BASIC CONCEPTS IN STRUCTURE-AIDED DRUG DESIGN

Acta Cryst. (1995). D**51**, 407–417

Small-Molecule Crystal Structures as a Structural Basis for Drug Design

BY CLAUDINE PASCARD

Cristallochimie, Institut de Chimie des Substances Naturelles, 91198 Gif-sur-Yvette, France

(*Received* 12 *July* 1994; *accepted* 12 *December* 1994)

Abstract

For a long time, the crystal structures of small molecules were regarded as useful only for establishing the stereochemical formulae of the crystallized compounds. Recently, chemists have realized that in the study of the environment in the solid state there exists valuable structural information on the binding characteristics of chemical groups. Numerous comparisons have been made which show the nearly perfect correlation between small-molecule structural results, and the observed binding in receptor–substrate complexes. Moreover, the observed conformations of flexible substrates interacting with the neighbouring molecules in their crystal structures, can lead to valuable hypotheses on their conformation when bound in the active site of a biological macromolecule.

Introduction

The starting point for drug design is a knowledge of the three-dimensional conformation of the active site. This can be found in two ways: the direct approach is based on a knowledge of the three-dimensional structure of the macromolecule, or better, on that of its complex with a molecule of known activity. Without the spatial structure of the active site, the indirect approach consists of deriving the stereochemical requirements of the unknown binding sites of the macromolecule by complementarity to the pharmacophore common to a family of active molecules.

In both methods, the search for a lead molecule starts from the spatial coordinates of the binding sites, that is the topology of the pharmacophore or of the 'antipharmacophore' (Murray-Rust & Glusker, 1984). This can be obtained by computing devices; for example, the program *CONCORD* (Pearlman, Rusinko, Skell, Banducci & McGarity, 1988) can be used to draw three-dimensional structures from two-dimensional chemical formulae with all the possible conformers. But it is obvious that the most convenient approach is, by far, the determination of the spatial coordinates of a molecule from NMR experiments in solution and from X-ray structure determination in the solid state. The relative spatial disposition of the necessary functional groups being known, the intermolecular interactions with the receptor have to be determined. In each case, a knowledge of the geometrical requirements of the different non-covalent bindings is a necessity.

The binding of a drug to its receptor is effected by the following types of interactions: electrostatic forces, hydrogen bonding, van der Waals forces (aromatic stacking, hydrophobic contacts). The hydrophobic forces are weak but numerous, and their large number will direct the substrate into the active site where stronger binding will take place.

It is only recently that chemical crystallographers have realized that a real wealth of structural information concerning the non-bonded interactions exists in the molecular packing observed in a crystal. What can we obtain from the crystal structure of a molecule which has a role in a biological system? The crystal state is an ideal medium in which to study the exact requirements of non-bonding interactions between chemical groups. A crystal can be visualized as a single supramolecule, every molecular fragment being bonded to the surrounding neighbours, which can be similar, or different, fragments, ions or solvent molecules. Their interactions are evidently of the same order as those between receptor and substrate. Therefore, not only shall we obtain a more precise geometrical description of the preferred binding mode of such a chemical group, but also the behaviour of the whole molecule as a result of its environment. The creation of the Cambridge Structural Database (CSD) (Kennard *et al.*, 1975; Allen *et al.*, 1979) opened a new era by making possible systematic searches among the structures stored.

If it can be proven that the results of these studies on small-molecule crystal structures can be reasonably extended to the receptor–ligand complex interactions, these studies will help to design the geometrical and chemical requirements of a lead molecule.

Systematic analyses of the binding parameters

The systematic analysis started nearly 20 years ago. One of the first examples is the analysis of 69 structures of sulfur-containing compounds in order to study non-bonded contacts around divalent sulfur (Rosenfeld,

Parthasarathy & Dunitz, 1977). A clear representation of the different directions of approach to a divalent sulfur by electrophiles and nucleophiles was derived (Fig. 1).

Then, as the Cambridge Structural Database became richer in structural information, the automatic retrieval of intermolecular contacts in organic molecules led to extensive studies of the environment of chemical functional groups. (Murray-Rust & Motherwell 1979; Rosenfeld & Murray-Rust, 1982; Murray-Rust, Stallings, Monti, Preston & Glusker, 1983; Taylor, Kennard & Versichel, 1983, 1984; Murray-Rust & Glusker, 1984; Vedani & Dunitz, 1985; Steiner & Saenger, 1992*a*). Most of these results are summarized in Jeffrey & Saenger's book (1993). A basic book on structure correlation has been published recently (Bürgi & Dunitz, 1994).

The most recent results, bringing new insights on binding geometry arise from the studies on the weakest interactions. The number of crystal structures stored in the Cambridge Structural Database is now sufficiently large (and the structures sufficiently accurate) to allow a systematic survey of aromatic stacking, of CH···O/N interactions, and of hydrophobic contacts (Allen *et al.*, 1991). Twelve years ago, a survey of 113 neutron diffraction structures showed good statistical evidence of an attractive interaction between C—H and O atoms, and also between C—H and N atoms (Taylor & Kennard, 1982). A systematic analysis on the basis of 30 neutron diffraction studies of carbohydrates (Steiner & Saenger, 1992*b*) led to a precise description of the geometrical characteristics of the CH···O interaction. Scientists such as Leiserowitz (Berkovitch-Yellin & Leiserowitz, 1984), Desiraju & Glusker (Desiraju, 1991; Desiraju, Kashino, Coombs & Glusker, 1993) have scrutinized many crystal structures to analyse their packing and see the extent of the influence of CH···O interactions. These authors showed that, although weak, this CH···O binding extends far beyond the van der Waals limit and, therefore, can have an orienting effect on molecules.

Desiraju (Sharma, Paneerselvam, Pilati & Desiraju, 1993) also compared the relative strengths of various weak bonds such as CH···O and π–π interactions. He pointed out that, in certain cases, the stacking of two aromatic rings can prevail over other types of hydrogen bonding (CH···O as well as OH···O). This should not be forgotten when modeling.

Christofer Hunter carefully analysed interactions involving aromatic compounds. He ascertained their electrostatic interactions and studied their relative binding energies (Hunter & Sanders, 1990; Hanton, Hunter & Purvis, 1992; Hunter, 1993). He found evidence that demonstrates the NH···π interaction, and that hydrogen bonding can occur by way of a change in hydridization (Fig. 2). Supramolecular chemists such as Stoddart (Anelli *et al.*, 1992) and Diederich (Klebe & Diederich, 1995) have used these findings to build their molecular Meccano.

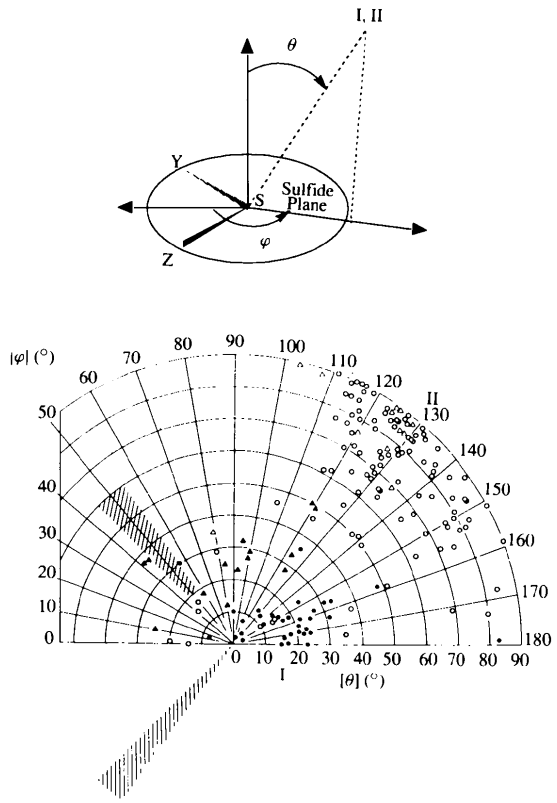


Fig. 1. Approach of a sulfide S atom by an electrophile (I) and a nucleophile (II). θ and φ define the direction. The observed sites are plotted, projected on the sulfide plane. Two distinct regions, (I): $0 < \theta < 40°$, $160 < \varphi < 180°$. (II): $60 < \theta < 80°$, $110 < \varphi < 150°$ (Rosenfeld, Parthasarathy & Dunitz, 1977).

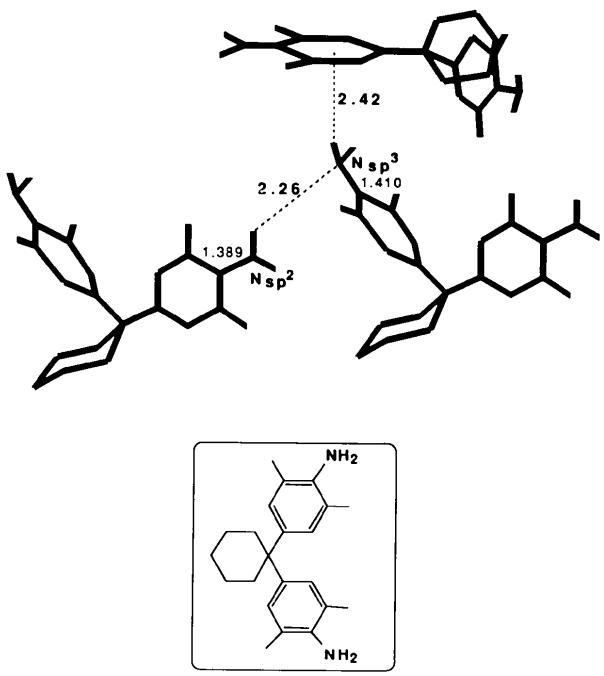


Fig. 2. NH···π interaction (Hanton, Hunter & Purvis, 1992).

Extension to the biological systems

All this structural information restricts the possible binding area around each functional group in the pharmacophore (or in the receptor). This is most useful in drug design as it lowers the number of derivatives to be tested. But is it realistic to extend these findings to the biological situation?

Recently, Tintelnot & Andrews (1989) published a geometrical analysis of the functional-group interactions observed in 18 enzyme–ligand complexes, the X-ray structure of which was known. This small number is due to the fact that only a few structures of complexes were known with sufficient accuracy. A summary of their results is shown in Fig. 3. For each functional group, there is a good equivalence between intermolecular geometries observed in those complexes, and the binding of the same groups deduced from statistical studies in small molecules.

A remarkable study has been carried out recently by Klebe (Klebe, 1994): like his predecessors, Klebe made a thorough search of the Cambridge Database for several functional groups found in protein sequences. For each one he studied the crystal-field environment and he drew histograms of their preferred interaction geometries. By similar searching, he observed the distribution of the environment for the same groups found in ligand–protein complexes structures selected among the most precise ones. The agreement between the histograms represented in Fig. 4 is striking. These geometrical restraints can be translated into rules which may serve as guidelines in drug design, and introduced as the 'composite crystal-field environments' into programs such as *SYBYL* (Tripos Associates, 1992) or *LUDI* (Böhm, 1992*a*,*b*).

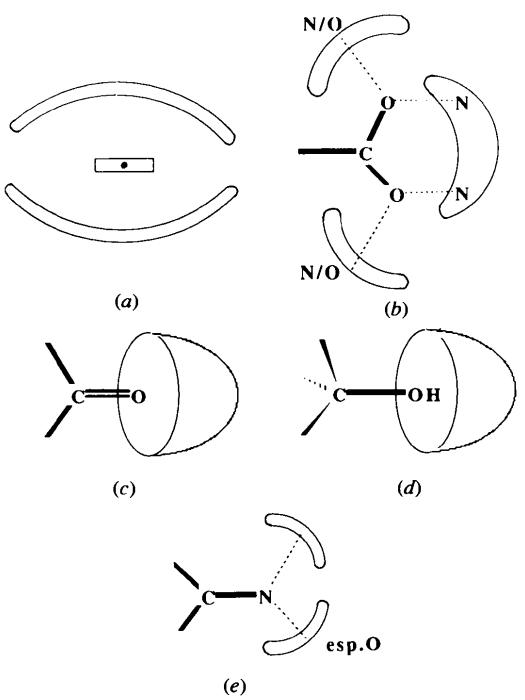


Fig. 3. Environment of the ligand functional groups in enzyme–ligand complexes. The clouds represent the observed surrounding sites. (*a*) phenyl ring seen along its link, (*b*) carboxyl, (*c*) carbonyl, (*d*) hydroxyl, (*e*) amino group (Tintelnot & Andrews, 1989).

Another correlation is provided *a posteriori* by a recent study of the water network in vitamin B_{12} coenzyme structure (Bouquiere, Finney, Lehman, Lindley & Savage, 1993). Steiner & Saenger (1993), using the information they have gathered from their investigation of hydrogen bonds in the environment of water molecules, showed that the water molecules localized in the protein structure are stabilized by coherent CH···O bonds which, therefore, play an important role in the hydration pattern of the coenzyme.

In conclusion, by using these geometrical constraints obtained by systematical analysis of small-molecule crystal structures, it is possible to define more accurately the binding areas of the pharmacophore of a drug.

Conformation analysis

I have shown so far how it is possible to derive the most probable binding geometry around certain functional groups. But what about the conformation of the molecule in the solid state compared to its conformation when bound to a receptor?

Usually the chemist does not trust the crystallographic results obtained in the solid state, and is more inclined to believe the gas-phase conformation obtained through energy minimization, or the conformation in solution obtained by NMR. For very small molecules, *ab initio* calculations can reduce the number of possible conformers to one or two, but as the number of degrees of freedom increases, it is usual to start from X-ray coordinates when available, and search for the conformation with the lowest energy.

It does not seem rational to dismiss the solid-state molecular conformation, because in the crystal the molecule is not isolated as in the gas phase, but is in the presence of other molecules. Consequently, it is confronted with external forces and its conformation will, therefore, be the result of the competition between intramolecular forces and its interactions with its environment. In a crystal cell, the molecule will combine its own shape with the binding requests of its surroundings: hydrogen-bond donors and acceptors, aromatic stacking and ionic bridges. Thus, the observed conformation will be a measure of the molecular flexibility and of the binding forces. Hydrophobic groups will be found in hydrophobic pockets, hydrogen-bond donors and acceptors will interact.

The situation is identical for a receptor–substrate complex, with a compromise between the ligand conformation and that of the receptor as, upon binding, some conformational change of the receptor may sometimes occur. In the crystal packing, the 'receptor'-binding sites are provided by the other molecules, solvent and ions.

Example 1

A recent conformational analysis was performed using a combination of X-ray analysis, NMR and molecular-mechanics simulation.

A cyclic peptidic opioid (Lomize, Flippen-Anderson, George & Mosberg, 1994) was found to have two conformers in solution, differing only at the S–S bridge (Fig. 5). The X-ray structure revealed two molecules per asymmetric unit which are indeed the same two conformers. Energy computations are consistent with two ring conformations which differ by 2.09 kJ (0.5 kcal). (Fig. 6). The flexible chain positions do not correspond to any of the energy minima. But they are controlled by the environment: a tyrosine, necessary for the activity and therefore bound to the receptor, is, in the solid state, tightly bound to the other molecules through hydrogen bonding by way of OH, NH and C=O groups. Its position is consequently the result of the environment (binding).

Example 2

The immunophilin, FKBP12, catalyzes the *cis–trans* isomerization of prolyl amide compounds. Rapamycin is a potent inhibitor of this enzyme. A group at SmithKline Beecham (Holt *et al.*, 1993) synthetized derivatives of rapamycin and of another natural inhibitor named FK506 (Fig. 7). The X-ray structures of the enzyme complexed with three inhibitors were solved, and the position of the inhibitors in the active site determined.

In the structure of the enzyme with compound (1), there are two complexes per asymmetric unit with an identical conformation of the inhibitor in the active site. Furthermore, the X-ray structure of the free ligand (1) exhibits the same conformation as that of the bound inhibitor. The unusual character of this inhibitor is its hydrophobic envelope: there are four carbonyl groups susceptible to act as receptors, and no other hydrogen-bond donors than carbon-bound H atoms. This property, added to the considerable mobility of the macrocycle skeleton should lead to a facile conformational change.

Interested by these results, we performed a molecular-mechanics conformational analysis of this flexible

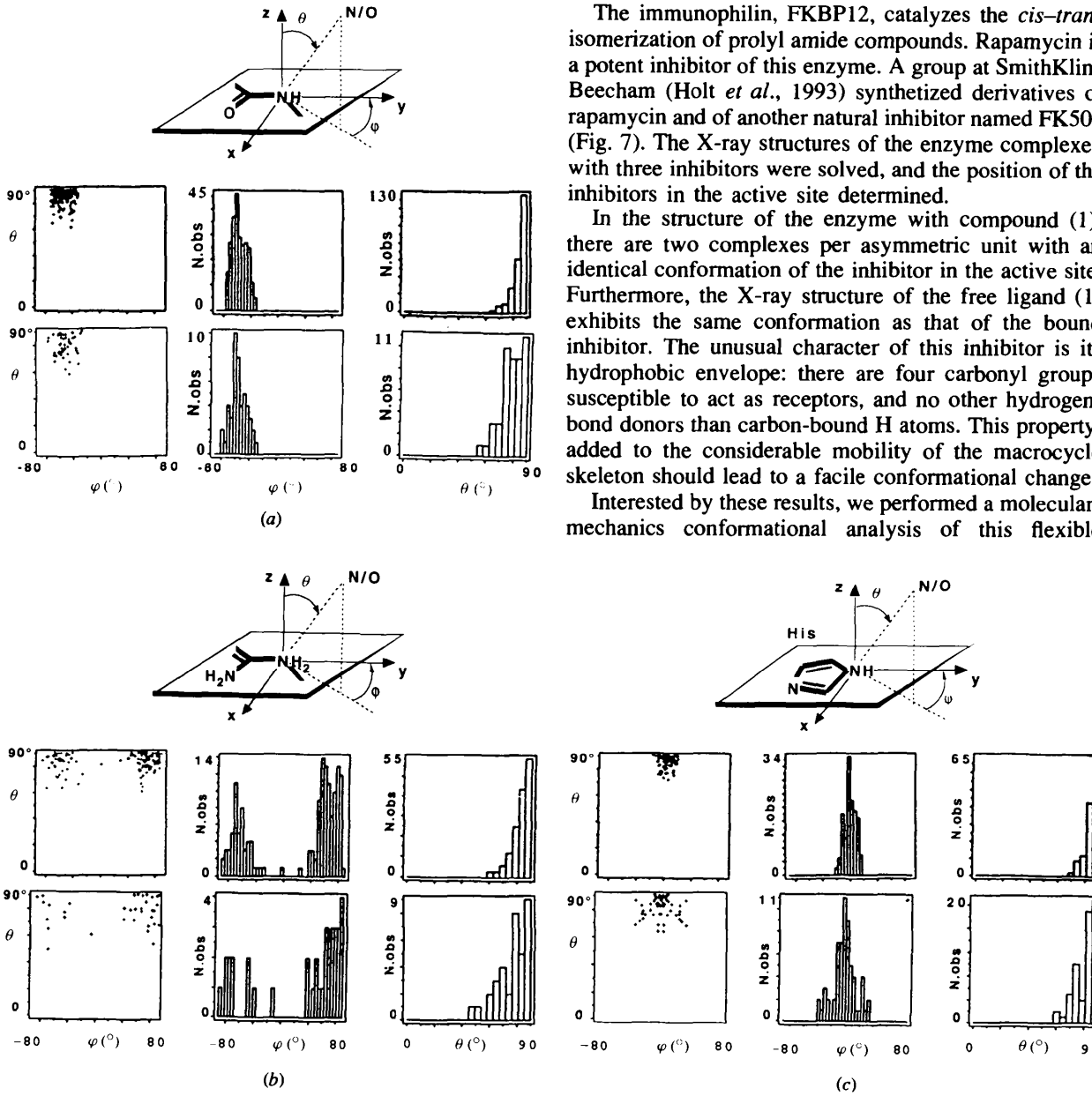


Fig. 4. Distribution of ligand acceptor groups around donor groups after Klebe (1994): correlation diagram θ *versus* φ, histograms of θ and φ. Top row: in small organic molecules; bottom row: in enzyme–ligand complexes. (*a*) Around the NH group of a peptidic backbone one single cluster $\varphi = -60°$ (average) $\theta = 90°$. (*b*) Around the NH group in a guanidinium residue: two clusters: $\varphi = +60$ and $-60°$, $\theta = 90°$, with a preference for *syn* position. (*c*) Around the NH group of an imidazole (histidine): one cluster: $\varphi = 0, \theta = 90°$.

macrocycle; among the eight lowest energies, within 4.5 kJ (~1 kcal) of the global energy minimum, we find two very different conformations with the same energy level, one of them being the solid-state conformation (Fig. 8).

We see that, without any external influence, this macrocycle can adopt many different conformations. But, in the solid-state experiments, the inhibitor is embedded in a molecular environment: the hydrophobic environment of molecule (1) in its crystal structure is strikingly similar to the surroundings of the same molecule (1) in the active site of the enzyme (Figs. 9*a*, 9*b* and 9*c*), with a high number of CH···O bonds (12) as well as carbon–carbon short contacts (13 less than 4.0 Å). Therefore, we suggest that the high number of weak interactions are sufficient to favor one conformation over the others. Then, small-molecule X-ray structure provides, interestingly, a reasonable hypothesis for the bound conformation of the ligand.

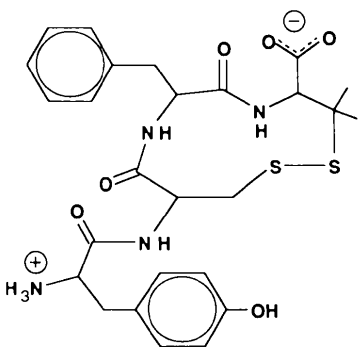


Fig. 5. Tyr-*cyclo*[D-Cys-Phe-D-Pen]OH.

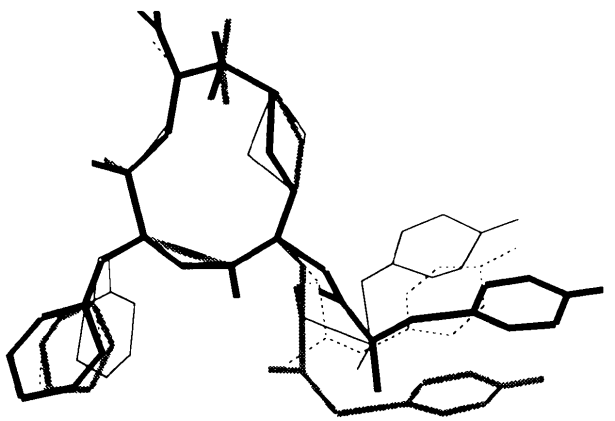


Fig. 6. Superimposition of the two solid-state conformations (I and II) and of the two minimum-energy conformations. Heavy lines (full and dashed): crystal data. Thin lines: computed conformations. Torsion angles χ of the D-Cys, RX (computed) Mol. I: χ_1 = –51 (–57)°, χ_2 = –141 (–148)°, χ_3 = 89 (93)°. Mol. II: χ_1 = 165 (178)°, χ_2 = 144 (150)°, χ_3 = –99 (103)°.

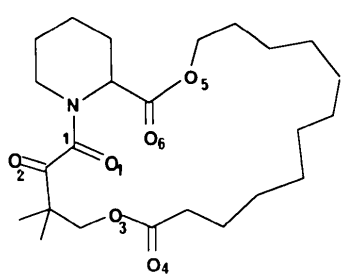


Fig. 7. Inhibitor of FKBP: (21S)-1-aza-4,4-dimethyl-6,19-dioxa-2,3,7,20-tetraoxobicyclo-[19.4.0]-pentacosane.

Example 3

The angiotensin-converting enzyme, or ACE, is an enzyme which cleaves the terminal dipeptide from the angiotensin I to give angiotensin II, a potent vasoconstrictor octapeptide. The enzyme ACE is a zinc metallopeptidase, the crystal structure of which is unknown.

Inhibitors of ACE are antihypertensive agents. Much work, aiming on the design of inhibitors has been carried out since the design of the first potent agent Captopril by Cushman & Ondetti (1980) from Squibb Laboratories. This led to the first model of the receptor site of ACE (Fig. 10*a*). Accumulated knowledge of the chemical and enzymatic properties led to the following requirements: a C-terminal proline, a hydrogen-bond acceptor group, *S* configuration of the peptide, a group liganding the zinc, and a hydrophobic group located in the *S*1 pocket. The relative position of the zinc ligand is not fixed.

In order to locate precisely the antipharmacophore of the ACE active site, Marshall (Mayer, Naylor, Motoc & Marshall, 1987) published a complete conformational study on 28 semi-rigid active molecules and two inactive ones with very different backbones. The geometrical data were taken from the CSD.

Recent work by Codding (Hausin & Codding, 1990) illustrates quite well the new actual approach, making

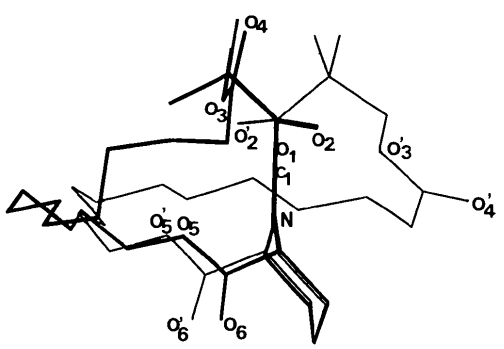


Fig. 8. The two lowest energy minimum conformations of FKBP were superimposed with the fragment (N—C1=O1) in common. This fragment is represented perpendicular to the figure. The two conformations, of same energy: 213.9 kJ, (global energy minimum: 212.5 kJ) are on each side of this common fragment (thick line = crystal data). Note the extreme positions of the O atoms O_2, O_3, O_4 and O'_2, O'_3, O'_4.

use of the more recent structural information in the CSD on the zinc coordination. With these new values (equivalent to those measured in the crystal structures of some metalloproteases), introduced in a structural and conformational study of ACE pseudopeptidic inhibitors, Codding could propose a new relative position of the zinc ligand.

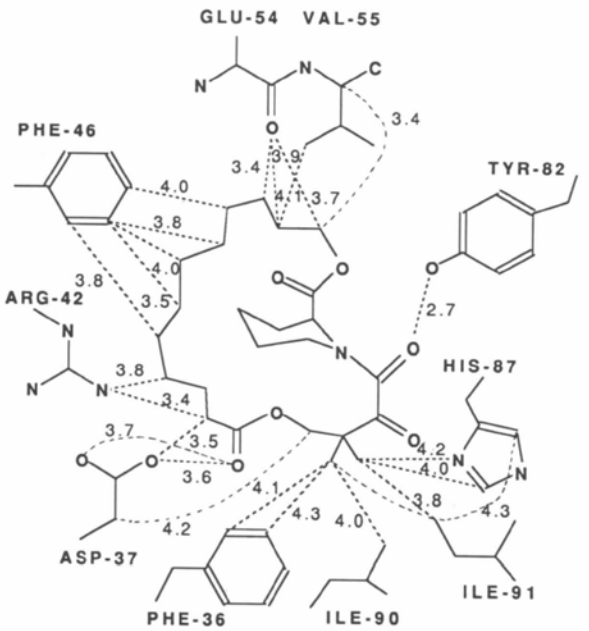
We had the opportunity to work on another series of inhibitors (Vincent, Rémond, Portevin, Serkiz, Laubie, 1982), the general backbone of which is represented Fig. 10(*b*). We solved the structure of one of them: perindoprilat, and we compared it to the previously published X-ray structures of seven inhibitors with similar pseudopeptidic skeletons (Pascard *et al.*, 1991) (Fig. 11). We noted that the conformations of all these molecules are very similar, with very good constancy in the torsion angles. This results in a unique orientation of the two carboxyl groups on each side of the amide plane. The C terminus and the zinc ligand make a relative dihedral angle of 140°.

This single conformation observed in eight crystal structures cannot be accidental, nor is it explained by packing considerations: indeed, three of these inhibitors have crystallized in different unit cells with different solvents: ethanol (perindoprilat), methanol (ramiprilat) (In *et al.*, 1986; Precigoux, Geoffre & Leroy, 1986) and water (elanaprilat) (Patchett, Harris & Tristram, 1980). Nevertheless, the binding directions are identical, involving the same functional groups (Fig. 12).

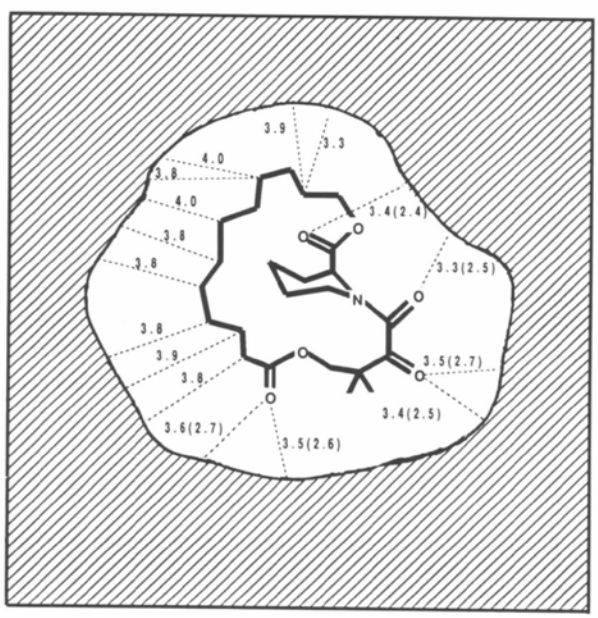
It does not seem rational to explain that this single conformation, observed in three different crystallographic unit cells with three different solvent molecules (some showing disorder), results from the external forces. Therefore, we assume that the observed conformation of that type of inhibitor is most probably its enzyme-bound conformation. This example stresses the importance of having several structures of a series of compounds with comparable backbone which enable us to select the constant conformation of a flexible molecule when it is interacting with binding sites.

Example 4

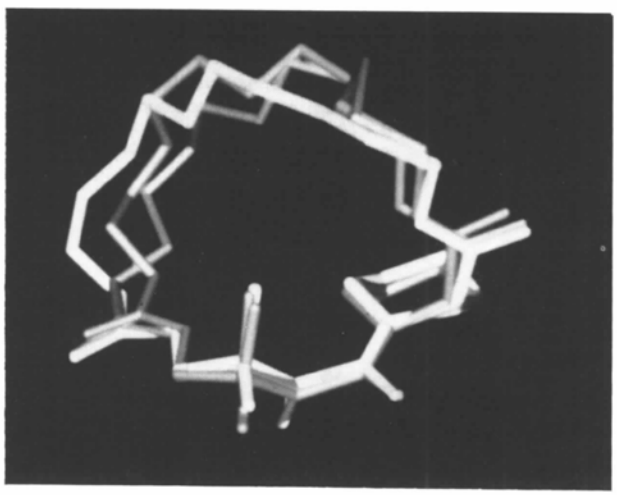
My last example will deal with drugs possessing anticonvulsant, anxiolytic and sedative properties, and interacting with the central nervous system. Their receptor is an oligomeric protein, the structure of which is



(*a*)



(*b*)



(*c*)

Fig. 9. The environment of the inhibitor of FKB12 (Holt *et al.*, 1993). (*a*) In the enzyme; (*b*) in its crystal structure; (*c*) superimposition of the unbound macrocycle (larger ribbon) and of bound ligand in the two complexes of the asymmetric unit (courtesy of Professor Egglestone).

unknown. The first relationship between structure and activity in this field comes from Camerman & Camerman (1972) who proposed, with the first crystal structure of a benzodiazepine (BDZ), a common pharmacophore fitting very well with the BDZ family (Fig. 13*a*).

Many BDZ derivatives were synthetized, but later, several new drugs, the β-carbolines (Fig. 13*b*), chemically unrelated to BDZ, were found to interact with high affinity with the BDZ receptor. They often display different pharmacological properties, ranging from full agonist to inverse agonist (Allen *et al.*, 1988). It has been assumed the biological response is related to the different receptor conformations they have induced. Agonist, inverse agonist and antagonist must have a common binding site for one part of the molecules, while the different properties arise from the second part of the binding groups.

It seems at first sight that very different molecules would simplify the determination of the common features of the pharmacophore. In fact, in this case, the molecules are so different that it is very difficult to superimpose the putative binding sites, and this has been the source of a considerable amount of work in molecular modeling and in crystallography.

The characteristics of the β-carboline family are: two aromatic rings which form a planar molecule, an indole NH, a pyridinic N atom, and a substituent in position 3 acting as hydrogen-bond acceptor. There are more hydrogen-bond acceptor sites than hydrogen-bond donors. How does such a molecule operate its binding to the environment?

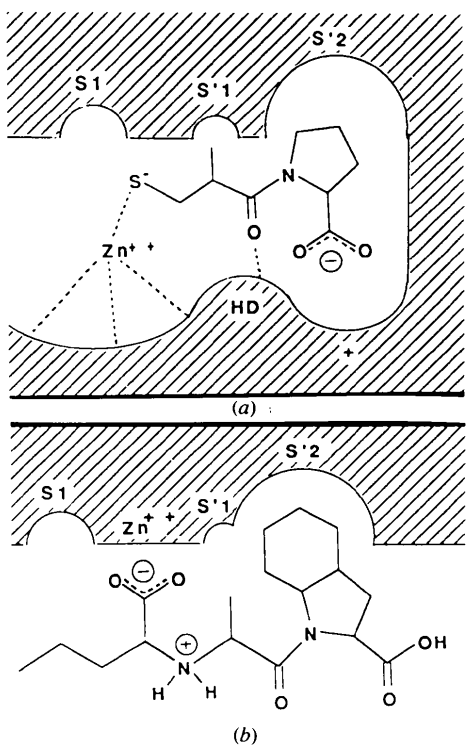
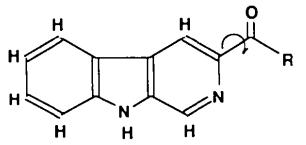


Fig. 10. (*a*) Inhibitor captopril in ACE active site. (*b*) Inhibitor perindoprilat.

Table 1. *Observed conformations of the side chain of β-carboline*

Conformation around (N)—C—C(=O) bond. From the Cambridge Structural Database.



	R = O—CH_3*	R = NH—CH_2—CH_3†
No. of entries	77	21
cis	55	1
trans	22	17‡
ΔE (kJ mol^{-1})	16.7	41.8

* β-CCM.
† β-CEA.
‡ A Previous search through the CSD by P. Codding (Muir & Codding, 1985) gave a proportion of 15 *trans* conformations out of 17 entries.

The crystal structures of two inverse agonists: β-CCM (Muir & Codding, 1985*a*; Bertolasi, Ferreti, Gilli & Borea, 1984) and DMCM (Bertolasi, Ferreti, Gilli & Borea, 1990) give clues on the possible binding to a receptor (Figs. 14 and 15). In the two structures, a three-center hydrogen bond exists between the indolic N atom, an O atom of the carbomethoxy group, and the pyridinic N atom alone. Moreover, there is evidence in the two structures of a CH···O interaction. The indole NH is a good hydrogen-bond donor; the N···HN bond is the minor bond of a three-center bond, the major one being NH···O. Thus, the pyridinic N atom is a good acceptor and should be included in the possible acceptor sites of the receptor hydrogen-bond donor. In both cases, the lack of hydrogen-bond donors constrains the structure to be stabilized by CH···O bonds. The C9 of the pyridine ring is willing to give its H atom, enhanced by the neighbouring position of the pyridinic N atom. This shows clearly in the crystal structures. This weak interaction will be useful in situating and fastening the drug inside the active site once the strong binding has taken place.

The conformation of the COOMe group is, however, different between the two crystal structures: *cis* in β-CCM and *trans* in DMCM. This calls for a comment: in the crystal structure of β-CEA, a carboline carbamoyl derivative (Muir & Codding, 1985*a*,*b*), the carbonyl group stands in *trans* position. Careful searches in the CSD showed that this was the preferred conformation taken by a CONH side chain by 15 structures out of 17.

We went over the CSD for the orientation of the COOMe side chain, and we found that the distribution was not so evident: there was only a majority of *cis* (63%) (see Table 1): the energy difference between the two conformations is 41.84 kJ mol^{-1} (10 kcal mol^{-1})

in the CONH case (β-CEA), and only 16.74 kJ mol^{-1} (4 kcal mol^{-1}) in the carbomethoxy case. In the latter case, it is reasonable to assume that there is an equilibrium in solution, and that the conformation is stabilized upon binding.

The crystallographic results on these two molecules and on other crystal structures of antagonists and agonists have been interpreted by Gilli & Borea (1991) to represent the antipharmacophore (Fig. 16) as the geometrical locus of the 'receptor'-binding sites localized in the crystal structures.

Influence of the medium

The influence of the medium on a molecular conformation can be studied in a small-molecule crystal structure. For example, the competition in binding between solvent and molecules is demonstrated by the packing differences of helical peptides (Karle, Flippen-Anderson, Sukumar & Balaram, 1992). A 15-residue apolar peptide crystallizes in different crystal forms, differing in the extent of hydration, in which completely parallel packing of helices occurs in one and antiparallel packing occurs in the other.

A synthetic receptor (Fig. 17*a*) crystallizes in two crystal forms with a high content of water in both, but with different crystal shapes (Figs. 17*b* and 17*c*) (Cesario *et al.*, 1993). We observe a 'hydrophobic collapse' (Rich, 1993), as result of the rearrangement of the negatively charged groups corresponding to a difference in the aqueous environment, disposed in a honeycomb or in layers.

The behaviour of cyclosporine (CsA) seems *a priori* to come against the general idea of this paper, with very different conformations between its crystal structure and in protein complexes. CsA is a well known

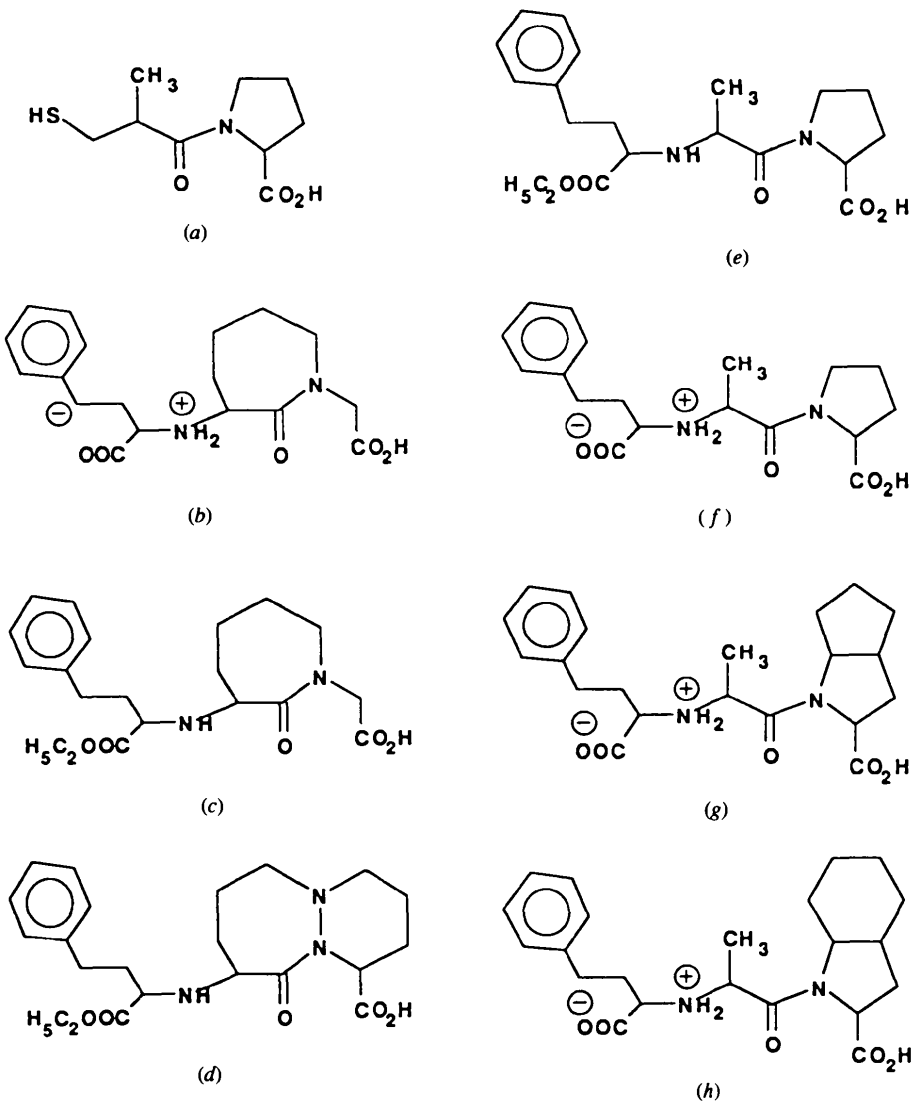


Fig. 11. ACE inhibitors with comparable backbone. (*a*) Captopril; (*b*) seven-member ring (Wyvratt *et al.*, 1984); (*c*) eight-member ring (Paulus, Hennings & Urbach, 1987); (*d*) cilazapril (Thorsett *et al.*, 1986); (*e*) elanapril (Fujinaga & James, 1980); (*f*) elanaprilat (Patchett *et al.*, 1980); (*g*) ramiprilat (In *et al.*, 1986); (*h*) perindoprilat (Pascard *et al.*, 1992).

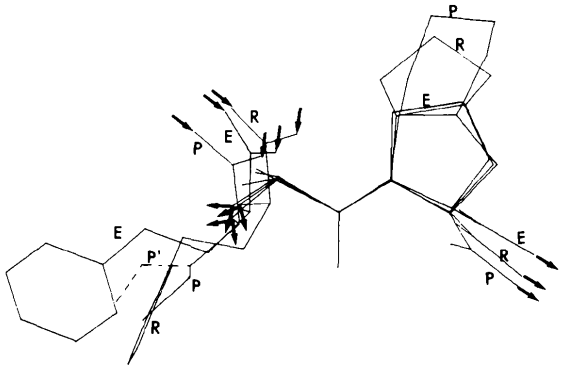
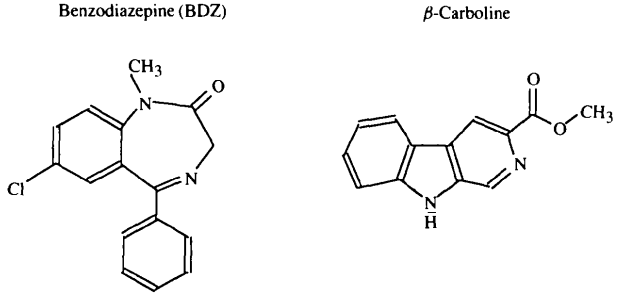


Fig. 12. Superimposition of enalaprilat (E), ramiprilat (R) and perindoprilat (P). The sites of the hydrogen-bond donors and hydrogen-bond acceptors of the 'antipharmacophore' are indicated by arrows.



(*a*) (*b*)

Fig. 13. (*a*) Benzodiazepine: diazepam. (*b*) β-carboline.

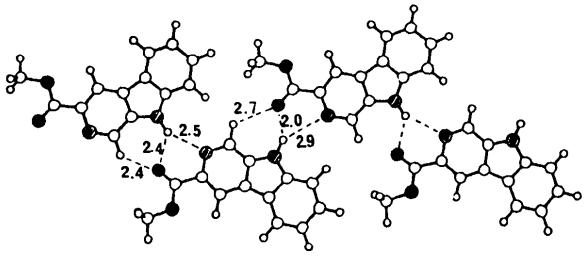


Fig. 14. Packing of β-CCM (Muir & Codding, 1985*a*,*b*; Bertolasi *et al.*, 1984). Short distances (Å) from H atoms to O atoms (black) and N atoms (hatched).

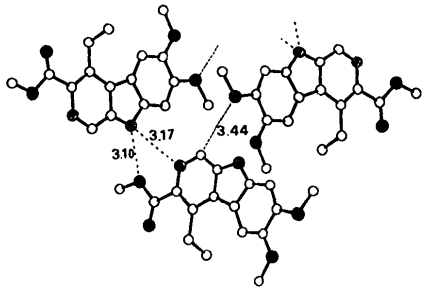


Fig. 15. Packing of DMCM (Bertolasi *et al.*, 1990). Hydrogen bonds with O atoms (black) and N atoms (hatched).

immunosuppressive drug. The crystal structure of the undecapeptide alone (Petcher, Weber & Ruegger, 1976), and the results of the conformational study by NMR in apolar solvent (Loosli *et al.*, 1985) show an identical conformation (Fig. 18*a*). All the polar groups are turned inwards, forming three intramolecular hydrogen bonds.

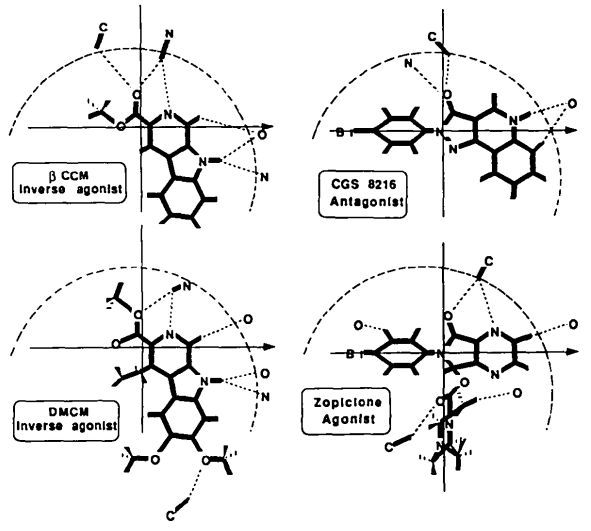
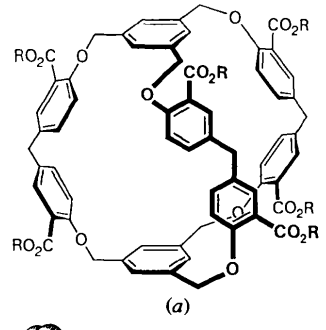
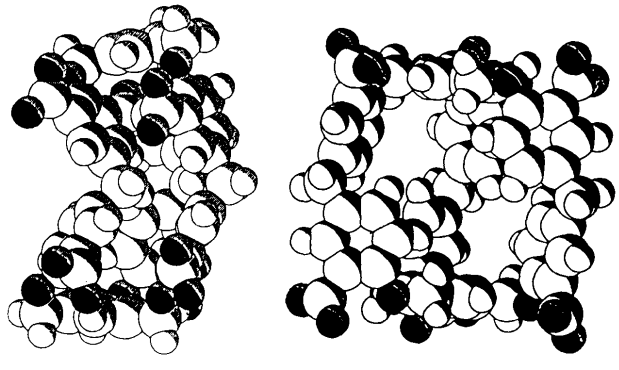


Fig. 16. Proposed binding-site repartition on the β-carboline receptor (Bertolasi *et al.*, 1990).



(*a*)



(*b*) (*c*)

Fig. 17. Macrobicycle (*a*) in two different crystal structures. (*b*) 'collapsed' conformation. Crystal calculated density: 1.32, relative energy = 343 kJ mol^{-1}, solvent: water in layers. (*c*) 'Inflated' conformation. Crystal calculated density: 1.16, relative energy = 423 kJ mol^{-1}, solvent: water in a honeycomb.

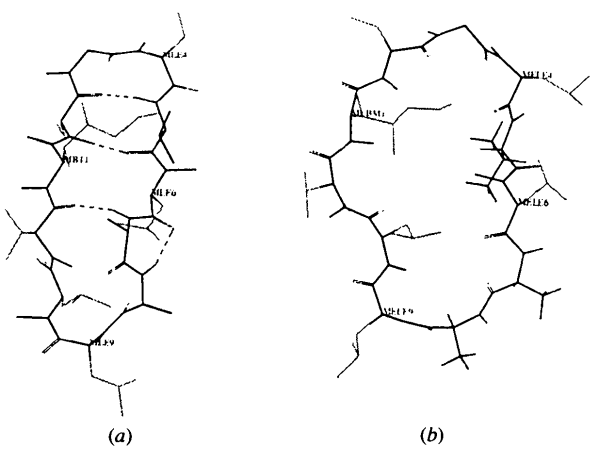
The external envelope is hydrophobic. This was believed to be near the active conformation.

Later, structural studies of complexes: CsA and cyclophilin by NMR in aqueous solvent (Weber *et al.*, 1991), and CsA and an Fab fragment by X-ray analysis (Altschuh, Vix, Rees & Thierry, 1992) show two similar conformations of the complexed CsA (Fig. 18*b*) but very different from the previous ones: not only does a *cis* peptidic bond become *trans*, but all the polar groups are turned outwards, breaking the internal hydrogen bonds, to form bonds with the receptor. The conformational change is so drastic that it is suggested that this bound conformation pre-exists in aqueous solvent (El Tayar *et al.*, 1993). It is then to be expected that the X-ray structure of CsA alone in aqueous solvent would give the answer. However, attempts to crystallize CsA in water have failed.

Concluding remarks

It has been shown that the structural information from small-molecule crystal structures provide accurate geometrical values necessary to design a more precise binding geometry of the ligand in the active site. It has also been shown that, in the solid state, the observed conformation correlates most often with the solution conformation, and is very near the computed energy-minimized conformation. When this is not the case, the environment of the ligand in its crystal cell is a possible image of the binding interactions existing in the active site. Thus, using the crystal conformation is a good start towards a reasonable hypothesis of the active conformation. The close inspection of the stereochemical relations between molecules in their crystal structures by crystallographers such as Duax, Glusker, Codding, Andrews, Karle, Vijayan and many others, brought a new insight to the binding between substrate and receptor.

These reflexions on the importance of the small-molecule crystal structures leading to structural information which can be of fundamental use in biology, are based on a plenary lecture given in Beijing (China) at the XV International Congress of Crystallography, in August 1993.



(*a*) (*b*)

Fig. 18. Cyclosporine CsA: (*a*) unbound (X-ray, and NMR in CCl_4), (*b*) bound to Cyclophilin (NMR).

The author is indebted to Dr Gerhard Klebe for communication of his work prior to publication, to Professor Gastone Gilli, Professor Peter Andrews, to the Institut de Recherche Servier, and to her coworkers: Dr Cesario, Dr J. Guilhem, Dr C. Riche and Dr Tchertanova. The author gratefully acknowledges Professors J. Clardy and D. S. Egglestone for providing Fig. 9(*c*) and Dr J. Glusker for her encouraging comments.

References

ALLEN, F. H., BELLARD, S., BRICE, M. D., CARTWRIGHT, B. A., DOUBLEDAY, A., HIGGS, H., HUMMELINK, T., HUMMELINK- PETERS, B. G., KENNARD, O., MOTHERWELL, W. D. S., RODGERS, J. R. & WATSON, D. G. (1979). *Acta Cryst.* B**35**, 2331–2339.

ALLEN, F. H., DAVIES, J. E., GALLOY, J. J., JOHNSON, O., KENNARD, O., MACREA, C. F., MITCHELL, E. M., MITCHELL, G. F., SMITH, J. M. & WATSON, D. G. T. (1991). *J. Chem. Inf. Comput. Sci.* **203** , 587–613.

ALLEN, M. S., HAGEN, T. J., TRUDELL, M. L., CODDING, P. W., SKOLNICK, P. & COOK, J. M. (1988). *J. Med. Chem.* **31** 1854–1861.

ALTSCHUH, D., VIX, O., REES, B. & THIERRY, J.-C. (1992). *Science*, **256**, 92–94.

ANELLI, P. L., ASHTON, P. R., BALLARDINI, R., BALZANI, V., DELGADO, M., GANDOLFI, M. T., GOODNOW, T. T., KAIFER, A. E., PHILP, D., PIETRASZKIEVICZ, M., PRODI, L., REDDINGTON, M. V., SLAWIN, A. M. Z., SPENCER, N., STODDART, J. F., VICENT, C. & WILLIAMS, D. J. (1992). *J. Am. Chem. Soc.* **114**, 193–218.

BERKOVITCH-YELLIN, Z. & LEISEROWITZ, L. (1984). *Acta Cryst.* B**40**, 159–165.

BERTOLASI, V., FERRETTI, V., GILLI, G. & BOREA, P. A. (1984). *Acta Cryst.* C**40**, 1981–1983.

BERTOLASI, V., FERRETTI, V., GILLI, G. & BOREA, P. A. (1990). *J. Chem. Soc. Perkin Trans.* II, pp. 283–289.

BÖHM, H.-J. (1992*a*). *J. Comput Aided Mol. Design*, **6**, 593–606.

BÖHM, H.-J. (1992*b*). *J. Comput. Aided Mol. Design*, **6**, 61–78.

BOUQUIERE, J. P., FINNEY, J. L., LEHMAN, M. S., LINDLEY, P. F. & SAVAGE, H. F. (1993). *Acta Cryst.* B**49**, 79–89.

BÜRGI, H. B. & DUNITZ, J. D. (1994). Editors. *Structure Correlation*. Weinheim: VCH.

CAMERMAN, N. & CAMERMAN, N. (1972). *J. Am. Chem. Soc.* **94** 268–272.

CESARIO, M., GUILHEM, J., LEHN, J.-M., MERIC, R., PASCARD, C. & VIGNERON, J.-P. (1993). *J. Chem. Soc. Chem. Commun.* pp. 540–543.

CUSHMAN, D. W. & ONDETTI, M. A. (1980). *Prog. Med. Chem.* **17**, 41–104.

DESIRAJU, G. R. (1991). *Acc. Chem. Res.* **24**, 290–296.

DESIRAJU, G. R., KASHINO, S., COOMBS, M. M. & GLUSKER, J. (1993). *Acta Cryst.* B**49**, 880–892.

EL TAYAR, N., MARK, A. E., VALLAT, P., BRUNNE, R. M., TESTA, B. & VAN GUNSTEREN, W. F. (1993). *J. Med. Chem.* **36** 3757–3764.

FUJINAGA, M. & JAMES, M. N. G. (1980). *Acta Cryst.* B**36** 3196–3199.

GILLI, G. & BOREA, P. A. (1991). *The Application of Charge Density Research to Chemistry and Drug-Design*, edited by G. A. JEFFREY & J. F. PINIELLA. New York: Plenum Press.

HANTON, L. R., HUNTER, C. A. & PURVIS, D. H. (1992). *J. Chem. Soc. Chem. Commun.* 1134–1136.

HAUSIN, R. J. & CODDING, P. W. (1990). *J. Med. Chem.* **33** 1940–1947.

HOLT, D. A., LUENGO, J. I., YAMASHITA, D. S., OH, H.-J., KONIALAN, A. L., YEN, H.-K., ROZAMUS, L. W., BRANDT, M., BOSSARD, M. J., LEVY, M. A., EGGLESTON, D. S., LIANG, J., SCHULTZ, W., STOUT, T. J. & CLARDY, J. (1993). *J. Am. Chem. Soc.* **115** 9925–9938.

HUNTER, C. A. (1993). *Angew. Chem. Int. Ed. Engl.* **32**(11), 1584–1586.

HUNTER, C. A. & SANDERS, J. K. M. (1990). *J. Am. Chem. Soc.* **112**, 5525–5534.

IN, Y., SHIBATA, M., DOI, M., ISHIDA, T., INOUE, Y., SASAKI, Y. & MORIMOTO, S. (1986). *J. Chem. Soc. Chem. Commun.* pp. 473–474.

JEFFREY, G. A. & SAENGER, W. (1991). *Hydrogen Bonding in Biological Structures*. Berlin: Springer Verlag.

KARLE, I. L., FLIPPEN-ANDERSON, J. L., SUKUMAR, M. & BALARAM, P. (1992). *J. Med Chem.* **35**, 3885–3889.

KENNARD, O., WATSON, D. G., ALLEN, F. H., MOTHERWELL, W. D. S., TOWN, W. G. & RODGERS, J. (1975). *Chem. Br.* **11**, 213–216.

KLEBE, G. (1994). *J. Mol. Biol.* **237**, 212–235.

KLEBE, G. & DIEDERICH, F. (1995). In preparation.

LOMIZE, A. L., FLIPPEN-ANDERSON, J. L., GEORGE, C. & MOSBERG , H. I. (1994). *J. Am. Chem. Soc.* **116**(2), 429–436.

LOOSLI, H.-R., KESSLER, H., OSCHKINAT, H., WEBER, H.-P, PETCHER. T. J. & WIDMER, A. (1985). *Helv. Chim. Acta*, **68**, 683–704.

MAYER, D., NAYLOR, C. B., MOTOC, I. & MARSHALL, G. R. (1987). *J. Comput. Aided Mol. Design*, **1**, 3–16.

MUIR, A. K. S. & CODDING, P. W. (1985*a*). *Can. J. Chem.* **63**, 2752–2756.

MUIR, A. K. S. & CODDING, P. W. (1985*b*). *Can. J. Chem.* **62**, 1803–1806.

MURRAY-RUST, P. & GLUSKER, J. P. (1984). *J. Am. Chem. Soc.* **106**, 1018–1025.

MURRAY-RUST, P. & MOTHERWELL, W. D. S. (1979). *J. Am. Chem. Soc.* **101** 4374–4376.

MURRAY-RUST, P., STALLINGS, W. C., MONTI, C. T., PRESTON, R. K. & GLUSKER, J. P. (1983). *J. Am. Chem. Soc.* **105**, 3206–3214.

PASCARD, C., GUILHEM, J., VINCENT, M., RÉMOND, G. PORTEVIN, B. & LAUBIE, M. (1991). *J. Med. Chem.* **34**, 663–669.

PATCHETT, A. A., HARRIS, E. & TRISTRAM, E. E. (1980). *Nature* (*London*), **288**, 280–283.

PAULUS, E. F., HENNINGS, R. & URBACH, H. (1987). *Acta Cryst.* C**43**, 941–945.

PEARLMAN, R. S., RUSINKO, A., SKELL, J. M., BANDUCCI, R. & MCGARITY, C. M. (1988). *CONCORD*. Tripos Associates, Inc, 1699 S. Hanley Road, Suite 303, St Louis, MO 63944, USA.

PETCHER, T. J., WEBER, H.-P. & RUEGGER, A. (1976). *Helv. Chim. Acta*, **59**, 1480–1488.

PRECIGOUX, G., GEOFFRE, S. & LEROY, F. (1986). *Acta Cryst.* C**42**, 1022–1024.

RICH, D. H. (1993). In *Perspectives in Medicinal Chemistry*, edited by TESTA, KYBURZ, FUHRER & GIGER. Basel: Verlag Helv. Chim. Acta.

ROSENFELD, R. E. & MURRAY-RUST, P. (1982). *J. Am. Chem. Soc.* **104**, 5427–5430

ROSENFELD, R. E.. PARTHASARATHY, R. & DUNITZ, J. D. (1977). *J. Am. Chem. Soc.* **99**(4), 860–862.

SHARMA, C. V. K., PANEERSELVAM, K., PILATI, T. & DESIRAJU, G. R. (1993). *J. Chem. Soc. Perkins Trans.* II, pp. 2209–2216.

STEINER, T. & SAENGER, W. (1992*a*). *Acta Cryst.* B**48**, 819–827.

STEINER, T. & SAENGER, W. (1992*b*). *J. Am. Chem. Soc.* **114**, 10146–10154.

STEINER, T. & SAENGER, W. (1993). *Acta Cryst.* D**49**, 592–593.

TAYLOR, R. & KENNARD, O. (1982). *J. Am. Chem. Soc.* **104** 5063–5070.

TAYLOR, R., KENNARD, O. & VERSICHEL, W. (1983). *J. Am. Chem. Soc.* **105**, 5761–5766.

TAYLOR, R., KENNARD, O. & VERSICHEL, W. (1984). *Acta Cryst.* B**40**, 280–288.

THORSETT, E. D., HARRIS, E. E., ASTER, S. D., PETERSEN, E. R., SNYDER, J. P., SPRINGER, J. P., HIRSFIELD, J., TRISTRAM, E. E., PATCHETT, A. A., ULM, E. H. & VASSIL, T. C. (1986). *J. Med. Chem.* **29**, 251–260.

TINTELNOT, M. & ANDREWS, P. (1989). *J. Comput. Aided Mol. Design*, pp. 67–84.

Tripos Associates Inc. (1992). *SYBYL, Molecular Modeling System, Version* 5.40, Tripos Associates Inc., 1699 S. Hanley Road, Suite 303, St Louis, MO 63944, USA.

VEDANI, A. & DUNITZ, J. D. (1985). *J. Am. Chem. Soc.* **107** , 7653–7658.

VINCENT, V., RÉMOND, G., PORTEVIN, B., SERFIZ, B. & LAUBIE, M. (1982). *Tetrahedron Lett.* **23**, 1677–1680.

WEBER, C., WIDER, G., VON FREYBERG, B., TRABER, R., BRAUN, W., WIDMER, H. & WÜTHRICH, K. (1991). *Biochemistry*, **30** 6563– 6574.

WYVRATT, M. J., TRISTRAM, E. E., IKELER, T. J., LOHR, N. S., JOSHA, H., SPINGER, J. P., ARISON, B. H. & PATCHETT, A. A. (1984). *J. Org. Chem.* **49**, 2816–2819.