Acta Cryst. (1998). D**54**, 25–31

Water Molecules Hydrogen Bonding to Aromatic Acceptors of Amino Acids: the Structure of Tyr-Tyr-Phe Dihydrate and a Crystallographic Database Study on Peptides

THOMAS STEINER,[a]* ANTOINE M. M. SCHREURS,[b] JAN A. KANTERS[b] AND JAN KROON[b]

[a]*Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-14195 Berlin, Germany, and* [b]*Bijvoet Center for Biomolecular Research, Department of Crystal and Structural Chemistry, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. E-mail: steiner@chemie.fu-berlin.de*

(*Received 4 February 1997; accepted 27 May 1997*)

Abstract

The crystal structure of Tyr-Tyr-Phe dihydrate contains a hydrogen bond formed between a water molecule and the Phe side chain. The geometry is centered with a distance of 3.26 Å between the water O atom and the aromatic centroid. In a database study on hydrated peptides, four related examples are found which exhibit a wide variability of hydrogen-bond geometries. The intermolecular surroundings of the water molecules are inspected, showing that they are typically involved in complex networks of conventional and non-conventional hydrogen bonds. Possible relevance for protein hydration is given.

1. Introduction

It has long been known from spectroscopic experiments that phenyl groups may act as acceptors of hydrogen bonds (Pimentel & McClellan, 1960), usually called '*X*—H···π(Ph) hydrogen bonds' or 'aromatic hydrogen bonds'. Organic small-molecule crystal structure were reported with a variety of *X*—H donors forming aromatic hydrogen bonds: N—H (*e.g.* Bakshi *et al.*, 1994), C—O—H (*e.g.* Ferguson, Gallagher, Glidewell & Zakaria, 1994), H_2O (*e.g.* Aubry, Protas, Moreno-Gonzales & Marraud, 1977), Cl—H (Deeg & Mootz, 1993), acidic C—H (Steiner, Starikov, Amado & Teixeira-Dias, 1995). In these crystals structures, *X*—H vectors are found which point (almost) exactly at the center of a Ph acceptor (Deeg & Mootz, 1993), but also considerably off-centered geometries are reported, and there are even examples with *X*—H pointing more or less linearly at one of the individual C atoms (Steiner, Starikov & Tamm, 1996). A soft hydrogen-bond geometry with slight preference for centred arrangements, as is observed in the crystalline state, is in line with gas-phase experiments on *X*—H···Ph bonded molecular dimers (Read, Campbell & Henderson, 1983). For the particular case of the dimer water–benzene, gas-phase data spectra indicate a donor–acceptor separation of 3.35 Å (from O_W to the benzene midpoint) and *ab initio* calculations predict a binding energy of 7.5 kJ mol^{-1} (~1.8 kcal mol^{-1}, Suzuki *et al.*, 1992).

For protein structures, aromatic hydrogen bonding with N—H donors became apparent early on from structural data (Perutz, Fermi, Abraham, Poyart & Bursaux, 1986; Burley & Petsko, 1986) and was supported by NMR experiments (Tüchsen & Woodward, 1987) and theoretical computations (Levitt & Perutz, 1988; Worth & Wade, 1995). Since for peptide and most side-chain N—H donors, H-atom positions can be calculated from the non-H atoms, analysis of N—H···Ph interactions can be performed reasonably well even for moderately resolved protein structures. For O—H donors (hydroxyl and water), H-atom positions cannot be calculated theoretically, and inferrence of O—H···Ph interactions from only non-H-atom positions is problematic. Therefore, only a few relevant cases were reported (Liu, Ji, Gilliland, Stevens & Armstrong, 1993; Engh *et al.*, 1996).

In structural biology, hydration phenomena are of utmost importance; structure as well as function of biomolecular systems depends crucially on the interactions with solvent molecules (Jeffrey & Saenger, 1991). Normally, water–biomolecule interactions are regarded only in terms of conventional O/N—H···O/N hydrogen bonds, whereas weaker interactions are neglected. It has been shown from neutron diffraction data that such simplified views are not justified: the weak C—H···O_W hydrogen bonds frequently contribute to water coordination, and this is true for hydrated small molecules as well as for macromolecular systems (Steiner & Saenger, 1993). Although aromatic hydrogen bonding with water donors has been observed in a few non-biological small-molecule structures (tetraphenylborate salts: Aubry *et al.*, 1977; Bakshi *et al.*, 1994; anionic calixarene: Atwood, Hamada, Robinson, Orr & Vincent, 1991), this phenomenon has, to our knowledge, never been discussed in detail for peptide hydration. For protein hydration, relevance of this interaction was seriously questioned (Flanagan, Walshaw, Price & Goodfellow, 1995), but on the other hand it was pointed out that it is probably of importance for water molecules which are often found included in internal hydrophobic cavities (Buckle, Cramer & Fersht, 1996).

In this contribution, we report the crystal structure of the hydrated phenyl-rich tripeptide Tyr-Tyr-Phe, which

contains a well defined O_W—H···Ph hydrogen bond. To complement the crystal structure, a database study on water–aromatic hydrogen bonding in small peptides is performed.

2. Experimental

2.1. *Crystal structure of Tyr-Tyr-Phe dihydrate*

A commercial sample (Sigma) of L-Tyr-L-Tyr-L-Phe acetate salt was dissolved in 0.1 *M* acetic acid; slow evaporation yielded prismatic crystals of Tyr-Tyr-Phe dihydrate (L-tyrosyl-L-tyrosyl-L-phenylalanine dihydrate, $C_{27}H_{29}N_3O_6.2H_2O$; M_r = 527.6). Crystals are stable under ambient condition. The space group is orthorhombic $P2_12_12_1$ with a = 9.306 (2), b = 12.12 (2), c = 23.53 (4) Å, V = 2654 (6) Å^3, Z = 4, D_c = 1.32 g cm^{-3}.

Intensity data were collected on an Enraf–Nonius FAST area detector at liquid nitrogen temperature (Mo $K\alpha$ X-rays with λ = 0.71073 Å, 0.85 × 0.40 × 0.40 mm crystal mounted on a glass pin, 16 403 reflections measured, 6055 unique, 5446 with $I > 2\sigma(I)$, μ = 0.098 mm^{-1}, no absorption correction). The data set is essentially complete to a crystallographic resolution of 0.61 Å, with some additional data measured to $\lambda/2\sin\theta_{max}$ = 0.54 Å. The structure was solved and refined with standard methods (*SIR*92, Altomare *et al.*, 1994; *SHELXL*93, Sheldrick, 1993). H atoms were treated in the riding model, with exception of the hydroxyl and water H atoms which were refined isotropically, and the N-terminal ammonium group which was allowed to rotate. Final R = 0.053, $wR(F^2)$ = 0.116 (for observed reflections).

2.2. *Database analysis*

In X-ray crystal structures of hydrated peptides, the water H-atom positions are normally the least reliably refined positions in the published set of atomic coordinates (if water H atoms are located at all). Therefore, care has to be taken in any analysis of published data to avoid inclusion of refinement artifacts. In consequence, restrictive quality criteria were used here, and all structures containing obvious or suspected dubious H-atom positions were included (see below). The underlying philosophy is that a small but reliable set of structural data is superior to a large one containing an uncontrollable fraction of inaccuracies or even artifacts.

Analyzed were peptide crystal structures archived in the Cambridge Structural Database (CSD, Allen *et al.*, 1991), spring 1996 update, with $R < 0.07$ (only ordered structures with all H-atom positions located). For hydrogen-bond analysis, X—H bond lengths were normalized to ideal bond lengths of O—H = 0.983, N—H = 1.009, C—H = 1.083 Å. In an initial step, all structures were retrieved which contain water–aromatic contacts with at least two H_W···C distances shorter than 3.0 Å (potential acceptors Phe, Tyr, Trp). This yielded seven crystal structures of potential relevance. In a second step, these structures were inspected in greater detail for possible dubious features. Excluded were the following cases which are regarded as unrealistic. (1) There is a potential hydrogen-bond partner closer than 2.9 Å to the water O atom, without a hydrogen bond being formed. (2) There is a potential hydrogen-bond partner closer than 2.7 Å to the water O atom without a fairly linear hydrogen bond (angle at H > 150°) being formed. (3) The water molecule is engaged in H_W···H contact(s) shorter than 2.0 Å. Four crystal structures passed these criteria, and are thus considered as containing well refined O_W—H···Ph interactions. The initially selected distance cutoff showed to be unnecessarily permissive here: the longest of the O_W—H···Ph contacts found has H···C separations well inside the selected cutoff distance.

3. Results and discussion

3.1. *Crystal structure of Tyr-Tyr-Phe dihydrate*

The crystal structure of Tyr-Tyr-Phe dihydrate was determined from low-temperature X-ray diffraction data at a resolution of 0.61 Å.† The structure is fully ordered, allowing location and refinement of all H-atom positions.

† Atomic coordinates, bond lengths and angles, and structure factors have been deposited with the IUCr (Reference: GR0724). Copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

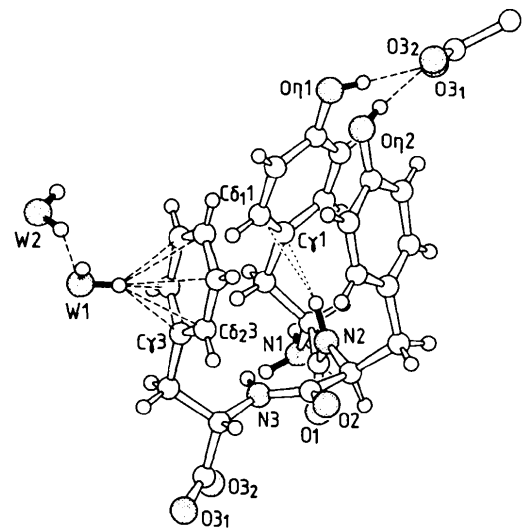


Fig. 1. Molecular structure of L-Tyr-L-Tyr-L-Phe as observed in the dihydrate crystal structure. Only atoms relevant to the discussion are labelled. Hydrogen bonds are indicated by dashed lines. O and N atoms are drawn shaded. Selection of torsion angles defining the molecular conformation: ψ_1 = −176.1 (2)°; ω_1 = −162.3 (2)°; φ_2 = −104.4 (2)°; ψ_2 = 23.7 (3)°; ω_2 = 167.2 (2)°; φ_3 = −106.2 (2)°; N3—C_α3—C3—O3$_1$ = 159.0 (2)°; χ_1^1 = −150.0 (2)°; $\chi_1^{2.1}$ = −125.8 (2)°; χ_2^1 = 52.7 (2)°; $\chi_2^{2.1}$ = 73.2 (3)°; χ_3^1 = −56.9 (2)°; $\chi_3^{2.1}$ = 118.7 (2)°.

Table 1. *Hydrogen-bond geometries (Å and °)*

Based normalized H-atom positions with O—H = 0.98, N– H = 1.03, C –H = 1.09 Å. *M* = centroid of the Phe3 phenyl ring. Of the C—H···O interactions in 1, only the most prominent one is given.

	H···*A*	*D*···*A*	Angle at H	Symmetry code
O/N/C—H···O bonds				
O_{W1}—H_1···$O3_2$	1.90	2.814 (3)	154	$\frac{1}{2}+x, \frac{1}{2}-y, 1-z$
O_{W2}—H_1···O_{W1}	1.87	2.784 (3)	154	$\frac{1}{2}+x, \frac{1}{2}-y, 1-z$
O_{W2}—H_2···$O3_1$	1.85	2.803 (3)	162	x, y, z
$O_\varepsilon1$—H···$O3_1$	1.74	2.714 (4)	175	$1\frac{1}{2}-x, -y, z-\frac{1}{2}$
$O_\varepsilon2$—H···$O3_2$	1.78	2.735 (4)	163	$1\frac{1}{2}-x, -y, z-\frac{1}{2}$
N1—H_1···O2	1.75	2.746 (2)	161	$x-1, y, z$
N1—H_2···O_{W2}	1.71	2.715 (4)	166	$x-1, y, z$
N1—H_3···$O3_2$	1.88	2.820 (5)	149	$x-\frac{1}{2}, -\frac{1}{2}-y, 1-z$
N3—H···O_{W1}	2.34	3.177 (3)	138	$x-\frac{1}{2}, \frac{1}{2}-y, 1-z$
$C_\alpha2$—H···O_{W2}	2.29	3.365 (4)	168	$x-\frac{1}{2}, -\frac{1}{2}-y, 1-z$
O_W—H···Ph bonds				
O_{W1}—H_2···$C_\gamma3$	2.90	3.573 (6)	127	x, y, z
O_{W1}—H_2···$C_{\delta_1}3$	3.11	3.680 (5)	119	x, y, z
O_{W1}—H_2···$C_{\delta_2}3$	2.53	3.419 (6)	151	x, y, z
O_{W1}—H_2···$C_{\varepsilon_1}3$	3.02	3.650 (4)	124	x, y, z
O_{W1}—H_2···$C_{\varepsilon_2}3$	2.42	3.396 (5)	175	x, y, z
O_{W1}—H_2···$C_\varphi3$	2.69	3.520 (5)	143	x, y, z
O_{W1}—H_2···*M*	2.42	3.258 (5)	143	x, y, z

Table 2. *Database analysis: O_W—H···Ph interactions in peptides*

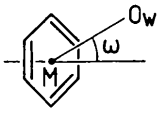
Geometries are given for normalized H-atom positions.

	This work	BIHXUL10†	JECYUL‡	SOJPAI§	TALVAD¶
Acceptor type	Phe	Phe	Tyr	Tyr	Phe
ω (°)	6.8	3.3	20.9	21.0	17.0
O_W···*M* (Å)	3.26	3.28	3.61	3.63	3.45
O_W···C range (Å)	3.40–3.68	3.49–3.63	3.38–4.31	3.41–4.31	3.33–4.07
O_W···C spread (Å)	0.28	0.14	0.93	0.90	0.74
H···*M* (Å)	2.42	2.48	2.87	2.90	2.47
H···C range (Å)	2.42–3.11	2.57–3.09	2.79–3.54	2.51–3.78	2.46–3.17
H···C spread (Å)	0.69	0.52	0.75	1.27	0.71
O_W—H···*M* (°)	143	138	132	132	171
O_W—H···C range (°)	118–172	112–165	106–158	115–152	141–159

† *Cyclo*-(L-Pro-L-Val-L-Phe-L-Phe-L-Ala-Gly) tetrahydrate (cycloamanide A tetrahydrate); Chiang *et al.* (1982). ‡ L-Asp-L-Arg-L-Val-L-Tyr tetrahydrate; Feldman & Eggleston (1990). § L-Pro-L-Tyr monohydrate; Klein *et al.* (1991). ¶ *Cyclo*-(L-Ser-L-Phe-L-Leu-L-Pro-L-Val-L-Asn-L-Leu) tetrahydrate (evolidine tetrahydrate); Eggleston *et al.* (1991).

The molecular structure of the tripeptide, which crystallized as a zwitterion, is shown in Fig. 1. Since the focus of the study is not on the peptide itself, but on the solvent–peptide interactions, the molecular conformation shall be described only briefly; relevant torsion angles are given in the legend of Fig. 1. The molecule adopts a folded conformation with all three aromatic side chains oriented roughly in the same direction. The tyrosine side chains are oriented such that their O_η—H hydroxyl groups approach to 3.175 (3) Å and form hydrogen bonds with the carboxylate end of a neighboring molecule. This roof-shaped arrangement shields the intermediate peptide N2—H from intermolecular interactions, and is associated with a short contact of N2—H with the Tyr1 aromatic moiety (H···$C_\gamma1$ = 2.47 Å, H···$C_{\delta1}1$ = 2.63 Å). Since this contact is possibly 'forced' by the particular conformation, we refrain here from speculating on its bonding or non-bonding nature. The other N—H donors are involved in conventional intermolecular hydrogen bonds, Table 1.

Of greater interest are the intermolecular interactions of the cocrystallized water molecules, which form a dimer (W1, W2). Molecule W1 donates an aromatic hydrogen bond to the side chain of Phe3, Fig. 1. The water O atom resides roughly above the aromatic centroid *M* with an O_W···*M* separation of 3.258 (5) Å and an angle ω of 6.8° (ω = angle between the line O_W—*M* and the normal of the Ph plane).



Because of this centered geometry, the individual O···C(Ph) separations fall in the narrow range 3.40 to 3.68 Å with a spread of only 0.28 Å (Table 2). The

orientation of the O_W—H vector is somewhat less centered with H···*M* = 2.42 Å and the individual H···C(Ph) distances in the range 2.42—3.11 Å. O_W—H appears to point more linearly at a C atom than at the centroid (O_W—H···$C_{\varepsilon 2}$3 = 175, O_W—H···*M* = 143°), but it should be remembered that in an X-ray diffraction study, H···Ph geometries are less reliably determined compared with O_W···Ph.

The intermolecular environment of the two water molecules is shown in Fig. 2. Apart from the O_W—H···Ph interation, W1 is engaged in three conventional N/O—H···O hydrogen bonds (one donated, two accepted). W2 is also engaged in three conventional hydrogen bonds (two donated, one accepted), and accepts an additional C_α—H···O_W interaction. The geometry of the latter is in the typical range for C_α—H···O hydrogen bonds, which have been shown to occur abundantly in amino acids and peptides (Jeffrey & Maluszynska, 1982) and also in proteins [Derewenda, Lee & Derewenda (1995); for background on C—H···O interactions, see Desiraju (1996); Steiner (1996, 1997)]. The hydrogen-bond coordination of both water molecules is therefore fourfold, which is the optimal case. For both water molecules, the fourfold coordination is achieved by participation of non-conventional hydrogen bonds: the surrounding of W2 is a typical example for the completion of a tetrahedral water coordination by a C—H···O_W hydrogen bond (discussed in detail by Steiner & Saenger, 1993), and for W1, a tetrahedral coordination is facilitated by formation of an aromatic hydrogen bond. In total, the intermolecular interactions of the water dimer are an exceptionally fine example of how ideal water coordination can be achieved by concerted efforts of O/N—H···O, C—H···O and O—H···Ph hydrogen bonding.

3.2. *Database analysis on O_W—H···Ph bonding in peptides*

To judge whether the above observations are exotic or relevant in a wider sense, it must be checked if similar configurations of water–peptide interactions are contained in crystal structures published earlier. The Cambridge Structural Database (§2.2) contains four well refined examples of O_W—H···Ph hydrogen bonds in peptide crystal structures; geometries and references are given in Table 2. This data sample is too small to allow general analysis with statistical significance. However, since none of these cases was discussed by the original authors, they deserve individual presentation in some detail here. It must of course be expected that also O_W—H···Ph bonds can exist which have quite different geometrical characteristics.

Looking at Table 2, it is obvious that the water–aromatic interactions can have very different geometries: in BIHXUL10, the geometry is close to centered with ω = 3.3° and O_W···*M* = 3.28 Å. The O_W···*M* and H···*M* distances are shorter than any of the individual O_W···C and H···C separations, respectively. This means that O_W—H seems to hydrogen bond with the entire Ph face. In SOJPAI, the situation is different: the water position is off-centered with ω = 21.0°, and O_W and H_W have much shorter distances to specific C atoms than to the aromatic midpoint. The individual O_W···C and H···C distances cover wide ranges, 0.90 and 1.27 Å, respectively; this means that not all parts of the π-system are equally engaged. The other three cases (including Try-Tyr-Phe dihydrate) are in between these extremes. It should be noted that aromatic hydrogen bonds can also be much more off-centered than that in SOJPAI, such as for an off-centered C—O—H···Ph bond with ω = 35.5° and O···C ranging from 3.34 to 4.85 Å, spread 1.51 Å, reported by Steiner *et al.* (1996).

Hydrogen bonding of water molecules cannot be reasonably discussed without looking at the complete intermolecular surrounding. This is shown below for three exemplary cases which possess different topologies. Compound SOJPAI (Pro-Tyr monohydrate, Klein *et al.*, 1991) contains only one symmetry-independent water molecule which is engaged in a very simple hydrogen-bond pattern, Fig. 3. It donates an O_W—H···O and an O_W—H···Ph hydrogen bond. The latter is not directed at the aromatic centroid, but rather at the midpoint of an aromatic C—C bond, with H···C distances of 2.51 and 2.63 Å, and the distance to the bond center 2.48 Å (the next shortest H···C distance is already 3.12 Å). The O_η hydroxyl group of the accepting tyrosine side chain donates a hydrogen bond to the water molecule in the next unit cell, so that an infinite chain is formed: O_W—H···Ph—O—H···O_W—H···Ph—O—H. One can speculate whether or not this chain is cooperative (*i.e.* the individual hydrogen bonds enhance each other by mutual polarisation of the constituents); such a question, unfortunately, cannot be answered from crystallography alone.

The tetrahydrate of the cyclic hexapeptide cycloamanide A (BIHXUL10, Chiang, Karle & Wieland, 1982)

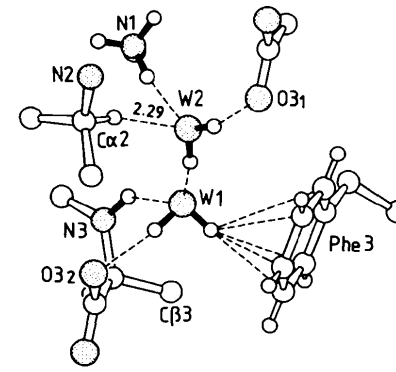


Fig. 2. Hydrogen-bond configuration of the two water molecules in Tyr-Tyr-Phe dihydrate. Distances are given in Å (based on normalized *X*—H bond lengths). Note that for both water molecules, favourable tetrahedral coordination is achieved by non-conventional hydrogen bonding: O_W—H···Ph for W1, and C—H···O_W for W2.

contains an infinite chain of interconnected water molecules, Fig. 4. The chain is formed by three of the four independent water molecules, the fourth one is isolated from this arrangement. Within the chain, two of the water molecules play conventional roles, whereas one donates a short and centered O_W—H···π hydrogen bond to a Phe side chain of the peptide. The hydrogen bonding in this water chain is dominated by conventional O/N—H···O interactions, but the O—H···Ph bond is clearly required for stabilization of the arrangement as a whole.

Most complex of the examples is the water structure of the tetrahydrated cyclic heptapeptide evolidine (TALVAD, Eggleston *et al.*, 1991), Fig. 5. In this crystal structure, an interstitial cavity is filled by a cluster of four interconnected water molecules. The wall of this cavity is formed by several polar groups (N—H, C—O—H, C=O), but also by apolar moieties like C_α—H, and Phe and Pro side chains. The enclosed water cluster is, therefore, in a partly polar and partly apolar environment. It is stabilized by a highly complex system of conventional and non-conventional hydrogen bonds, involving O/N—H···O, C—H···O and O—H···Ph interactions. The water molecules do not avoid contact with the 'apolar' groups, which might here be better characterized as 'only weakly polar', but forms weak hydrogen bonds with them. This is apparently a more favourable situation than water molecules suffering from completely unsatisfied hydrogen-bond potentials (and is nicely in line with previous results on the role of C—H···O_W bonds in the coordination of water molecules, Steiner & Saenger, 1993).

3.3. *Can O_W—H···Ph bonding be predicted without knowledge of the H-atom position?*

The above observations show that in the hydration of peptides, water–aromatic interactions on occasions play roles that deserve attention. There is no reason why this should not be valid for proteins also. It would therefore be highly desirable to detect this type of interaction from data of the kind which is available from standard protein structure refinements, *i.e.* from only non-H atom positions. A simple approach is to perform a search for short water–aromatic contacts with O_W···C or O_W···*M* distances shorter than a suitable cutoff distance, similar to that commonly carried out for 'possible O—H···O hydrogen bonds'. Following this approach, several tentative searches were performed in the amino acid and peptide subset of the CSD, which were based on different criteria using only O_W and C(Ph) positions, ignoring the (published) H-atom positions. A successful

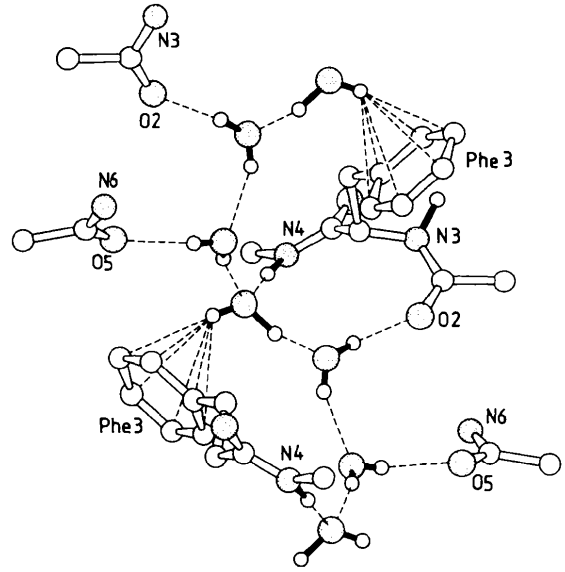


Fig. 4. Section of the hydrogen-bond pattern in the cyclic hexapeptide cycloamanide A hydrate, BIHXUL10 (drawn using coordinates from Chiang *et al.*, 1982).

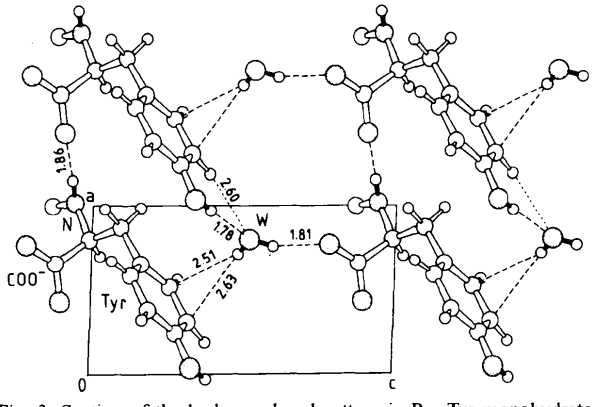


Fig. 3. Section of the hydrogen-bond pattern in Pro-Tyr monohydrate, SOJPAI (drawn using coordinates from Klein *et al.*, 1991). Distances are given in Å (based on normalized *X*—H bond lengths). Of the peptide, only the Tyr residue is shown. For the long contact C_ε—H···O_W it is questionable whether or not it should be regarded as a weak hydrogen bond.

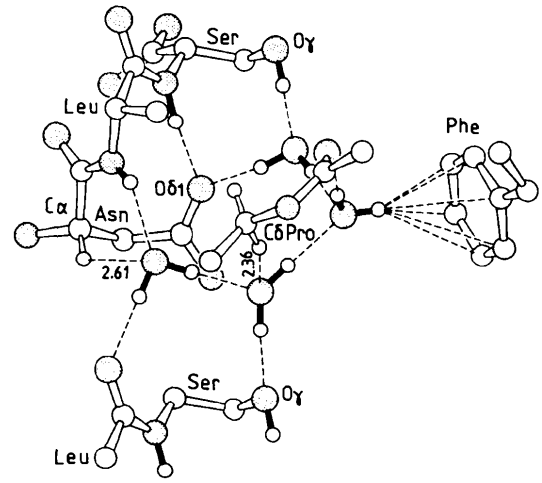


Fig. 5. Section of the hydrogen-bond pattern in the cyclic heptapeptide evolidine tetrahydrate, TALVAD (drawn using coordinates from Eggleston *et al.*, 1991). For C—H···O_W contacts, H···O_W distances are given in Å (based on normalized *X*—H bond lengths).

search algorithm should yield the cases listed in Table 2. For pragmatic reasons, criteria should be as simple as possible, using only parameters which are readily calculated by standard crystallographic programs; unfortunately, this excludes the angle ω which is a non-standard parameter.

Criteria which are based on only one $O_W\cdots C(Ph)$ separation yield numerous contacts which are clearly non-hydrogen bonds; they are therefore unsuitable. The simple criterion that in an O_W—H···Ph hydrogen bond, all six $O_W\cdots C$ separations must be < 4.6 Å, and at least one must be shorter than 3.6 Å, yielded the examples in Table 2 and the 'dubious' cases sorted out previously (see §2.2), but no case which is clearly not an O—H···Ph hydrogen bond. Unfortunately, because of the small quantity of data, the predictive power of this result is very limited; we cannot regard it as a basis for introducing reasonably realiable search criteria for the identification of water–aromatic hydrogen bonds. Furthermore, it might be necessary to define more permissive criteria in order not to omit possible longer and more off-centered aromatic hydrogen bonds than those in Table 2. In any way, the searches performed show the relevant result that in the peptide crystal structures published until now, there is not a single case of short water approach to the face of the Ph group without the formation of an aromatic hydrogen bond.

4. Concluding remarks

The crystal structure of Tyr-Tyr-Phe dihydrate contains a centered O—H···Ph hydrogen bond from a water molecule to the face of the Phe aromatic moiety. The distance from O_W to the aromatic center is only 3.26 Å. In a database study on well refined hydrated peptides, four previous examples of water–aromatic hydrogen bonds were found. Comparison of these data shows a wide variability of O_W—H···Ph hydrogen-bond geometries: there are almost ideally centered arrangements with the water molecule residing over the aromatic centroid, and there are very off-centered arrangements.

When inspecting the intermolecular environment of the relevant water molecules, some common features become apparent. Typically, the water molecules are engaged in complex patterns of conventional and non-conventional hydrogen bonds, normally involving O/N—H···O, C—H···O and O—H···Ph interactions. If only the conventional hydrogen bonds are looked at and the non-conventional ones are neglected, the resulting water coordination is highly unfavourable or even inconceivable. If the non-conventional C—H···O_W and O_W—H···Ph interactions are considered as potential constituents of water coordination, arrangements are obtained which can be easily rationalized.

If water molecules find themselves in a partly apolar surrounding, they tend to form as many conventional hydrogen bonds as possible, and then 'fill up' the remaining hydrogen-bond potential with weak non-conventional hydrogen bonds. Water molecules accept C—H···O_W interactions if O—H and N—H donors are not available (Steiner & Saenger, 1993), and they donate O_W—H···π interactions rather than 'donating nothing'.

It can be expected that the above observations are relevant not only for peptide, but also for protein hydration. Aromatic amino-acid side chains at the surface of proteins are potential hydrogen bond acceptors for water molecules. Water molecules close to such aromatic faces should tend to make the best out of this situation, and that is forming an aromatic hydrogen bond. As a matter of fact, O_W—H···Ph interactions should be less stable than conventional hydrogen bonds, and their average lifetime at the protein–solvent interface should be shorter. In X-ray structure analyses, these water molecules would be more difficult to locate and refine than well coordinated water molecules, and possibly remain unobserved. Internal water molecules of proteins are often found poorly coordinated by conventional hydrogen-bond partners (Baker & Hubbard, 1984). Apart from protein C—H donors, aromatic residues might give an opportunity to satisfy their hydrogen-bond potentials (Buckle *et al.*, 1996).

TS thanks Wolfram Saenger (Freie Universität Berlin) for giving him the opportunity to carry out part of the study in his laboratory, and the Research Network 'Molecular Recognition Phenomena' of the European Union for supporting a stay at the Bijvoet Center of the Utrecht University.

References

Allen, F. H., Davies, J. E., Galloy, J. E., Johnson, J. J., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187.

Altomare, A., Cascarano, G., Giacovazzo, C., Guagliardi, A., Burla, M. C., Polidori, G. & Camalli, M. (1994). *J. Appl. Cryst.* **27**, 435.

Atwood, J. L., Hamada, F., Robinson, K. D., Orr, G. W. & Vincent, R. L. (1991). *Nature (London)*, **349**, 683–684.

Aubry, A., Protas, J., Moreno-Gonzales, E. & Marraud, M. (1977). *Acta Cryst.* B**33**, 2572–2578.

Baker, E. N. & Hubbard, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.

Bakshi, P. K., Linden, A., Vincent, B. R., Roe, S. P., Adhikesavalu, D., Cameron, T. S. & Knop, O. (1994). *Can. J. Chem.* **72**, 1273–1293.

Buckle, A. M., Cramer, P. & Fersht, A. R. (1996). *Biochemistry*, **35**, 4298–4305.

Burley, S. K. & Petsko, G. A. (1986). *FEBS Lett.* **203**, 139–143.

Chiang, C. C., Karle, I. L. & Wieland, T. (1982). *Int. J. Pept. Protein Res.* **20**, 414–420.

Deeg, A. & Mootz, D. (1993). *Z. Naturforsch.* **486**, 571–576.

Derewenda, Z. S., Lee, L. & Derewenda, U. (1995). *J. Mol. Biol.* **252**, 248–262.

Desiraju, G. R. (1996). *Acc. Chem. Res.* **29**, 441–449.

Eggleston, D. S., Baures, P. W., Peishoff, C. E. & Kopple, K. D. (1991). *J. Am. Chem. Soc.* **113**, 4410–4416.

Engh, R. A., Brandstetter, H., Sucher, G., Eichinger, A., Baumann, U., Bode, W., Huber, R., Poll, T., Rudolph, R. & von der Saal, W. (1996). *Structure*, **4**, 1353–1362.

Feldman, S. H. & Eggleston, D. S. (1990). *Acta Cryst.* C**46**, 678–682.

Ferguson, G., Gallagher, J. F., Glidewell, C. & Zakaria, C. M. (1994). *Acta Cryst.* C**50**, 70–73.

Flanagan, K., Walshaw, J., Price, S. L. & Goodfellow, J. M. (1995). *Protein Eng.* **8**, 109–116.

Jeffrey, G. A. & Maluszynska, H. (1982). *J. Biol. Macromol.* **4**, 173–185.

Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*. Berlin: Springer.

Klein, C. L., Cobbinah, I., Rouselle, D., Malmstrom, M. C. Sr & Stevens, E. D. (1991). *Acta Cryst.* C**47**, 2386–2388.

Levitt, M. & Perutz, M. F. (1988). *J. Mol. Biol.* **201**, 751–754.

Liu, S., Ji, X., Gilliland, G. L., Stevens, W. J. & Armstrong, R. N. (1993). *J. Am. Chem. Soc.* **115**, 7910–7911.

Perutz, M. F., Fermi, G., Abraham, D. J., Poyart, C. & Bursaux, E. (1986). *J. Am. Chem. Soc.* **108**, 1064–1078.

Pimentel, G. C. & McClellan, A. L. (1960). *The Hydrogen Bond*. San Francisco: Freeman.

Read, W. G., Campbell, E. J. & Henderson, G. (1983). *J. Chem. Phys.* **78**, 3501–3508.

Sheldrick, G. A. (1993). *SHELXL93. Program for the refinement of crystal structures*. University of Göttingen, Germany.

Steiner, Th. (1996). *Cryst. Rev.* **6**, 1–57.

Steiner, Th. (1997). *Chem. Commun.* pp. 727–734.

Steiner, Th. & Saenger, W. (1993). *J. Am. Chem. Soc.* **115**, 4540–4547.

Steiner, Th., Starikov, E. B., Amado, A. M. & Teixeira-Dias, J. J. C. (1995). *J. Chem. Soc. Perkin Trans. 2*, pp. 1321–1326.

Steiner, Th., Starikov, E. B. & Tamm, M. (1996). *J. Chem. Soc. Perkin Trans. 2*, pp. 67–71.

Suzuki, S., Green, P. G., Bumgarner, R. E., Dasgupta, S., Goddard, W. A. III & Blake, G. A. (1992). *Science*, **257**, 942–945.

Tüchsen, E. & Woodward, C. (1987). *Biochemistry*, **26**, 1918–1925.

Worth, G. A. & Wade, R. C. (1995). *J. Phys. Chem.* **99**, 17473–17482.