Journal of Computer-Aided Molecular Design, 3 (1989) 67–84
ESCOM

J-CAMD 038

Geometries of functional group interactions in enzyme–ligand complexes: Guides for receptor modelling

Marina Tintelnot and Peter Andrews*

School of Pharmaceutical Chemistry, Victorian College of Pharmacy Ltd, 381 Royal Parade, Parkville, Victoria, Australia 3052

Received 16 May 1988
Accepted 29 October 1988

Key words: Receptor modelling; Enzyme–ligand complexes; X-ray data; Functional group orientations; Geometry of intermolecular interactions

SUMMARY

An approach is described which makes use of X-ray structural data from enzyme–ligand complexes in order to obtain information for application in receptor modelling. The atomic surroundings of five different ligand functional groups were determined for all complex structures recorded in the Brookhaven Protein Data Bank. These atomic surroundings were then superimposed with respect to the atoms of the functional groups of the ligands in order to obtain clouds of neighbouring atoms. General principles were sought to describe the orientiation or favoured position of groups or atoms around those functional groups when bound to a macromolecule. Some simple conclusions and leads for further modelling were thus derived.

INTRODUCTION

Knowledge about drug receptors at a molecular level is still limited. X-ray structures are as yet unobtainable, and receptors are generally described in terms of their function rather than their structural characteristics. Hence, approaches to receptor modelling based on structure–activity relationships have been developed by various research groups [1].

In our own work, we have built receptor models by looking at a series of related drug molecules supposedly acting at the same receptor site, studying their conformational and physicochemical properties, and combining this information with known biological activities. In particular, we have investigated common topographical arrangements of fundamental functional groups within different classes of drugs which are assumed to have a common pharmacophore structure. This method has recently been applied to building a general receptor model for CNS-active drugs [2]. In this model, the phenyl ring common to most CNS-active drug classes was assumed to interact

*New address: Bond University, Private Bag 10, Gold Coast Mail Centre, Queensland, Australia 4217.

0920-654X/89/$3.50 © ESCOM Science Publishers B.V.

with a relatively planar surface in the receptor, and the orientation of the ring was consequently regarded as paramount. But is this assumption justified?

Another modelling method used by one of the authors (MT) was Höltje and Kier's 'receptor-mapping by model interaction calculations' [3], in which interaction energies are calculated between particular pharmacophoric groups of drug molecules and amino acids as binding-site models. Good correlations between interaction energies, calculated for a constant orientation of the drugs with respect to the receptor model, and biological activities led to the conclusion that these calculations, performed without graphics facilities and using very simple 'grids' for model orientations [4, 5], may simulate the actual receptor environment. But how close are they to reality?

The present study is an attempt to include more information about protein–ligand complexes in the receptor modelling procedure. As already mentioned, there is little possibility of looking at the 'molecular reality' of drug–receptor complexes. But why shouldn't we apply information about enzyme–substrate, –coenzyme, or –inhibitor complexes, which are likely to have parallels with the complexes we want to describe, to receptors? Perhaps in this way we can obtain a more realistic idea of spatial orientations of interacting groups and thus assess the accuracy of the assumptions commonly made in receptor modelling. There are two main classes of studies on non-covalent binding, especially hydrogen bonding, of different functional groups. On the one hand, there are studies on small molecules (mainly based on crystal structures retrieved from the Cambridge Crystallographic Database); special bond types; and particular classes of molecules. On the other hand, there are studies on macromolecules using data from the Brookhaven Protein Data Bank.

The small-molecule group includes the work of Taylor et al. on the N—H···O=C hydrogen bond [6–9] and the C—H···O, C—H···N, and C—H···Cl hydrogen bonds [10]. There are also studies by Jeffrey and co-workers on the O—H···O hydrogen bond [11]; hydrogen bond geometries in amino acids [12, 13]; hydrogen bonding in nucleosides and nucleotides [14]; and hydrogen bonds in barbiturates, purines, and pyrimidines [15]. Murray-Rust and co-workers have studied NH_3^+ groups [16], carbon-halogen bonds [17, 18], and sp^2- and sp^3-hybridized oxygen atoms [19]. The macromolecular area is most extensively covered by Baker and Hubbard's comprehensive work on hydrogen bonding in globular proteins [20]. There is also a series of specific studies on aromatic interactions, which include analyses of the geometry of phenyl–phenyl interactions [21–23], as well as studies of oxygen and sulphur atoms in the phenyl environment [24, 25].

However, a collective description of macromolecule–ligand complexes is lacking. Naturally, we might expect intermolecular geometries (e.g., the enzyme–ligand interaction) not to differ fundamentally from the orientations of interactions formed within small molecules or protein molecules. Nevertheless, it may be useful to study the characteristics of the atomic environments of individual ligand functional groups in complexes with macromolecules. Ideally, this could result in a pattern of preferred orientations which could serve as a template for hypothetical drug binding-site models.

MATERIALS AND METHODS

In order to base this study of protein–ligand complexes on material from a uniform source, all enzyme–ligand complexes held in the Brookhaven Protein Data Bank were taken as basic data. Unfortunately, there are only about 40 entries for protein–small molecule complexes in the Data

Bank (January 1987 issue), and even with this limited data base there are two more problems. First, among these complexes there are many structures of the same enzyme from different species, or the same enzyme with different inhibitors, rather than the variation in data we would expect for a collection of different proteins. Second, the structures in the Data Bank are sometimes the first relatively low-resolution refinements of X-ray structures which have subsequently been described with better resolution. The enzyme complexes finally included in the study were those with reasonable resolution ($\leq$2.5 Å), with R-factors promising reasonably well-refined structures ($R \leq 0.25$), and without structural disorder or covalent binding where this appeared likely to influence the interaction geometries. Eighteen structures (Table 1) met all these requirements.

To collect information on the three-dimensional surroundings of functional groups of the bound ligands in the different enzymes, the Brookhaven coordinate information was transferred to the Victorian College of Pharmacy's molecular modelling system MORPHEUS. This system was then used to focus on the ligand molecule and its surrounding atoms, measure all distances within 3.8 Å from the ligand molecule, store fragments containing the atomic coordinates within this range of the functional groups of the ligand, superimpose these fragments gained from different enzymes, and plot them as atom clouds from various points of view. The distance cut-off of 3.8 Å between ligand and protein was calculated by adding the approximate van der Waals' radii of interacting atoms plus an additional 0.5 Å. Considering, for example, 2.8 or 2.9 Å for mean O···O or O···N hydrogen-bonding distances, 1 Å more should include the position of the main contacts in the structures. Covering a larger zone would result in less directionality of the interacting groups in studying geometries of interaction in either small molecules [6–19] or in highly refined protein structures [20–25].

Water molecules are included in the study when they are part of the Brookhaven data. But no specific examination was performed for water-mediated interactions, partly due to the chosen cut-off distance. This would be an interesting separate study.

Because of the relatively small amount of data, there was no statistical analysis of the results, and we restricted ourselves to a visual evaluation of the results using projections of the coordinates in three perpendicular planes to gain a perception of the three-dimensional image. Phenyl, carboxyl, carbonyl, amino, and hydroxyl groups and their environments were treated. Structures for which there was uncertainty about the reliability of the atomic positions – as in the case of the citrate carboxyl groups in 5-LDH – were omitted. To this end, data files were carefully checked for notes and literature references. Nonplanar RR'C-N groups (three structures out of 25 containing an amino group) were also excluded to obtain a uniform arrangement of carbon atoms for superimposition.

The final results consist of superimpositions of the collected data and are represented as atom clouds containing the information from a number of different enzyme–ligand complexes, each cloud representing the surroundings of one type of functional group. We then looked for general rules or favoured orientations within these clouds (i.e., an accumulation of atoms in a particular direction).

RESULTS AND DISCUSSION

Basic information about the collected data is given in Table 2. Figures 1–7 represent 'slices' of the clouds described above. With respective functional groups positioned in the xy-plane, slices 2 Å

TABLE 1
ENZYME–LIGAND COMPLEXES IN THE BROOKHAVEN PROTEIN DATA BANK (JANUARY 1987)

Code	Enzyme	Ligand	No. of residues	No. of molecules	Solvent molecules	Resolution (Å)	R-factor	Structure included in superimposition
5ADH	Alcohol dehydrogenase	ADP-ribose, 2-Methyl-2,4-pentanediol	374	1	Yes	2.9	0.22	No
6ADH	Alcohol dehydrogenase	NAD, DMSO	374	2	No	2.9	0.38	No
3CPA	Carboxypeptidase A	Glycyl-L-tyrosine	307	1	No	2.0		Yes
4CPA	Carboxypeptidase A	Potato carboxypeptidase A inhibitor	307	1	No	2.5	0.196	Yes
7CAT	Catalase	NADPH	506	1	Yes	2.5	0.212	Yes
8CAT	Catalase	NADPH	506	2	Yes	2.5	0.191	Yes
2CHA	Alpha-chymotrypsin A	Tosyl (covalent)	245	1	Yes	2.0		Yes
2CTS	Citrate synthase (pig)	Co*A, Citrate	437	1	Yes	2.0	0.161	Yes
3CTS	Citrate synthase (chicken)	Co*A, Citrate	437(unk.)	1	Yes	1.7	0.192	Yes
4CTS	Citrate synthase	Oxaloacetate	437	2	Yes	2.9	0.182	No
3DFR	Dihydrofolate reductase (*L.casei*)	NADPH, Methotrexate	162	1	Yes	1.7	0.152	Yes
4DFR	Dihydrofolate reductase (*E.coli*)	Methotrexate	159	2	Yes	1.7	0.155	Yes
2GRS	Glutathione reductase	FAD	478	1	No	2.0		Yes
1GPD	D-Glyceraldehyde-3-phosphate dehydrogenase	NAD	334	2	No	2.9		No
2YHX	Hexokinase B	O-Toluoylglucosamine	457	1	No	2.1		Yes
2MCP	Immunoglobulin	Phosphocholine	220	1	No	3.1	0.185	No
2PKA	Kallikrein A	Benzamidine	232	2	Yes	2.05	0.220	Yes
2KAI	Kallikrein A	Pancreatic trypsin inhibitor	232	1	Yes	2.5	0.224	No
3LDH	Lactate dehydrogenase (dogfish)	NAD, Pyruvate	330	1	No	3.0		No
5LDH	Lactate dehydrogenase (pig)	S-Lac-NAD, Citrate	334	1	No	2.7	0.196	Yes
2MDH	Malate dehydrogenase	NAD	324/325	2	No	2.5	0.324	No
1PPD	Papain	2-Mercaptoethanol (covalent)	212	1	Yes	2.0	0.145	No
3PGK	Phosphoglyceratekinase	ATP, 3-Phosphoglycerate	416	1	No	2.5		No
2SNS	Staphylococcal nuclease	Deoxy-diphosphothymidine	149	1	No	1.5		Yes
4TLN	Thermolysin	L-Leucyl-hydroxylamine	316	1	Yes	2.3	0.169	Yes

5TLN	Thermolysin	HONH-benzylmalonyl-L-alanyl-glycine-*p*-nitroanilide	316	1	Yes	2.3	0.179	Yes
7TLN	Thermolysin	$CH_2CO(N\text{-}OH)Leu\text{-}OCH_3$ (covalent)	316	1	Yes	2.3	0.170	No
3PTB	Beta-trypsin	Benzamidine	223	1	Yes	1.7	0.182	Yes
1TPP	Beta-trypsin	*p*-Amidino-phenyl-pyruvate	223	1	Yes	1.4	0.191	Yes
3PTP	Beta-trypsin	Diisopropylphosphoryl (covalent)	223	1	Yes	1.5	0.154	No
2PTC	Beta-trypsin	Pancreatic trypsin inhibitor	223	1	Yes	1.9	0.187	No
1TPA	Anhydro-trypsin	Pancreatic trypsin inhibitor	223	1	Yes	1.9	0.175	No

TABLE 2
ANALYSIS OF ENZYME–LIGAND COMPLEXES IN THE BROOKHAVEN DATA BANK

	No. of available structures	No. of atoms within 3.8 Å	No. of atoms within 3.8 Å/structure	Superimposition	Closest distance (Å)	Dominant amino acids
Phenyl	12	72	6.0	Plane 6 atoms	3.1/3.2 H_2O ~3.6	Mostly hydrophobic
Carboxyl	17	134	7.9	Plane 3 atoms	Arg 2.6/2.9 ~2.6	*No* hydrophobic
Carbonyl	23	84	3.65	Plane 4 atoms	from 2.6 ~2.8	Mostly hydrophilic
Hydroxyl	37	165	4.46	4 atoms	~2.4/2.5/2.6	2/3–1/3 Hydrophilic-hydrophobic
Amino	22[a]	118	5.36	Plane 4 atoms	~2.8/2.9	1/2–1/2 Hydrophilic-hydrophobic[b]

[a]Planar structures only.
[b]Including chain.

above and below the atoms of each functional group in the xy-plane and ± 2 Å cuts in the xz- and yz-planes (measuring from the heteroatoms of the group, or the centre of the ring in the phenyl case) are shown. Figure 8 explains this procedure. With the superimposed group positioned in the xy-plane, we have in each figure a view perpendicular to the xy-plane, a view horizontal to this plane viewed from the y-direction (i.e., looking at the xz-plane), and another from the x-direction looking at the yz-plane.

This type of representation was chosen in order to give the reader a clearer overview. The cloud representations could only be clarified by using different colours, the original Brookhaven labelling and connectivities for all structures combined with listings of the measured distances, or by comparing the cloud representations with all the single superimposed structures.

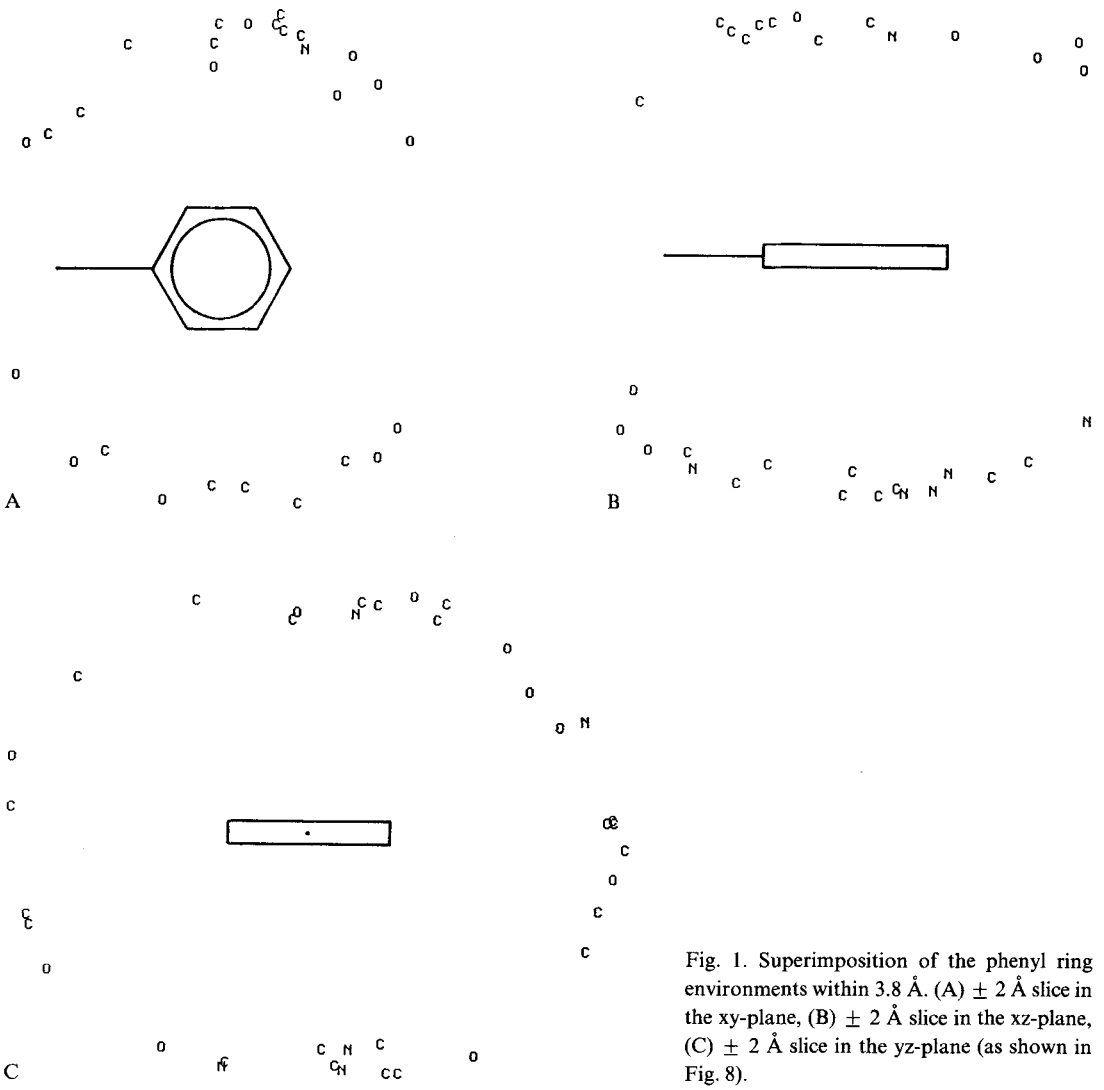

Fig. 1. Superimposition of the phenyl ring environments within 3.8 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

The phenyl environment (Figs. 1A–C, 2A–C)

The atoms surrounding the phenyl ring were examined not only within 3.8 Å (Fig. 1), where we find a very low atom density, but also within 4.5 Å (Fig. 2) to obtain a slightly better idea of the total environment. It is relatively difficult to decide on any clear-cut interpretation of these data, but some observations are simple and obvious:

- there are no other groups in a real coplanar arrangement with the ligand phenyl ring;
- there are both polar and nonpolar amino acid residues around the ring;
- the surrounding atoms are further away (indicating weaker interactions) from the phenyl group than from other functional groups studied;
- there is an average of six atoms interacting with each ring, and only single atoms, not whole functional groups, are in the closer distance range.

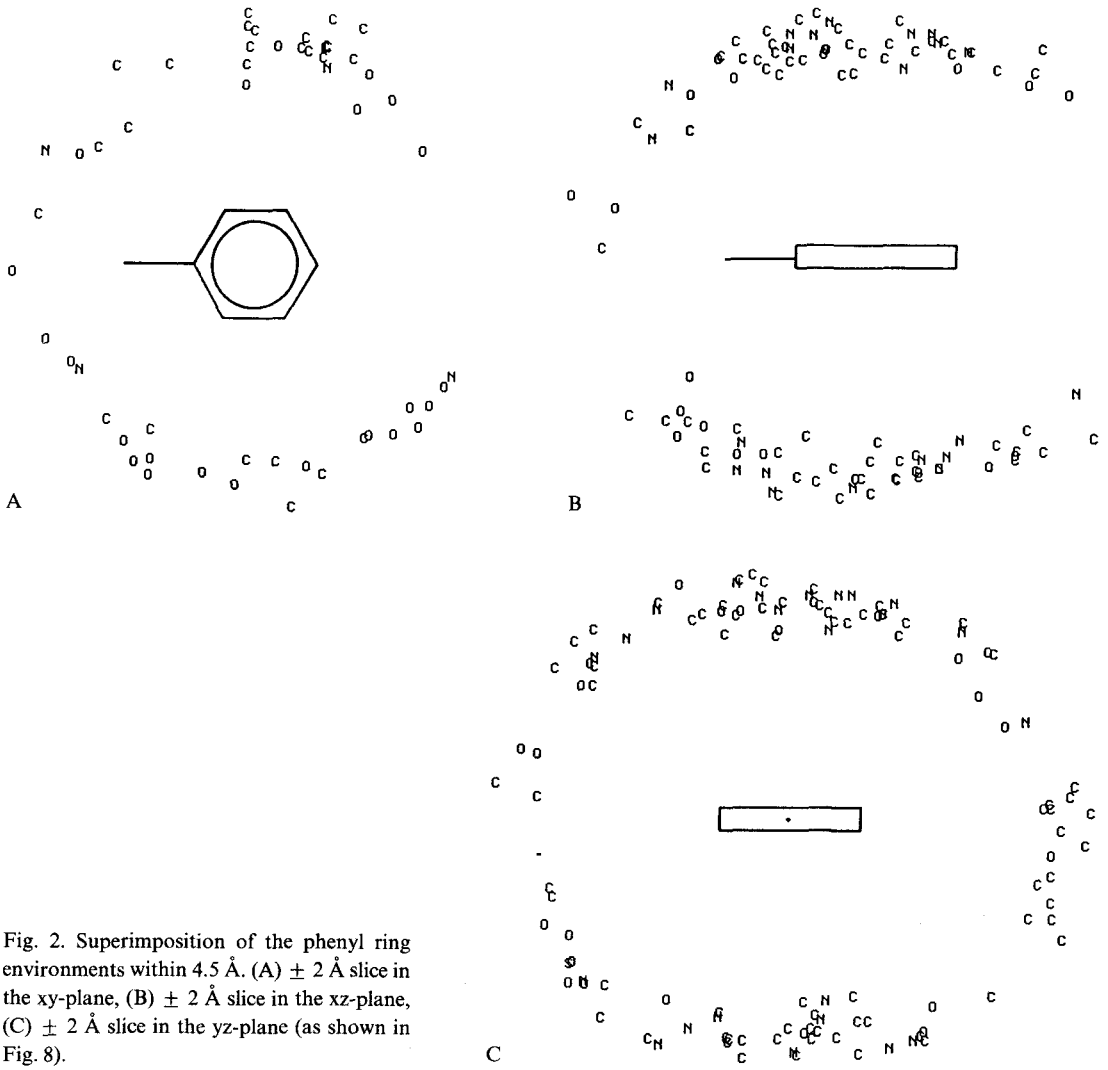

Fig. 2. Superimposition of the phenyl ring environments within 4.5 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

Of most interest to us is the actual shape of the atom cloud surrounding the phenyl ring. If, as we frequently assume, the orientation of the phenyl ring is important for binding, we would expect the phenyl ring environment to be a flattened shape with an ellipsoidal cross-section. If, on the other hand, the ring is performing no more than a hydrophobic space-filling role, the orientation should be less important and a hemispherical volume with a circular cross-section might be expected.

On balance, our data support the former alternative, although the number of structures available is clearly insufficient to draw unambiguous conclusions. It is apparent, however, that phenyl rings and their protein 'receptor groups' do not form coplanar layers, which is how most geometries were built in our theoretical modelling [3, 4]. It is also clear that in this distance range the perpendicular phenyl-phenyl orientations, as proposed in studies at longer distances [21], are unimportant. The results, however, are in accordance with other studies on the phenyl ring environments in proteins which showed a preferred interaction between oxygen atoms and the edge of planar aromatic rings [24]. This may be seen by comparing Fig. 2A with 2B and 2C.

The carboxyl environment (Fig. 3A–C)

Contrary to the phenyl case, we find here two quite distinct kinds of binding: either chelate-like with both oxygen atoms involved, or sideways involving one oxygen only. The carboxyl group sits – as expected – in an almost totally polar environment and it can be seen that functional groups of the enzyme, rather than single atoms, are involved in binding. There is no obvious preference for chelate or lateral binding.

For chelate-type binding, the guanidino group of arginine is invariably the binding group. These arginine interactions are consistently found in completely unrelated enzymes. In addition to the five amino acid residues showing 'frontal' interaction with the carboxyl group, there are seven that approach from the side. Those binding in the sideways manner are three histidines, an arginine, a glutamic acid, a tyrosine, a glycine-peptide chain, and others. That is, for lateral bonding there is no preference for a special group of amino acid residues, whereas in frontal binding we find exclusively arginine. This is an important finding for receptor modelling and, perhaps even more so, for receptor-based drug design.

Further evidence for the role of arginine in carboxyl binding is provided by complexes not included in our analyses. For example, structure 4TPI in the Protein Data Bank [26], a bovine trypsinogen in a ternary complex with a valine-valine dipeptide and a semisynthetic homologous bovine pancreatic trypsin inhibitor, has the reactive site lysine replaced by an arginine residue. (We excluded this structure from our initial data set because of its similarity to other serine proteases which basically show the same arrangement in the active site.) In the original enzyme–inhibitor complex, a water molecule mediates hydrogen bonding between the reactive site lysine and an aspartic acid, as in the case of trypsin. In the mutant, where this lysine is replaced by arginine, the water molecule is absent and a direct chelate is formed.

There are many more examples of arginine–carboxyl interactions that appear to have key functional roles in enzymes, including some in which an arginine-like structure in the ligand interacts with a carboxyl group from the binding site (i.e., the reverse of that described above). In dihydrofolate reductase, for example, an aspartic acid forms a complex with the 2,4-diaminopteridine ring of the powerful inhibitor methotrexate. Directed mutagenesis of this single amino acid to ser-

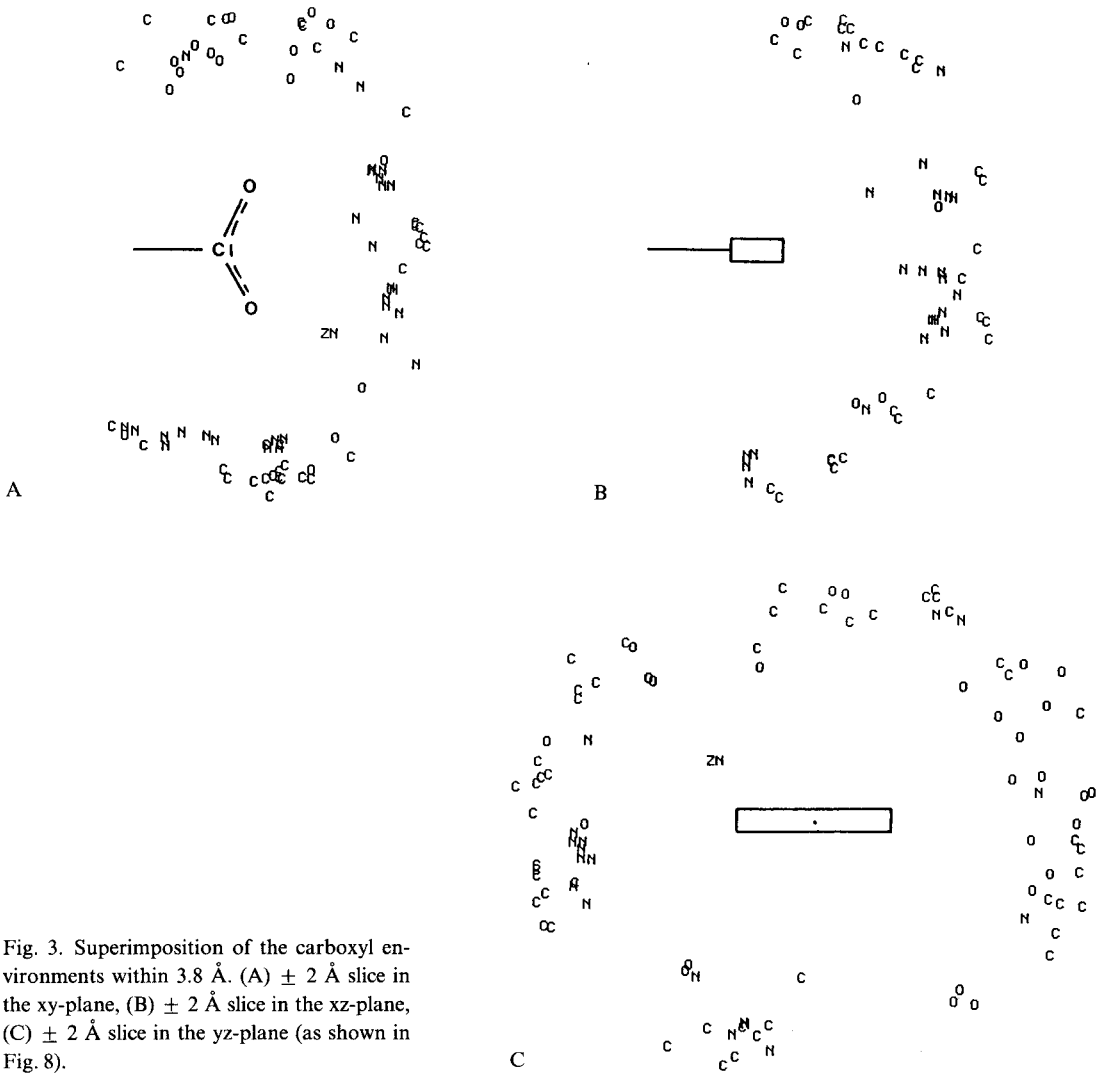

Fig. 3. Superimposition of the carboxyl environments within 3.8 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

ine results in an interaction mediated by an additional water molecule [27]. This sort of complex achieves activity only through a decrease in pH. It is therefore assumed that aspartic acid has its function in protonation of the substrate. Another example proving the functional importance of the basic part of the arginine residue is that of staphylococcal nuclease [28], in which two arginine residues form two hydrogen bonds each with a phosphate group. Here they seem to have binding and catalytic roles; they polarize the phosphate group to make it more susceptible to nucleophilic attack by a water molecule which is itself bound to a glutamic acid residue. A similar situation around a phosphate group is found in the lactate dehydrogenase complex 5LDH and in other glycolytic enzymes, such as alcohol dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase, where the arginine residue again interacts with negatively charged phosphate or carboxylate moieties of substrates and cofactors [29].

Finally, Baker and Hubbard attest to the same situation in their enormous evaluations of hydrogen bonding in proteins [20]. Arginine is almost exclusively hydrogen-bonded – sometimes via water molecules – to either main-chain carbonyl groups or carboxyl groups. In 29% of total interactions, it ion-pairs with carboxyl groups. A very recent, accurate analysis of the geometries of these intramolecular arginine–carboxyl interactions in 37 proteins is reported by Singh et al. [30].

Clearly, the generality of the arginine–carboxyl interaction should be kept in mind for application in further receptor-modelling studies.

The carbonyl environment (Figs. 4A–C, 5A–C)

Contrary to the carboxyl group, where interactions are partially or entirely ionic, the directionality of hydrogen bonding through the influence of the oxygen sp^2 lone pairs should be more ob-

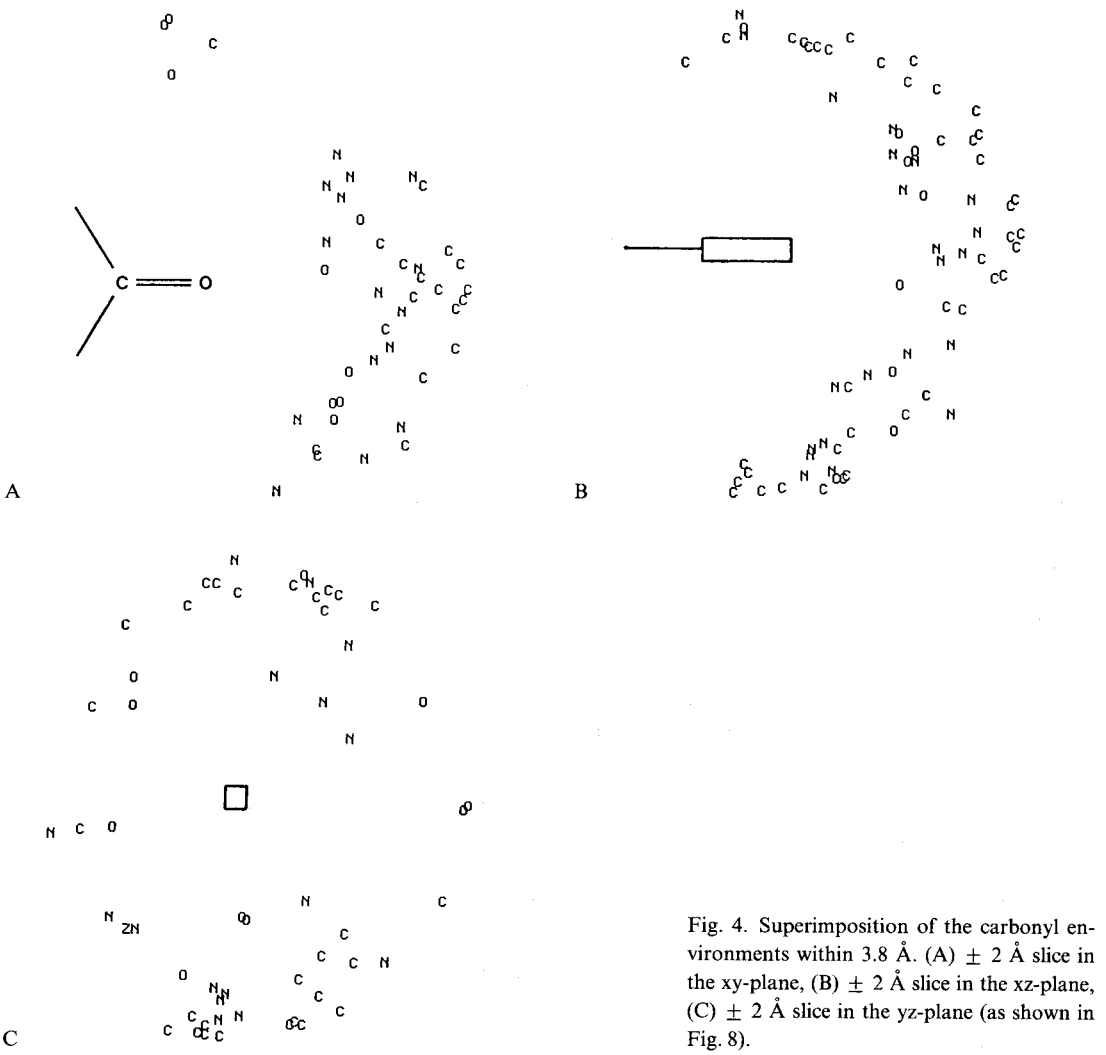

Fig. 4. Superimposition of the carbonyl environments within 3.8 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

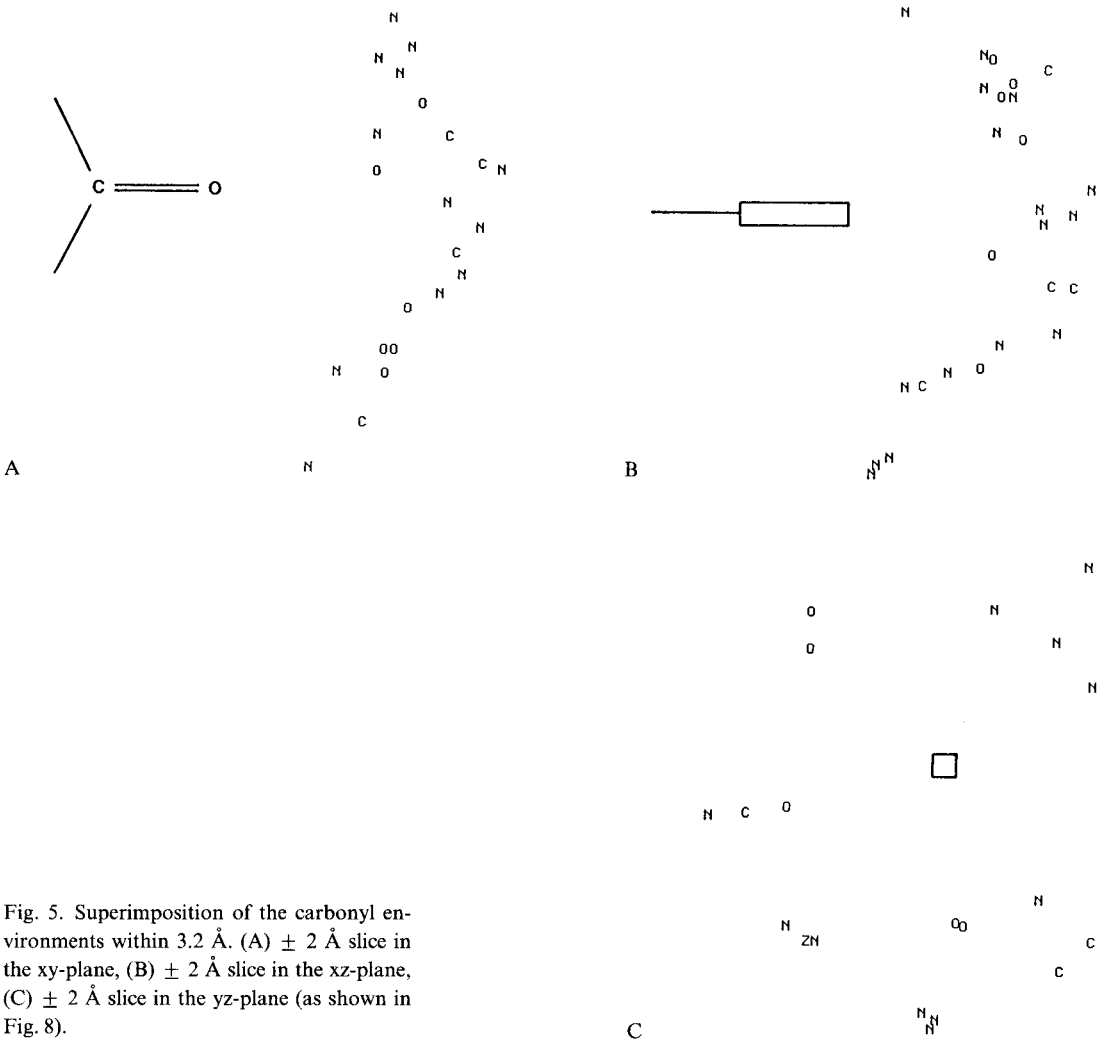

Fig. 5. Superimposition of the carbonyl environments within 3.2 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

vious in the case of carbonyl groups. This influence should increase with decreasing H···O distance.

Taylor et al [6] studied 889 small molecules, containing a total of 1 509 N—H···O=C hydrogen bonds, from the Cambridge Crystallographic Database and found mean values of 120–130° for the C=O···N(H) angle and 0° or 180° for the R—C=O···N(H) dihedral angle (i.e., the dominant orientation is in plane, as expected theoretically). The findings of Murray-Rust and Glusker [19] in Cambridge Crystallographic Database searches for O···X (X=O, N) intermolecular contacts of less than 3 Å led to the same conclusion.

For our enzyme–ligand structures, data were collected for both 3.8 Å (Fig. 4) and 3.2 Å (Fig. 5) distance cut-offs, since the directionality should be more marked at closer distances. The results within 3.8 Å show a more or less spherical shape, with at least as many atoms above or below the plane of the carbonyl group as in plane with the expected oxygen lone-pair positions. The results

at less than 3.2 Å distance show slightly more concentration in a cone around the elongated C=O axis with its peak at the oxygen atom, but there is no accumulation in distinct planar lone-pair positions. This is best seen from a head-on view of the carbonyl function (Figs. 4C, 5C), where we would expect atoms in a plane but find apparently random orientation instead. Orientations of receptor groups for the carbonyl group are thus less in accordance with theory than in the case of the carboxyl group. It remains to be seen whether this behaviour is dependent on the atoms covalently attached to the carbonyl function (i.e., whether the carbonyl type is a ketone, an amide, etc.).

The hydroxyl environment (Fig. 6A–C)

In the hydroxyl case, even less directionality is observed and atoms seem to be equally distributed in a clear hemisphere. Although there are more atoms within the range studied than in the

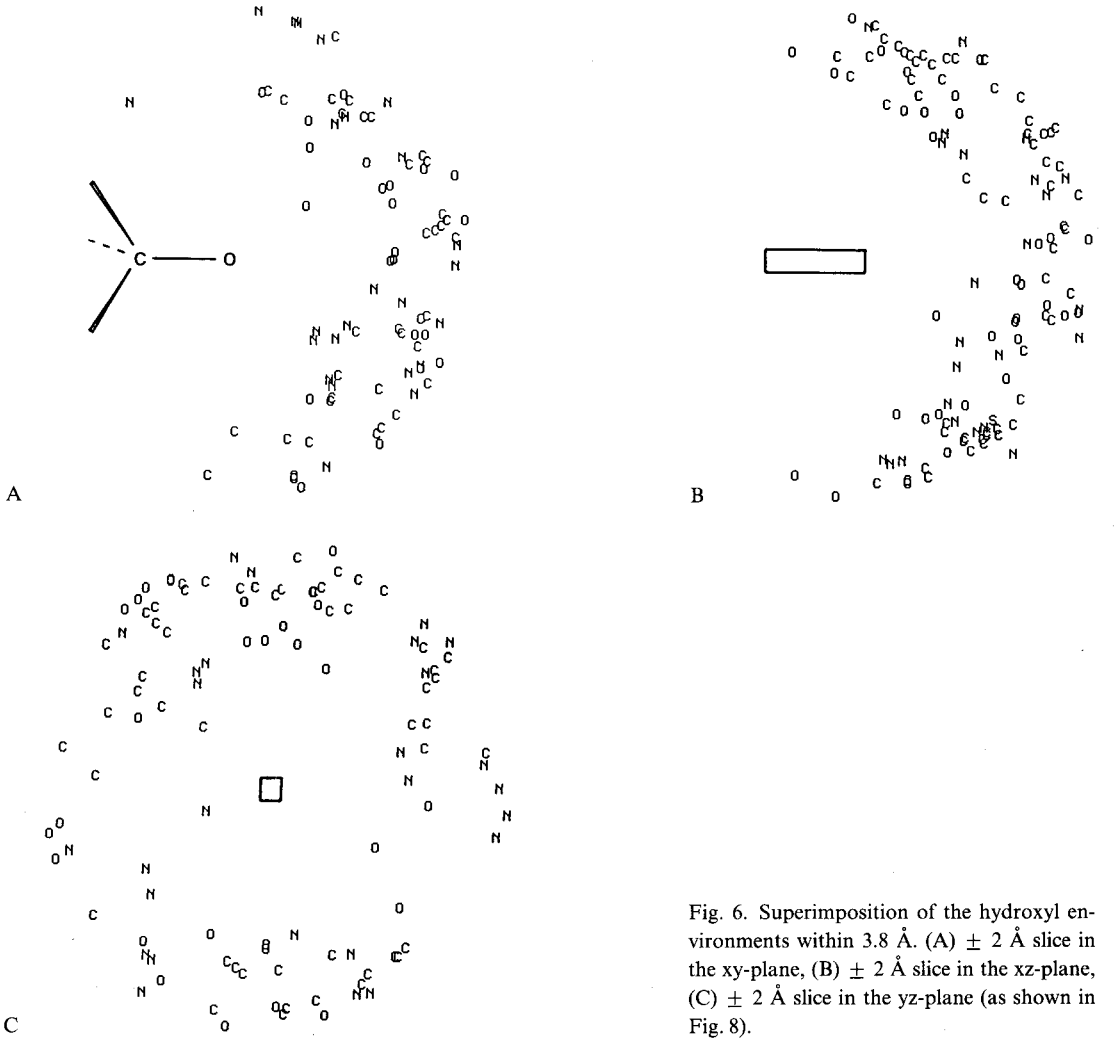

Fig. 6. Superimposition of the hydroxyl environments within 3.8 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

carbonyl case, there is no spatial orientation corresponding to electron densities of nonbonded electrons in this case of sp^3 hybridization. This is entirely as expected, in view of the almost free rotation of the hydroxyl group, as well as the different donor and acceptor capabilities of the hydroxyl group. More information would clearly be avaiblable if the H-coordinates were included in the X-ray structures, but there is not enough resolution in the protein crystal structures studied to be able to localize H-positions. There are, however, studies by Vedani and Dunitz [31] analysing geometries of hydrogen bonds in crystal structure data retrieved from the Cambridge Crystallographic Database that include hydroxyl groups. Those results are similar to what we found in the complexes. A nearly homogenous distribution of donor hydrogen atoms around acceptor hydroxyl oxygen atoms, without any significant preference of the hydrogen-bond donor for tetrahedral lone-pair directions, is shown.

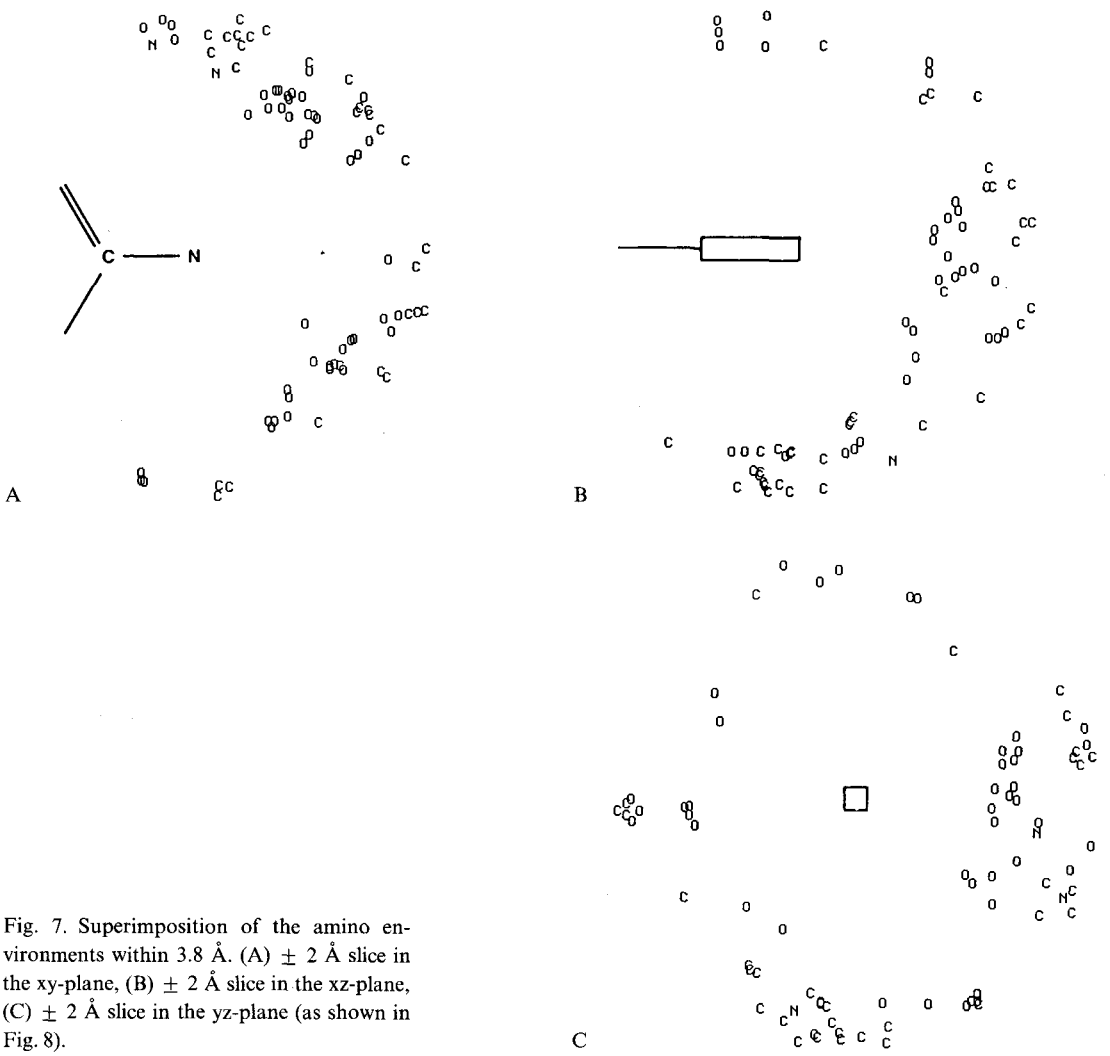

Fig. 7. Superimposition of the amino environments within 3.8 Å. (A) ± 2 Å slice in the xy-plane, (B) ± 2 Å slice in the xz-plane, (C) ± 2 Å slice in the yz-plane (as shown in Fig. 8).

The amino environment (Fig. 7A–C)

The results show more directional preference for the amino group than was seen in the carbonyl case. Groups containing oxygen, including aspartic or glutamic acid as well as parts of the peptide backbone, are situated in the amino plane in locations anticipated for interactions with the hydrogen atoms. A second, less significant accumulation of oxygen atoms appears in the 90° position above the nitrogen, where perhaps interactions with the lone pair or partial double bonds take place [32].

Figure 9 schematically summarizes the results described above. Figures 10–12 give a more spatial impression of the probability for certain atom types and locations in the carboxyl, amino and hydroxyl environments and show the respective clouds in a view along the bond to the functional group studied, demonstrating the prevalence of N or O atoms in these environments.

CONCLUSION

We are aware of the many simplifying assumptions made in this study and of several problems not taken into consideration. First, the material currently available is not sufficient to be statistically significant, and inclusion of more structures in the future will obviously be useful. Second, as more data become available the topic can be studied in more detail by taking into account the

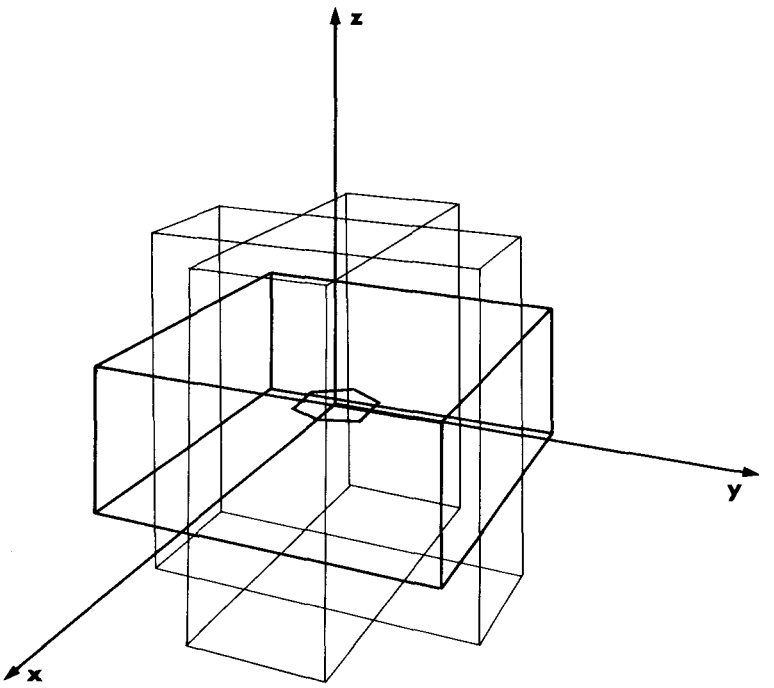

Fig. 8. Geometry of the depicted parts of the atomic environments of single functional groups.

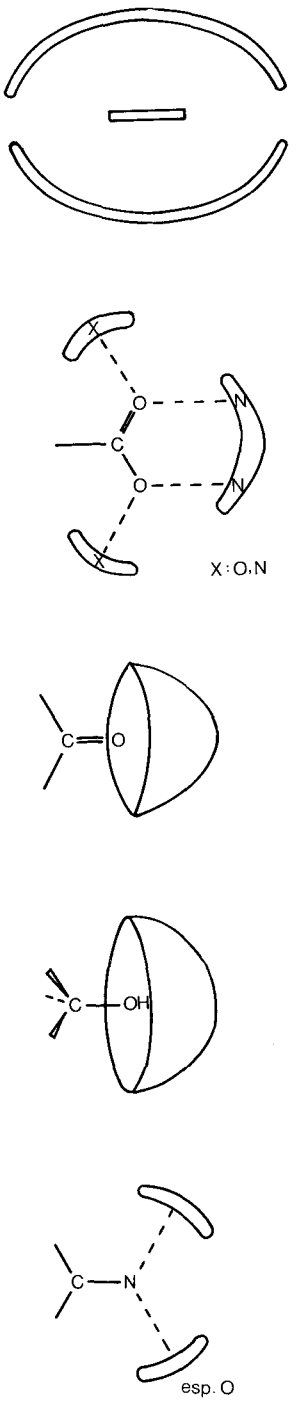

Fig. 9. Synopsis of the results for all superimpositions.

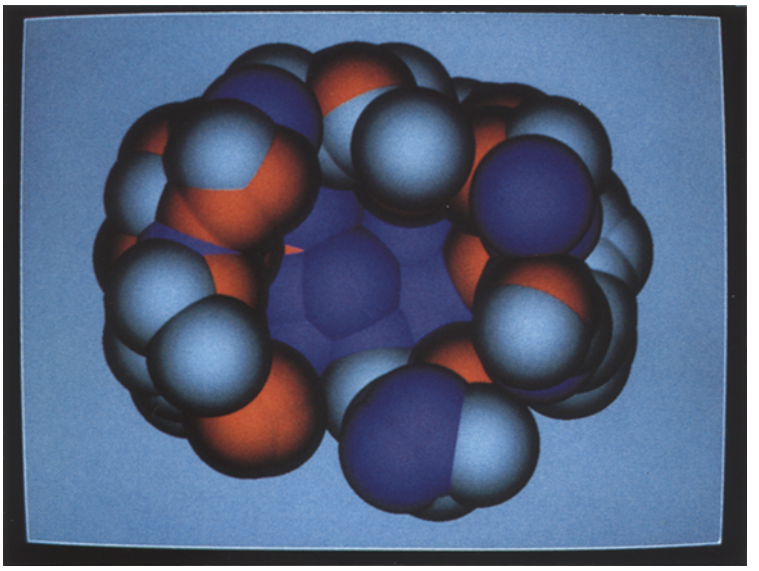

Fig. 10. Space-filling head-on view of the carboxyl environments.

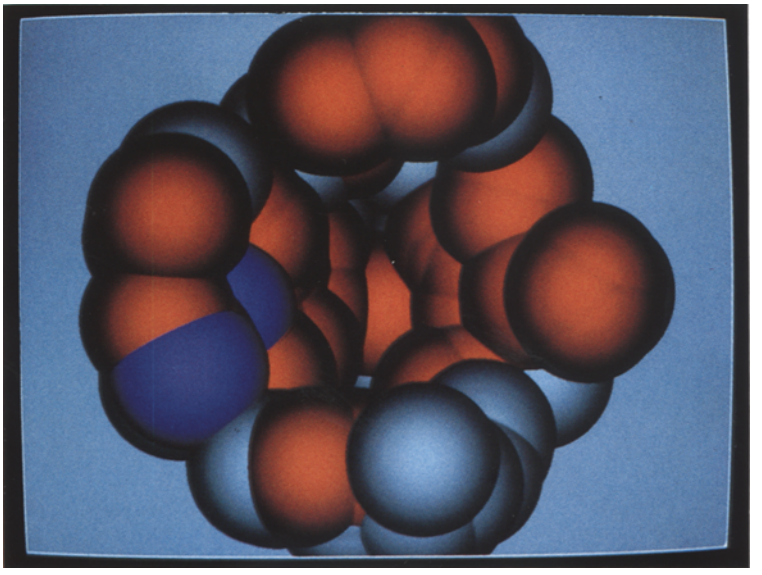

Fig. 11. Space-filling head-on view of the amino environments.

role of neighbouring atoms and groups. Third, there is the fact that proteins are not static and the data used represent only one of several possible favoured states. Finally, it has to be kept in mind that a number of weaker bonds may contribute significantly to the overall interaction [33], and that the orientation of particular groups should, therefore, not be overemphasized.

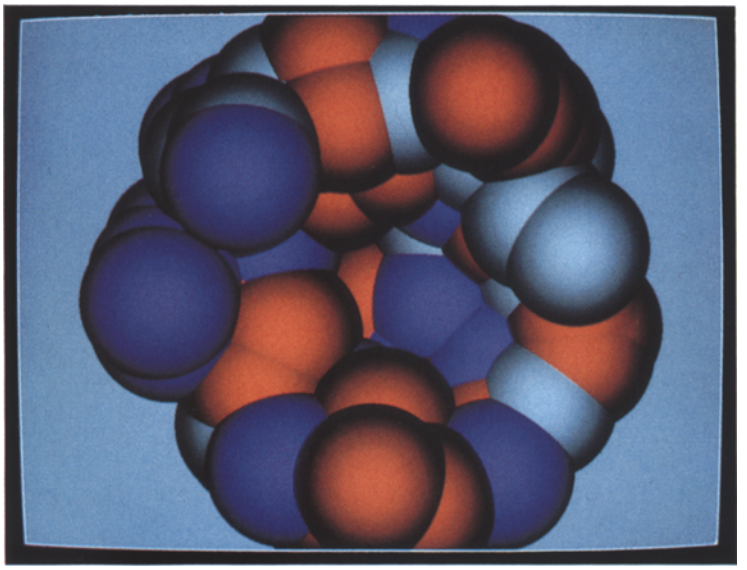

Fig. 12. Space-filling head-on view of the hydroxyl environments.

There are, however, a number of clear and useful conclusions from this study. First, it can be seen that there is an equivalence between intermolecular geometries found in enzyme–ligand complexes and the orientation of groups in small molecules. This means we should be able to extrapolate from small molecules to complex modelling with, perhaps, the exception of the carbonyl group, where the geometries seem to differ.

Second, some obvious geometric design rules emerge, especially the chelate-like orientation between carboxyl groups and arginine residues (or, conversely, between guanidine- or amidine-like structures and glutamic or aspartic acid residues) and the generally nonaromatic, nonplanar environment of aromatic groups in this series of enzyme–inhibitor interactions.

We intend to combine these results with theoretical studies of preferred group orientations by standard interaction calculations and to use the finding in ongoing studies of drug structure–activity relationships and receptor modelling.

Note: Machine-readable data of the figures in this article are available on request. These are provided on IBM (5.25 inch) or Macintosh (3.5 inch) diskettes. Please contact the publisher for details.

REFERENCES

1 Franke, R., Theoretical Drug Design Methods, Elsevier, Amsterdam, 1984, pp. 315–363.
2 Lloyd, E.J. and Andrews, P.R., J. Med. Chem., 29 (1986) 453–462.
3 Höltje, H.-D. and Kier, L.B., J. Pharm. Sci., 63 (1974) 1722–1725.
4 Höltje, H.-D. and Tintelnot, M., Quant. Struct.-Act. Relat., 3 (1984) 6–9.
5 Tintelnot, M., Arch. Pharm. (Weinheim), 320 (1987) 275–279.
6 Taylor, R., Kennard, O. and Versichel, W., J. Am. Chem. Soc., 105 (1983) 5761–5766.
7 Taylor, R. and Kennard, O., Acta Crystallogr., Sect. B, 39 (1983) 133–138.

8 Taylor, R., Kennard, O. and Versichel, W., J. Am. Chem. Soc., 106 (1984) 244–248.
9 Taylor, R., Kennard, O. and Versichel, W., Acta Crystollogr., Sect. B, 40 (1984) 280–288.
10 Taylor, R. and Kennard O., J. Am. Chem. Soc., 104 (1982) 5063–5070.
11 Ceccarelli, C., Jeffrey, G.A. and Taylor, R., J. Mol. Struct., 70 (1981) 255–271.
12 Jeffrey, G.A. and Maluszynska, H., Int. J. Biol. Macromol., 4 (1982) 173–185.
13 Jeffrey, G.A. and Mitra, J., J. Am. Chem. Soc., 106 (1984) 5546–5553.
14 Jeffrey, G.A., Maluszynska, H. and Mitra, J., Int. J. Biol. Macromol., 7 (1985) 336–348.
15 Jeffrey, G.A. and Maluszynska, H., J. Mol. Struct., 147 (1986) 127–142.
16 Rosenfield Jr., R.E. and Murray-Rust, P., J. Am. Chem. Soc., 104 (1982) 5427–5430.
17 Murray-Rust, P., Stallings, W.C., Monti, C.T., Preston, R.K. and Glusker, J.P., J. Am. Chem. Soc., 105 (1983) 3206–3214.
18 Ramasubbu, N., Parthasarathy, R. and Murray-Rust, P., J. Am. Chem. Soc., 108 (1986) 4308–4314.
19 Murray-Rust, P. and Glusker, J.P., J. Am. Chem. Soc., 106 (1984) 1018–1025.
20 Baker, E.N. and Hubbard, R.E., Prog. Biophys. Mol. Biol., 44 (1984) 97–179.
21 Burley, S.K. and Petsko, G.A., Science, 229 (1985) 23–28.
22 Blundell, T., Singh, J., Thornton, J., Burley, S.K. and Petsko, G.A., Science, 234 (1986) 1005.
23 Singh, J. and Thornton, J.M., FEBS Lett., 191 (1985) 1–6.
24 Thomas, K.A., Smith, G.M., Thomas, T.B. and Feldmann, R.J., Proc. Natl. Acad. Sci. U.S.A., 79 (1982) 4843–4847.
25 Reid, K.S.C., Lindley, P.F. and Thornton, J.M., FEBS Lett., 190 (1985) 209–213.
26 Bode, W., Walter, J., Huber, T., Wenzel, H.R. and Tschesche, H., Eur. J. Biochem., 144 (1984) 185–190.
27 Howell, E.E., Villafranca, J.E., Warren, M.S., Oatley, S.J. and Kraut, J., Science, 231 (1986) 1123–1128.
28 Cotton, F.A., Hazen Jr., E.E. and Legg, M.J., Proc. Natl. Acad. Sci. U.S.A., 76 (1979) 2551–2555.
29 Riordan, J.F., McElvany, K.D. and Borders Jr., C.L., Science, 195 (1977) 884–886.
30 Singh, J., Thornton, J.M., Snarey, M. and Campbell, S.F., FEBS Lett., 224 (1987) 161–171.
31 Vedani, A. and Dunitz, J.D., J. Am. Chem. Soc., 107 (1985) 7653–7658.
32 Kolb, V.M. and Scheiner, S., J. Pharm. Sci., 73 (1984) 389–392.
33 Fersht, A.R., Leatherbarrow, R.J. and Wells, T.N.C., Trends Biochem. Sci., 11 (1986) 321–325.