bell, S.M., Zheng, J.H., Teneyck, L.F., Xuong, N.H., Taylor, S.S., and Sowadski, J.M. 1993. 2.0 Å refined crystal-structure of the catalytic subunit of CAMP-dependent protein-kinase complexed with a peptide inhibitor and detergent. *Acta Crystallogr. D* **49**: 357–361.
- Ladbury, J.E. 1996. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem. Biol.* **3**: 973–980.
- Laederach, A., Dowd, M.K., Coutinho, P.M., and Reilly, P.J. 1999. Automated docking of maltose, 2-deoxymaltose, and maltotetraose into the soybean β -amylase active site. *Proteins* **37**: 166–175.
- Lam, P.Y.S., Jadhav, P.K., Eyermann, C.J., Hodge, C.N., Ru, Y., Bacheler, L.T., Meek, J.L., Otto, M.J., Rayner, M.M., Wong, Y.N., et al. 1994. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **263**: 380–384.
- Matthews, B.W. 1993. Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* **62**: 139–160.
- McCarter, J.D. and Withers, S.G. 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **4**: 885–892.
- McDonald, I.K. and Thornton, J.M. 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**: 777–793.
- Mikami, B., Degano, M., Hehre, E.J., and Sacchettini, J.C. 1994. Crystal structures of soybean β -amylase reacted with β -maltose and maltal: Active site components and their apparent roles in catalysis. *Biochemistry* **33**: 7779–7787.
- Mikami, B., Hehre, E.J., Sato, M., Katsube, Y., Hirose, M., Morita, Y., and Sacchettini, J.C. 1993. The 2.0 Å resolution structure of soybean β -amylase complexed with α -cyclodextrin. *Biochemistry* **32**: 6836–6845.
- Niefind, K., Putter, M., Guerra, B., Issinger, O.G., and Schomburg, D. 1999. GTP plus water mimic ATP in the active site of protein kinase CK2. *Nat. Struct. Biol.* **6**: 1100–1103.
- Parthasarathy, S. and Murthy, M.R.N. 1997. Analysis of temperature factor distribution in high-resolution protein structures. *Protein Sci.* **6**: 2561–2567.
- Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.* **85**: 2444–2448.
- Pujadas, G. and Palau, J. 1997. Anatomy of a conformational transition of β

- strand 6 in soybean β -amylase caused by substrate (or inhibitor) binding to the catalytical site. *Protein Sci.* **6**: 2409–2417.
- Pujadas, G. and Palau, J. 1999. TIM barrel fold: Structural, functional and evolutionary characteristics in natural and designed molecules. *Biologia [Bratisl]* **54**: 231–254.
- Pujadas, G., Ramírez, F.M., Valero, R., and Palau, J. 1996. Evolution of β -amylase: Patterns of variation and conservation in subfamily sequences in relation to parsimony mechanisms. *Proteins* **25**: 456–472.
- Raymer, M.L., Sanschagrin, P.C., Punch, W.F., Venkataraman, S., Goodman, E.D., and Kuhn, L.A. 1997. Predicting conserved water-mediated and polar ligand interactions in proteins using a K-nearest-neighbors genetic algorithm. *J. Mol. Biol.* **265**: 445–464.
- Ringe, D. and Petsko, G.A. 1986. Study of protein dynamics by X-ray diffraction. *Methods Enzymol.* **131**: 389–433.
- Rowland, P., Björnberg, O., Nielsen, F.S., Jensen, K.F., and Larsen, S. 1998. The crystal structure of *Lactococcus lactis* dihydroorotate dehydrogenase A complexed with the enzyme reaction product throws light on its enzymatic function. *Protein Sci.* **7**: 1269–1279.
- Rowland, P., Nielsen, F.S., Jensen, K.F., and Larsen, S. 1997. The crystal structure of the flavin containing enzyme dihydroorotate dehydrogenase A from *Lactococcus lactis*. *Structure* **5**: 239–252.
- Rowland, P., Nørager, S., Jensen, K.F., and Larsen, S. 2000. Structure of dihydroorotate dehydrogenase B: Electron transfer between two flavin groups bridged by an iron-sulphur cluster. *Structure* **8**: 1227–1238.
- Sayle, R. and Milner-White, E.J. 1995. RasMol: Biomolecular graphics for all. *Trends Biochem. Sci.* **20**: 333–379.
- Shaltiel, S., Cox, S., and Taylor, S.S. 1998. Conserved water molecules contribute to the extensive network of interactions at the active site of protein kinase A. *Proc. Natl. Acad. Sci.* **95**: 484–491.
- Shoichet, B.K., Baase, W.A., Kuroki, R., and Matthews, B.W. 1995. A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci.* **92**: 452–456.
- Takano, K., Funahashi, J., Yamagata, Y., Fujii, S., and Yutani, K. 1997. Contribution of water molecules in the interior of a protein to the conformational stability. *J. Mol. Biol.* **274**: 132–142.
- Taylor, S.S. and Radzio-Andzelm, E. 1997. Protein kinase inhibition: Natural and synthetic variations on a theme. *Curr. Opin. Chem. Biol.* **1**: 219–226.
- Tronrud, D.E. 1996. Knowledge-based B-factor restraints for the refinement of proteins. *J. Appl. Crystallogr.* **29**: 100–104.
- Wallace, A.C., Laskowski, R.A., and Thornton, J.M. 1995. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**: 127–134.
- Zhang, X.-J. and Matthews, B.W. 1994. Conservation of solvent-binding sites in 10 crystal forms of T4 lysozyme. *Protein Sci.* **3**: 1031–1039.
- Zheng, J., Trafny, E.A., Knighton, D.R., Xuong, N., Taylor, S.S., Ten Eyck, L.F., and Sowadski, J.M. 1993. 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr. D* **49**: 362–365.