ELSEVIER

Biophysical Chemistry 95 (2002) 195–201

Biophysical Chemistry

www.elsevier.com/locate/bpc

Hydrogen bonds from water molecules to aromatic acceptors in very high-resolution protein crystal structures

Thomas Steiner*

Institut für Chemie–Kristallographie, Freie Universität Berlin, Takustr. 6, D-14195 Berlin, Germany

Received 20 June 2001; accepted 2 July 2001

Abstract

Short contacts of water molecules with the π-faces of aromatic residues were studied in a set of 75 very high resolution (<1.1 Å) protein X-ray crystal structures. For 18 water molecules found at distances to aromatic midpoints <3.5 Å, it was attempted to assign the hydrogen bond configuration (without experimental knowledge of the H-atom positions) by inspection of the surrounding. For approximately one-quarter of the cases, evidence for an O–H···π hydrogen bond was found, another one-quarter does not form such a hydrogen bond, and for the remaining half, no conclusive assignment could be made. The results confirm the relatively frequent occurrence of aromatic hydrogen bonding in biomolecular hydration, but also underline difficulties in hydrogen bond assignment without reliable knowledge of the H-atom positions.

Keywords: Aromatic hydrogen bonds; Biomolecule–water interactions; Crystal structure; Database analysis

1. Introduction

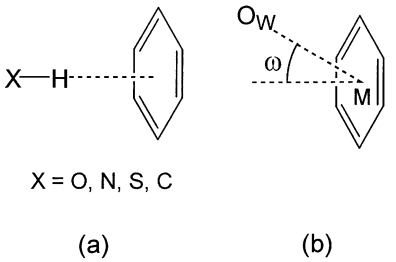
Structure and function of biomolecules depend critically on the interactions with the surrounding solvent, in particular with water molecules. By far the most important of these interactions is the hydrogen bond [1]. Formerly, studies of biomolecule–water hydrogen bonds have been restricted to the conventional types *X*–H···*Y*, where *X* as well as *Y* are very electronegative atoms (O, N). Only recently, it became clear that also weaker hydrogen bond types like C–H···O_W [2–4] and O_W–H···π [5,6] occur in biomolecular hydration (surveyed in Desiraju and Steiner [7]). They play important roles at biomolecule–solvent interfaces where relatively hydrophobic parts of a biomolecule are in contact with the solvent, and also for interior water molecules in proteins which are often only poorly co-ordinated with normal hydrogen bonds.

The so-called 'aromatic hydrogen bonds' *X*–H···π are hydrogen bonds with the π-electron cloud of aromatic groups acting as the acceptor (Scheme 1a) [7]. In proteins, potential acceptors are the side chains of Phe, Tyr, Trp (both rings), uncharged His, and suitable bound ligands. Typical energies are approximately half as high as for hydrogen bonds with conventional acceptors, *X*–H···O. With the important water donors, numerous examples are known from crystallographic studies

*Fax: +49-30-838-56702.

E-mail address: steiner@chemie.fu-berlin.de (T. Steiner).

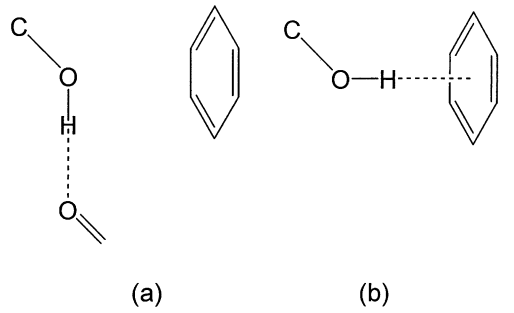
0301-4622/02/$ - see front matter
PII: S0301-4622(01)00256-3



Scheme 1. Aromatic hydrogen bonding. (a) Overall arrangement. (b) Definition of ω as the angle between the $O_W\cdots M$ line and the ring normal; *M* is the aromatic centroid.

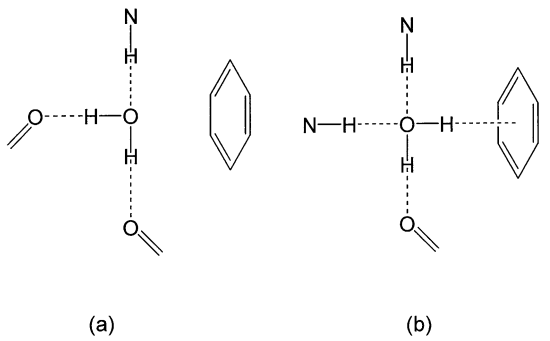
of organic molecules [8,9], and also of hydrated peptides [5]. The $O_W\cdots M$ distances (M = aromatic centroid) are typically in the range of 3.1–3.7 Å. Such distances are compatible also with van der Waals contacts, and therefore, hydrogen bond assignment critically requires to show that an O_W–H vector is actually pointing at the π-face. In most small molecule crystal structures, H-atom positions are determined experimentally, so that O_W–H$\cdots\pi$ hydrogen bonds can be clearly assigned. In protein X-ray crystal structures, on the other hand, the water H-atoms can neither be seen, nor can their theoretical positions be calculated from the non-H framework. In consequence, only 'possible' O_W–H$\cdots\pi$ hydrogen bonds can be assigned on the basis of $O_W\cdots\pi$ contact geometries (i.e. short $O_W\cdots M$ distance, small angle ω), Scheme 1b; examples in practical studies can be found in Kryger and co-workers [10,11]. Only in one case, in the ultra-high resolution X-ray crystal structure of concanavalin A, a different electron density peak possibly representing a H-atom directly suggested the formation of an O_W–H$\cdots\pi$ hydrogen bond [6].

The ambiguities in assigning X–H$\cdots\pi$ hydrogen bonds to short $X\cdots\pi$ contacts can be further specified for different types of X–H groups. For amide and peptide donors, it was shown that there are two kinds of short 'N-over-π' contacts, one representing N–H$\cdots\pi$ hydrogen bonds, and the other one stacked interactions with N–H oriented parallel to the aromatic plane [12,13]. These two interaction types have very similar N$\cdots\pi$ geometries, but they can be distinguished if the H-atom



Scheme 2. Short 'O-over-π' contacts of hydroxy groups that allow assignment of the hydrogen bond pattern even without location of the H-atom. (a) With a strong acceptor binding the single donor capacity of O–H. An O–H$\cdots\pi$ bond can no more be formed. (b) In the absence of any other potential acceptors, formation of an O–H$\cdots\pi$ bond is the only possibility to satisfy the donor potential.

position is at least roughly known. This is usually the case for amide and peptide N–H. With hydroxy donors, the situation is more unpleasant because the H-atom position cannot be calculated. Nevertheless, inspection of the surrounding often allows to assign the hydrogen bond configuration. If, for example, there is a strong carbonyl or carboxylate acceptor at a short distance, it certainly binds the



Scheme 3. Short 'water-over-π' contacts that allow assignment of the hydrogen bond pattern even without location of the H-atoms. (a) With two strong acceptors binding the double donor potential of H_2O. An O_W–H$\cdots\pi$ bond can no more be formed. (b) If only one other potential acceptor is present, formation of an O_W–H$\cdots\pi$ bond is the only possibility to satisfy the double water donor potential. If the water molecule contacts other waters, related interpretation is usually not possible because the direction of the hydrogen bond(s) is not unambiguous.

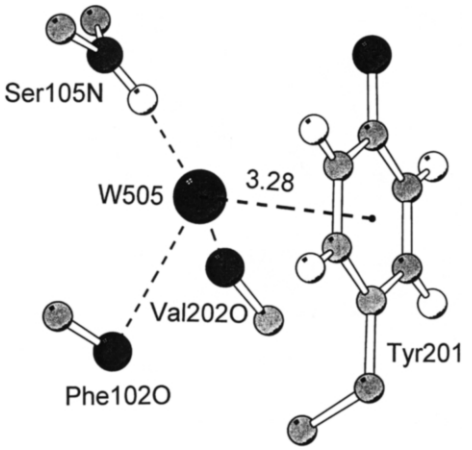


Fig. 1. Example of a water molecule that does not form an aromatic hydrogen bond despite a very short $O_W\cdots\pi$ contact: the donor capacity is fully bound by two C=O acceptors (in hydroxynitrile lyase [23], No. 15 in Table 1), compare Scheme 3a.

single donor capacity of the hydroxy group, and the H-atom is no more free to engage in an O–H$\cdots\pi$ interaction (Scheme 2a). The O$\cdots M$ distance may still be short around and below 3.5 Å (for an example, see Steiner and Koellner [14]). If, on the other hand, no conventional hydrogen bond acceptor at all is available for the hydroxy donor, one may assume quite safely that an aromatic hydrogen bond is formed (Scheme 2b); otherwise the O–H group would be dangling. In many instances, no conclusive interpretation of this kind is possible

Also with water molecules contacting π-faces, there are some configurations that allow assignment of the hydrogen bond scheme. If there are two carbonyl or carboxy groups at hydrogen bond distance, they probably bind both water H-atoms in O_W–H$\cdots$O=C hydrogen bonds, and no donor capacity is left for formation of an O_W–H$\cdots\pi$ interaction (Scheme 3a). If there is only one conventional acceptor available, formation of the O_W–H$\cdots\pi$ is very likely to avoid unsatisfied donor potential (Scheme 3b). In many other configurations, no clear assignment can be made.

In a recent statistical investigation of X–H$\cdots\pi$ hydrogen bonding in 592 protein crystal structures (resolution $\leq$1.6 Å), we have found 735 $O_W\cdots\pi$ contacts with $O_W\cdots M<3.8$ Å and $\omega<25°$ that might represent aromatic hydrogen bonds [14]. Because of the problems in assigning hydrogen bond character, these contacts were not discussed in greater detail. However, the large absolute number suggests to investigate the matter in more detail, prompting the present database study using only X-ray crystal structures of highest resolution ($<$1.1 Å). In this study, the environment of all relevant water molecules is inspected individually, leading to a higher level of interpretation compared to the previous work.

2. Methodology

All protein and peptide X-ray crystal structures with resolutions $<$1.1 Å (released until 20 April 2001) were retrieved from the Protein Data Bank (PDB, [15]). After sorting out multiply determined structures, simple mutants, and proteins/peptides without π-acceptors, 75 crystal structures remained. Compared to the previous study [14], tighter geometric criteria for 'possible' O_W–H$\cdots\pi$ hydrogen bonds were selected, $O_W\cdots M<3.5$ Å and $\omega<25°$. For structures with more than one symmetry-independent protein molecule forming the same kind of $O_W\cdots\pi$ contacts, only one was considered. This yielded 18 structurally independent water–π contacts in 11 crystal structures [6,16–25]. The co-ordination of these water molecules was then inspected one by one, and putative conventional hydrogen bonds were assigned ($O_W\cdots O/N<3.5$ Å, H$\cdots O_W<3.0$ Å). C–H$\cdots O_W$ hydrogen bonding was also considered, but will be discussed below only where necessary.

3. Results

The data sample consists of 18 $O_W\cdots\pi$ contacts that might represent aromatic hydrogen bonds. Numerical and bibliographical information is compiled in Table 1 (protein names and crystallographic resolutions given in the table footnote). The $O_W\cdots M$ distances are in the range of 3.08–3.5 Å, with the lower value close to the shortest O_W–H$\cdots\pi$ hydrogen bond found as yet (3.07 Å [9]). The conventional hydrogen bond co-ordination ranges from one- to fourfold (fourfold is optimal),

Table 1

Nr	PDB	Subst	Wat.	π	O···*M*	ω(O)	Water-co-ordination	O–H···π bond
1	1bkr	(a)	1004	Phe67	3.45	19.3	Leu80O [Lys78Cα]	Likely
2	1bkr	(a)	1027	Tyr103	3.46	16.2	W1202, W1073	Undecidable
3	1bs9	(b)	365	Tyr110	3.21	3.5	Asn112N, Thr111N, Thr111Og1	Likely
4	1cc8	(c)	1059	Tyr53	3.49	10.4	W1049	Likely
5	1ds1	(d)	485	Phe275	3.39	5.7	Tyr222OH, Asn305Nd2, W511, W742	Unlikely
6	1exr	(e)	2140	Phe12	3.45	24.6	Glu11Oe1, W2165	Undecidable
7	1fsg	(f)	1550	Tyr27	3.21	21.9	W1218, W1585, W1008	Undecidable[a]
8	1fsg	(f)	1144	Tyr205	3.22	6.8	Asp147Od1, Cys203O, Gly80O, W1049	No[a]
9	1ic6	(g)	433	Tyr61	3.08	9.3	W524, W686	Undecidable
10	1ic6	(g)	349	Tyr236	3.27	8.9	Tyr23OH, Asn275Nd2	Likely
11	1nls	(h)	375	Tyr12	3.22	8.8	W348, W255, W264	Undecidable
12	1qj4	(i)	546	Phe64	3.35	12.3	Phe64O, W590, W653, W673	Unlikely
13	1qj4	(i)	649	Trp128(6-ring)	3.46	11.3	Tyr133Oη, W619	Undecidable
14	1qj4	(i)	984	Phe150	3.47	13.5	W559, W580, W922	Undecidable
15	1qj4	(i)	505	Tyr201	3.28	12.2	Phe102O, Ser105N, Val202O	No
16	1qj4	(i)	964	Tyr222	3.35	19.7	Lys223O, W628, W834	Undecidable
17	1ql0	(j)	286	Trp123(6-ring)	3.40	20.8	Lys172Nz, Trp123O	Likely[a]
	1ql0	(j)	286	Trp123(5-ring)	3.31	16.8	As above	
18	4lzt	(k)	1134	Trp62(5-ring)	3.18	13.1	W1024	Undecidable[(k)]

[a] There is an analogous contact in a symmetry-independent monomer. (a) Calponin homology (CH) domain, 1.1 Å [16]; (b) acetylxylan esterase, 1.1 Å [17]; (c) Atx1 metallochaperone protein, 1.02 Å [18]; (d) clavamitic acid synthase, 1.08 Å [19]; (e) calmodulin, 1.0 Å [20]; (f) hypoxanthine–guanine phosphoribosyltransferase, 1.05 Å [21]; (g) proteinase K, 0.98 Å [22]; (h) concanavalin A, 0.94 Å [6]; (i) hydroxynitrile lyase, 1.1 Å [23]; (j) SM endonuclease, 1.1 Å [24]; (k) triclinic HEW lysozyme, 0.95 Å; in the closely related structure 3lzt, additional water molecules around the relevant water have been found [25].

and the groups involved are given in the second column from the right.

O_W–H···π interactions can be excluded almost certainly for two of the 18 water molecules, No. 8 and 15 of Table 1. For both, C=O acceptors bind the full donor capacity of the water molecules, analogous to Scheme 2a; one example is illustrated in Fig. 1. Nevertheless, the O_W···*M* distances are very short with 3.22 and 3.28 Å, and the approach to the π-face is close to perpendicular with ω = 6.8 and 12.2°, respectively. In two other cases (No. 5 and 12), the water molecule is fourfold co-ordinated with conventional hydrogen bond partners, three of them potential acceptors, so that formation of an O_W–H···π hydrogen bond seems unnecessary and unlikely (not impossible, though). Therefore, in total four of the 18 water molecules (=22%) probably do not form an aromatic hydrogen bond despite a short O_W···*M* contact. This demonstrates that criteria based only on favourable values of O_W···*M* and ω are completely unsuitable for hydrogen bond assignment.

'Likely' O_W–H···π hydrogen bond formation can be assigned to five of the 18 water molecules involved in short O_W···π contacts (=28%), i.e. No. 1, 3, 4, 10 and 17. In all cases, the water co-ordination is characterised by a deficiency of conventional acceptors, so that the double donor potential of the water molecule can be satisfied only if the O_W–H···π hydrogen bond is actually

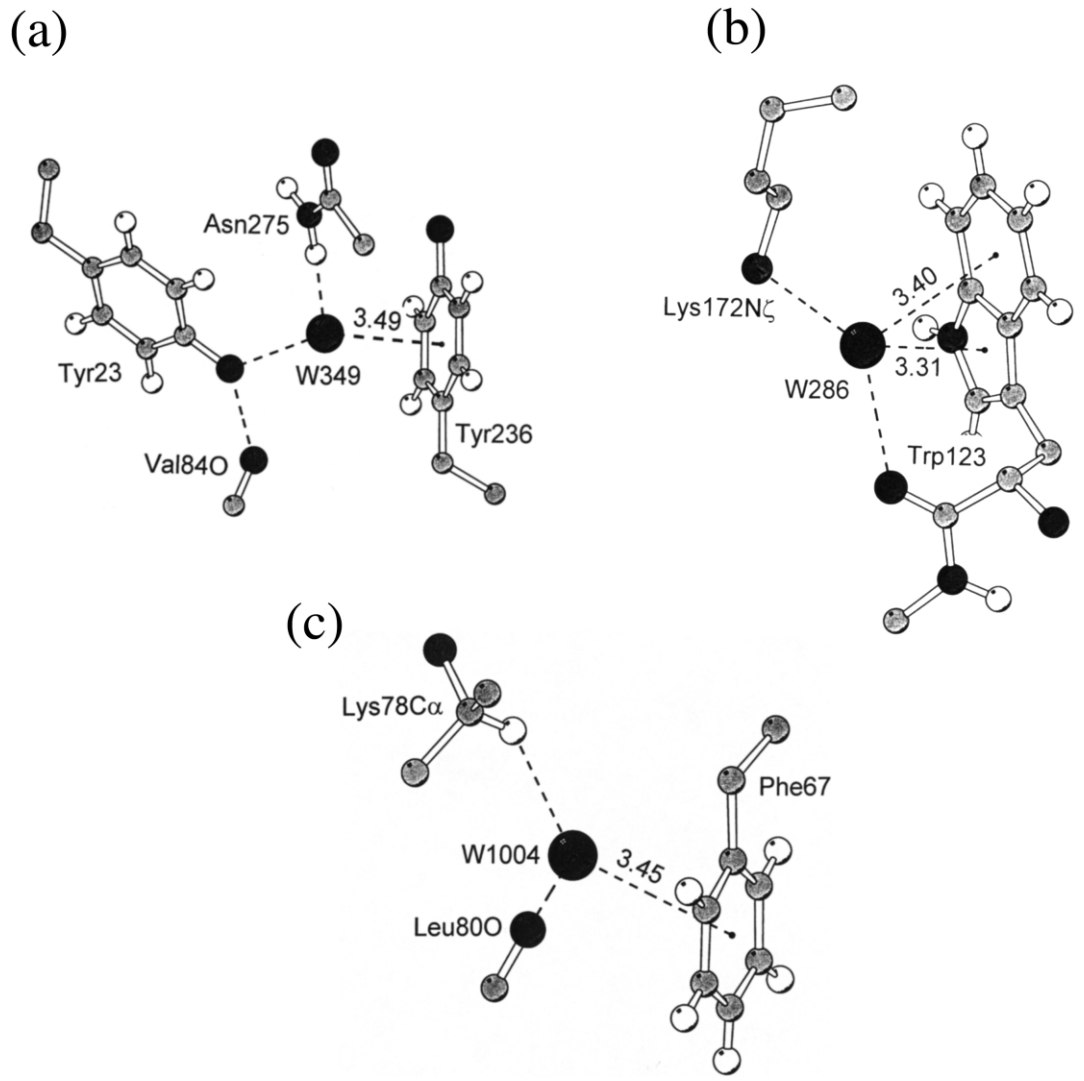
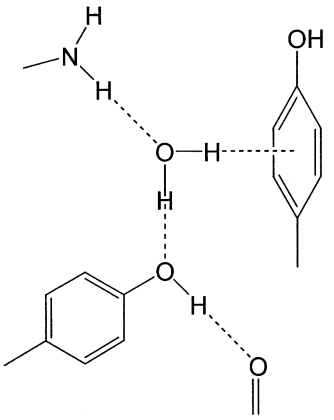


Fig. 2. Three examples of water molecules that probably form an aromatic hydrogen bond. Otherwise, they would have unsatisfied donor capacity; compare Scheme 3b. (a) In proteinase K [22] (No. 10 in Table 1); the probable hydrogen bond array is depicted in Scheme 4. (b) In *Serratia marcescens* endonuclease [24] (No. 17 in Table 1); either of the ring systems, or both at the same time, could serve as the acceptor. (c) In calponin homology (CH) domain [16] (No. 1 in Table 1).

formed. There are significant variations in the details of the co-ordination, so that three different examples are shown in Fig. 2. In the first example, Fig. 2a and Scheme 4, the water is in hydrogen bond distance with an $AsnN_{\delta}$ and a $TryO_{\eta}$ group. The latter satisfies its donor capacity with a peptide C=O acceptor, and serves as an acceptor to the water molecule. The other water donor functionality has only the neighbouring π(Tyr) acceptor available as a hydrogen bond partner. In the second example, Fig. 2b, the only two acceptors available for the water molecule are a peptide C=O and the aromatic side chain of the same residue. It cannot be decided, however, if the putative O_W–H$\cdots\pi$ interaction is donated to the five- or the six-membered ring, or to both. Finally, in the third



Scheme 4. Assignment of hydrogen bonds for the water molecule W349 in proteinase K [22] (No. 10 in Table 1), Fig. 1a.

example, Fig. 2c, the water molecule has only one conventional hydrogen bond partner, a peptide C=O acceptor. The other potential acceptor is a Phe side chain, and the only donor is a C_α–H group ($C_\alpha\cdots O_W = 3.26$ Å). This is an interesting example of a water molecule that does not accept any conventional hydrogen bonds, and is involved in C–H$\cdots O_W$ interactions instead [26]. However, some ambiguity remains with all these examples, and also those not explicitly discussed here. The reason is that there is always a certain possibility that a neighbouring water molecule has not been located in the X-ray study, and that the true co-ordination is different from the one given in Table 1. This would lead to different hydrogen bond arrays, possibly without an O_W–H$\cdots\pi$ interaction.

The 'undecidable' cases dominate in the present statistical study. Out of the 18 water molecules, no conclusive assignment of the hydrogen bond array was possible for nine (=50%), even if the second co-ordination shell around the water molecules was considered. Typically, these cases involve one or more water–water contacts for which the orientation of the hydrogen bond could not be derived. For water molecules that are in contact with a π-face and three other waters (e.g. No. 7, 14), attempts of such an assignment without experimental knowledge of the H-atom positions is more or less hopeless.

4. Summary and conclusions

Short 'water-over-π' contacts were retrieved from a set of 75 very high-resolution protein X-ray crystal structures. The resolution limit of 1.1 Å includes only the best of the currently available data. Eighteen structurally independent $O_W\cdots\pi$ contacts with O$\cdots M$ distances <3.5 Å and angles $\omega<25°$ were identified. By inspecting the surrounding, it was attempted to assign the hydrogen bond array, and to conclude if an O_W–H$\cdots\pi$ hydrogen bond is formed or not. For approximately one-quarter of the cases, the conclusion was positive, for one-quarter it was negative, and for the rest one-half, no decision could be made.

These results carry several implications. One is that O_W–H$\cdots\pi$ hydrogen bonding is a relatively common phenomenon in protein hydration, confirming corresponding assumptions made previously [5,14]. On the other hand, there are also numerous short $O_W\cdots\pi$ contacts *without* formation of an aromatic hydrogen bond. This parallels observations made for amide and peptide N–H, and for hydroxy groups, and it means that hydrogen bond assignment based only on short $O_W\cdots\pi$ distances and small angles ω is inadequate. Such contacts can only be considered as 'possible' hydrogen bonds, with a high probability that they do not represent hydrogen bonds in fact.

The attempt to analyse O_W–H$\cdots\pi$ hydrogen bonds without experimental knowledge of the H-atom positions leads to interesting results. On the other hand, it also is to some degree disappointing. The fraction of at least 50% 'undecidable' cases is without a doubt unbearably high. Proper and more complete description of water–biomolecule interactions obviously requires to determine H-atom positions experimentally. This is a challenge to modern structural biology, and points in particular at the emerging technique of neutron diffraction using large area sensitive detectors [27] that might become very helpful in this context.

References

[1] G.A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures, Springer, Berlin, 1991.
[2] T. Steiner, W. Saenger, Role of C–H$\cdots$O hydrogen bonds in the coordination of water molecules. Analysis

of neutron diffraction data, J. Am. Chem. Soc. 115 (1993) 4540–4547.
[3] T. Steiner, W. Saenger, C–H···O hydrogen bonds as integral part of water coordination. The example of the ordered water cluster in vitamin B_{12} coenzyme at 15 K, Acta Crystallogr., Sect. D 49 (1993) 592–593.
[4] Z.S. Derewenda, L. Lee, U. Derewenda, The occurrence of C–H···O hydrogen bonds in proteins, J. Mol. Biol. 252 (1995) 248–262.
[5] T. Steiner, A.M.M. Schreurs, J.A. Kanters, J. Kroon, Water molecules hydrogen bonding to aromatic acceptors of amino acids. The structure of Tyr–Tyr–Phe dihydrate and a crystallographic database study on peptides, Acta Crystallogr., Sect. D 54 (1998) 25–31.
[6] A. Deacon, T. Gleichmann, A.J. Kalb (Gilboa), et al., The structure of concanavalin A and its bound solvent determined with small-molecule accuracy at 0.94 Å resolution, J. Chem. Soc., Faraday Trans. 93 (1997) 4305–4312.
[7] G.R. Desiraju, T. Steiner, The Weak Hydrogen Bond in Structural Chemistry and Biology, Oxford University Press, Oxford, 1999.
[8] J.L. Atwood, F. Harnada, K.D. Robinson, G.W. Orr, R.L. Vincent, X-Ray diffraction evidence for aromatic π-hydrogen bonding to water, Nature 349 (1991) 683–684.
[9] T. Steiner, A.M.M. Schreurs, M. Lutz, J. Kroon, Making very short O–H···Ph hydrogen bonds: the example of tetraphenylborate salts, New J. Chem. 25 (2001) 174–178.
[10] G. Kryger, I. Silman, J.L. Sussman, Structure of acetylcholinesterase complexed with E2020 (Aricept®): implications for the design of new anti-Alzheimer drugs, Structure 7 (1999) 297–307.
[11] G. Koellner, G. Kryger, C.B. Millard, I. Silman, J.L. Sussman, T. Steiner, Active-site gorge and buried water molecules in crystal structures of acetylcholinesterase from *Torpedo californica*, J. Mol. Biol. 296 (2000) 713–735.
[12] J.B.O. Mitchell, C.L. Nandi, I.K. McDonald, J.M. Thornton, S.L. Price, Amino/aromatic interactions. Is the evidence stacked against hydrogen bonding?, J. Mol. Biol. 239 (1994) 315–331.
[13] G.A. Worth, R.C. Wade, The aromatic (i+2) amine interaction in peptides, J. Phys. Chem. 99 (1995) 17473–17482.
[14] T. Steiner, G. Koellner, Hydrogen bonds with π-acceptors in proteins: frequencies and role in stabilizing local 3D structures, J. Mol. Biol. 305 (2001) 535–557.
[15] H.M. Berman, J. Westbrook, Z. Feng, et al., The Protein Data Bank, Nucl. Acid. Res. 28 (2000) 235–242.
[16] S. Banuelos, M. Saraste, K.D. Carugo, Structural comparisons of calponin homology domains: implications for actin binding, Structure 6 (1998) 1419–1431.
[17] D. Ghosh, M. Erman, M. Sawicki, et al., Determination of a protein structure by iodination: the structure of iodinated acetylxylan esterase, Acta Crystallogr., Sect. D 55 (1999) 779–784.
[18] A.C. Rosenzweig, D.L. Huffman, M.Y. Hou, A.K. Wernimont, R.A. Pufahl, T.V. O' Halloran, Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution, Structure 7 (1999) 605–617.
[19] Z.H. Zhang, J.S. Ren, D.K. Stammers, J.E. Baldwin, K. Harlos, C.J. Schofield, Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase, Nat. Struct. Biol. 7 (2000) 127–133.
[20] M.A. Wilson, A.T. Brünger, The 1.0 Angstrom crystal structure of Ca^{2+}-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity, J. Mol. Biol. 301 (2000) 1237–1256.
[21] A. Heroux, E.L. White, L.J. Ross, A.P. Kuzin, D.W. Borhani, Substrate deformation in a hypoxanthine–guanine phosphoribosyltransferase ternary complex: the structural basis for catalysis, Structure 8 (2000) 1309–1318.
[22] C. Betzel, S. Gourinath, P. Kumar, et al., Structure of a serine protease proteinase K from Tritirachium album limber at 0.98 Å resolution, Biochemistry 40 (2001) 3080–3088.
[23] K. Gruber, M. Gugganig, U.G. Wagner, C. Kratky, Atomic resolution crystal structure of hydroxynitrile lyase from *Hevea brasiliensis*, Biol. Chem. 380 (1999) 993–1000.
[24] S.V. Shlyapnikov, V.V. Lunin, M. Perbandt, et al., Atomic structure of the *Serratia marcescens* endonuclease at 1.1 Å resolution and the enzyme reaction mechanism, Acta Crystallogr., Sect. D 56 (2000) 567–572.
[25] M.A. Walsh, T.R. Schneider, L.C. Sieker, Z. Dauter, V. Lamzin, K.S. Wilson, Refinement of triclinic hen egg-white lysozyme at atomic resolution, Acta Crystallogr., Sect. D 54 (1998) 522–546.
[26] T. Steiner, Water molecules which apparently accept no hydrogen bonds are systematically involved in C–H···O interactions, Acta Crystallogr., Sect. D 51 (1995) 93–97.
[27] N. Niimura, Y. Minezaki, T. Nonaka, et al., Neutron Laue diffractometry with an imaging plate provides an effective data collection regime for neutron protein crystallography, Nat. Struct. Biol. 4 (1997) 909–914.