Packing of aromatic rings against tryptophan residues in proteins

Uttamkumar Samanta, Debnath Pal and Pinak Chakrabarti*

Department of Biochemistry, Bose Institute, P-1/12 CIT Scheme VII-M, Calcutta 700 054, India

Correspondence e-mail: pinak@boseinst.ernet.in

The geometry of the interaction of the aromatic side chains of phenylalanine (Phe), tyrosine (Tyr), tryptophan (Trp) and histidine (His) with the indole ring of Trp has been analyzed using the structures in the Protein Data Bank in order to understand the dependence of the packing behaviour on the size and chemical nature of the aromatic rings. The Phe ring prefers to interact either perpendicularly, with its edge pointing towards the Trp face, or in an offset-stacked arrangement. The edge-to-face motif is typical of a Trp–Trp pair. While parallel stacking is the dominant feature of Trp–His interaction, Tyr packs in a more uniform manner around Trp with a higher than expected occurrence at the edge and a few cases of possible OH–π interaction.
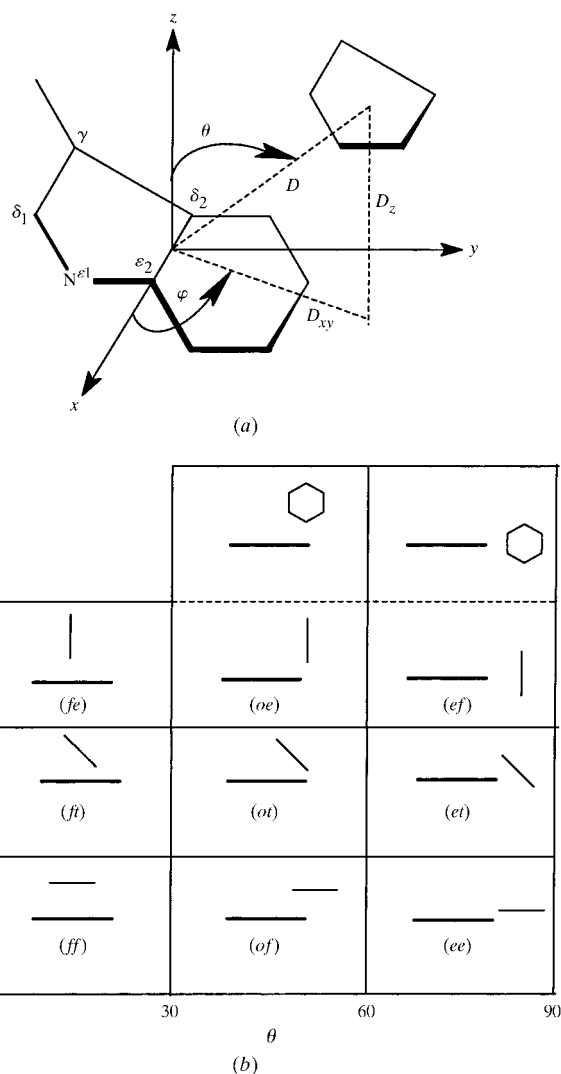
1. Introduction

Despite the fact that aromatic rings stack against each other in DNA (Dickerson, 1983), in proteins they show a preference for being perpendicular to each other (Burley & Petsko, 1985; Singh & Thornton, 1985). These observations led to theoretical studies on the energetics of these two binding modes (Burley & Petsko, 1986; Linse, 1992; Jorgensen & Severance, 1990; Hobza *et al.*, 1994), analyses of their prevalence in the crystal structures of small organic molecules (Cox *et al.*, 1958; Desiraju & Gavezzotti, 1989; Dahl, 1994) and their use in the design of supramolecular assemblies and molecular-recognition studies (Beeson *et al.*, 1994; Newcomb & Gellman, 1994; Zhang & Moore, 1992; Seel & Vögtle, 1992; Ferguson *et al.*, 1991; Cochran *et al.*, 1992; Paliwal *et al.*, 1994; Moody *et al.*, 1987; Grossel *et al.*, 1993; Benzing *et al.*, 1988; Muehldorf *et al.*, 1988; Hunter, 1991; Zimmerman *et al.*, 1991). The preferred mode of interaction between aromatic rings is usually stacked offset (face-to-face with the rings in a staggered arrangement) or T-shaped (edge-to-face); the latter mode, in which the H atoms (with a partial positive charge) on one ring points toward the π-electron cloud on a second, appears to be energetically favourable (Hunter & Sanders, 1990; Gould *et al.*, 1985; Hunter, 1994). However, if two aromatic rings are different (in size and aromaticity) the question arises whether edge-to-face orientation is still the preferred mode and, if so, whether one of the rings prefers to interact *via* its face (rather than its edge) with the other. The answer may be sought from an analysis of the geometry of interaction of the tryptophan (Trp) ring with other aromatic rings [of phenylalanine (Phe), tyrosine (Tyr) and histidine (His)] in high-resolution globular protein structures. Being a heteroaromatic fused-ring system, Trp is quite distinct from all other protein rings and is, in a way, similar to purine bases in DNA. Moreover, Trp is located in the binding sites of many proteins such as lysozyme (Maenaka

et al., 1995), transcription factor (Kodandapani et al., 1996) and streptavidin (Chilkoti et al., 1995), the latter binding its substrate, biotin, with an affinity which is among the highest displayed for non-covalent interactions between a ligand and a protein ($K_a \simeq 10^{13}\,M^{-1}$). Additionally, Trp clusters are found in many globular proteins and particularly in membrane proteins (Artymiuk et al., 1994; Schiffer et al., 1992), for which only a few structures of low resolution are available. In order to understand the features of such interactions we have also included the Trp–Trp pair in our analysis.

2. Methods and calculations

The protein structures were selected from the Protein Data Bank (PDB, 1996 version; Bernstein et al., 1977) with constraints of 25% maximum sequence identity upon pairwise

alignment (Hobohm & Sander, 1994), crystallographic resolution ≤ 2.0 Å and an R value ≤ 0.20. So as to consider only the well ordered Trp residues, those with more than two ring atoms having a thermal factor >30 Å^2 were excluded; this gave a list of 719 residues. An aromatic residue (Ar) with one or more of its ring atoms (including, in the case of Tyr, the hydroxyl O atom) at a distance ≤ 4 Å from any Trp ring atom was considered as a packing partner; contacts which are brought about by the application of crystal symmetry operators were also counted.

In order to characterize the geometry of the Trp–Ar interaction, two parameters independent of the coordinate system were calculated: D, the distance between the two ring centres (which for Trp was taken as the mid-point of the $C^{\delta 2}$—$C^{\varepsilon 2}$ bond) and P, the angle between the ring planes. Additionally, the coordinates were expressed in a molecular-axes system (Fig. 1a) defined with the origin at Trp centre, the x axis along $C^{\delta 2}$—$C^{\varepsilon 2}$, the z axis perpendicular to the ring plane and the y axis in the plane of the ring, orthogonal to x and z. The spherical polar angles specifying the partner centre were determined: θ, the co-latitude ($0 \leq \theta \leq 90°$, i.e. no distinction is made between the $+z$ and $-z$ directions) and φ, the longitude ($-180 < \varphi \leq 180°$). D was also resolved into two components, D_{xy} (the distance of the projection of the partner centre on the xy plane from the origin) and D_z (the vertical distance).

To facilitate the visualization of the relative orientation between the rings, the combination of θ and P angles have been grouped into grids of $30 \times 30°$ and the idealized motif in each box is depicted, along with its designation, in Fig. 1(b). The convention used for naming is as follows. The first letter indicates whether the partner is interacting with the face (f) or

Figure 1
(a) The coordinate axes based on Trp and the definition of the geometric parameters. (b) Schematic representations and their nomenclature (see text for details) for ring orientations corresponding to various combinations of P and θ values (in °). Lines signify aromatic planes (the longer one for Trp and the shorter one for the partner) perpendicular to the paper; in two cases, however, the aromatic group, shown as a hexagon, is in the plane of the paper.

Table 1
Statistics of interplanar geometries.

Ranges of θ and P in each grid are as shown in Fig. 1(b). Expected numbers (from a random distribution) are shown in parentheses. Observed values significantly higher or lower than the expected values are given in bold or italics, respectively. In case of Tyr, the observed numbers of hydrogen-bonding and OH–π interactions involving the side-chain hydroxyl group are given in square brackets. To be identified as a hydrogen bond, the O atom must be along the edge of the Trp ring at a distance <3.5 Å from $N^{\varepsilon 1}$, whereas for an OH–π interaction the atom has to be along the face at a distance less than 4 Å from any Trp atom; the value of θ defined for O with the origin placed on the contact atom in the Trp ring (Fig. 1a) is used to determine whether it is located on the Trp face ($\theta < 45°$) or edge ($\theta > 45°$).

P range		θ range		
		0–30°	30–60°	60–90°
60–90°	Phe	**43** (33.1)	49 (59.7)	69 (68.2)
	Tyr	11 (9.8) [0, 1]	23 (33.3) [4, 2]	**70** (60.9) [5, 2]
	Trp	**15** (8.5)	18 (21.5)	29 (32.0)
	His	6 (9.4)	11 (12.4)	**18** (13.2)
30–60°	Phe	14 (18.9)	36 (34.1)	42 (39.0)
	Tyr	9 (7.9) [0, 2]	29 (26.5) [0, 4]	45 (48.6) [7, 4]
	Trp	2 (6.6)	17 (16.6)	29 (24.8)
	His	7 (7.8)	10 (10.3)	12 (10.9)
0–30°	Phe	5 (10.1)	**27** (18.2)	17 (20.8)
	Tyr	1 (3.3) [0, 0]	**19** (11.2) [0, 7]	15 (20.5) [0, 0]
	Trp	0 (1.9)	8 (4.9)	6 (7.2)
	His	**12** (7.8)	12 (10.3)	5 (10.9)

the edge (*e*) of Trp, or in an intermediate situation where the partner centre is on the face but offset (*o*) relative to the Trp centre. The second letter indicates whether the partner ring is tilted (*t*) relative to the Trp ring (30 < *P* < 60°) or (for other values of *P*) if the closest point of contact involves the face (*f*) or the edge (*e*) of the partner.
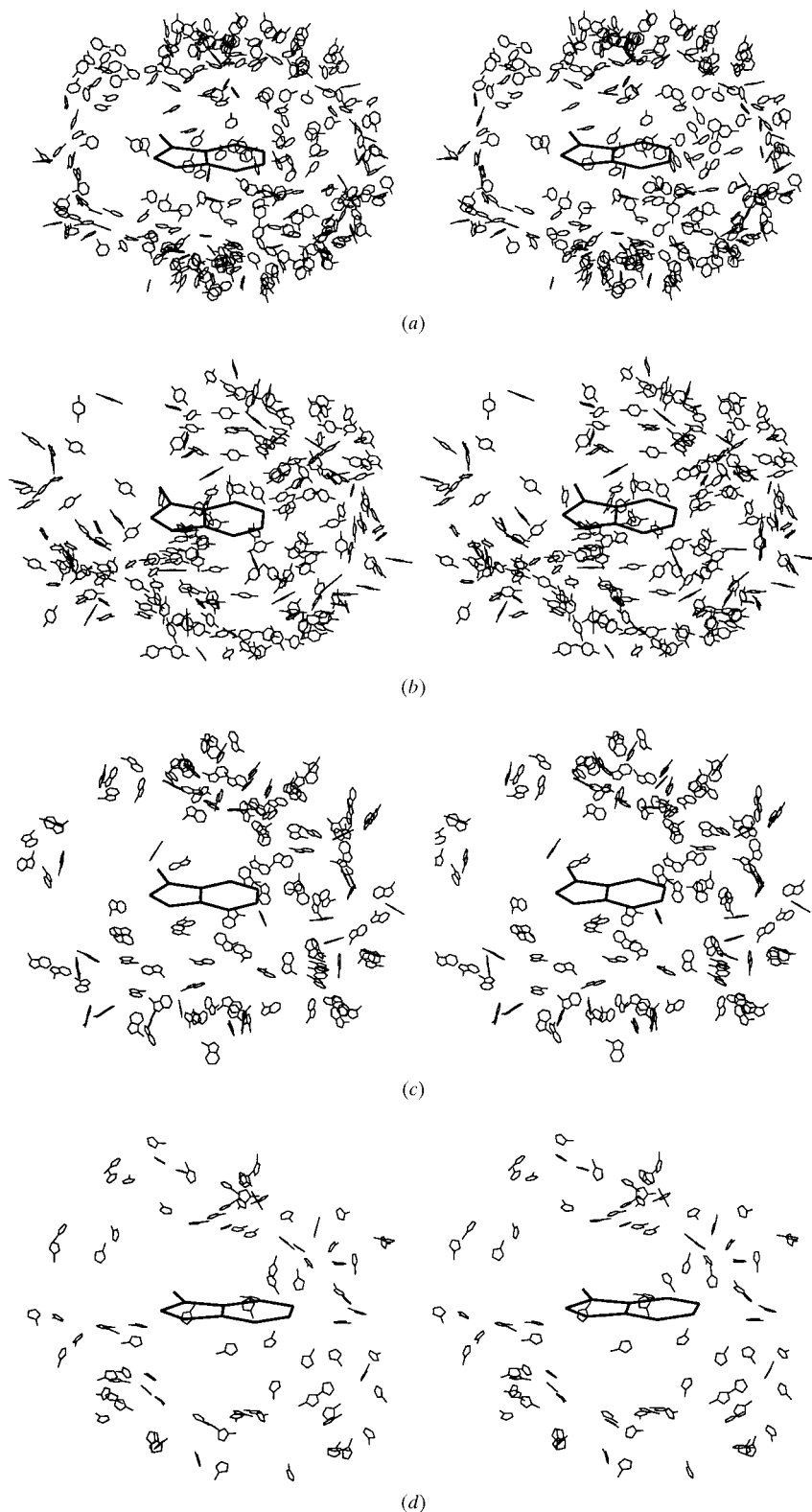
When the two rings are at 90°, a rotation about the vertical axis of the partner ring does not change the values of *P* and θ, although the relative orientations may be altered. As a result, two further limiting orientations are shown in two extra boxes at the top of Fig. 1(*b*). In the top right-hand grid, the additional structure involving the two rings should be named as *ee* (as in the bottom right-hand grid); but as this relative orientation was hardly ever observed (only two cases), the grid is assumed to contain only the *ef* structure. For the complete specification of the spatial organization of the two interacting aromatic groups, one also needs to consider the spherical polar angles (θ', φ') of the Trp centre relative to the molecular axes for the partner (Singh & Thornton, 1985; Brocchieri & Karlin, 1994). However, as we are mainly concerned with the orientation of the partner ring relative to Trp, we have used only two parameters, *P* and θ, thereby facilitating the visualization of the spatial disposition of the rings (Fig. 1*b*).

A random distribution of *P* or θ varies having a density equal to the sine of the angle (Singh & Thornton, 1985). For the joint distribution of *P* and θ, the expected value (E_{ij}) in a grid (Fig. 1*b*) was calculated using

$$E_{ij} = \sum_j O_{ij} \sum_i O_{ij} / \sum_i \sum_j O_{ij},$$

where O_{ij} is the observed frequency of occurrence in the grid corresponding to the *i*th row and *j*th column (*i* and *j* varying from 1 to 3). When the observed value in a grid differs significantly from the expected value (the absolute value of the difference is greater than the root-mean-square difference of all the grid points of a Trp–Ar pair), it is highlighted in Table 1. The expected value can also be approximated (data not shown) by the function sin*P*sinθ, and a similar result was obtained. The χ^2 test was used to determine the statistical significance in the difference between the observed and expected distributions for each Trp–Ar pair.

3. Results

Details of the PDB files, residues involved and the geometry of interactions are given

Figure 2
Scatterplot (in stereo) of the packing of (*a*) Phe, (*b*) Tyr, (*c*) Trp and (*d*) His residues against Trp. For clarity, the partner rings are reduced (by a scale factor of 0.2) in size. The atoms C$^\beta$ (for all residues) and the hydroxyl O (Tyr) are included in molecular diagrams.

Table 2
Average D_{xy} and D_z distances (with their standard deviations) (Å).

P range		θ range		
		0–30°	30–60°	60–90°
60–90°	Phe	1.4 (7), 4.9 (2)	4.4 (8), 4.1 (5)	6.1 (7), 1.4 (10)
	Tyr	1.4 (7), 5.0 (2)	4.6 (9), 4.2 (7)	6.0 (7), 1.5 (9)
	Trp	1.8 (8), 5.3 (4)	4.5(10), 4.6 (7)	6.0 (8), 1.6 (10)
	His	1.3 (6), 4.6 (2)	3.7 (9), 3.9 (5)	5.8 (6), 1.1 (6)
30–60°	Phe	1.7 (4), 4.6 (2)	4.6 (9), 3.9 (5)	6.1 (7), 1.5 (8)
	Tyr	1.9 (4), 4.6 (2)	4.3 (8), 3.8 (5)	6.2 (8), 1.8 (9)
	Trp	1.5 (7), 5.2 (1)	4.6 (11), 4.4 (6)	6.6 (9), 1.8 (12)
	His	1.5 (5), 4.2 (2)	3.9 (10), 3.5 (5)	5.8 (7), 1.5 (8)
0–30°	Phe	1.8 (2), 3.9 (2)	3.8 (10), 3.5 (4)	6.2 (6), 2.3 (8)
	Tyr	0.5, 3.6	3.9 (10), 3.4 (3)	6.2 (10), 2.1 (8)
	Trp		4.0 (7), 3.4 (5)	6.7 (8), 2.3 (9)
	His	1.4 (5), 3.6 (2)	3.7 (8), 3.2 (3)	6.0 (8), 1.8 (12)

in the supplementary material.[1] The numbers of various residues interacting with Trp are 302 (for Phe), 222 (Tyr), 93 (His) and 124 (Trp); the number in the last case is twice the number (62) of Trp–Trp pairs, as each interaction can be analyzed from the perspective of both residues. The packing of the partner rings against Trp can be seen in Fig. 2. The geometry of packing as measured by P and presented in Fig. 3 shows that there are variations in the most-preferred mode for different residues. When the partner is Phe, Tyr or Trp, the two rings are nearly perpendicular in about 50% of cases. His has the highest inclination to pack in a parallel fashion (31%), while Trp has the least (only 11%). The variation of P with θ is shown in Fig. 4 and the frequency of occurrence in the individual $30 \times 30°$ grids is presented in Table 1. The χ^2 values for each partner residue, 14.04, 13.86, 13.90 and 9.12 for Phe, Tyr, Trp and His, respectively, indicate the distributions to be different from a random distribution: when $\chi^2 = 9.50$ ($df = 4$), the probability of the distribution arising by chance is <5%.

Referring to Fig. 1(b) and Table 1, it can be said that while stacking (ff or of geometry) is avoided by Trp, it is preferred by His. With Phe or Trp as the partner, the interaction in the perpendicular fashion (fe) is greater than expected. For the former, a substantially higher population is also observed in the of geometry. Tyr stands apart from the other three aromatic residues as most of its points have $\theta > 40°$ (Fig. 4).

An idea of the vertical spacing (D_z) between the two rings as the partner moves over the face of Trp (variation in D_{xy} keeping the interplanar angle constant) can be obtained from Fig. 5; the values of D_{xy} and D_z in individual grids are given in Table 2. For the face-to-face stacked orientation (ff), the average value of D_z is 3.6 (2) Å for His [the value is slightly lower, 3.2 (3) Å, for the of geometry]. In comparison, for the Phe-Phe stacked pair the spacing is 3.1–3.4 Å (Gould et al.,

1985) and in B-DNA it is 3.4 (4) Å (Dickerson, 1983). The corresponding distances for the face-to-edge (fe) interaction involving His is 4.6 (2) Å, and the distance increases with the size of the partner ring.

An idea of the location of the partner centroid can be obtained from its projection on the Trp ring, measuring its distance from the x axis using the angle φ (Fig. 1a). The numbers of cases with negative and positive values of φ are 104 and 198, respectively, for Phe; 85 and 137, repectively, for Tyr; 43 and 81, respectively, for Trp and 38 and 55, respectively, for His. A reason for the significantly smaller number of occurrences with negative φ values could be a consequence of the smaller space available on the side of the five-membered ring resulting from the presence of the C^β atom on this side (Fig. 2).

4. Discussion

Studies using protein structures have brought out the salient features of aromatic–aromatic interaction. However, these have not touched on the effect of the difference in size and chemical nature of the interacting aromatic rings on their spatial disposition. For example, with dissimilar rings there are two distinct edge-to-face orientations (ef and fe in Fig. 1b) which may not be isoenergetic. The wealth of crystallographic data on organic solids cannot be used to address this question, as binary complexes between aromatic hydrocarbons are rarely encountered. Protein structures have a rich repertoire of four types of distinct aromatic residues (Phe, Tyr, His and Trp) and we have analyzed the geometry of packing of these residues against the indole ring of Trp. About 85% of the partner residues are more than four residues away from the Trp ring (see supplementary material). Consequently, the geometry observed is not dictated by the steric constraints of having the two residues close in sequence along the polypeptide chain and should reflect the inherent features of binding. Indeed, the results should complement the catalogue of non-bonded interactions involving heterocyclic ring systems

Figure 3
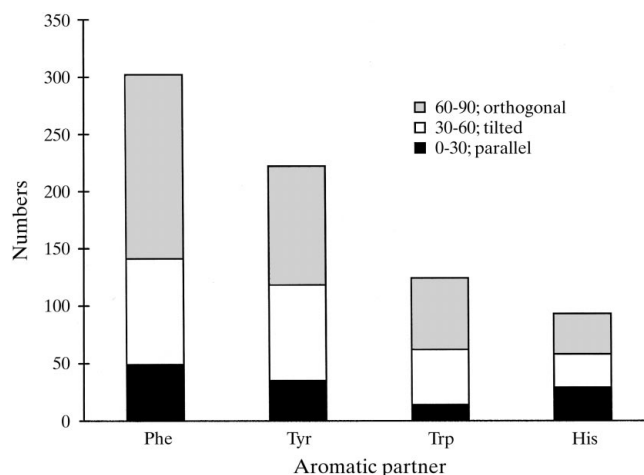Distribution of the angle P in the ranges 0–30, 30–60 and 60–90° among various aromatic partners.

---

[1] Supplementary material listing PDB file names, interacting aromatic residues and their sequence numbers, the closest contact distances and the atoms involved, and various geometrical parameters (including hydrogen bonding in the case of Tyr) describing the packing of the rings is available from the IUCr electronic archive (Reference: ad0074). Services for accessing these data are described at the back of the journal.

compiled in the IsoStar (1998) database for use in rational molecular design.

4.1. Phe

With two rings perpendicular to each other, two distinct orientations are possible for a Trp–Phe pair: Trp-face interacting with Phe-edge (*fe*) and Trp-edge with Phe-face (*ef*) (Fig. 1*b*) and the former geometry is expected to occur about 50% less often than the latter (Table 1) [if approximated by $\sin\theta$ (based on geometric considerations only), the expected occurrence should be even smaller]. However, *fe* is found to occur at a higher frequency (Fig. 2), suggesting that the binding energy between the rings in this relative orientation is greater than that provided by the *ef* geometry. Recent *ab initio* studies involving heteroaromatic ring systems have shown that a benzene molecule interacting in a perpendicular fashion with the π face of a pyridine molecule provides a binding energy which is even greater than that obtained when the former molecule is interacting with the edge and along the lone pair on the pyridine N atom (Samanta *et al.*, 1998). Another region in the P, θ plot which is distinctly populated has the *of* geometry (Fig. 4). This offset-stacked interaction has been shown to be attractive for the Phe-Phe pair (Hunter *et al.*, 1991).

4.2. Trp

For Trp−Trp interaction, the *ef* and *fe* orientations are equivalent. Most contacts have an orthogonal or tilted geometry. [59% of points are in grids corresponding to *fe*, *ef* and *et*. As has been mentioned in §3, for Trp–Trp interactions there are 62 pairs of geometries which are correlated. To verify whether any bias has been introduced by using all the 124 geometries, we randomly chose one geometry from each pair (10^5 times) and found that the average distribution in the above three grids remains the same, 59 ± 4%.] The stacking interaction is rare (Fig. 2); indeed, there are hardly any points within a radius of 50° in Fig. 4 or with values of D_z and D_{xy} within 5 Å (Fig. 5).

### 4.3. His

Of all the residues, His is the only one showing marked preference for the stacked arrangement (*ff* and *of*, Fig. 4), with the horizontal displacements between the two ring centres varying from ∼0.5 to ∼5 Å (Fig. 5). With a pK_a value of around 6.5, the imidazole ring of the His side chain is positively charged under physiological conditions and for such a ring the stacking interaction appears to be energetically favourable. A line of evidence supporting this comes from the study of the helical content of peptides in which the Trp–His pair was placed with three residues in between, so that on helix formation the rings can overlap; these peptides have the highest helical content when His is protonated (Fernandez-Recio *et al.*, 1997). Mutational studies have also shown that a Trp–charged His pair can stablize a protein by more than 4 kJ mol^{-1} (Loewenthal *et al.*, 1992).

Like His, hexafluorobenzene has also been shown to stack parallel with benzene and naphthalene (Laatikainen *et al.*, 1995; Coates *et al.*, 1998). The stacked geometry arises from the interaction between the highest occupied molecular orbital (HOMO) of the donor ring (Trp) and the lowest unoccupied molecular orbital (LUMO) of the acceptor ring (His) and, using an indole–adenine system, it has been shown that the protonation of the nucleic acid base lowers the LUMO energy, leading to a strengthening of the HOMO–LUMO mixing and thus a reinforcement of the stacking interaction (Ishida *et al.*, 1988). Parallel-stacked Trp–His appears to be a particularly resilient packing motif and a similar interaction between imidazole and adenine or guanine rings may also be important in the intercalation of imidazole-based drugs between DNA bases (Geierstanger & Wemmer, 1995; Trauger *et al.*, 1998). The only other grid with more than the expected occupancy is *ef* (Table 1), which suggests that contrary to what has been observed for the Trp–Phe pair, if the two rings are perpendicular, it is the edge of Trp which prefers to interact with the His face.

4.4. Tyr

Tyr provides an unique situation in which there are very few points (only 16.7%) with $\theta <$ 40° (Fig. 4). The *of* and *ef* geometries have more
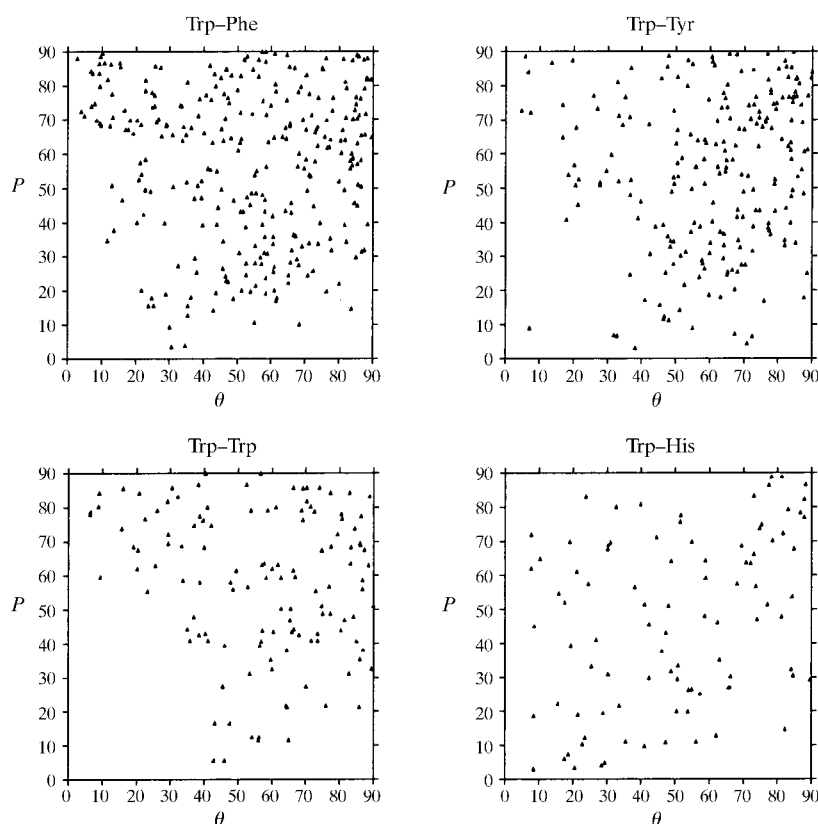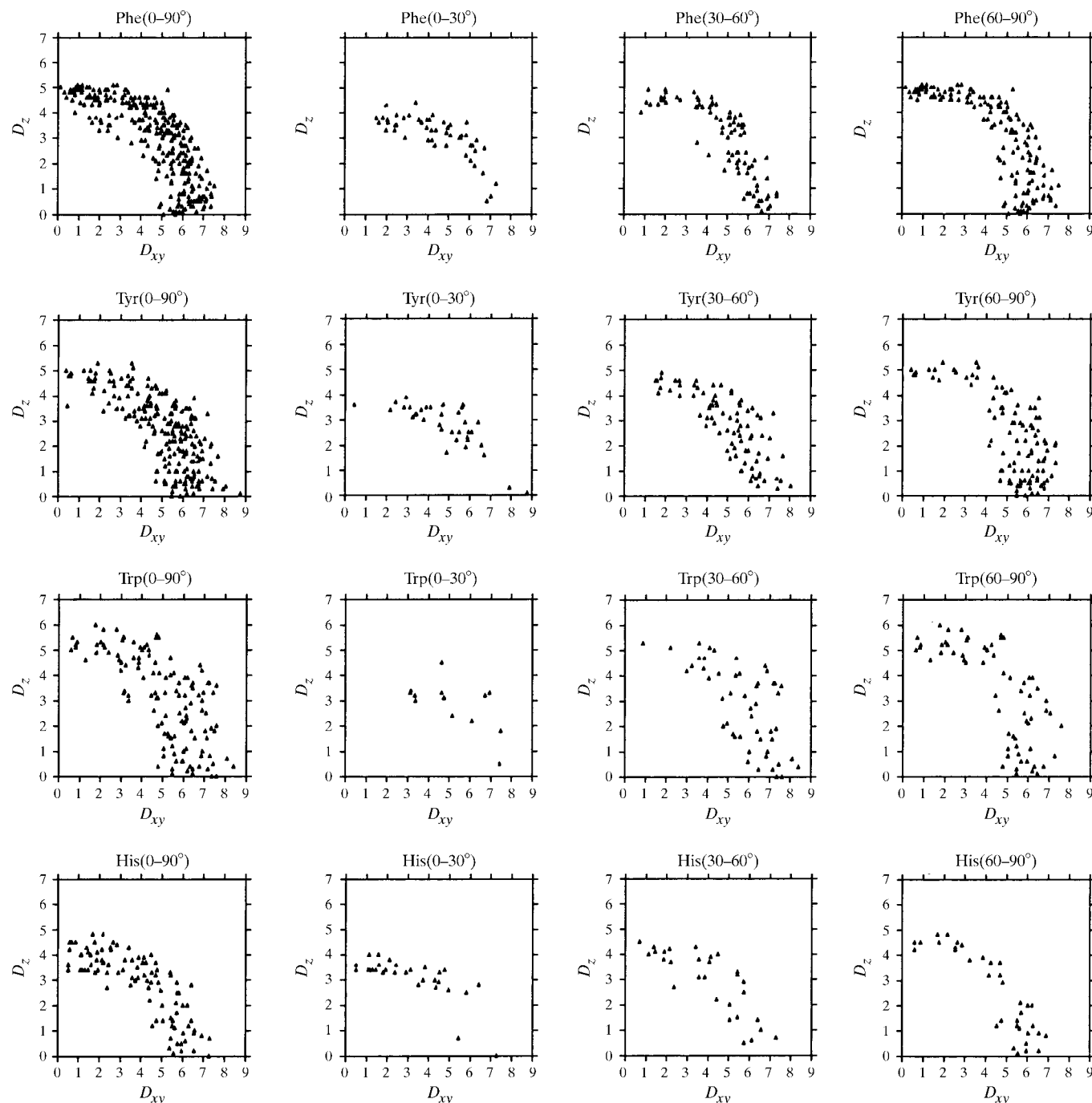

Figure 4
Relationship between the angles P and θ (in °).

than the expected numbers of points, whereas *oe* has less. It appears that the points are distributed about a line joining the *of* and *ef* grids. As a result, different Tyr rings, when superposed on the reference Trp ring (Fig. 2), cover the whole surface uniformly, whereas with Phe the points are less dense at Trp edge.

If the positively charged imidazolium ion is at one end of the chemical spectrum representing the aromatic residues, the electron-rich phenolic group is at the other. There are additional forces which are brought into play when Tyr is involved. The hydroxyl group is close to the π face of Trp in 22 cases, the mean of the distances from the nearest ring atom being 3.6 (2) Å (Table 2 and supplementary material Table S3). These may constitute what is known as the OH$-\pi$ interaction (although in the absence of hydroxyl hydrogen coordinates one cannot convincingly conclude whether or not the proton is pointing towards the ring), which is increasingly being identified in protein and small-molecule structures (Malone *et al.*,

Figure 5
Correspondence between the distances D_{xy} and D_z (in Å) for different pairs of residues, grouped depending on the range (full 0–90° and broken into ranges of 30°) of the angle P.

1997; Suzuki *et al.*, 1992). Additionally, in 16 cases the hydroxyl group forms a hydrogen bond with $N^{\varepsilon 1}$ of Trp at a distance of 3.0 (1) Å, and these cases are located in the three top right-hand grids.

5. Conclusions

Each of the four aromatic residues interacting with Trp has a signature motif (Table 1). Of all the packing modes, face-to-face stacking is the least observed, with the exception of His for which it is the most favoured. The parallel alignment could be a general feature of recognition of a protonated ring by a Trp residue. For the Phe–Phe pair the edge-to-face and offset-stacked geometries have been observed (Gould *et al.*, 1985; Hunter *et al.*, 1991). However, for the Trp–Trp pair analyzed here, only the former is prevalent. With dissimilar residues, there are two distinct edge-to-face possibilities; for the Trp–Phe pair, the edge of the latter interacting with the face of the former is more stable. For Tyr, the preferred mode is its face interacting with Trp-edge or Trp-face in an offset fashion, and a few cases of OH–π interaction can be observed. Besides providing stability to protein structures and being useful for recognizing molecules at some of their binding sites, the judicious placement of aromatic residues (so that they may interact favourably) should be an attractive tool in *de novo* protein design (Hill & DeGrado, 1998). The derived geometric parameters (Table 2 and Fig. 5) should be useful in modelling such interactions in proteins.

References

Artymiuk, P. J., Poirrette, A. R., Grindley, H. M., Rice, D. W. & Willett, P. (1994). *J. Mol. Biol.* **243**, 327–344.
Beeson, J. C., Fitzgerald, L. J., Gallucci, J. C., Gerkin, R. E., Rademacher, J. T. & Czarnik, A. W. (1994). *J. Am. Chem. Soc.* **116**, 4621–4622.
Benzing, T., Tjivikua, T., Wolfe, J. & Rebek, J. Jr (1988). *Science*, **242**, 266–268.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Brocchieri, L. & Karlin, S. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 9297–9301.
Burley, S. K. & Petsko, G. A. (1985). *Science*, **229**, 23–28.
Burley, S. K. & Petsko, G. A. (1986). *J. Am. Chem. Soc.* **108**, 7995–8001.
Chilkoti, A., Tan, P. H. & Stayton, P. S. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 1754–1758.
Coates, G. W., Dunn, A. R., Henling, L. M., Ziller, J. W., Lobkovsky, E. B. & Grubbs, R. H. (1998). *J. Am. Chem. Soc.* **120**, 3641–3649.
Cochran, J. E., Parrott, T. J., Whitlock, B. J. & Whitlock, H. W. (1992). *J. Am. Chem. Soc.* **114**, 2269–2270.
Cox, E. G., Cruickshank, D. W. J. & Smith, J. A. S. (1958). *Proc. R. Soc. London Ser. A*, **247**, 1–21.

Dahl, T. (1994). *Acta Chem. Scand.* **48**, 95–106.
Desiraju, G. R. & Gavezzotti, A. (1989). *J. Chem. Soc. Chem. Commun.* pp. 621–623.
Dickerson, R. E. (1983). *Sci. Am.* **249**, 94–111.
Ferguson, S. B., Sanford, E. M., Seward, E. M. & Diederich, F. (1991). *J. Am. Chem. Soc.* **113**, 5410–5419.
Fernandez-Recio, J., Vazquez, A., Civera, C., Sevilla, P. & Sancho, J. (1997). *J. Mol. Biol.* **267**, 184–197.
Geierstanger, B. H. & Wemmer, D. E. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 463–493.
Gould, R. O., Gray, A. M., Taylor, P. & Walkinshaw, M. D. (1985). *J. Am. Chem. Soc.* **107**, 5921–5927.
Grossel, M. C., Cheetham, A. K., Hope, D. A. O. & Weston, S. C. (1993). *J. Org. Chem.* **58**, 6654–6661.
Hill, R. B. & DeGrado, W. F. (1998). *J. Am. Chem. Soc.* **120**, 1138–1145.
Hobohm, U. & Sander, C. (1994). *Protein Sci.* **3**, 522–524.
Hobza, P., Selzle, H. L. & Schlag, E. W. (1994). *J. Am. Chem. Soc.* **116**, 3500–3506.
Hunter, C. A. (1991). *J. Chem. Soc. Chem. Commun.* pp. 749–751.
Hunter, C. A. (1994). *Chem. Soc. Rev.* **23**, 101–109.
Hunter, C. A. & Sanders, J. K. M. (1990). *J. Am. Chem. Soc.* **112**, 5525–5534.
Hunter, C. A., Singh, J. & Thornton, J. M. (1991). *J. Mol. Biol.* **218**, 837–846.
Ishida, T., Doi, M., Ueda, H., Inoue, M. & Scheldrick, G. M. (1988). *J. Am. Chem. Soc.* **110**, 2286–2294.
IsoStar (1998). *Library of Intermolecular Interactions, Version* 1.1, Cambridge Crystallographic Data Centre.
Jorgensen, W. L. & Severance, D. L. (1990). *J. Am. Chem. Soc.* **112**, 4768–4774.
Kodandapani, R., Pio, F., Ni, C.-Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R. A. & Ely, K. R. (1996). *Nature (London)*, **380**, 456–460.
Laatikainen, R., Ratilainen, J., Sebastian, R. & Santa, H. (1995). *J. Am. Chem. Soc.* **117**, 11006–11010.
Linse, P. (1992). *J. Am. Chem. Soc.* **114**, 4366–4373.
Loewenthal, R., Sancho, J. & Fersht, A. R. (1992). *J. Mol. Biol.* **224**, 759–770.
Maenaka, K., Matsushima, M., Sunada, H. S. F., Watanabe, K. & Kumagai, I. (1995). *J. Mol. Biol.* **247**, 281–293.
Malone, J. F., Murray, C. M., Charlton, M. H., Docherty, R. & Lavery, A. J. (1997). *J. Chem. Soc. Faraday Trans.* **93**, 3429–3436.
Moody, G. J., Owusu, R. K., Slawin, A. M. Z., Spencer, N., Stoddart, J. F., Thomas, J. D. R. & Williams, D. J. (1987). *Angew. Chem. Int. Ed. Engl.* **26**, 890–892.
Muehldorf, A. V., Van Engen, D., Warner, J. C. & Hamilton, A. D. (1988). *J. Am. Chem. Soc.* **110**, 6561–6562.
Newcomb, L. F. & Gellman, S. H. (1994). *J. Am. Chem. Soc.* **116**, 4993–4994.
Paliwal, S., Geib, S. & Wilcox, C. S. (1994). *J. Am. Chem. Soc.* **116**, 4497–4498.
Samanta, U., Chandrasekhar, J. & Chakrabarti, P. (1998). *J. Phys. Chem. A*, **102**, 8964–8969.
Schiffer, M., Chang, C.-H. & Stevens, F. J. (1992). *Protein Eng.* **5**, 213–214.
Seel, C. & Vögtle, F. (1992). *Angew. Chem. Int. Ed. Engl.* **31**, 528–549.
Singh, J. & Thornton, J. M. (1985). *FEBS Lett.* **191**, 1–6.
Suzuki, S., Green, P. G., Bumagarner, R. E., Dasgupta, S., Goddard, W. A. & Blake, G. A. (1992). *Science*, **257**, 942–945.
Trauger, J. W., Baird, E. E. & Dervan, P. B. (1998). *J. Am. Chem. Soc.* **120**, 3534–3535.
Zhang, J. & Moore, J. S. (1992). *J. Am. Chem. Soc.* **114**, 9701–9702.
Zimmerman, S. C., Wu, W. & Zeng, Z. (1991). *J. Am. Chem. Soc.* **113**, 196–201.