The Role of Aromatic Rings as Hydrogen-Bond Acceptors in Molecular Recognition

M. F. Perutz

Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

The role of aromatic rings as hydrogen-bond acceptors in molecular recognition

By M. F. Perutz

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

Observations of phenol–benzene and ammonia–benzene complexes in the gas phase show that hydrogen bonds link their proton donors to the π electrons of the benzene with a bond energy of between 2 and 4 kcal mol⁻¹, large enough to be biologically significant. Intramolecular hydrogen bonds between OH and NH donors and aromatic acceptors have also been found in crystal structures of organic compounds. NH–aromatic interactions stabilize α-helices if donors and acceptors occur at successive turns of the helix. These interactions also contribute to the stability of several proteins and play an important part in cellular and synaptic signal transmission.

Some years ago my colleagues and I studied the stereochemistry of drug binding by haemoglobin (Perutz *et al.* 1986). We found that haemoglobin has a high affinity for the antilipidaemic drug bezafibrate. X-ray analysis showed that bezafibrate bound to the walls of the water-filled cavity that runs through the centre of the haemoglobin molecule. Among the nine amino acid residues in contact with the drug was an asparagine whose amino group pointed at the plane of one of the drug's two benzene rings, suggestive of a hydrogen bond with the ring's π electrons, but the resolution was not good enough to measure the N-···C distances sufficiently accurately to be sure of their being shorter than the van der Waals distance (figure 1).

All the same, the observation was interesting enough to make me ask whether such hydrogen bonds are known to exist, what their energy is, what role they play in the recognition of substrates and transmitters, and in stabilizing protein structures.

Unsurprisingly, evidence for aromatic rings acting as hydrogen bond acceptors was first discovered by that superb chemist, the late George Pimentel. He measured the lowering of the infrared NH-stretching frequency of N-ethylacetamide in various solvents and found that 'benzene appears to interact with the amide hydrogen to a considerable extent, since the free NH frequency is lowest where benzene is used as a solvent'. He suggested that the π electrons of benzene act as a base (Klemperer *et al.* 1954).

In recent years clear evidence for the acceptance of hydrogen bonds by aromatic rings has come from spectroscopic studies of the association of benzene with phenol and ammonia in the gas phase and from the crystal structures of compounds containing intramolecular hydrogen bonds between benzene rings and proton donors.

McPhail & Sim (1965) reported an intramolecular hydroxyl-benzene hydrogen bond at an oxygen to benzene-plane distance of 3.1 Å in a cyclic peptide structure
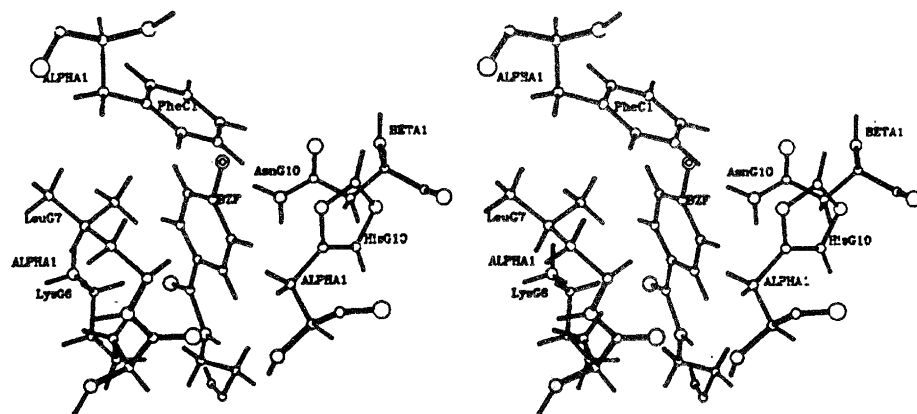
105

Figure 1. Stereoview of the interactions of AsnG10(108)β with the chlorobenzene ring of the drug bezafibrate bound to human deoxyhaemoglobin. (Reproduced, with permission, from Perutz *et al.* (1986).)
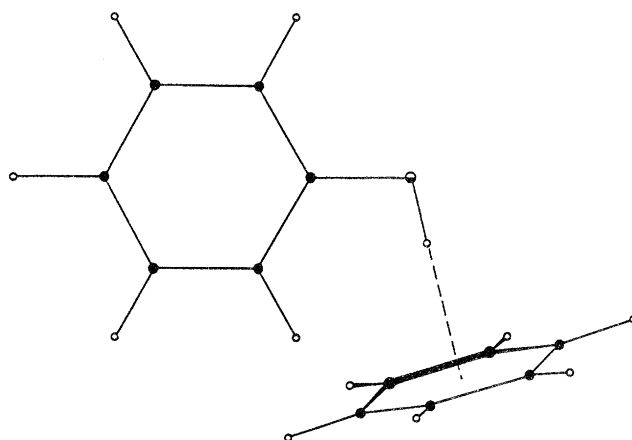

Figure 2. Phenol–benzene binding geometry in the gas phase. The phenol donates a proton to the π electrons of the benzene ring. (Reproduced, with permission, from Knee *et al.* (1987).)

determined by X-ray crystallography. Zewail and others at CalTech studied the dissociation of the tyrosine–benzene complex by picosecond photofragment spectroscopy and estimated the energy of the hydrogen bond between the hydroxyl and the π-electrons at room temperature *in vacuo* to be about 4 kcal mol^{-1} (figure 2; Knee *et al.* 1987). In carbon tetrachloride solution, measurements of the association constant of the tyrosine–benzene complex led to an enthalpy of association of -4 kcal mol^{-1}, but the loss of rotational and translational entropy on association reduced the binding energy to about 1 kcal mol^{-1}.

High resolution optical and microwave spectra of benzene–ammonia dimers in the gas phase showed that ammonia forms a hydrogen bond with the aromatic ring with a bond energy of 2.4 kcal mol^{-1}, and with a nitrogen to plane of the ring distance of 3.6 Å. The ammonia molecule rotates freely about its threefold symmetry axis (figure 3; Rodham *et al.* 1993).

A search of the crystallographic data base for NH–benzene interactions by Michael Levitt revealed a peptide that contains two intramolecular NH–benzene bonds with
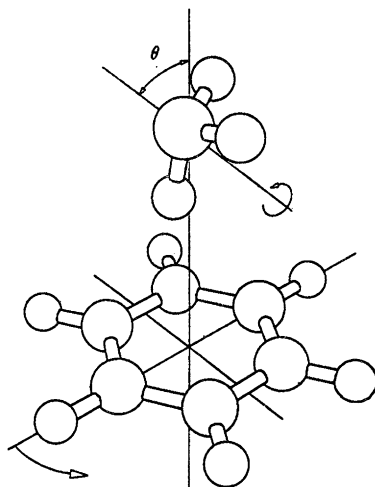
Figure 3. Ammonia–benzene binding geometry in the gas phase. (Reproduced, with permission, from Rodham *et al.* (1993).)

(benzyloxycarbonyl-1-amino-
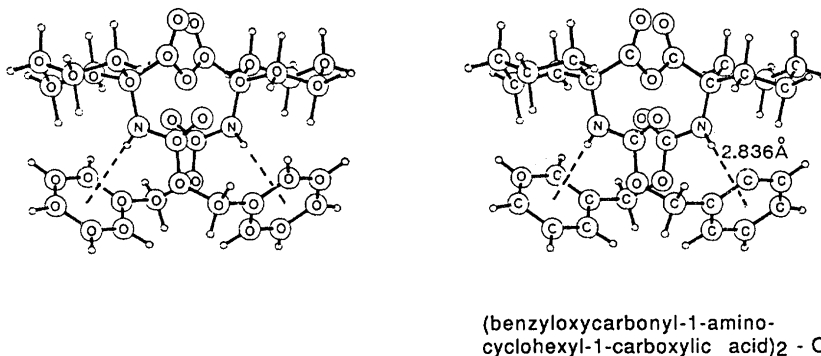cyclohexyl-1-carboxylic acid)2 - O

Figure 4. Stereoview of internal hydrogen bonds between two NH groups and two benzene rings in (benzyloxycarbonyl-1-aminocyclohexyl-1-carboxylic acid)$_2$-O. (By courtesy of Dr David Watson, University Chemical Laboratory, Cambridge, U.K.)

distances from the nitrogen to the plane of the benzene of only 2.8 Å, much shorter than the sum of the van der Waals radii which would lie between 3.5 and 4.0 Å (figure 4; Valle *et al.* 1988).

These data leave no doubt that the bond exists and has physiologically significant strength. How large a role does it play in protein structure and function? Burley & Petsko (1986) made a computer search for interactions between amino groups and aromatic rings in 33 highly refined protein structures and concluded that they occur more frequently than random and may contribute to the enthalpy of protein stability. J. M. Thornton *et al.*s (this volume) detailed structural analysis has shown that few of these interactions really include hydrogen bonds from the donors to the aromatic rings, because competition from nearby stronger hydrogen bond acceptors draws them away even if they are close to the rings. All the same, several well-documented cases of such stabilizing interactions have come to light.

The first such evidence came from the high resolution structure of the pancreatic trypsin inhibitor, where Tüchsen & Woodward (1987) found the benzene ring of tyrosine 37 sandwiched between the NH of Gly 37 and the NH$_2$ of Asn 44, each of
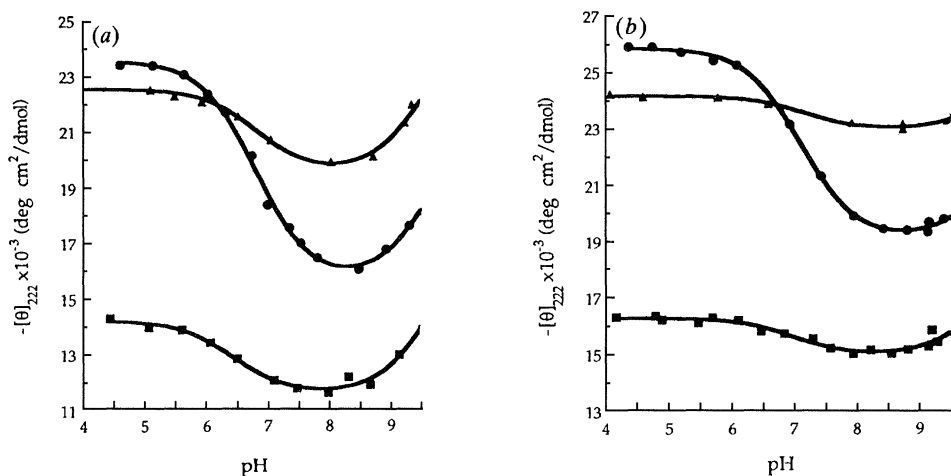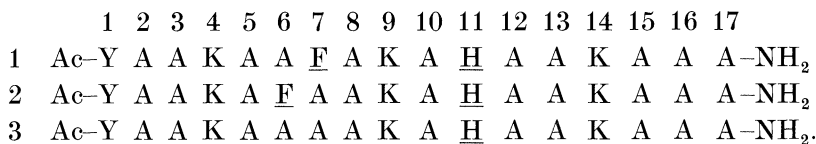
M. F. Perutz

Figure 5. pH dependence of the mean residue ellipticity of peptides (1) F12 H16 (●), (2) F11 H16 (■) and (3) H16 (▲), (a) in 10 mM NaCl and (b) in 1 M NaCl. (Reproduced, with permission, from Armstrong *et al.* (1993).)

them closer to the ring than the van der Waals distance. Nuclear magnetic resonance spectra of the inhibitor show the resonances of the peptide proton of Gly 37 and the N_ζ proton of the amide of Asn 44 to be shifted upfield by the ring current of the Tyr 35 aromatic ring. Tüchsen & Woodward's accurate measurements enabled Levitt & Perutz (1988) to calculate the likely free energy of such a bond from the partial negative charges centred on the ring carbons and the partial positive charges centred on the amide hydrogens. They arrived at an energy minimum of about 3 kcal mol^{-1}, at a nitrogen to ring centre distance of 3.4 Å, similar to the ammonia–benzene bond in the vapour.

Armstrong *et al.* (1993) measured the interaction between a histidine and a phenylalanine separated by one turn of an α-helical peptide. They synthesized the peptides with the sequence:

```
        1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16 17
1   Ac–Y  A  A  K  A  A  F  A  K  A  H  A  A  K  A  A  A–NH₂
2   Ac–Y  A  A  K  A  F  A  A  K  A  H  A  A  K  A  A  A–NH₂
3   Ac–Y  A  A  K  A  A  A  A  K  A  H  A  A  K  A  A  A–NH₂.
```

(1) contains a phenylalanine and a histidine four residues apart, optimally spaced for interaction in an α-helix. In (2) they are spread five residues apart, too far for interaction. (3) lacks the phenylalanine.

Circular dichroism at 222 nm showed peptide 1 to be more helical than 2; the helicity of (1) rises strongly at low pH as its histidine becomes positively charged, while the helicities of (2) and (3) rise only weakly as their α-amino groups become charged; after that the curves remain flat. Curves (1) and (2) show less helicity than (3), except below pH 6.2 (figure 5). This is due to the intrinsic destabilizing effect of histidines in the middle of α-helices; they only confer stability at the C-termini of α-helices where they compensate the helix dipole. When Armstrong *et al.* replaced the phenylalanine with the non-aromatic cyclohexylphenylalanine, the histidine no longer stabilized the α-helix. Increasing ionic strength has little effect on the helix-
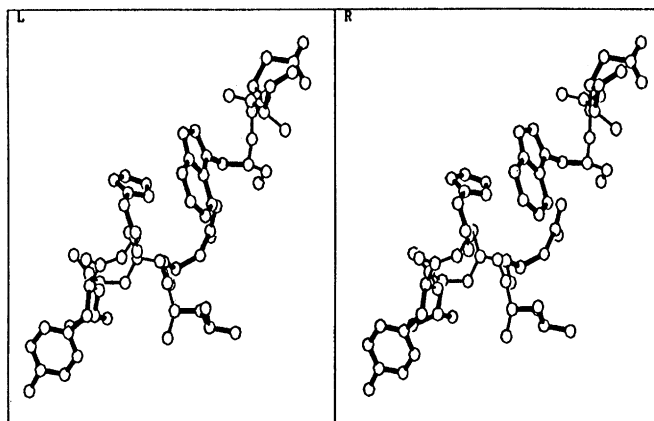
Figure 6. Stereoview of the interaction between His 18 and Trp 94 in Barnase; J. Cameron & H. Kendrick, unpublished crystal structure at pH 6. (Reproduced, with permission, from Loewenthal *et al.* (1992).)
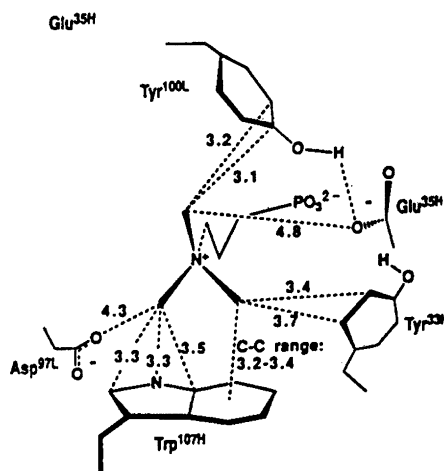

Figure 7. Binding of phosphocholine to the mouse immunoglobulin McPC603, showing the contacts between the tertiary ammonium group and the aromatic rings surrounding it. (Reproduced, with permission, from Dougherty & Stauffer (1990).)

stabilizing effect of the histidine–phenylalanine interaction. These results proved that the histidine stabilizes the α-helix by acting as a hydrogen bond donor to the phenylalanine, but they left the free energy of stabilization open.

Barnase (bacterial ribonuclease) is a small protein from *Bacillus amyloliquefaciens*. It contains a four-turn α-helix stabilized at its C-terminus by histidine 18. Facing the edge of the histidine ring is tryptophan 94, such that its $N_\delta H$ points at the π-electrons of the indol ring (figure 6). Loewenthal *et al.* (1992) determined the interaction between the two rings (1) by measuring the change of pk_a of the histidine when the tryptophan is replaced by other residues, (2) by measuring the difference in stability of barnase at pH 5.8 when the histidine is protonated and at pH 9 when it is unprotonated, and (3) by double mutant cycles: measuring the free energy of unfolding in urea of the wild-type protein, of the mutants Trp94 → Leu and His18 → Gly, and of the double mutant Trp94 → Leu and His18 → Gly. The results
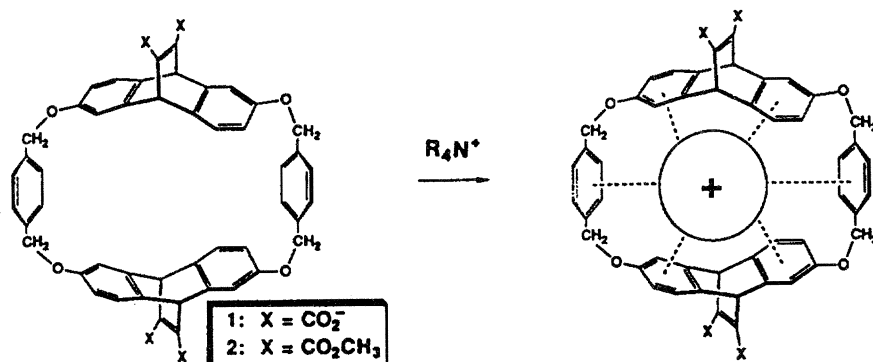
Figure 8. Synthetic model for the binding of a positively charged tertiary ammonium group $(R_4N)^+$ to a synthetic acceptor made of six aromatic rings joined in a circle. The dotted lines indicate the stabilizing cation–π interactions. (Reproduced, with permission, from Dougherty & Stauffer (1990).)

show that the free energy of interaction between the tryptophan and the protonated histidine amounts to $1.4\ \text{kcal mol}^{-1}$, and with the unprotonated histidine to $0.4\ \text{kcal mol}^{-1}$, relative to solvation with water.

Other observations concern the role of aromatic hydrogen bond acceptors in the binding of choline to immunoglobulins, enzymes or synaptic receptors. Phospho-choline is bound as a hapten by the F_{ab} fragment of a mouse lymphoma protein (Satow *et al.* 1986). Its positively charged tertiary ammonium group is not compensated by anionic side chains; instead it donates hydrogen bonds to the aromatic rings of a tryptophan and two tyrosines (figure 7).

This observation stimulated Dougherty and Stauffer to synthesize a quaternary ammonium-binding model made of six aromatic rings joined in a circle (figure 8). This model binds acetylcholine with a dissociation constant $K_D = 50\ \mu\text{M}$, a value comparable to those of biological recognition sites. Its primary binding force must come from cation–π interactions, because the positively charged tertiary ammonium group was found to be bound in preference to an uncharged tertiary butyl group.

Dougherty and Stauffer therefore proposed that the electron-rich aromatic rings of tyrosine and tryptophan, rather than anionic groups, bind acetylcholine in acetylcholinesterase and in cholinergic receptors. Their prediction, made in 1990, was confirmed by the recently solved structure of acetylcholinesterase (Sussman *et al.* 1991). Its active site lies at the head of a 20 Å long gorge whose floor and walls are lined with aromatic side chains. The active site itself lacks any anionic amino acid to compensate the positive charge of the tertiary ammonium group. Instead the group is well placed to form hydrogen bonds with the indole ring of a tryptophan (figure 9).

Opening and closing of the ion channel in the nicotinic acetylcholine receptor is effected by an allosteric transition of its five subunits. The opening is triggered by the binding of acetylcholine to a site that is widely separated from the channel. Photo-affinity labelling of the acetylcholine-binding site led to the labelling of two cysteines, two tryptophans and possibly two tyrosines, but of neither a glutamate nor an aspartate. This suggests that here again the tertiary ammonium group donates hydrogen bonds to aromatic rings (Dennis *et al.* 1988; Galzi *et al.* 1990).

Tyrosine kinases are vital in cellular signal transduction. Cellular forms of the Rous sarcoma (*rsc*) tyrosine kinase are either transmembrane growth hormone receptors or cytosolic non-receptor proteins associated with the inner plasma
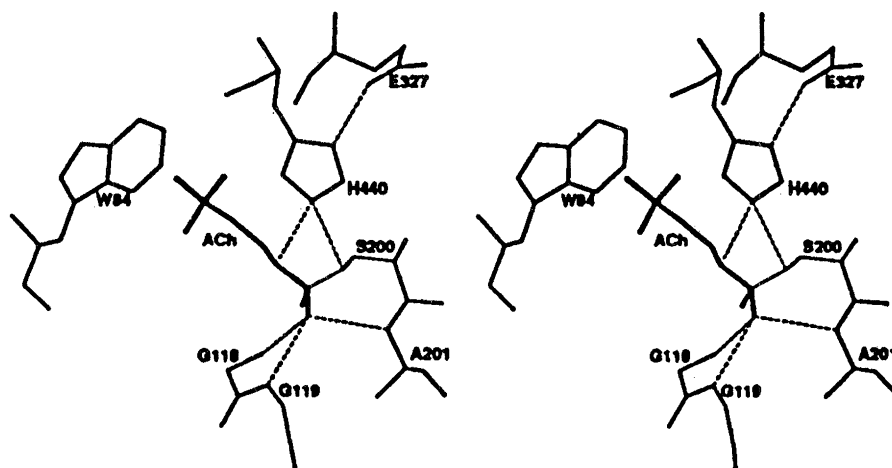
Figure 9. Stereoview of a theoretical model of acetylcholinesterase showing the hydrogen bonding of the acetyl group to polar groups of the enzyme and the probable interaction of the tertiary ammonium group with Trp 84. (Reproduced, with permission, from Sussman *et al.* (1991).)
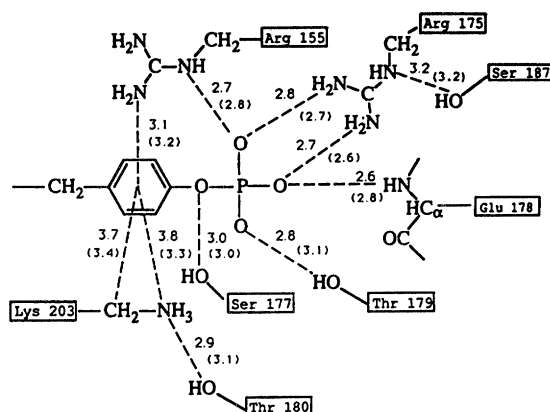

Figure 10. Diagrammatic view of the phosphotyrosine-binding site in the SH2 domain of the viral Rous sarcoma oncogene which forms part of a tyrosine kinase. Note the short distance between N of Arg 155 and the plane of the benzene ring. (Reproduced, with permission, from Waksman *et al.* (1992).)

membrane. Activation of the growth hormone receptor causes autophosphorylation of tyrosines on its cytoplasmic side. This is recognized as a signal by protein domains known as *rsc*-homology-2, or SH2 domains, which then transmit the signal to other proteins in the signal transduction pathway.

J. Kuriyan and his colleagues have crystallized the complex of the SH2 domain of the viral Rous sarcoma protein kinase with a pentapeptide from the platelet-derived growth factor. This peptide has a phosphotyrosine at its amino terminus. The phosphotyrosine is bound to the protein by hydrogen bonds from the guanidinium group of an arginine the plane of the phenol ring at the short distance of 3.1 Å, clearly indicating hydrogen bonding. There is also a more distant interaction with the ammonium group of a lysine (figure 10) (Waksman *et al.* 1992).

The chemical literature contains many more examples of interactions between proton donors and aromatic acceptors, but the ones quoted here are sufficient to

show that such interactions are of biologically significant strength, that they play an important role in synaptic and cellular signal transduction and also contribute to the stability of certain protein structures.

References

Armstrong, K. M., Fairman, R. & Baldwin, R. L. 1993 The (i,i+4) Phe–His interaction studied in an alanine-based α-helix. *J. mol. Biol.* **230**, 284–291.

Burley, S. K. & Petsko, G. A. 1986 Amino–aromatic interactions in proteins. *FEBS Lett.* **203**, 138–143.

Dennis, M., Giraudat, J., Kotzyba-Hilbert, F., Goeldner, M., Hirth, C., Chang, J.-Y., Lazure, C., Chrétien, M. & Changeux, J. P. 1988 Amino acids of the Torpedo marmorata acetylcholine receptor α subunit labeled by a photo affinity ligand for the acetylcholine binding site. *Biochemistry* **27**, 2346–2357.

Dougherty, D. A. & Stauffer, D. A. 1990 Acetylcholine binding by a synthetic receptor: implications for biological recognition. *Science, Wash.* **250**, 1558–1560.

Galzi, J. L., Revah, F., Black, D., Goeldner, M., Hirth, C. & Changeux, J. P. 1990 Identification of a novel amino acid: α-tyrosine 93 within the cholinergic ligand-binding sites of the acetylcholine receptor by photo affinity labelling. *J. biol. Chem.* **265**, 10430–10437.

Klemperer, W., Cronyn, M. W., Maki, A. M. & Pimentel, G. C. 1954 Infrared studies of the association of secondary amides in various solvents. *J. Am. chem. Soc.* **76**, 5846–5848.

Knee, J. L., Khundkar, L. R. & Zewail, A. H. 1987 Picosecond photofragment spectroscopy. III. vibrational predissociation of van der Waals' clusters. *J. chem. Phys.* **87**, 115–127.

Levitt, M. & Perutz, M. F. 1988 Aromatic rings as hydrogen bond acceptors. *J. mol. Biol.* **201**, 751–754.

Loewenthal, R., Sancho, J. & Fersht, A. R. 1992 Histidine–aromatic interactions in barnase: elevation of histidine pk_a and contribution to protein stability. *J. mol. Biol.* **224**, 759–770.

McPhail, A. T. & Sim, G. A. 1965 Hydroxyl-benzene hydrogen bonding: an X-ray study. *Chem. Commun.* 125–127.

Perutz, M. F., Fermi, G., Abraham, D. J., Poyart, C. & Bursaux, E. 1986 Hemoglobin as a receptor of drugs and peptides: X-ray studies of the chemistry of binding. *J. Am. chem. Soc.* **108**, 1064–1078.

Rodham, D. A., Suzuki, S., Sueram, R. D., Lovas, F. J., Dasgupta, S., Goddard, W. A., III & Blake, G. A. 1993 Hydrogen bonding in the benzene–ammonia dimer. *Nature, Lond.* **362**, 735–737.

Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. 1986 Phosphocholine binding immunoglobulin F_{ab} McP603: an X-ray diffraction study at 2.7 Å. *J. mol. Biol.* **190**, 593–604.

Sussman, J. L., Harel, M., Frolow, F., Oeffner, C., Goldman, A., Toker, L. & Silman, I. 1991 Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. *Science, Wash.* **253**, 872–879.

Tüchsen, E. & Woodward, C. 1987 Assignment of Asn44 side chain primary amide ^{1}HNMR resonances and the peptide amide resonance of Gly 37 in basic pancreatic trypsin inhibitor. *Biochemistry* **26**, 1918–1925.

Waksman, G., Kominos, D., Robertson, S. C. *et al.* 1992 Crystal structure of the phosphotyrosine recognition domain SH2 of *v-src* complexed with tyrosine-phosphorylated peptides. *Nature, Lond.* **358**, 646–653.