PROTEIN CHEMISTRY AND STRUCTURE:

π-Stacking Interactions: ALIVE AND WELL IN PROTEINS

Georgia B. McGaughey, Marc Gagné and Anthony K. Rappé

J. Biol. Chem. 1998, 273:15458-15463.
doi: 10.1074/jbc.273.25.15458

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 25, Issue of June 19, pp. 15458–15463, 1998

Printed in U.S.A.

π-Stacking Interactions

ALIVE AND WELL IN PROTEINS*

(Received for publication, January 26, 1998)

Georgia B. McGaughey‡, Marc Gagné§, and Anthony K. Rappé¶‖

From the ‡Wyeth-Ayerst Research, Structural Biology, Princeton, New Jersey 08540, §Joint Institute for Laboratory Astrophysics, University of Colorado, Boulder, Colorado 80309-0440, and ¶Colorado State University, Department of Chemistry, Fort Collins, Colorado 80523

A representative set of high resolution x-ray crystal structures of nonhomologous proteins have been examined to determine the preferred positions and orientations of noncovalent interactions between the aromatic side chains of the amino acids phenylalanine, tyrosine, histidine, and tryptophan. To study the primary interactions between aromatic amino acids, care has been taken to examine only isolated pairs (dimers) of amino acids because trimers and higher order clusters of aromatic amino acids behave differently than their dimer counterparts. We find that pairs (dimers) of aromatic side chain amino acids preferentially align their respective aromatic rings in an off-centered parallel orientation. Further, we find that this parallel-displaced structure is 0.5–0.75 kcal/mol more stable than a T-shaped structure for phenylalanine interactions and 1 kcal/mol more stable than a T-shaped structure for the full set of aromatic side chain amino acids. This experimentally determined structure and energy difference is consistent with *ab initio* and molecular mechanics calculations of benzene dimer, however, the results are not in agreement with previously published analyses of aromatic amino acids in proteins. The preferred orientation is referred to as parallel displaced π-stacking.

Attractive nonbonded interactions between aromatic rings are seen in many areas of chemistry, and hence are of interest to all realms of chemistry. Porphyrin aggregation (1), the conformation of diarylnaphthalenes (2) and phenylacetylene macrocycles (3), and the strength of Kevlar (4) can be attributed, at least in part, to aromatic-aromatic interactions. Aromatic-aromatic interactions have been implicated in catalytic hydroformylation (5), the catalytic formation of elastomeric polypropylene (6), and the asymmetric *cis* dihydroxylation of olefins (7). The vast majority of medicinal agents contain aromatic substituents and their differential recognition by proteins is likely dominated by aromatic-aromatic interactions (8). In biologically related areas of chemistry, aromatic-aromatic interactions are crucially involved in protein-deoxynucleic acid complexes where interactions between aromatic residues and base pairs are seen in x-ray crystal structures (9, 10).

Because aromatic-aromatic interactions are so prevalent across chemistry, a large body of experimental and theoretical work has focused on determining the gas phase structure of the prototype, benzene dimer (11–14, 30). As summarized recently by Sun and Bernstein (15), the experimentally observed structure depends heavily upon the observation technique. Off-centered parallel displaced, **1p**, and T-shaped, **1t**, structures are the most commonly cited orientations (Structure 1).

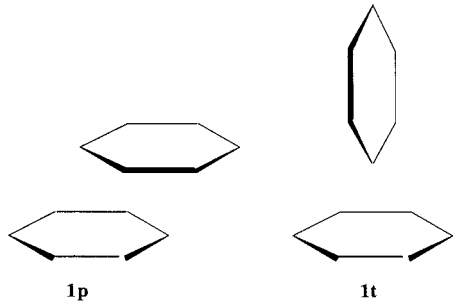
Large scale *ab initio* electronic structure theory suggests that the off-centered parallel displaced and T-shaped structures are nearly isoenergetic (11–14, 30). As reported by Sun and Bernstein (15), empirical force field studies favor either off-centered parallel displaced or T-shaped structures depending upon the magnitude of the partial charges ($q_H = -q_C$) used in the electrostatic model. Small charges (<0.153) favor parallel displaced geometries; large partial charges (>0.3) favor T-shaped structures. Sun and Bernstein (15) suggest further that the intermolecular potential surface is quite soft and that "one must view the dimer as a dynamic system rather than one with a well defined structure."

It is our thesis that in the hydrophobic core of a protein, in the solid state, the dynamical properties of benzene-benzene are quenched and a preferred structure does prevail, albeit the preferred structure of benzene dimer in a hydrophobic environment. The methodologies used are summarized under "Materials and Methods." Results and discussion are provided under "Results," and conclusions are drawn under "Discussion."

MATERIALS AND METHODS

Brookhaven Protein Data Bank—To determine the nature of aromatic-aromatic interactions in the hydrophobic cores of proteins the Brookhaven Protein Data Bank has been analyzed. A previously defined (16, 17) representative subset of proteins containing only nonhomologous proteins and only proteins with high x-ray crystallographic resolution (18, 19) was used. The subset contained 505 proteins. Noncovalent interactions between the side chains of the aromatic amino acids phenylalanine (Phe), tyrosine (Tyr), histidine (His), and tryptophan (Trp) amino acids were examined.

To distinguish between configurations **1p** and **1t** the relative orientations of the aromatic side chains need to be cataloged. The shape of axially symmetric aromatic rings can most naturally be represented in terms of the center of mass of the ring, the ring centroid, and the unique axis perpendicular to the ring plane, the surface normal vector (see Fig. 1*a*). The intermolecular orientational information of one aromatic ring



STRUCTURE 1

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed: Dept. of Chemistry, Colorado State University, Fort Collins, CO 80523. Tel.: 970-491-6292; Fax: 970-491-1801; E-mail: trappe@mail.chm.colostate.edu.

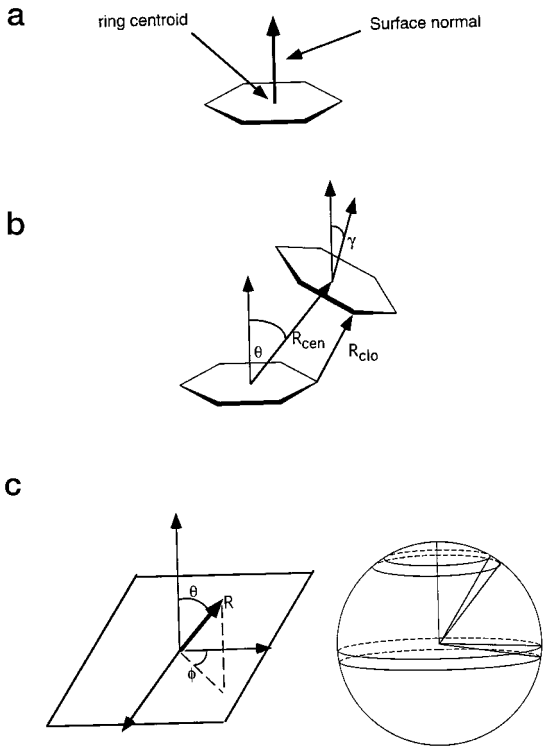


FIG. 1. *a,* essential structural features of axially symmetric systems such as benzene. *b*, spherical polar pair orientational coordinates. *c*, Euler angle pair orientational coordinates and θ, γ unit surface area near $\theta < 30°$ compared with θ, γ unit surface area near $\theta = 90°$.

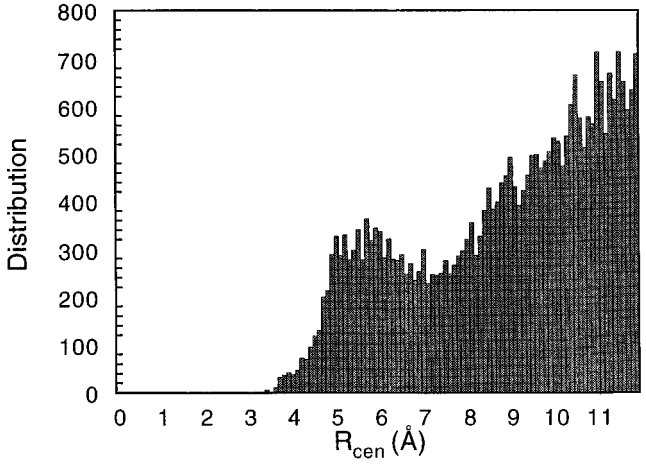


FIG. 2. **R^{cen} distribution for 30,444 pairs of aromatic-aromatic amino acid side chains.** $R_{cen} < 12.0$ Å.

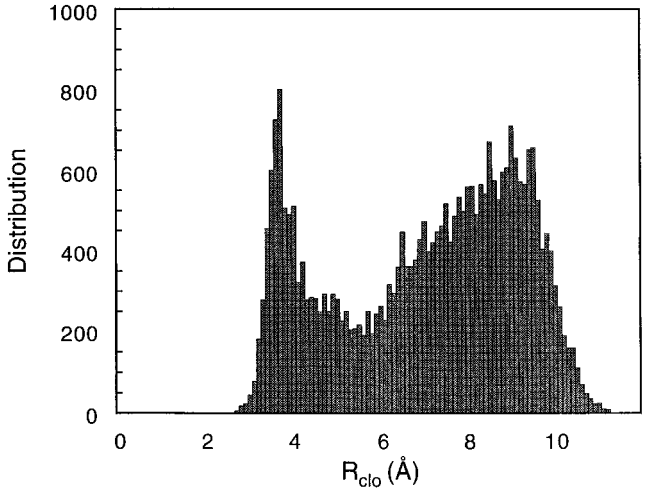


FIG. 3. **R_{clo} distribution for 30,444 pairs of aromatic-aromatic amino acid side chains.** $R_{cen} < 12.0$ Å.

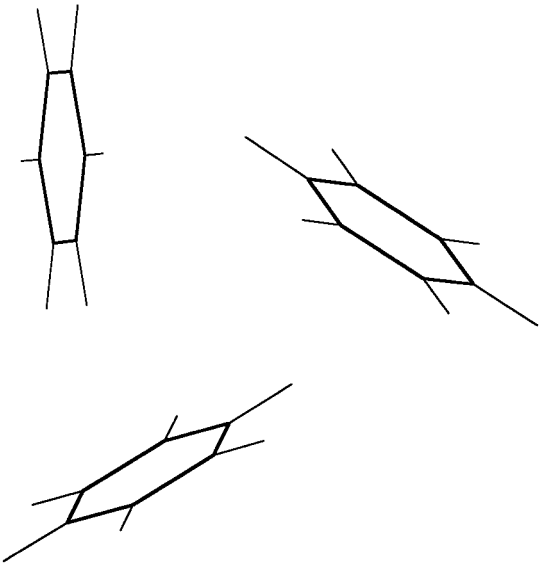


FIG. 4. **RFF1 optimized benzene trimers.**

with respect to another, the pair orientation, is described by the centroid-centroid separation, R_{cen}, a center-normal angle, θ, and a normal-normal angle, γ (see Fig. 1*b*). The angles, θ and γ, correspond to solid body azimuthal angle rotation and Euler angle yaw, respectively (21).

Pairs of aromatic residues were identified based on $R_{cen} < 12.0$ Å. For both Phe and Tyr, the six carbons constituting the phenyl ring were used to determine the centroid; in Trp, only the five atoms in the five-membered portion of the indole ring were used, and the five atoms in the imidazole ring of His were used. A total of 30,444 centroid matches were found for all possible combinations of Phe, Tyr, Trp, and His (*e.g.* Phe-Phe, Phe-Tyr, Phe-Trp, Phe-His, etc.). The number distribution of R_{cen} values, shown in Fig. 2, was found to be bimodal with a minimum at ~7.5 Å.

In addition, for each pair of aromatic residues ($R_{cen} < 12.0$ Å), closest contact distances (R_{clo}) between the respective carbon and nitrogen atoms were calculated. The number distribution of R_{clo} values was found to also be bimodal with a prominent minimum between 4.5–5 Å (see Fig. 3). We interpret the minima in R_{cen} and R_{clo} distributions as representing the distance at which the interaction between the aromatic rings drops below the Boltzman temperature factor (~0.6 kcal/mol at 300 K). Inside the minimum in the distribution there is a binding interaction between the rings; outside the minimum any direct ring-ring interaction is lost because of random thermal motion. Thus, residue pairs with $R_{clo} < 4.5$ Å and/or $R_{cen} < 7.5$Å contain information about the pair orientation preferences of aromatic side chains. A total of 1,682 aromatic-aromatic amino acid dimer pairs with R_{clo} values less than 4.5 Å were found with 13% below 3.4 Å, the minimum value of the interatomic distance between two aromatic rings (20). R_{cen}, θ, and γ were determined for these pairs.

As is true for any axially symmetric system, the probability distribution of solid angle space is asymmetric in both θ and γ. The source of this asymmetry is shown in Fig. 1*c* for the spherical polar angle θ. θ is naturally peaked around 90° because of an increase in meridianal angle ϕ space as θ progesses from 0° to 90°. This leads to an increase in surface area and thus an increase in probability of occurrence. To properly identify the intrinsic energetic preferences of aromatic-aromatic interactions the angle probability distributions discussed below have been normalized so that a distribution without an angular energetic preference would appear flat.

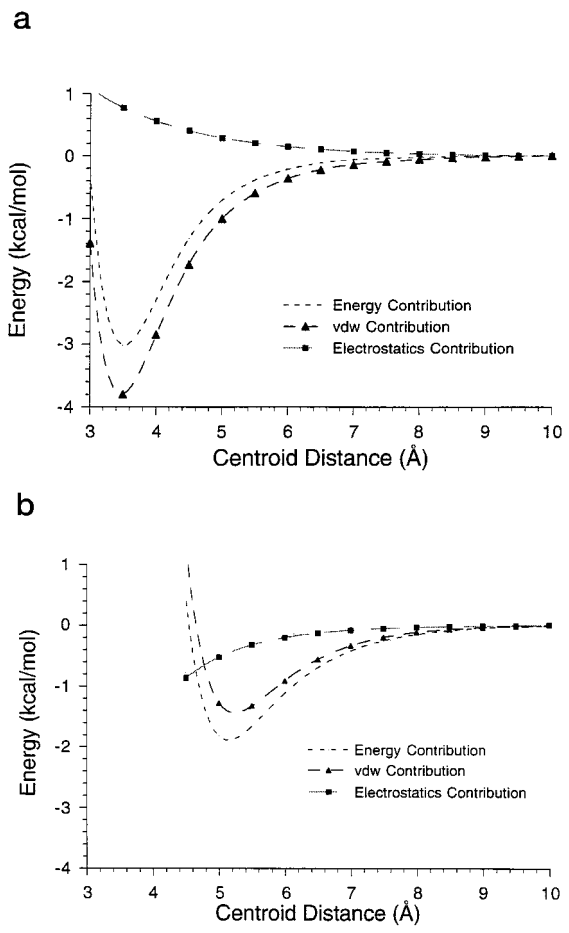


FIG. 5. *a*, total energy, electrostatic, and van der Waals potential surfaces for parallel-stacked benzene dimer. *b*, total energy, electrostatic, and van der Waals potential surfaces for T-shaped benzene dimer.

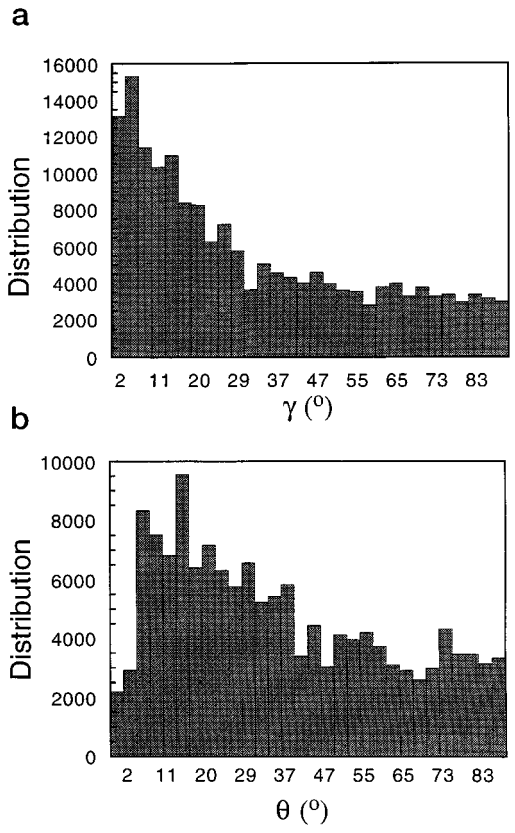


FIG. 6. *a*, γ angle distribution of 1,682 dimer clusters of aromatic-aromatic amino acid side chains. *b*, θ angle distribution of 1,682 dimer clusters of aromatic-aromatic amino acid side chains.

Molecular Mechanics Studies—Another complicating factor in the analysis of the structural preferences of aromatic side chains is that aromatic molecules tend to form higher order clusters. In these clusters intra-cluster orientation is dictated by the cluster rather than discrete pairwise interactions. For example, three spheres can pack with the pair distances retained but three disks will adopt a pinwheel arrangement to maximize the individual interactions. This pinwheel arrangement was confirmed for benzene trimer by molecular mechanics (RFF1) (see Fig. 4). To ascertain the importance of higher order clustering effects, isolated dimers and isolated trimers need to be and have been analyzed separately in the present study.

Further, because the vast majority of aromatic side chain residues are in the hydrophobic interior of proteins, the isolated dimers are present in a hydrophobic sea. The structural impact of this hydrophobic sea was investigated by placing a parallel-displaced dimer in a droplet of methane. The RFF1 isolated dimer and methane droplet structures were virtually identical.

For reference, the RFF1 parallel-displaced and T-shaped binding energies of 2.75 and 1.95 kcal/mol are in reasonable accord with the MM3(95)[1] values of 2.57 and 1.88 kcal/mol, respectively.

To further characterize the nature of the nonbonded interaction between benzene rings and to find out when the "bond" between them drops below the Boltzman temperature factor, parallel-stacked and T-shaped potential energy surfaces were constructed by incrementally increasing the centroid distance by 0.5 Å starting at a R_{cen} of 3.5 Å and stopping at a R_{cen} of 10.0 Å. The binding energies were determined at each centroid distance and the van der Waals and electrostatics contributions to the binding energy are plotted in Fig. 5. In the parallel-stacked case, the van der Waals contribution is the dominating effect and the electrostatics contribution is actually repulsive, although small (<1 kcal/mol). On the other hand, the van der Waals contribution in the T-shaped case is not overwhelming, and it is the attractive electrostatics contribution that results in the overall binding of ~2.0 kcal/mol. Significantly, for both parallel-stacked and T-shaped structures the binding energy drops below the Boltzman temperature factor (0.592 kcal/mol at 300 K) at roughly 7.5Å.

RESULTS

Population distributions for the inter-ring orientational angles θ and γ shown in Figs. 6–8 were generated considering only dimers of aromatic side chains and correcting for spherical polar and Euler angle probability bias. If there was no intrinsic angular energetic preference, the profiles in Figs. 6–8 would appear flat; instead the θ distribution (Fig. 6*b*) has a peak near 30° and the γ distributions (Figs. 6*a* and 7) have peaks around 0°. This combination of θ and γ, determined from experimental data, corresponds to an off-centered parallel configuration, in accord with most *ab initio* and empirical force-field structural estimates of gas phase benzene dimer. To directly compare the preferred conformation of aromatic amino acids in the Protein Data Bank with the *ab initio* and molecular mechanics results of benzene dimer, we focus on the γ distribution for Phe-Phe interactions. The distribution for Phe-Phe interactions (Fig. 7*b*) is less peaked than the distribution for all aromatic side chains (Figs. 6*a* and 7*a*). Further, because there is six times less Phe-Phe data there is more scatter in the plot. For both Phe-Phe interactions and the full dataset the shape of the γ distribution can be fit to a Boltzmann distribution assuming 1) the parallel-displaced structure is more stable that the T-shaped

[1] The program MM3(94) is available from the Quantum Chemistry Program Exchange, Dept. of Chemistry, University of Indiana, Bloomington, IN 47401 and from Tripos Associates, Inc., 1699 South Hanley Rd., Suite 303, St. Louis, MO 63144.

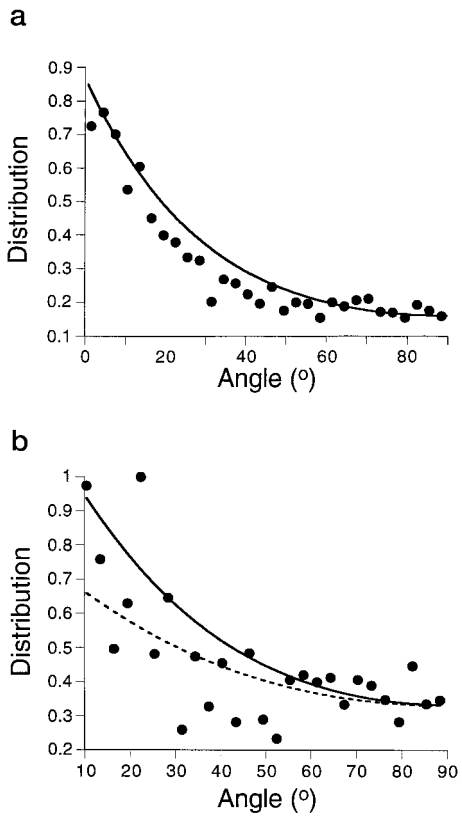


FIG. 7. *a*, γ angle distribution of all dimer cluster pairs; a Boltzmann distribution with an energy factor of 1 kcal/mol is superimposed over the experimental data. *b*, γ angle distribution of dimer clusters of Phe-Phe pairs; Boltzmann distributions with energy factors of 0.5 kcal/mol (*dashed line*) and 0.75 kcal/mol (*solid line*) are superimposed over the experimental data.

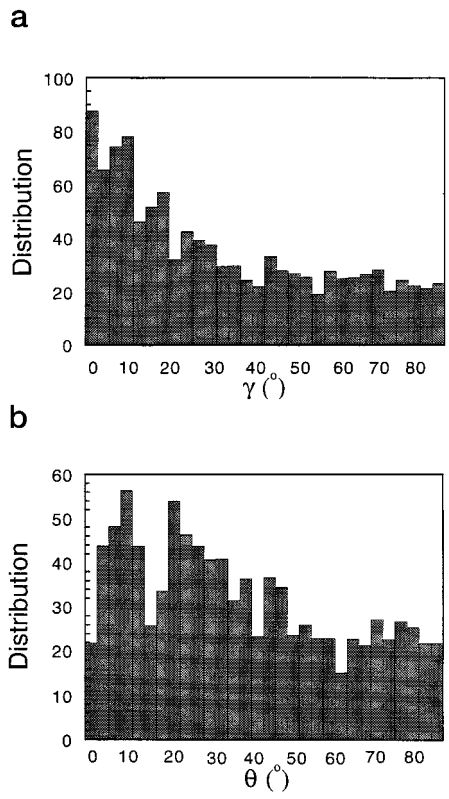


FIG. 8. *a*, γ angle distribution of trimer clusters of 1,144 aromatic-aromatic amino acid side chains. *b*, θ angle distribution of 1,144 trimer clusters of aromatic-aromatic amino acid side chains.

structure, 2) that the energy difference has a sin γ dependence, and 3) that the temperature is 300 K. For the full dataset the parallel-displaced structure is found to be more stable by 1.0 kcal/mol, as indicated by the *solid line* in Fig. 7*a*. For the Phe-Phe pairs the parallel-displaced structure is found to be more stable by 0.5–0.75 kcal/mol. The 0.5–0.75 kcal/mol distributions are shown as *dashed* and *solid lines*, respectively, in Fig. 7*b*. The Phe-Phe energy difference is consistent with *ab initio* electronic structure as well as molecular mechanics estimates of the energy difference.

In Fig. 8 we show that the orientational effects of π-stacking are less apparent in the probability distribution of trimers. Normalized θ and γ distributions for aromatic side chain amino acid pairs with $R_{clo} < 7.5$ Å do not show as pronounced peaks for trimers as for dimers. The peak in θ of ~25°, which was seen in the dimers, does not manifest itself as clearly in the trimer clusters. Rather, there appears to be two peaks between a θ value of 10° and 30° as opposed to one distinct value of 20° seen in the dimer cluster.

Normalized θ and γ distributions from the full set of data for aromatic-aromatic amino acid pairs with $R_{clo} > 7.5$ Å do not show any pronounced peaks. As discussed above, this is presumably because of the absence of thermally significant binding at this large distance.

When homopairs of side chains where examined, 4,716 matches were found for Phe-Phe, 3,050 for Tyr-Tyr, 1,124 for His-His, and 688 for Trp-Trp. The only structural difference between Phe and Tyr is the presence of the *para*-OH on Tyr. A plot of the centroid distance *versus* the closest contact distance represents this finding (Fig. 9). Because Phe lacks the *para*-OH, there are a greater number of centroid contacts found that are less than 6.5 Å the minimum in the Phe-Phe plot; 4,716 Phe-Phe interactions less than 12.0 Å were tabulated, and 1,226 or 26% pair orientations were found less than 6.5 Å. In contrast, only 3,050 Tyr-Tyr interactions less than 12.0 Å were found and 556 or 18% pair orientations existed less than 6.5 Å. This difference can be seen by comparing Fig. 9*a* with 9*c* and Fig. 9*b* with 9*d*. Even though in the parallel shaped Tyr dimer there are no steric interactions inhibiting a R_{clo} between the two amino acids, in other orientations (such as the T-shaped), the *para*-OH does reduce the number of R_{clo}.

Additional information is found in plots of intermolecular distances (R_{cen}, R_{clo}) *versus* the interplanar angle, γ. If R_{clo} is plotted *versus* the interplanar angle, γ, a near constancy in R_{clo} is found (see Fig. 10). Regardless of angle, the aromatic side chains orient in a fashion to minimize R_{clo} between the two rings and thus maximize the van der Waals attraction. Further, as also shown in Fig. 2, the number density drops off at a R_{clo} of ~4.5Å. When R_{cen} is plotted *versus* the interplanar angle, γ, the bottom of the distribution is linearly dependent upon angle (Fig. 11). This is because of parallel orientations (small γ) having shorter R_{cen} than T-shaped orientations.

DISCUSSION

By using a nonhomologous set of proteins, correcting for probability distribution bias, and including only isolated dimer pairs we find aromatic side chain amino acids do have a preferred intermolecular structure. The preferred parallel-displaced orientation is found to be more stable than a T-shaped structure by 0.5–0.75 kcal/mol for Phe-Phe dimers and by 1.0 kcal/mol for the full set of dimers.

Other authors (22–29) have examined the orientation be-

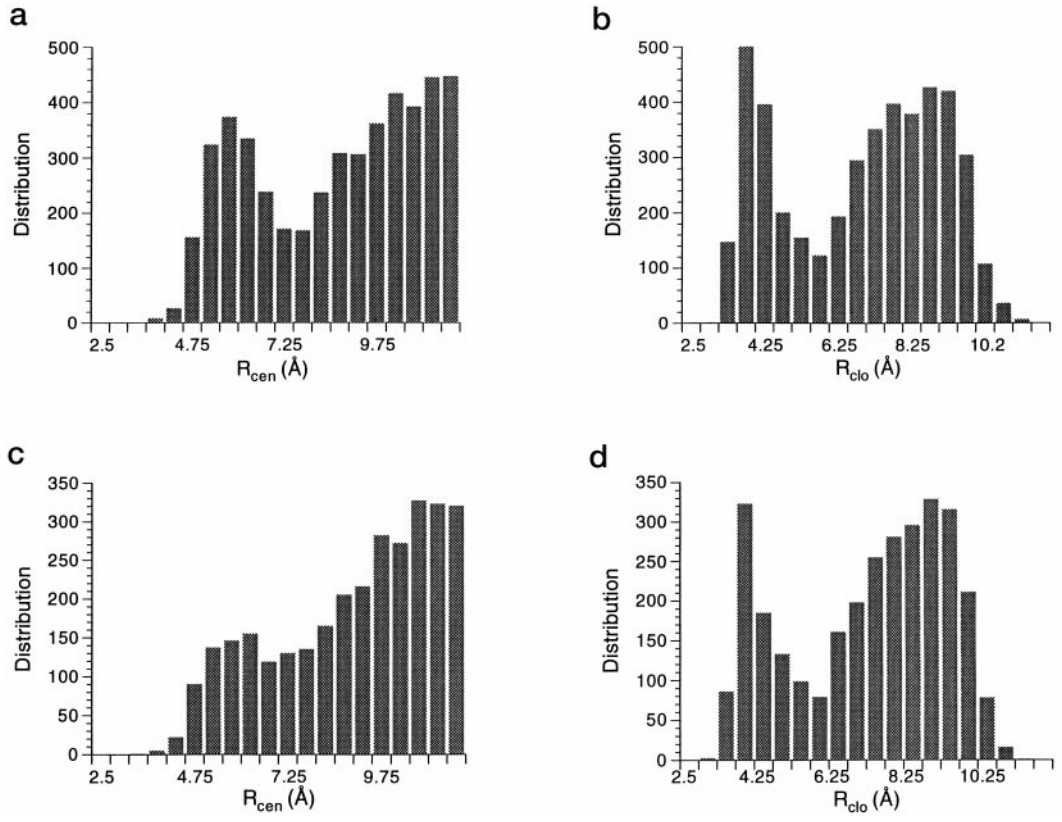


FIG. 9. *a*, R_{cen} distribution for 4,716 Phe-Phe pairs. *b*, R_{clo} distribution for 4,716 Phe-Phe pairs. *c*, R_{cen} distribution for 3,050 Tyr-Tyr pairs. *d*, R_{clo} distribution for 3,050 Tyr-Tyr pairs.

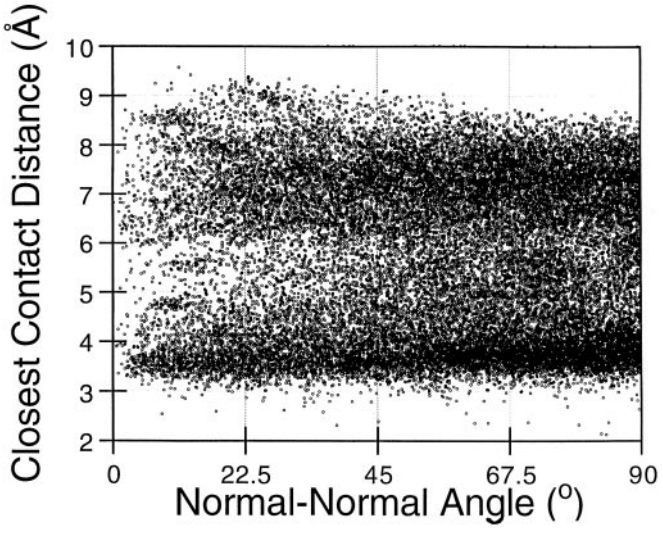


FIG. 10. **Plot of intermolecular distance, R_{cen}, *versus* interplanar angle, γ.**

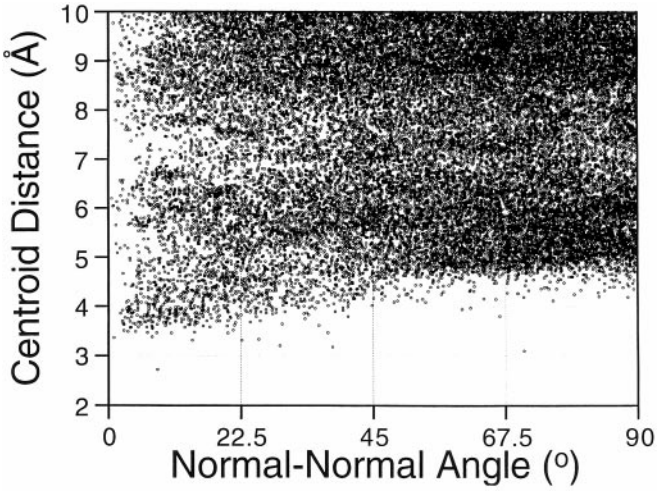


FIG. 11. **Plot of intermolecular distance, R_{clo}, *versus* interplanar angle, γ-.**

tween aromatic-aromatic side chain amino acids. They suggest that the majority of aromatic-aromatic interactions can be attributed to T-shaped configurations and that parallel displaced orientations are not generally found in proteins, in contrast to the present study. As has been pointed out by Thornton *et al.* (27) this may largely be a result of neglecting the inherent bias in the probability distribution of angles. The present study uses a more extensive, more representative sample of nonhomologous proteins than previous investigations. Moreover, clustering appears to dilute the effect of π-stacking. Future studies will focus on the role of π-stacking in determining tertiary structure and its possible impact on structure-based drug design.

REFERENCES

1. Abraham, R. J., Eivazi, F., Pearson, H., and Smith, K. M. (1976) *J. Chem. Soc. Chem. Comm.* 699–701
2. Cozzi, F., Cinquini, M., Annuziata, R., and Siegel, J. S. (1993) *J. Am. Chem. Soc.* **115,** 5330–5331
3. Shetty, A. S., Zhang, J., and Moore, J. S. (1996) *J. Am. Chem. Soc.* **118,** 1019–1027
4. Tanner, D., Fitzgerald, J. A., and Phillips, B. R. (1989) *Prog. Rubber Plast. Technol.* **5,** 229–251
5. Castonguay, L. A., Rappé, A. K., and Casewit, C. J. (1991) *J. Am. Chem. Soc.* **113,** 7177–7183
6. Pietsch, M. A., and Rappé, A. K. (1996) *J. Am. Chem. Soc.* **118,** 10908–10909

7. Kolb, H. C., Andersson, P. G., and Sharpless, K. B. (1994) *J. Am. Chem. Soc.* **116,** 1278–1291
8. Gilman, A. G., Rall, T. W., Mies, A. S., and Taylor, P. (1993) *The Pharmaceutical Basis of Therapeutics*, 8th Ed., McGraw Hill, Inc., New York
9. Ishida, T., Doi, M., Ueda, H., Inoue, M., and Scheldrick, B. M. (1988) *J. Am. Chem. Soc.* **110,** 2286–2294
10. Kamiichi, K., Danshita, M., Minamino, N., Doi, M., Ishidia, T., and Inoue, M. (1986) *FEBS Lett.* **195,** 57–60
11. Hobza, P., Selzle, H. L., and Schlag, E. W. (1994) *J. Am. Chem. Soc.* **116,** 3500–3506
12. Tsuzuki, S., Uchimaru, T., Mikami, M., and Tanabe, K. (1996) *Chem. Phys. Lett.* **252,** 206–210
13. Jaffe, R. L., and Smith, G. D. (1996) *J. Chem. Phys.* **105,** 2780–2788
14. Chipot, C., Jaffe, R., Maigret, B., Pearlman, D. A., and Kollman, P. A. (1996) *J. Am. Chem. Soc.* **118,** 11217–11224
15. Bernstein, E. R., and Sun, S. (1996) *J. Phys. Chem.* **100,** 13348–13366
16. Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., and Weng, J. (1987) in *Protein Data Bank. Crystallographic Databases-Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., and Sievers, R., eds) Data Commission of the International Union of Crystallography, Bonn, Cambridge, Chester, pp. 107–132, 29 May 1996
17. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Mayer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasuri, M. (1977) J. Mol. Biol. **112,** 535–542
18. Hobohm, U., Scharf, M., Schneider, R., and Sander, C. (1992) *Protein Sci.* **1,** 409–417
19. Hobohm, U., and Sander, C. (1994) *Protein Sci.* **3,** 522–524
20. Hunter, C. A., Singh, J., and Thornton, J. M. (1991) *J. Mol. Biol.* **218,** 837–846
21. Fan, C. F., Olafson, B. D., Blanco, M., and Hsu, S. L. (1992) *Macromolecules* **25,** 3667–3676
22. Blundell, T., Singh, J., Thornton, J., Burley, S. K., and Petsko, G. A.. (1986) *Science* **234,** 1005
23. Hunter, C. A., and Sanders, J. K. M. (1990) *J. Am. Chem. Soc.* **112,** 5525–5534
24. Burley, S. K., and Petsko, G. A. (1985) *Science* **229,** 23–28
25. Burley, S. K., and Petsko, G. A. (1986) *J. Am. Chem. Soc.* **108,** 7995–8001
26. Hunter, C. A. (1993) *Philos. Trans. R. Soc. Lond. A Math. Phys. Sci.* **345,** 77–85
27. Thornton, J. M., Singh, J., Campbell, S., and Blundell, T. L. (1988) *Biochem. Soc. Trans.* **16,** 927–930
28. Singh, J., and Thornton, J. M. (1985) *FEBS Lett.* **191,** 1–6
29. Singh, J., and Thornton, J. M. (1992) *Atlas of Protein Side-Chain Interactions*, Vol. 1–2, Oirl Press at Oxford University Press, Oxford, UK
30. Hobza, P., Selzle, H. L., and Schlag, E. W. (1996) *J. Phys. Chem.* **100,** 18790–18794