PINSA, **65, A**, No. 6, November 1999, pp. 655-681

PHARMACEUTICAL APPLICATIONS OF NMR SPECTROSCOPY

A V S SARMA AND A C KUNWAR
Indian Institute of Chemical Technology, Hyderabad - 500 007 (India)

(Received 17 June 1999; Accepted 19 July 1999)

Early pharmaceutical applications of NMR spectroscopy were in the identification and characterization of chemically synthesized drugs or biologically active molecules derived from a variety of natural sources, which formed the basis of pharmaceutical products. However, several major advances in the technique has elevated the status of NMR spectroscopy as one of the most important tools in the area of rational drug design. Studies of drug-receptor interactions, more specifically protein-ligand interactions, provide the basic framework for understanding the mode of action of the drugs and in the design of more potent analogues. Immune system is a marvel of nature and immunosuppression by drugs or pathogens are topics of tremendous contemporary interest. Drugs that suppress the immune system are used to prevent the rejection of the organ transplant in major graft surgeries, while human immunodeficiency virus (HIV), the causative agent of the dreaded disease acquired immunodeficiency syndrome (AIDS), plays havoc with the immune system. Several breakthroughs have been made in these two areas of immunology in the recent past. NMR spectroscopy has provided the structure of immunosuppressant drugs, their receptors and the drug-receptor complex, providing handle in deciphering the various steps involved in the immune response. The AIDS drugs inhibiting the enzyme, reverse transcriptase, which is involved in the transcription of the viral RNA to DNA, compete with the natural nucleosides, due mainly to the similarity of their molecular geometries and curtail the formation of viral DNA. HIV protease, a homodimer, is inhibited by small molecules. These drugs have transformed the therapy of AIDS. The structure obtained by NMR spectroscopy has thrown light on the role of flexibility of protease-inhibitor complex in the function of the protease. The development of a vaccine against AIDS is a challenging task due to high mutation rates of the virus. The peptides in main binding site of the HIV neutralizing antibodies have been investigated in detail, to obtain the immunogenic features. The recent applications of NMR in drug design using combinatorial chemistry have been very impressive. The presentation discusses the role of NMR spectroscopy in these specific pharmaceutical applications.

Key Words: NMR; Drugs; Immunosuppressant; AIDS; HIV; Immunophilins; Reverse Transcriptase Inhibitors; HIV Protease Inhibitors; Combinatorial Chemistry

Introduction

Chemical sciences have profoundly influenced the development of modern biology. Biological cofactors, metabolites, hormones etc. have been characterized and synthesized using the tools of organic chemistry. Similarly physical techniques like X-ray crystallography and NMR spectroscopy have been used to elucidate the structure of bio-molecules. The impact of biology in chemistry however, can not be undermined. The complementary relationship between chemistry and biology is very well exemplified in the area of drug discovery where interactions involving the ligands and their receptors play a crucial role. Chemists and biologists have generated medicines of enormous importance to mankind. In spite of the increasing sophistication available to chemists, they are far from matching nature's efficiency in producing complex molecules that carry out the intriguing process of life. Most of the leads of drug molecules have, therefore come from natural biological sources like bacteria, fungi, plants and marine organism. So much so, nearly all 'wonder drugs' in use presently have been derived from natural products. Further refinements and fine tuning is then done by organic chemists and medicinal chemists for effecting desirable changes for an efficacious and useful drug. To understand the biological activity of these molecules a knowledge of the three-dimensional structure of the drugs, receptors and the drug receptor complex is of great importance and NMR spectroscopy has a crucial role to play in such studies[1,2].

Historically applications of NMR in pharmaceutical sciences started in late 1950s. First applications were in assisting in the identification and characterization of chemically synthesized drugs or biologically active molecules derived from

a variety of natural sources. The role of NMR as essentially an analytical tool has now been supplemented by a more fundamental application in the design of new drugs, an outcome of major advances in NMR instrumentation and methodology in recent years. Recent publications of several books, monographs and reviews amply demonstrate the importance of NMR in drug design[3-7].

Real pharmaceutical applications of NMR spectroscopy started in 1960s and 1970s, with the determination of structure and conformation of biologically important molecules with molecular masses upto about 1 kDa. Conformational information is obtained from the chemical shifts (δ), indirect spin-spin couplings (J) and nuclear Overhauser enhancements (*nOe*). The increasing availability of higher field spectrometers and sophistication in instrumentation enhanced several fold the complexity of molecules amenable to NMR study after mid 1970s. The development of two dimensional NMR (2D NMR)[1], a big-leap in NMR methodology, has enabled a wide range of new applications in the area of drug design. These applications have mainly exploited the usefulness of biologically active peptides and peptidomimetics as potential targets for analogue based drug design and proteins in receptor based drug design. The advent of 3D and 4D NMR[8,9] coupled with very high field NMR spectrometers (^{1}H frequency of upto 800 MHz) has extended the size of biomolecules that can be studied at molecular level to about 40 kDa, though recently a homodimer of *Escherichia coli* chemotasin kinase, having a molecular weight of 142 kDa, has been studied[10]. The ease of uniformly labelling the proteins with ^{13}C and ^{15}N has also contributed very significantly in these advancements.

NMR spectroscopy is being used very widely in the area of drug design. Such studies usually aim at obtaining: a) the structure of the drug molecule, its receptor bound conformation as well as its dynamics at the receptor site, b) the structure of the receptor and its complex with the drug, which provides the information on the binding site and functional group interaction participating in the ligand recognition and binding.

Immunosuppressants

Importance of some of the aspects mentioned above is very well demonstrated in the two problems of tremendous interest in contemporary immunology. In this article, we discuss some specific structural studies involving immunosuppressant drugs and drugs for the treatment of acquired immunodeficiency syndrome (AIDS) in detail. Most of the discussion that follows pertains to NMR data, additionally supported by molecular dynamics calculations [11,12] and X-ray diffraction studies.

The immune system is a marvel of nature. It is not controlled by any central organ and functions through an exotic information network which is constantly fighting and engaged in imperceptible war against foreign invaders, like bacteria, viruses, fungi and parasites. Already a large body of information exists about the immune system and more and more fascinating details are becoming available. Immunosuppressants, used for preventing the rejection of the transplanted organs in surgery, have played an important role in our understanding of the immune system. Tremendous efforts in finding a cure for AIDS, the dreaded disease, caused by infection with human immunodeficiency virus (HIV) are being made by scientists all over the world for several years. Understanding the finer details of the functioning of the immune system may help in the prevention and cure of AIDS.

Immunosuppression has enormous medical relevance. The success of organ transplantation surgery is based to a large extent on the miracle drug cyclosporin A (CsA)[13] (Fig. 1A) obtained from fungal sources. With the use of CsA we can traverse the fine line between the graft rejection and infection of the patients due to a necessary suppression of the immune system. CsA binds with high affinity (dissociation constant, K_d=30 nM) to an abundant cellular protein ($\approx$ 5 nM, upto 1-2% of the total cellular protein), called cyclophilin (CyP)[14]. Subsequently another drug, FK506[15] (Fig. 1B), which is about 10 times more active than CsA, was discovered. FK506 binds to another ubiquitous protein, referred to as FK binding protein (FKBP)[16,17], with very high affinity (K_d =0.4 nM). Rapamycin[18] (Fig. 1C), which is structurally very similar to FK506, is another immunosuppressant compound (not yet approved as a drug) with even larger activity than FK506 and binds with greater affinity (K_d = 0.2 nM) to FKBP. Both cyclophilin and FKBP are enzymes (referred to as rotamases) with peptidyl-prolyl *cis-trans*

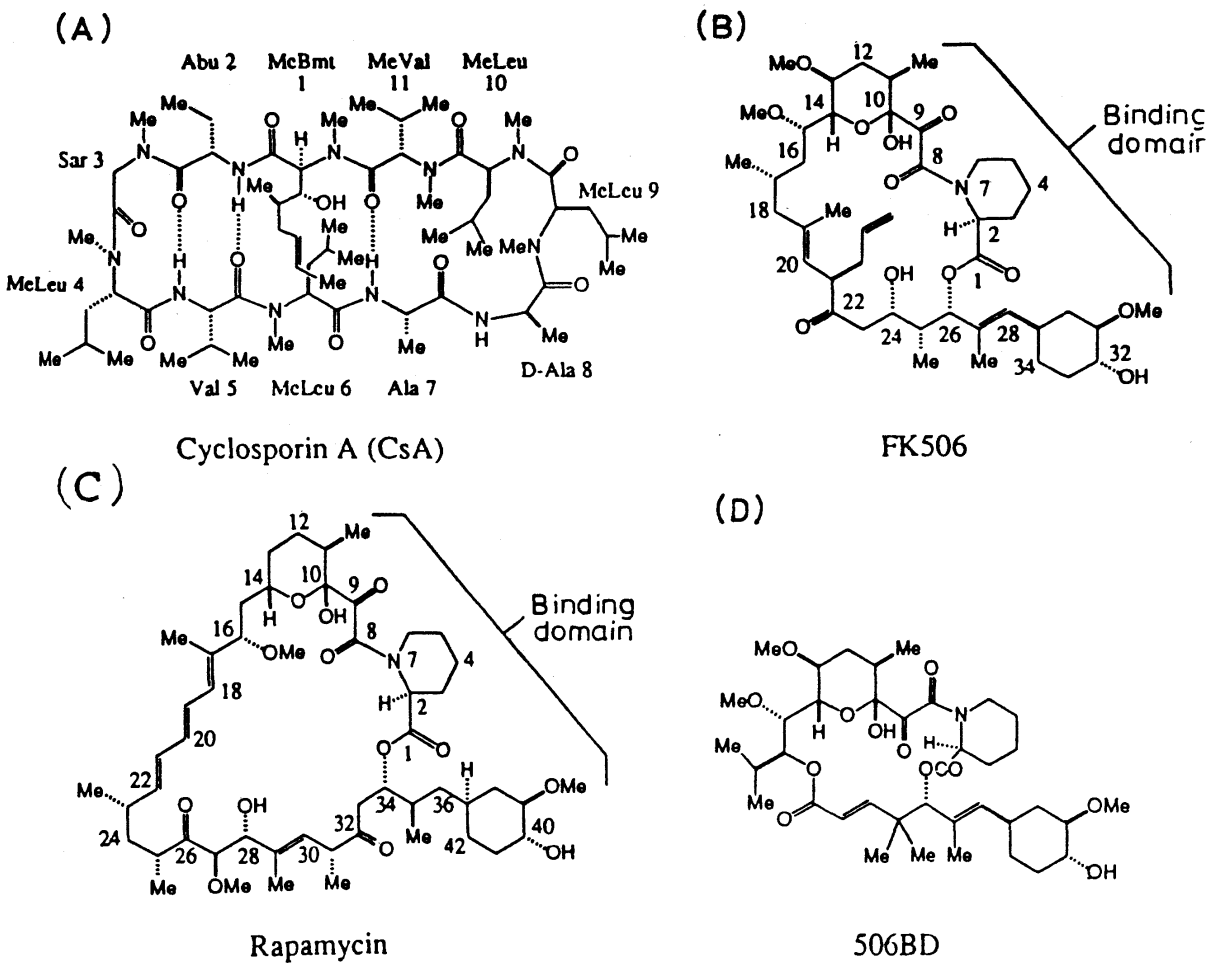

Fig. 1 Chemical formula of (A) cyclosporin A (CsA), (B) FK506 (C) rapamycin (D) 506BD. Binding domain of FK506 and rapamycin is shown. The region opposite to the binding domain is the effector domain.

isomerase activity [14,16,17]. The enzymatic activity of CyP is inhibited by CsA (inhibition constant, K_i=6 nM)[14] and that of FKBP by FK506 (K_i=0.4 nM)[14,15]. These enzymes collectively called immunophilins speed up the isomerization of peptide-proline bonds in polypeptides and thereby facilitate protein folding. Immunophilins may be thus helping to refold proteins that might have changed their conformation during transport through membranes[19-21]. Importantly the binding of these active molecules blocked the isomerase activity of immunophilins, thus inhibiting their normal cellular functions. The availability of these drugs has led to a better understanding at the molecular level of the biochemical signal transduction involved in the immune response. All these molecules inhibit cell proliferation, but exert their actions at different stages of T cells activation. CsA and FK506 inhibit Ca^{2+} dependent signalling pathways, which result in the prevention of transcription of several genes, including those encoding for interleukin-2 (IL-2), which are normally activated by stimulation of T cell receptors (TCR). Rapamycin, on the other hand, intervenes in the Ca^{2+} independent pathway, interfering in events more closely related to DNA synthesis, by inhibiting the differentiation and proliferation of T cell, normally mediated by the IL-2 receptor (IL-2R)[22].

The similarity in biological functions, coupled with the inhibition of two different rotamases by these drugs, resulted in a belief that the immunosuppression is due to the inhibition of rotamase. Strong evidence against the "rotamase hypothesis" came from rapamycin and another structurally similar molecule 506BD[23] (Fig. 1D), a synthetic ligand which binds to FKBP with high affinity (K_d=20 nM). Due to their structural

similarity with FK506, rapamycin and 506BD bind tightly to FKBP and inhibit the rotamase activity of the enzyme (K_i=0.2 nM for rapamycin and 5 nM for 506BD)[23,24]. However, while other drugs lead to immunosuppression, 506BD does not, implying that inhibition of rotamases is not enough for a molecule to be immunosuppressant.

It was subsequently shown that calcineurin (CN), also called protein phosphatase 2B, a calmodulin dependent serine/threonine protein phosphatase[25], possesses all the predicted biochemical properties of a target for both CyP-CsA and FKBP-FK506 complexes and gets specifically inhibited by them[26]. CN is a heterodimeric protein composed of two subunits calcineurin A (CNA), which contains the calmodulin-binding and phosphatase active sites and calcineurin B (CNB), which is a Ca^{2+} binding protein. Calmodulin binding to CNA enhances the phosphatase activity of the enzyme by about ten times. The picture emerging from a series of elegant experiments shows that, when FKBP-FK506 and CyP-CsA encounter the target protein CN, the exposed domain of the complex binds to it. Thus in the pentapartide complex, drug-immunophilin-CNA-CNB-calmodulin, CsA and FK506 behave like molecular glue. Structural dissimilarities of CsA and FK506 as well as the unrelated sequences of CyP and FKBP make these results all the more remarkable. To understand the basis of molecular recognition within these pentameric complexes some detailed structural analyses have been undertaken and lot more experiments are still required to be performed. The question pertaining to the role of the complex has been addressed later. A series of elegant NMR experiments have provided information on the ligand receptor interaction which are additionally supported by X-ray diffraction and molecular dynamics (MD) calculations. The discussions on the structure in this paper represent overall view from these techniques.

Structure of Cyclosporin A, Cyclophilin and their Complex

Cyclosporin A (Fig. 1A) is a neutral cyclic undecapeptide consisting of several unnatural amino acids, seven of them N-methylated. All the amino acids are lipophilic, making it water insoluble. A large number of synthetic and natural analogues have been used for structure-activity relationships. Kessler *et al.* have performed elegant NMR studies on CsA[27-34]. In solvents of low-polarity like $CDCl_3$, C_6D_6 and THF-d_8[27-31,34], CsA takes very similar conformations. On the other hand, in solvents of higher polarity, the changes in the conformations are directly visible in the NMR spectra; many conformations exist in equilibrium inter-converting slowly on the NMR time scale. In DMSO, at least seven conformations can be observed[30]. Because CsA is insoluble in water, as a first step in understanding the function and interactions of CsA with CyP, detailed structural studies have been undertaken in $CDCl_3$. The first such study of cyclosporin[27] could not clear some ambiguities and subsequently some very high quality spectra were obtained at 600 MHz[33]. Using 117 distance constraints, structure refinement and modelling studies revealed a structure shown in Fig. 2 as a stereo plot. The structure is very similar to that obtained from X-ray diffraction[29] with some of the notable features as:

1 The amide bond between MeLeu 9 and MeLeu 10 is *cis*.
2 All four NHs are involved in internal H-bonds, with Val 5 NH and Ala 7 NH partially participating in three center H-bonds (H-bonds are very similar to that obtained from crystal structure). The MeBmt(4(R)-4-[(E)-but-2-enyl]-4,N-dimethyl-L-threonine) side chain is folded over the backbone, like the crystal structure, as against the earlier NMR observation, where extended structure was obtained and MeBmt 1 OH is involved in H-bond with the MeBmt 1 CO.
3 Existence of the type II' β turn involving Abu 2-Sar 3-MeLeu 4-Val 5 was definitely confirmed, unlike earlier study where such details could not be obtained.
4 The residues MeVal 11-MeBmt 1-Abu 2-Sar 3-MeLeu 4-Val 5-MeLeu 6-Ala 7 form a β sheet structure.
5 Six of the seven N-Me point outside the ring and thus expose to the solvent. Only MeVal 11 is directed towards the centre of the macrocycle.
6 The residues 7-11 form an irregular structure stabilized by hydrogen bond between D-Ala 8 NH-MeLeu 6 CO and several van der Waals interactions.

Several studies on the structure of CyP have been undertaken in aqueous solutions[35-38]. CyP

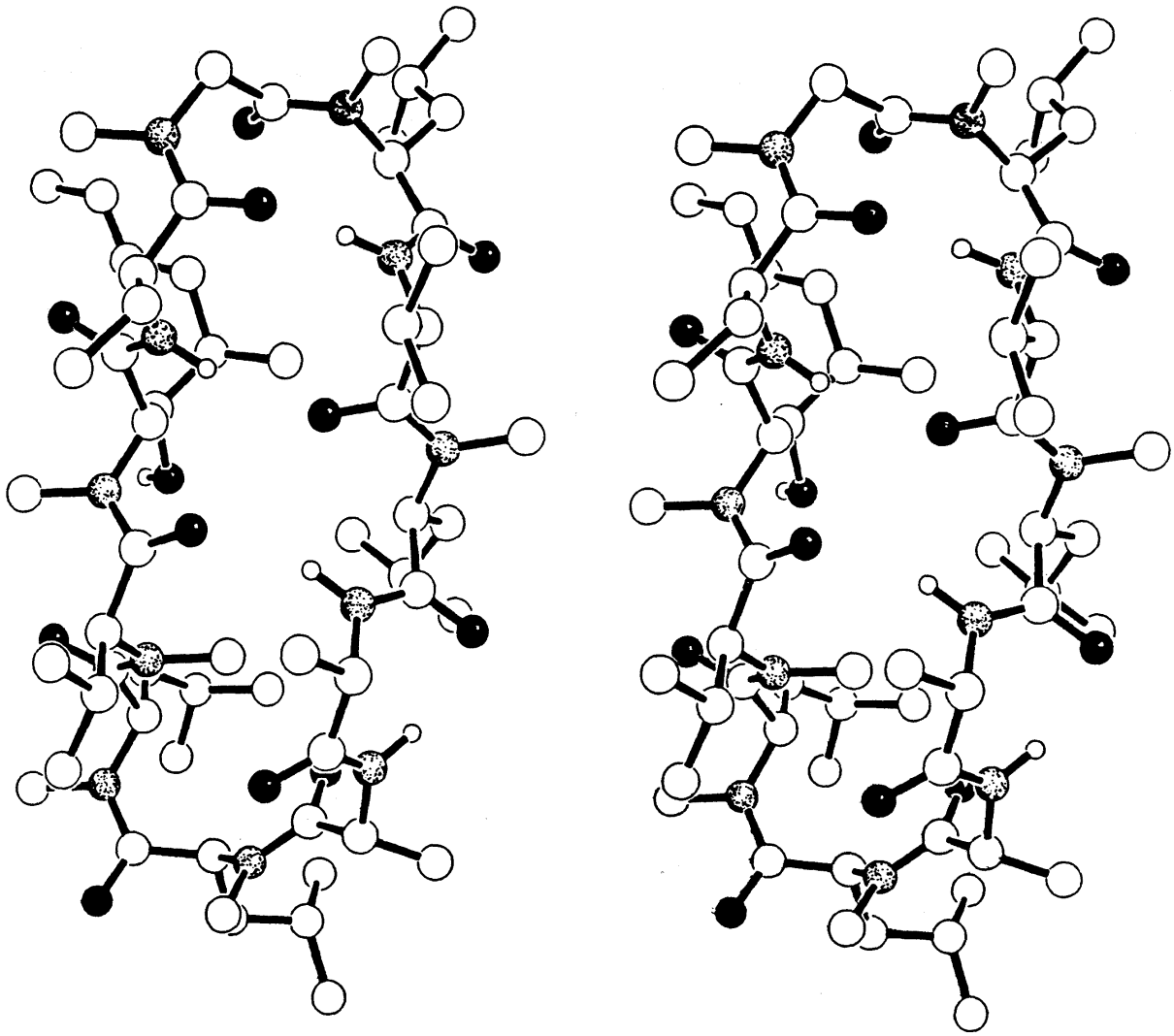

Fig. 2 Stereoplot of CsA structure obtained by restrained molecular dynamics calculations. The type II' β turn between Abu 2 and Val 5 is at the top of the molecule. The O-atoms are filled and the N atoms are stripled (Adapted with premission from H Kessler *et al. Helv Chim Acta* 73 (1990) 1669 Copyright ©1990 by the Neuc Schweizerische Chemische Gesellschaft

consists of single polypeptide chain with 165 amino acids having a molecular mass of 17.8 kDà. It is a roughly spherical molecule with a radius of about 17Å. CyP has two α helices and eight stranded β-barrel that has a +1,-3,-1,-2,-2,-3 topology (Fig. 3B). The barrel consists of two roughly perpendicular four stranded β sheets connected by short junctions at residues Leu 98-Gly 13 and Phe 53-Ile 156. Inside the barrel, a tightly packed core contains most of the hydrophobic side chains. Other hydrophobic residues are located in the contact region of the two amphipathic helices with the β-barrel and in the CsA binding site (to be discussed later).

There is a structural similarity between CyP and the super family of β-proteins family[39] involved in ligand transport including retinol binding protein (RBP)[40], bilin binding protein[41], insecticyanin[42] and β-lactoglobulin[43]. Most of these molecules encapsulate their ligands in the β barrel core. By contrast the barrel with hydrophobic residues and the ligand binding site in CyP is outside the barrel. The topology of cyclophilin also differs from the simple $(+1)_n$ up-and-down fold in RBP class of proteins or the (-3,+1,+1) Greek key topology that is most frequently found in antiparallel β-barrel proteins. The two cross over connections Gly 64-Ile 97 and Thr 116-Gly 130 (Fig. 3C), in particular

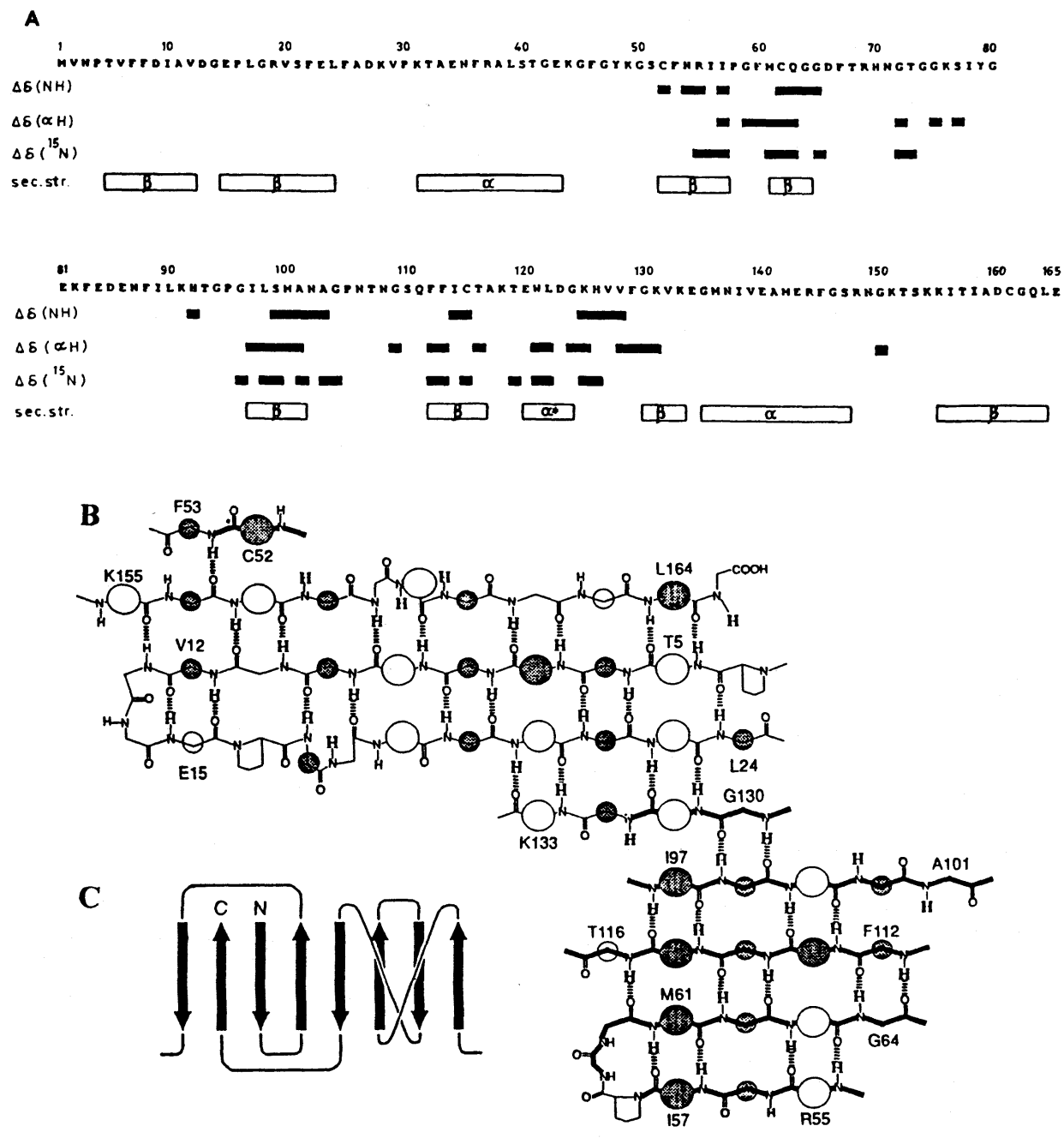

Fig. 3 Summary of NMR results with CsA binding site and the secondary structure and topology of CyP (A) Amino acid sequence, chemical shift data used for deleniating CsA binding site and the regular α-helical and β-sheet structures (α^* depicts a short α helix or a turn like structure). The residues displaying significant backbone chemical shift variation on binding of CsA are shown with black bars. (B) Schematic representation of the antiparallel β-sheet in CyP showing a global (+1,-3,-1,+1,-2,-3) topology of the β-sheet. (C) Drawing showing the sequence positions of the individual β-strands, the hydrogen bonds, the distribution of the hydrophobic side chains and the location of residues with large chemical shift changes on binding to CsA. The presence of hydrogen bonds is identified with dashed lines. The amino acid side chains are represented with circles at C_α carbon positions using the following notation: big circles, side chains toward the reader, small circles, side chains pointing away from the reader; shaded circles, hydrophobic residues; empty circles, polar and charged residues. For Pro a five-membered ring is drawn. Ala, Gly, Tyr and His have no symbol at the C_α position. Essentially all side chains below the plane of the β sheet are hydrophobic (adapted with permission from J Kallen *et al. Nature* **353** (1991) 276. Copyright © 1991 by the Macmillan Magazines Ltd)

represent an unusual topological feature. As both loops lie on the outside of the barrel, the (-2,+1,-2) topology requires that the two loops cross each other. The unusual left handed connection of the Thr 116-Gly 130 loop may be rationalized by the length of the loop which may accommodate a variety of local conformations. Fig. 4, shows a ribbon diagram of CyP clearly depicting the secondary and tertiary structure.

The study of CyP-CsA complex in solution shows very small changes in the structure of CyP on complexation with CsA[44]. The CyP-CsA complex structure is also similar to that observed in the solid state for CyP[45] and CyP-tetrapeptide complex[35]. The structure of the complex is stabilized by several hydrophobic interactions. Fig. 3A depicts the residues which show significant chemical shift variation on binding to CsA. Chemical shift differences are indicated by black bars when they exceed the limits $\Delta\delta$(NH) = 0.1 ppm $\Delta\delta(\alpha H)$ = 0.05 ppm and $\Delta\delta(^{15}N)$ = 0.5 ppm. The side chains of CyP residues Trp 121, Phe 60, Ile 57, Leu 122, Phe 113, His 126, Ala101, Ala 103 and Thr 73 form a hydrophobic pocket that interacts with the hydrophobic surface of the CsA residues 9-11 and 1-3. A large number of intermolecular hydrogen bonds also stabilize the structure of the complex. The NMR data show hydrogen bonds between Arg 55(ηNH)-MeLeu 10(CO), Trp 121(εNH)-Leu 9(CO), and Asn 102(CO)-Abu 2 NH. The Cα-C' bond of MeVal 11 appears to mimic a twisted amide bond, which is proposed as a mechanism of the rotamase catalysis of CyP and will be discussed in greater detail later.

The structure of CsA in the complex was revealing[46]. This information has been obtained by various groups[47-49]. The earlier studies used isotope labelling [47,48], to suppress the spectral lines due to CyP, by either filtering them by heteronuclear editing[50] or uniform ^{2}H labelling of the CyP[49]. These manipulations provided information only on the bound CsA which of course agreed with complete structure of CsA-CyP complex obtained subsequently[46]. The isotope labelling followed by

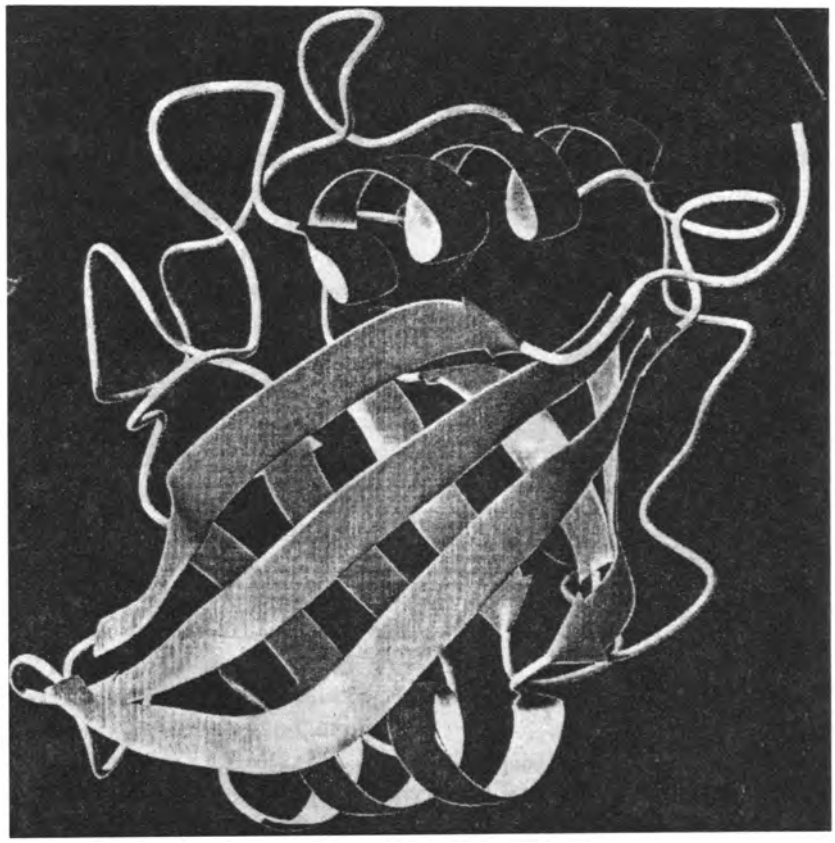

Fig. 4 A ribbon diagram of cyclophilin (CyP) (adapted with permission from K Wuthrich *et al. J Mol Biol* **272** (1997) 64 Copyright © (1997) Academic Press Limited)

heteronuclear editing experiments demonstrate the power of NMR spectroscopy for drug-receptor studies. In an unlabelled system the large number of protons from the receptor protein would interfere with the observation of the resonance lines of the ligand. Binary complexes are ideally suited for studies with efficient labelling schemes, as the two components can be labelled separately with ^{13}C or ^{15}N before complex formation. Suitably designed sequences can be used for heteronuclear editing to separate the lines from the two components of the complex. Complete deuteration of one of the molecules in the complex enables the editing in a rather obvious and simple manner. Such studies represent an attractive avenue for the use of NMR and related fields on drug design. A very important finding emerged that showed the structure of CsA is strikingly different from that in $CDCl_3$ solution as well as the one obtained in the crystalline state by X-ray diffraction. The salient features of the structure shown in Fig. 5 are:

1 The bound CsA lacks regular secondary structure; all amide bonds taking a *trans* geometry.
2 It has no intramolecular H-bonds.
3 But for MeLeu 9 and MeVal 11, rest of the five NMe groups are directed towards the centre of the macrocycle enabling them to make large number of van der Waals contacts. This implies that compared to the free CsA, the bound CsA structure has essentially been turned inside out.
4 Ten of the C=O and four of the NH are either

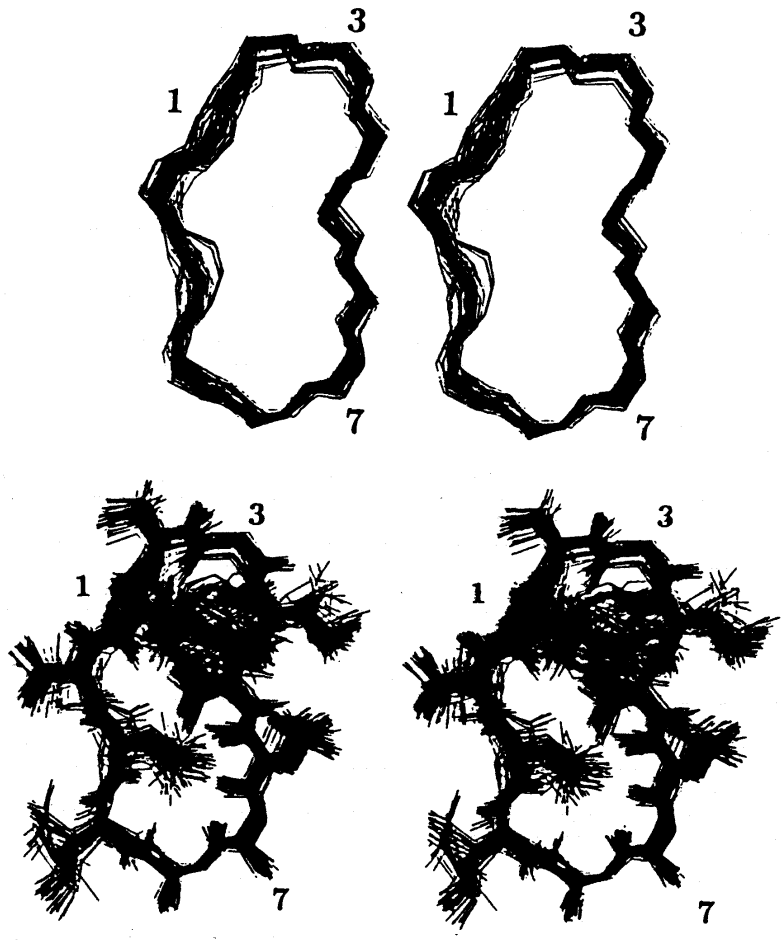

Fig. 5 Structure of CsA bound to CyP stereoviews are shown of a superposition of several best energy minimized structures. (A) The bonds connecting the backbone atoms (B) the bonds connecting all heavy atoms with the same orientation as in (A) (adapted with permission from K Wuthrich *et al. Biochemistry* **30** (1991) 6563 Copyright © American Chemical Society).

partially or fully exposed enabling them to interact with CyP.

As already mentioned residues 9-3 show many internuclear contacts with CyP and constitute the binding domain, while residues 4-8 do not interact with CyP. The residues 4-8 possibly bind with CN when CsA is bound to CyP and should correspond to effector domain of the CsA. It is interesting to note that CsA backbone from Sar 3 NH to Ala 7 CO, with the exception of MeLeu 4 resembles one strand of pleated β sheet. It is suggestive that the interaction between CN and CyP-CsA may be similar to those between strands of a β sheet. Residues 3-7 present a regular surface to CN and may make several hydrophobic contacts with it.

Recent work by Wuthrich *et al.*[38] on CyP is an elegant study which has in addition reviewed the three dimensional structures of CyP and CyP-CsA complex (Fig. 4). Detailed comparison of free CyP revealed subtle but significant conformational differences that can be related to lattice constants in the crystal structures in most cases. ^{15}N relaxation times and NMR line shapes analyses for CyP in the free form and Cyp-CsA complex revealed transitions of polypeptide loops surrounding the ligand-binding site from locally flexible conformations in the free protein, some of which includes well defined conformational equilibria, to well defined spatial arrangements in the CyP-CsA complex. The NMR structure presents a highly relevant reference for studies of changes in structure and internal mobility of the binding pocket upon ligand binding, and possible consequences of this conformational variability for CN recognition by the CyP-CsA complex.

NMR results on the structure of the CyP-CsA complex are consistent with the previously observed structure-activity relation of modified CsA analogues. For example, the marked decrease in CyP affinity for CsA analogues substituted with larger side chains at the residue 11 can be rationalized by the tight fit of the MeVal 11 in the middle of the binding pocket[51,52]. In contrast MeLeu 9 and Abu 2 are located near the sides of the binding pocket and a variety of other substituted amino acids can be accommodated without substantially influencing the affinity to CyP. Similarly since the residues 5 to 8 do not interact with CyP, and as expected, can tolerate the modifications. Some analogues, like (MeAla 6) CsA and (D-Abu) CsA, with modifications at these sites, despite binding tightly to CyP are inactive. This is consistent with their decreased affinity to CN. Though the pentapartide complex, CsA-CyP-CNA-CNB-Calmodulin, is too big to study by NMR, some details are available from X-ray diffraction in crystalline state[53].

CsA has several undesirable effects arising due to CsA mediated inhibition of CN in cells, outside the immune system. Recently Schreiber *et al.*[54] have synthesized a modified CsA (CsA*) that does not bind to CyP and consequently does not inhibit CN in cell. They also modified the CyP (CyP*) in its CsA binding pocket which promoted high affinity complexation with the CsA* (K_d = 9 nM). The modified drug-receptor complex (CyP*-CsA*) inhibited CN (IC_{50} = 25 nM). The ability to inhibit CN in a tissue specific fashion, for example, through the expression of CyP*, may provide a method to understand the role of CN in different cell types. These studies have raised some interesting questions and may play a role in development of new immunosuppressant.

Structure of FK, Rapamycin, FKBP and FKBP-Ligand Complex

NMR studies have revealed that the structure of free FKBP is characterized by a five stranded antiparallel concave sheet with a novel +3,+1,-3,+1 loop topology[55,56]. The strands of sheet correspond

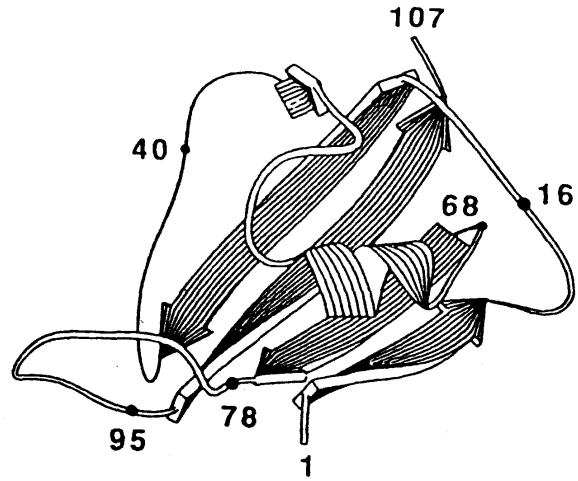

Fig. 6 Carbon ribbon diagram of human FKBP. The five stranded sheet, an α-helix and the connecting loops are indicated. Also shows the N- & C-termini and other residues in the protein are also shown (adapted with permission, from S W Michnick *et al. Science* **252** (1991) 836. Copyright © 1991 by the American Association for the Advancement of Science).

to residues 2-8, 21-30, 35-38 with 46-49 (interrupted by a loop at residues 39-45), 71-76 and 97-106 and run almost at 90° to the long molecular axis (Fig. 6). The only helix, containing residues 57-63, is amphipathic and runs along the long axis of the enzyme and lies against the sheet, forming a tightly packed hydrophobic core. The core consists of entirely aliphatic and aromatic residues. The conserved aromatic and aliphatic side chains of Tyr 26, Phe 36, Phe 46, Val 55, Ile 56, Trp 59, Tyr 82 and Phe 99 line a shallow cleft at the N terminus of α helix, forming the FK506 and rapamycin binding sites[57,58]. The side chains of these residues are well defined in both the free and complexed (with ligand) FKBP. Both FK506 and rapamycin fit into the binding pocket tightly and have a number of similar hydrophobic and hydrogen-bonding interactions with the protein. The FKBP binding domain of FK506 extends from the C24 hydroxyl group through the diketo pipecolinyl moiety and the pyranose ring to the C15 methoxy group. This finding was anticipated on the basis of structural analysis of FK506 and rapamycin and resulted in the design of 506BD and other interesting analogues[59]. The unsubstituted cyclohexyl ring (C26-C34) and the carbon atoms C18-C23 of FK506 do not make contact with FKBP and remain exposed to the solvent. This solvent exposed portion is referred to as the effector domain and should bind with calmodulin-CN to form a pentapartide complex. For rapamycin the orientation of the cyclohexyl ring (C34-C42) is different from that in FK506 and makes several contacts to FKBP. From extremely high field shifts of pipecolinyl methylenes and several NOE crosspeaks, it was concluded that pipecolinyl moiety makes van der Waals contact with the indole ring of Trp 59, at the back of the binding pocket, and with the side chains of Tyr 26, Phe 46, Val 55, Ile 56 and Phe 99 at the sides of the pocket. The pyranose ring comes in close proximity to residues Tyr 26, Asp 37, Tyr 82. His 87 and Ile 90. The C24-C26 region of FK506 and C28-C32 of rapamycin also makes close contact with Phe 46 and Glu 54. So many complimentary hydrophobic contacts support the large binding constant of FK506 and rapamycin with FKBP.

In addition to large number of van der Waals interactions, several intermolecular hydrogen bonds are formed between the ligands and FKBP. For both ligands, the hydrogen bonds between the C1 ester carbonyl-Ile 56 NH, the C8 amide carboxyl-Tyr 82 phenolic OH, C10 hemiketal OH-Asp 37 carboxylate are observed. Glu 54 CO makes hydrogen bonds with C24 hydroxyl of FK506 and C28 hydroxyl of rapamycin. An additional hydrogen bond between C40 hydroxyl-Glu 53 side chain CO is observed for rapamycin, but does not exist in case of FK506. The hydrogen bonds involving Ile 56 NH, Glu 54 CO and Glu 53 CO resemble antiparallel β-sheet-like contacts, often seen between protein and peptide ligands. These hydrogen bonds then extend the portions of FK506 and rapamycin, which can be treated as peptidomimetic, to include the regions between C1 ester carbonyl group and C24 (C28 for rapamycin) hydroxyl group[7].

Since the inhibition of phosphatase activity of the CN takes place only with the binding to FKBP-ligand complex, it is important to analyze the differences between the free and bound form of both the protein and ligands. Majority of residues in FKBP-ligand complex are not appreciably perturbed by ligand binding. Even the aliphatic and aromatic residues making direct contact to ligand (Tyr 26, Phe 46, Val 55, Ile 56, Trp 59 and Phe 99) show only minor conformational changes between the bound and free protein. The regular structures involving sheets and α helix also show very small variations in the two structures. However, there are several loops in FKBP, located at the rim of the binding pocket, that seem to show significant changes in their structure and dynamics on ligand binding. For free FKBP the seven residue bulge, Ser 39 to Pro 45, appears to be disordered, while in FKBP-FK506 this region shows large variations of chemical shifts, implying that many of the residues in this region form hydrogen bonds resulting in change in conformation and mobility of the bulge. Interestingly on ligand binding the region Lys 34 to Lys 44 is the most charged, suggesting that it plays an important role in binding to the CN.

Significant changes are also observed on ligand binding in the sequence Tyr 82 to His 94. This loop in free FKBP is again poorly defined and the region appears to have atleast two distinct conformations that inter-convert slowly on the NMR time scale. On ligand binding, this loop shows a distinct well defined conformation, stabilized by several hydrogen bonds within the protein and has two type II β turns involving residues 87-90 and 92-95 and van der Waals

contacts between the ligand and enzyme. These changes again are suggestive of the fact that residues Tyr 82 to His 94 may be involved in the CN recognition along with the residues Ser 39 to Pro 45. This would then result in a CN binding site that spans almost 20 Å and contain regions of high hydrophobicity from the effector domains of the ligands and high positive charge density from the protein. To find a definitive answer more detailed experiments would be required.

Though the differences in the protein structure on ligand binding are small, the ligand itself shows major structural changes. In organic solvents, NMR studies show that FK506 exists as a 2:1 mixture of *cis-trans* amide bond rotamers[15,60-62]. NOE restrained molecular dynamics studies have revealed structure of both the rotamers[61,62] (Fig. 7). These structures differ considerably from those in the solid state[15].. The *cis* structure, in the binding domain is very similar to the structure in the crystalline phase. However, there are large differences in the torsional angles in the effector domain, involving the C-C bonds of C20 to C27 chain. This produces an extended and relatively flat

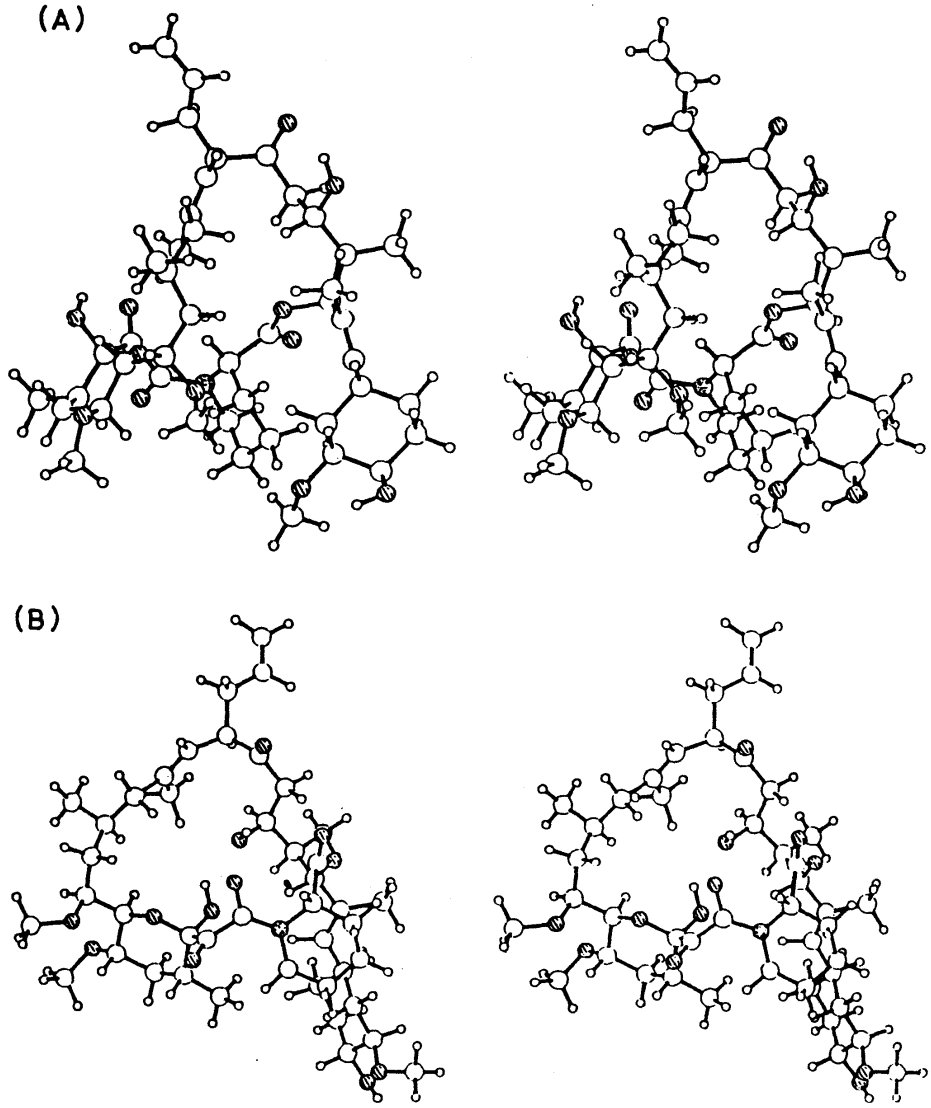

Fig. 7 Stereostructure of low-energy conformations of FK506 from restrained molecular dynamics (A) conformation containing a *cis* peptide bond between atoms 7 and 8. (B) conformations with a *trans* peptide bond between atoms 7 and 8 (adapted with permission) from P Karuso *et al. J Amer Chem Soc* **112** (1990) 9434. Copyright © 1990 by the American Chemical Society).

structure and cyclohexyl residue extends away from the macrocycle in the solid state. In solution it is wrapped under the macrocycle due to C25-C26-C27-C28 dihedral angle of 91°. Interestingly the *trans* isomer in solution has a different structure than that in the solid state at the FKBP binding domain, but the rest of the molecule is almost similar. The differences in the torsional angles at C1-C2 and C9-C10 accommodate the change from *cis* to *trans* bond. Thus only other major changes exist around C16 to C19.

The structure of bound FK506 is drastically different from that of the free ligand (both in the solution and solid state[63]). The major change in solid state is the *trans* amide bond in the bound ligand while it is *cis* (the major component) in solution. Large changes are observed in the torsional angles around C1-C2, C14-C15, C15-C16 and C16-C17 bonds in solid as well in solution state. These changes result in a more compact shape for the binding domain and forms a hydrophobic core of the ligand.

As against FK506, the free and bound conformations of rapamycin are very similar[7,64]. Rapamycin has a *trans* amide bond in crystalline state, while in organic solvents it exists as a 4:1 mixture of *trans* and *cis* amide bond rotamers. It also has the pyranose moiety folded under the macrocycle and cylcohexyl moiety extended outward from the body of the molecules. In fact, the conformation of the FKBP binding domain of rapamycin is almost identical to that of FK506. The implicit greater preorganization of rapamycin might be thought to provide greatly enhanced binding relative to FK506, which in practice has a binding constant only twice as large.

The structural studies of FKBP-FK506 attempted to provide a model for the rotamase inhibition by ligands[60]. One mechanism suggested was the formation of tetrahedral enzyme substrate adduct. The rotation about the C-N bond in the adduct followed by the expulsion of the enzyme results in an amide bond isomerization[65]. The second mechanism suggested the binding of a transition state structure, which contains a twisted or distorted amide bond[66] (Fig. 8). The energy for isomerization in this model becomes available due to favourable noncovalent interactions between the enzyme and a peptide substrate.

The mechanism involving the tetrahedral intermediate was ruled out from some elegant NMR experiments which showed the non-existence of tetrahedral adduct expected due to nucleophilic attack of a side chain of the enzyme on one of the electrophilic carbonyl in FK506 at C8 or C9[60]. These studies also showed that FK506 binds to FKBP in a single *trans* conformation, although the unbound ligand exists in organic solution as a 2:1 mixture of *cis* and *trans* amide bond rotamers.

The support for twisted amide bond comes also from the solid state studies on FK506 and rapamycin where the dihedral angle between C8-

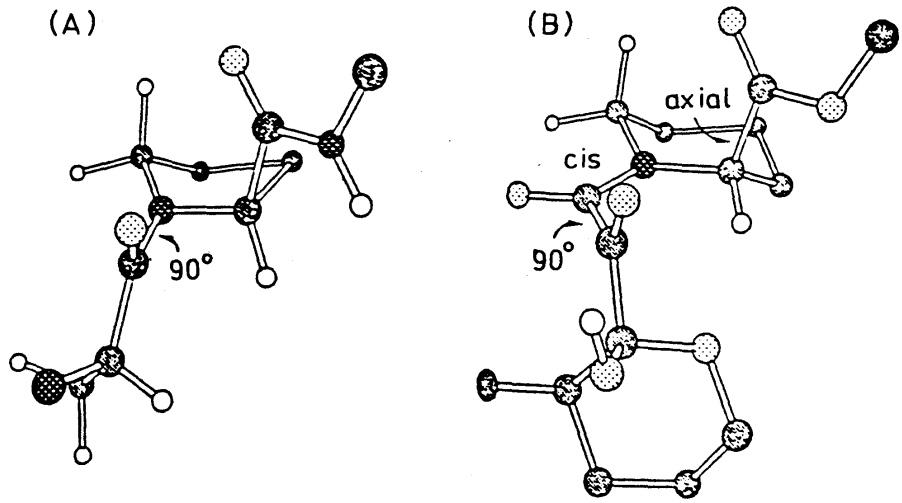

Fig. 8 (A) Model of twisted peptido-prolyl amide bond portion of FK506/rapamycin showing a 90° dihedral angle between N7-C8-C9-C10. (B) Substructure of FK506. (adapted with permission from S L Schreiber *et al. J Org Chem* **55** (1990) 4984 Copyright © 1990 American Chemical Society).

C9 bond is approximately 90°. This conformation is also maintained in the complexes of FKBP with both the ligands. A planar N7-C8 amide group along with the 90° dihedral angle between C8 and C9 places the keto carbonyl roughly perpendicular to the plane of the pipecolinyl ring. The pipecolinyl ring most probably mimics the proline in natural peptide substrates. Therefore, the keto carbonyl of FK506 or rapamycin is in the same position as would be a twisted amide carbonyl group of a peptide, undergoing rotamase catalysis. Thus the perpendicular keto carbonyl group of FK506 and rapamycin allows the ligand to mimic a transition state structure involving a twisted amide bond. The view of FK506 and rapamycin as twisted amide peptidomimetics was further elaborated by studies of substrate, specifically of FKBP to peptide, series succinyl-Ala-Xaa-Pro-Phe-(p-nitro) anilide[66a]. It was found that peptides having Xaa with branched hydrophobic amino acids, Leu, Ile, Val were more favoured than other amino acids. Analysis of FK506 and rapamycin data shows that they bind to FKBP because of their ability to mimic a Leu-Pro dipeptide with twisted amide bond. Similar observation have been made for CsA, where Cα-C1 bond of MeVal 11 appears to mimic a twisted amide bond.

The change in the structure of CsA and FK506 in free state and when bound to immunophilins is intriguing. Wuthrich *et al.* in their commentary[67], note these observations on CsA as "Receptor induced conformation change of the immunosuppressant, cyclosporin A". Such changes have been observed in other systems, yet one is led to think that the structure of the drug is getting perturbed due to the requirement that they should fit tightly into the receptor binding sites of the immunophilins. These results seem to confirm the expectation arising from the accumulated knowledge that flexible molecules change their shapes to form optimal interactions with binding partners.

A practical consequence of such a situation is the frustration in attempting to design drugs by analogy to the structure of flexible unbound active substrate. Thus Jorgenson has commented on these results as "Rusting of the lock and key model for protein ligand binding"[68]. Results on the structural studies of the analogues of the drugs and the receptors have been striking. These studies brought some more surprises to the fore and produced more comprehensible results. Availability of water soluble analogues of both CsA and FK506 and their NMR investigation in aqueous solvents were revealing. Interestingly both [D-MeSer 3-D-Ser(O-Gly) 8]Cs[69] and FK506 analogue, [32-Arg] ascomycin[70], have structures in water that are very similar to the structure of CsA and FK506 bound to their respective immunophilins. These results suggest that the immunophilin bound conformation of these drugs largely preexist in aqueous solutions and is not induced by the protein as previously postulated. These studies also brought out a well known, yet forgotten fact (in the context of studies with these immuno-suppressants), that solvents do play a major role in deciding the structure of molecules and one should keep this in perspective before comparing the structure of the free drug molecule, with that bound to its receptor.

The structural studies of FK506 bound to a mutant FKBP have also been undertaken[71]. This mutant protein causes the ligand backbone near the C16 to move by several angstroms. FK506 appears to undergo shifts in the binding pocket, with the largest change occuring at the effector region of the drug, which is involved in the binding with CN. Denesyuk *et al.*[72] made a computer modelling study to see why despite such dissimilar structures CsA-CyP and FK506-FKBP perform the same biochemical function. It was found that nine amino acids of CyP and FKBP have similar arrangements at their active site pockets. Also, FK506 and CsA have structural similarities where they come in contact with the active site. The resemblance is specially striking around MeLeu 9, MeLeu 10 in CsA and pipecolinyl moiety in FK 506.

It has already been mentioned in the beginning of this paper that HIV-1[73], which is a retrovirus, carrying its genetic information in RNA, causes the suppression of the immune system. One is, therefore, tempted to look for similarities between the mode of action of immunosuppressant drugs and HIV. Recent discovery, that CyP binds to the Gag poly protein (the poly proteins are cleaved by protease, which produce functional proteins of the virus) of HIV-1[73] and is specifically incorporated in HIV-1 virions[74], being required for their infectious activity, has sparked great interest. NMR spectroscopy has shown[75] that the heptapeptide, Ser-Glu-Asn-Tyr-Pro-Ile-Val, a model of an HIV-1 protease cleavage site in the Gag poly protein is a substrate of CyP. The *cis-trans* conversion rate of

the heptapeptide is enhanced by CyP. It raises the interesting possibility that CyP may be functioning in the HIV virion to catalyze the generation of the conformers in the poly proteins necessary for cleavage by HIV protease. These observations have raised many questions on several aspects of immunology, protein folding and retrovirology, specifically the question whether Gag protein of HIV-1 is also an immunosuppressant. CsA was found to compete with Gag for binding to CyP and led to a speculation whether the Gag protein has effects similar to that of CsA. It was shown that CyP-Gag complex does not bind to calcineurin, suggesting HIV-1 protein may mediate immunosuppression by other mechanism. One possibility is Gag protein disrupts the binding of CyP to a natural ligand[21] and thereby influences T cell activation. Another important finding was that the virus that contained less CyP, due either to treatment with drugs or mutation, was less infectious and thus gives scope for therapeutic intervention. This leads to a suggestion that for a cyclophilin depleted virus the replication gets blocked between the reverse transcriptions of the RNA and integration of the DNA into the host cell genome[76]. It seems reasonable to investigate whether cyclosporins, in particular non-immunosuppressant analogues, can be used for treatment of AIDS. With this notable observation that immunosuppressants, the drugs as well as the HIV, compete with each other, it would be interesting to discuss therapies of AIDS and the applications of NMR spectroscopy in the advancement in this area.

AIDS and Structure of Reverse Transcriptase Inhibitors

HIV infects the T4 lymphocytes, which play a central role in regulating the immune system[77]. The resulting depletion of T4 cells, the hallmark of AIDS, leave the patients vulnerable to opportunistic infection that would not harm a healthy person. Even today, almost after two decades of the first observation of AIDS, efforts are going on to find an effective vaccine for the prevention or an efficacious drug for the cure of the disease. The triple combination therapy proposed in 1995 appears to be extremely successful and has brought tremendous cheer in the AIDS community. Several studies have shown that the potent combination of the drugs can reduce the amount of HIV in blood to levels almost below the detectable limit of most sensitive assays. It has recently been discovered that this therapy also dramatically reduces the amount of HIV RNA in lymph nodes, the hiding place of the virus. These results are dramatic. Though, HIV is even more impervious to attack when it transforms its genetic materials from RNA to DNA, weaves itself into the host genes, and does not make new viruses. This strategy allows it to escape from the drugs. Duration of the disease would depend on how long cells harbouring HIV DNA survive, how many infected cells exist in an infected person and how much new viruses they can produce. Based on the computer model, Ho *et al.*[78] suggest that several years of treatment with a 100% inhibiting anti-HIV regimen might be able to eliminate all remaining viruses. As of now however, it would be wrong to believe that a cure of AIDS is close by. The understanding of how HIV causes diseases can be converted into treatment strategies. Several strategies in use to inhibit the HIV are discussed below and various sites of attack by the drugs are[77]:

1 Absorption of the virus to the cell membrane, which depends on the interaction of the viral envelope glycoprotein gp120 and two cellular receptors CD4 and CCR5 (cysteine-cysteine chemokine receptor 5) which are present on the surface of a subset of T-lymphocytes and certain other immune cells.
2 Transcription of the viral RNA to proviral DNA by the RT and degradation of the residual RNA by ribonuclease H.
3 Integration of the retroviral DNA into the host DNA. This is a two step process. The first one, termed as 3' processing, involves removal of two nucleotides from each 3' end of the reverse transcribed viral DNA. In the second step, called the DNA strand transfer, the resulting 3'-OH end of the viral DNA are covalently joined to newly created 5' ends in the target DNA.
4 Expression of the viral genes by the regulatory proteins such as the *trans* acting transcriptional activator (tat).
5 Proteolytic cleavage of the viral precursor proteins by the viral proteinase, which itself gets auto-catalytically cleaved from precursor protein.
6 Glycosylation of the viral glycoproteins, an

intermediate step between the first and second rounds of glycosylation requiring trimming of the terminal glucose and mannose moieties by glycosidases and mannosidases respectively.

The target that has received maximum attention is the synthesis of viral DNA by the enzyme RT. This strategy is attractive because it involves a step that is unique to retroviruses. The large size of RT (≈50 kDa) forbids the study of the enzyme structure by detailed NMR spectroscopy. However structural study on the inhibitors in solution have been of great relevance. The natural nucleosides and nucleotides, that are essential for the viral DNA chain to grow, have hydroxyl groups at 3' position. 3'-hydroxyl group acts as the site of attachment of the next nucleotide of the growing DNA chain. Any break in the chain or in chain growth results in the inhibition of viral replication. Action at this stage generated the first AIDS drug 3'-azido-2',3'-dideoxythymidine (AZT) and several subsequent FDA approved drugs come from the 2',3'-dideoxynucleoside (ddN) class of compounds[77]. In these molecules 3'-hydroxy group is absent and is replaced by a group that does not permit formation of phosphodiester linkages. AZT, approved as the first AIDS drug about a decade back, was found to reduce the morbidity and mortality involved with HIV infection. It is a prodrug and like other nucleosides, gets phosphorylated to AZT triphosphate (AZTTP) by a series of kinases (that usually phosphorylate thymidine). AZT-monophosphate (AZTMP) is first formed by thymidine kinase and two phosphates are then added by the sequential action of thymidylate kinase and nucleotide diphosphate kinase to form AZTTP, the active moiety of AZT.

AZT inhibits the action of RT by two mechanisms (i) AZTTP preferentially attaches to the nucleotide binding site and competitively inhibits the binding of deoxythymidine-5'-triphosphate, thereby preventing its attachment to the growing chain of viral DNA; (ii) AZT acts as a chain terminator after AZTTP is added to the growing viral DNA chain. Because of the presence of azido group at the 3' position 5',3'-phosphodiester linkage cannot be formed and viral DNA synthesis stops.

All ddN analogues possibly act in a similar fashion as AZT, implying that they must be phosphorylated intracellularly to their 5'-triphosphate analogues before interacting with the RT. ddNTPs have about 50 fold greater affinities for the HIV RT than the corresponding natural 2'-deoxynucleoside- 5'-triphosphate (dNTP). Thus ddNTPs may act as either competitive inhibitors, i.e. prevent the incorporation of the natural substrate (dNTP) or alternate substrates and thus get incorporated in the growing viral DNA chain, leading to the DNA chain termination.

The difference in the anti HIV activity of different ddNs could be attributed to (i) difference in the rate (or extent) of phosphorylation to their 5'-triphosphates; (ii) difference in the affinity of these 5'-triphosphates for the HIV RT; (iii) difference in their ability to influence pool sizes of the substrates (dNTPs), they compete with. Increasing evidence suggests that the relative ability by which 2',3'-dideoxynucleosides generate their triphosphates intracellularly, correlates well with their activity. Recent experiments have thrown more light on the mechanism of AZT inhibition of the RT[79]. Although the AZT conversion to AZTTP is facile, conversion of AZTMP to AZTTP is inefficient resulting in the accumulation of AZTMP. As AZTMP binds to the thymidylate kinase with high affinity, an increase in concentration of AZTMP antagonizes the normal activity of thymidylate kinase and thymidine monophosphate (TMP) conversion to thymidine triphosphate (TTP) becomes inefficient.

Cellular toxicity and viral resistance limits the efficacy of AZT. Subsequently other nucleoside based RT inhibitors like 2',3'-dideoxyinosine (ddI), 2',3'-dideoxycytidine (ddC), 2',3'-didehydro-2',3'-dideoxythymidine (d4T) and 3'-thiacytidine (3TC) were approved as drugs making them the first of the five AIDS drugs. These drugs and some other protease inhibitor drugs are often to be used in combination with AZT in the therapy for best results[79].

Various methods like X-ray diffraction, NMR spectroscopy, Laser-Raman spectroscopy and theoretical calculations have been used to study the molecular structure of ddNs. NMR study of several ddNs have been performed in solution and certain trends have emerged[81-90]. Nucleosides show high conformational flexibility primarily involving rotations about single bonds and the sugar pucker[91]. The important rotations are about the exocyclic C4'-C5' bond and the glycosyl bond. The dihedral angles γ and χ characterize the orientation of

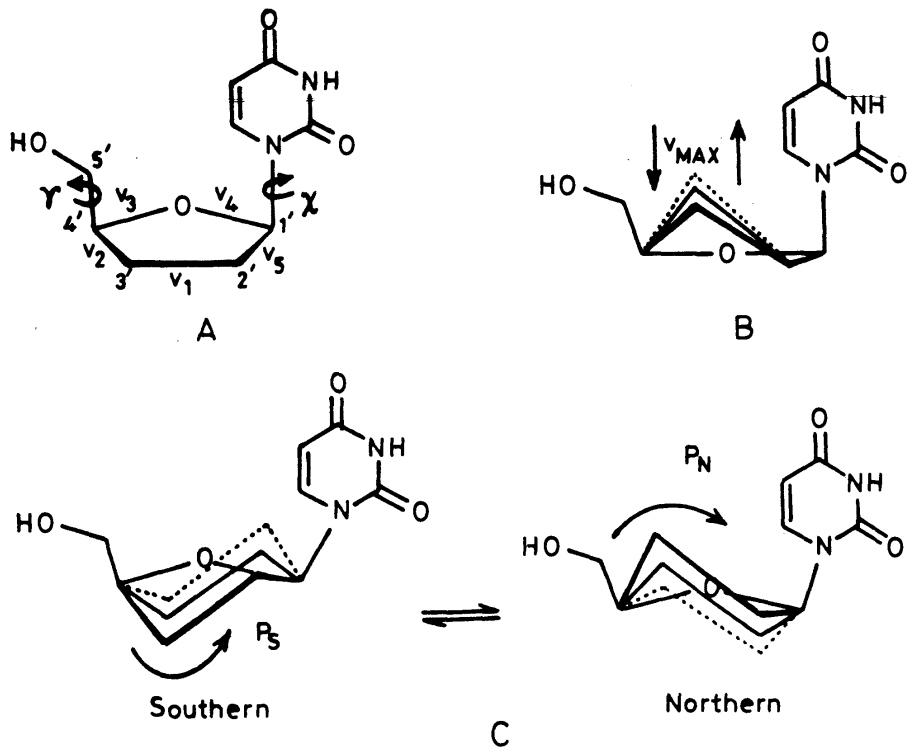

Fig. 9 Conformation of a dideoxynucleosides. (A) Definition of dihedral angles γ and χ. The angles ν_1-ν_5 are the five endocyclic torsion angles which are related to the exocyclic torsion angles derived from NMR coupling constants. (B) Definition of the maximum puckering angle ν_{max}. This is the maximum deviation from planarity for the ring and is equal to the largest value of the five endocyclic torsions. (C) Definition of the pseudorotational angle P. P varies from 0-360°, and as it increases the ring pucker progresses through alternating twist and envelope forms.

hydroxymethyl group and the base respectively. The sugar ring geometry is described by using the concept of pseudorotation[92], in which the five endocyclic torsion angles are related to puckering amplitude (ν_{max}) and phase angle of pseudorotation (P) (Fig. 9). NMR studies of sugar rings of nucleosides and nucleotides show the presence of a fast equilibrium between two broad classes of conformers. These two classes are referred to as the N-type, with phase angle $P(N) = 0°\pm90°$, puckering amplitude $\nu_{max}(N)$ and population X_N and the S-type with phase angle $P(S) = 180°\pm90°$, puckering amplitude $\nu_{max}(S)$ and population X_S (X_S=1-X_N). The observed couplings are averaged over two sugar puckers, characterized by P(N), P(S), $\nu_{max}(N)$, $\nu_{max}(S)$, X_N and X_S.

It is found that conformation around C4'-O5' have a preference for γ^+ conformers while the least populated conformer is γ^- and the glycosidic angle prefers an *anti*-conformation. These results are not very different from those for natural nucleosides. X-ray diffraction study also provide qualitatively the same results. The sugar ring geometries show interesting variations depending on the substituents in the ring. In case of 3' substituted ddNs the population of the two conformers depend on the electronegativity of the substituent. While AZT shows almost equal populations of N-type and S-type conformers, ddI, ddC and other similar nucleosides (without a substituent at 3') show predominance of N-type conformers. Natural nucleosides and 3' fluoro-substituted ddNs, on the other hand show predominance of the S-type sugar conformation. These results follow a pattern consistent with *gauche* effect[93], which predicts the preference of S-type conformers for 2'-deoxyribose sugars. When the substituents at 3' are more electronegative, the *gauche* effect is reinforced and preference for S-type sugar pucker is enhanced, while a less electronegative group leads to predominance of N-type sugar pucker, as is

observed in unsubstituted ddNs. X-ray results of Van Roey *et al.* on several of the ddNs have shown the presence of unusual sugar pucker of C3' exo/C4' endo[94]. They have implicated the anti HIV activity of some of the ddNs to this sugar ring conformation. NMR studies, however, do not support such a conclusion. The value of ν_{max} from NMR and X-ray diffraction studies for these nucleosides is around 35°.

From these results no definitive conclusions about the geometrical requirements of a RT inhibiting AIDS drug could be drawn. Since ddNs are prodrugs, they are converted to the dideoxynucleotides, which in turn interact with the RT. In view of these, the NMR studies of dideoxynucleotides have also been undertaken[95]. AZTTP did show some significant variation in the population of the sugar ring conformation compared to AZT. While for AZT the value of X_S was 0.52, which increased to 0.75 for AZTTP, a value very similar to that for natural nucleotides. Such a variation in sugar conformation was, for example, not observed for other ddNTP. The ddNTPs investigated showed predominance of *anti* conformers about the glycosidic bond, γ^+ about the C4'-O5' bond and β^t about C5'-O5' bond. These conformations are similar to those obtained for other natural nucleotides and reflect more rigid structure for the dideoxynucleotides compared to the dideoxynucleosides.

Due to large size of the receptors, difficulty in getting labelled receptors or low binding affinity of the drug to the receptor, it is not always feasible to determine the complete structure of the protein-ligand complex. In such cases it is often possible to derive information about the bound ligand conformation using transferred NOE (TRNOE) technique[96]. This is suitable when exchange between free and bound ligand is fast on the relaxation time scale. Exchange of the nOe from bound ligand to free ligand occurs, allowing the information on bound ligand to be obtained from the readily detected free ligand signal. The application of TRNOE has been made on the RT bound conformation of AZTTP[97]. These studies show that the bound conformation of sugar ring is not the normal C2'-endo or C3'-endo conformer, instead it is C4'-exo. However the RT bound conformation of thymidine triphosphate was also identical to the RT bound AZTTP conformation. This leads to a suggestion that the similarity in molecular geometries enable AZTTP to competitively inhibit the binding of TTP to RT and may be responsible for the success of AZT as an AIDS drug.

Structure of HIV Protease-Inhibitor Complex

The design and development of HIV protease inhibitors, the other class of AIDS drugs, have revolutionized AIDS therapy and brought a ray of hope in the gloomy world of AIDS patients. In the late 1995 and early 1996 some of the protease inhibitors were approved as drugs (bringing the total number of drugs approved till date for AIDS to 11) and a combination therapy using two RT inhibitors, AZT and 3TC, and a protease inhibitor, ritonavir (one of the four approved by FDA, others being saquinavir, indinavir and nelfinavir), has shown excellent results[79]. This invariably brought the number of HIV in the blood almost below the detectable limit. The full implications of this therapy, which can fail for a variety of reasons, specially due to the intricacies of administering the drugs at a strictly prescribed time, over a long duration, are yet to be realized. Keeping HIV at bay is a daunting task. Even the best treatments have a difficult time completely suppressing viral replication, which gives drug resistant mutants a chance to appear. Thus fully rebuilding an HIV ravaged immune system is a tall order[98].

The development and design of protease inhibitor drugs have used the lessons learned from research in renin inhibitors and their crystal structures. Three dimensional structures obtained from X-ray diffraction[99] investigation showed gross structural similarities between HIV protease and pepsin family of mammalian proteases. It was obvious from detailed structures that there are significant structural differences between them, which could be exploited to design protease inhibitors, specific to the viral enzyme. Crystal structure of several HIV protease and HIV protease-inhibitor complexes have been obtained. Such studies were mainly responsible for the developments of the protease inhibitor drugs for AIDS [100,101].

HIV protease, from the family of aspartyl protease, a homodimer with 99 amino acids in each monomer and molecular mass of 22.2 kDa, is ideal for NMR investigations. Structure of HIV protease in solution by NMR is not available as yet, despite the fact that NMR technique is suitable for this

enzyme. The difficulty has been due to the rapid autolysis of the enzyme. The structure of a potent cyclic urea based inhibitor (DMP323)-protease complex, however has been obtained by NMR recently[102,103]. The target of an inhibitor design is based on an important difference between the viral and mammalian proteases, that a structural water molecule is present, which coordinates the tips (residues 51, 52) of the viral protease flaps to the inhibitor. The DMP323 was rationally designed as the inhibitor of protease, (K_i=0.27 nM), as this cyclic urea analogue displaced the water molecule. Indeed the conserved water molecule was not seen in the HIV protease- DMP323 crystal structures[101] as well as in the NMR studies[104].

The stereospecific assignments of several methylene and methyl protons were obtained by the use of various 3D and 4D techniques[103]. The ^1H spectra of DMP323 bound to protease were obtained by a combined use of ^{13}C edited 2D NMR spectra. Intramonomer nOes were distinguished from intermonomer nOes by comparing the results from NMR spectra of fully ^{13}C, ^{15}N labelled protease with spectra of a isotopic heterodimer which has 50% of the monomer molecules labelled with ^{13}C, ^{15}N. Special editing techniques were used for specially selecting intermonomer nOes[105,106].

The tertiary structure of the complex was determined by using a systematic three step approach[103]. First step was to obtain an approximate monomer structure followed by the dimer structure and lastly the structure of full protease-inhibitor complex was derived. A preliminary monomer structure was obtained using intramonomer distance and dihedral angle constraints. For finding the dimer structure intermonomer nOe and hydrogen bonds were obtained. Fig. 10 shows the topology of the four stranded antiparallel β sheet, formed by the *N*-terminal β strand (the first β strand, β1) and C-terminal β strand (the tenth β stand, β10) of the monomer units along with the nOe and hydrogen bonds. This β1-β10'-β10-β1' (prime used for distinguishing one monomer from the other, as well the amino acids in one residue are numbered 1 to 99 while in the other from 101 to 199) arrangement is same as observed in the X-ray crystal structure[100]. In addition to the terminal regions, intermonomer nOes are observed between residues at 25-27 and 49-54, showing the proximities of these residues of the two monomers. Using several of the intramonomer and intermonomer distance constraints, the dimer structure without inhibitor was obtained.

The last step in obtaining the structure of protease-inhibitor complex was to include the data on the inhibitor. Conformational studies[101] predicted that the 7-membered cyclic urea ring conformation of DMP323 (Fig. 11A) prefers the pseudo-diequitorial hydroxyl groups and pseudo-diaxial benzyl groups (P1-P1'). This was confirmed by X-ray diffraction studies of DMP323 and protease-DMP323 complex. The NMR data on protease-DMP323 complex is in agreement with the X-ray single crystal structure of DMP323 and protease-DMP323 complex[101].

The dimer structure derived by NMR, without the inhibitor, was used to obtain the final structure by positioning the urea oxygen along the symmetry axis of the molecule and the inhibitor diol moieties close to Asp 25 and Asp 125 carbonyls in conformity with previous X-ray[101] and NMR[102,104,107,108] studies. This crucial step enabled distinguishing several drug-dimer nOes and resolved many ambiguous assignments. The structure of protease-inhibitor complex obtained using distance and dihedral angle constraints is shown as 23 superimposed structures in Fig. 11B. Fig. 11C shows the structure in the form of ribbon diagram.

This study provided the first solution structure of HIV protease (with or without the inhibitor) which closely resembled the X-ray single crystal structures of several protease and protease-inhibitor complexes. The first four residues form the outer part of the interfacial four stranded β-sheet (β1, β1'), involving terminal residues, that stabilizes the dimeric structure. There are ten β strands involving residues 1-4, 9-15, 19-25, 30-35, 45-49, 53-60, 62-65, 70-78, 83-86 and 96-99 and a single helix of the monomer at residues 87-95. Several turns and hair-pins join the regular secondary structures. The agreement between NMR and X-ray structure [100,101] is generally good, though distinct variations exist locally in several regions of the sequence that appeared to be important for functions. While the polypeptide chain is found to be disordered at residues 66-69, which show both type I' and type II' β turns, the residues 38-42 show large amplitude internal motions on 10^{-9} - 10^{-12} s time scale.

In contrast to rapid motions for residues 38-42,

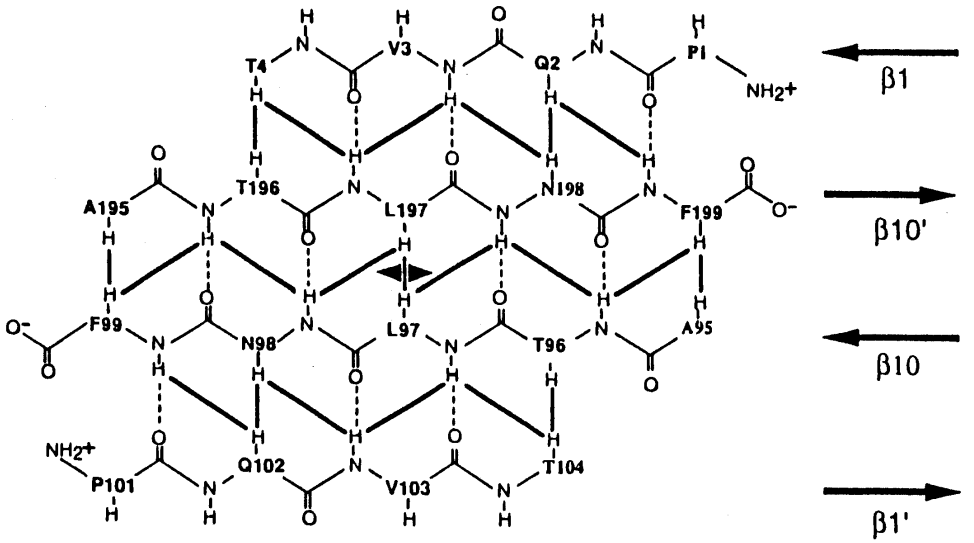

Fig. 10 Intermonomer β sheet topology of the HIV protease. Residues in the first monomer unit are numbered from 1-99 while in second monomer from 101-199 (the strands in this monomer are labelled by the prime superscripts) (adapted with permission from T Yamazaki *et al. Protein Science* **5** (1996) 495 Copyright © 1996 The Protein Society).

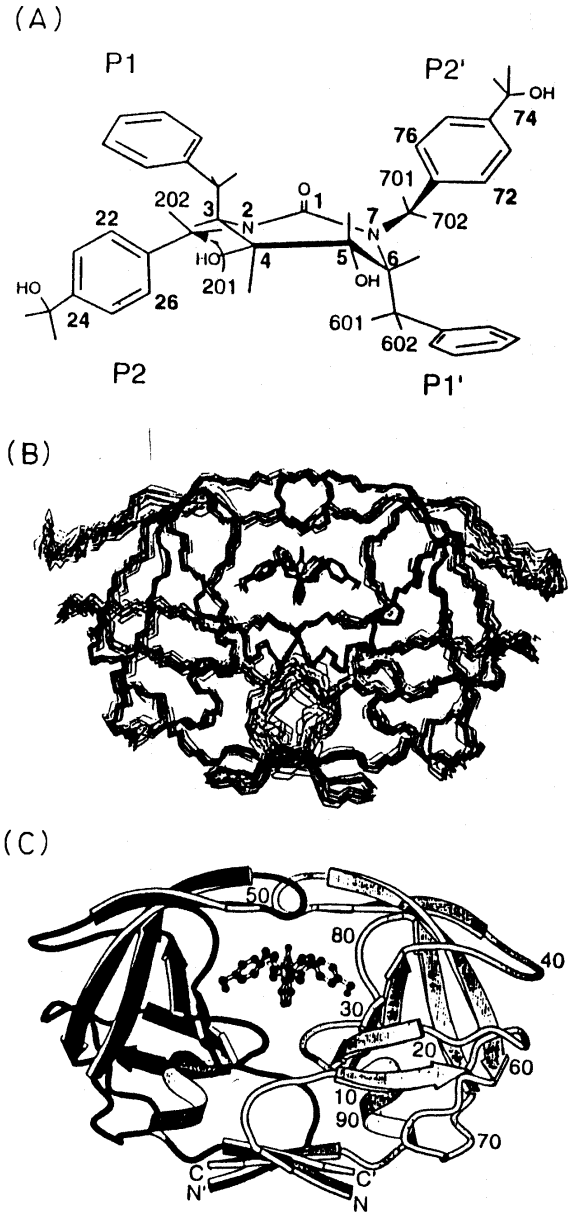

Fig. 11 (A) Chemical structure and conformation of DMP 323. (B) View of the superimposed heavy atom coordinates of 23 NMR structures of protease-DMP323 complex. (C) Ribbon diagram of the restrained minimized average NMR structure of protease-DMP323 complex. (adapted with permission from T Yamazaki *et al. Protein Science* **5** (1996) 495 Copyright © 1996 The Protein Society).

Thr 4 and Leu 5 execute motions in 10^{-3} -10^{-6} s time scale[108]. The primary site of autolysis of the protease is Leu 5-Trp 6 peptide bond, and flexibility of this region may be important for viral function, specially because of the suggestion that the rate of cleavage of this bond may regulate the activity of the protease[109]. The tips of the flaps of protease dimer, linking Ile 50 and Gly 51 with β turn, also execute slow motions on 10^{-3} -10^{-6} s time scale. This region is also involved in an important functional role in stabilizing the enzyme-substrate complex. The NMR data is not consistent with a single type of turn at residues 50-53. Low temperature studies also failed to freeze into a single conformation. It is presumed that, on crystal formation a single conformation is stabilized in the solid state. The flexibility of the residues 38-42 and 50-51 may be important in facilitating substrate binding and product release, because these residues make up the respective hinges and tips of the protease flaps.

The flexibility of the protease, so significantly evident in the protease-DMP323 complex, suggests that the HIV protease alone (without the drug) is very flexible. Such a suggestion has already been made by the X-ray studies of uninhibited protein, which shows open flaps that permit access to the active site involving Asp 25. Recently Kent *et al.* carried out NMR studies[110] on chemically synthesized HIV protease labelled with ^{13}C at the catalytically essential Asp 25 in the presence of an inhibitor pepstatin, a mimic of the tetrahedral intermediate formed in the enzyme catalysis. It was interesting to note that in the protease dimer, one of the Asp 25 was protonated while the other was deprotonated. Similar observation was made by Torchia *et al.* using an asymmetric protease inhibitor[111]. The direct observation of the chemical properties of the catalytic apparatus of the enzyme provides concrete information on which to base the design of improved HIV-1 protease inhibitors. Although autolysis precludes studies of the wild type protease by NMR spectroscopy, mutants with full activity and much higher stability of the enzyme are now available[112]. This will enable the characterization of the structure and dynamics of the free enzyme by NMR methods.

Structure of Peptides from V3 Loop of gp120

The best way to combat a disease is to prevent it. Vaccination is the simplest, safest and most effective form of prevention of the disease. Vaccines have achieved great success against viruses. Stupendous effort is on to develop an effective vaccine against the HIV and it remains, perhaps, the most formidable challenge for the virologist today[98]. Such efforts have not succeeded so far mainly because of the devious nature of the virus, which can hide in cells, change the composition of its coat protein and install its own genes within the genes of its host. Lack of a good animal model also does not permit investigations to combat these ploys of the HIV.

Most of the vaccine development effort has focused on gp120, the envelope glycoprotein with a mol. weight of about 120 kDa. Due to heavy glycosylation, most of gp120 is obscured from immunological sight by the cloud of sugars. The immunogenic properties of gp120 bring out two of its features; one is the binding site of the host T4 cell receptor molecule CD4 and other a loop, which is highly immunogenic[113]. The main binding site of the HIV-1 neutralizing antibodies, the principal neutralizing determinant (PND), is located within the disulfide linked third hypervariable loop in the V3 region of the gp120[114-117]. Fusion between the cell surface and the virus also results due to a proteolytic cleavage of a peptide bond within the V3 loop. Sequence of PND from several HIV-1 isolates revealed that certain amino acid residues are relatively conserved[118]. The sequence Gly-Pro-Gly-Arg-Ala-Phe (GPGRAF) occurs in many PNDs and the antisera directed against it were shown to neutralize multiple isolates[119]. The conserved sequence has, therefore, been the target of extensive studies for peptide based vaccine development[120-122]. Several gp120 peptide based vaccines are already at various stages of clinical trials. Despite the fact that the results are disappointing so far[79], possibly due to the hypervariability, existence of several conformationally distinct forms or low accessibility of the envelope protein, very serious efforts are being made to find a vaccine for AIDS. The recent discoveries of the chemokines and their receptors, CCR5 and CXCR4 for the HIV will definetely open up new avenues for the development of drugs and vaccines against AIDS[123,124].

Study and comparison of the conformation of these immunogenic peptides with those of the antibody bound peptides and of the corresponding parts of the proteins could be of great help in

understanding the molecular basis for the cross reactivity of antipeptide antibodies with the native proteins. Such studies are also likely to throw light on the contribution of various structural features to their antigenicities. V3 loop is typically 35 amino acid long, with residue number 1 and 35 as invariant cysteine involving disulfide bridge. Conformational studies of several PND peptides in the V3 loop have been carried out by NMR[125-133], specially to understand the nature of the conserved GPGRAF sequence and the role of the disulfide linked loop. The following synthetic peptides containing upto 40 amino acid residues (referred to with single letter code of amino acids), mostly with the sequence from the common HIV-1 isolates, have been studied (numbers inside the brackets are references):

RP70 - INCTRPNYNK RKRIHIGPGR AFYTTKNIIG TIRQAHCNIS[125]
RP142 - YNKRKRIHIG PGRAFYTTKN IIGC[125]
RP342 - IHIGPGRAFY TC[125]
RP135 - NNTRKSIRIQ RGPGRAFVTI GKIG[126]
P547 - NNTRKSIRIQ RGPGRA[126]
P344 - GPGRAFVTIG KIGGKK[126]
T1RF - KQIINMWQEV GKAMYATRPH NNTRKSIHMG PGKAFYTTG[127]
T1CANO - KQIINMWQEV GKAMYATRPN NNTRKSITKG PGRVIYATG[127]
P1 - CTRPNNNTRK SIHIGPGRAF YTTGEIIGDI RQAHC[128]
P2 - P1(D-Ala17)[128]
P3 - P1(D-Pro16)[128]
P4 - P1(N6-Glucose NAc)[128]
P5 - CTRPNYNKRK SIHIGPGRAF YTTGEIIGDI RQAHC[129]
P6 - C(P5)[129]
P7 - Cc(CTRPSNNTR TSITIGPGQV FYRTGDIIGD IRKAYC)[130]
P8 - Cc(CTRPNDNTR KSIPMGPGKA FYATGDIICN IRQAHC)[131]
P9 - Cc(CTRPNNNTR KSITKGPGRV IYATGQIIGD IRKAHC)[131]
P10 - GPGRAF[132]
P11 - GPGRAFC[132]
P12 - GPGRAFGPGRAF[132]
P13 - GPGRAFResin)[133]

For peptides RP70, RP142, Chandrasekhar *et al.*[125] found a significant population of conformations containing a β-turn at the highly conserved sequence GPGR, while for the shorter peptide RP342 the population of β-turn conformation was not significant. The other segments of these peptides showed conformational averaging. Zvi *et al.*[126] studied peptides RP135, P547 and P344. RP135, a 24 amino acid peptide corresponds to PND of IIIB isolates of HIV-1. Other 16 residue peptides include GPGRA sequence with N-terminal part of RP135 (P547) and C-terminal sequence of RP135 with GPGRA (P344). In RP135 the C-terminal part was found to exist in several transient turn like conformations referred to as "nascent helix" with a transient turn in the GPGR sequence. The conformation of shorter peptides was similar to those of the corresponding portions of RP135. Spicer *et al.*[127] studied two synthetic peptides taken from C4 and V3 domains of HIV-1 gp120. The segment taken from C4 domain in both peptides exhibits helical character with a reverse turn in the conserved GPGX (X=R or K) sequence. Griesinger *et al.*[128] studied synthetic peptide antigens, P1, P2, P3 and P4, containing 35 amino acids, from V3 region of the glycopeptide of HIV-1. In P2 a D-Ala was introduced at residue 17, instead of Gly, the i+2 position, for stabilizing the type II β turn. They also observed high propensity of type II β turn in P1, P2 and P4 (a glycopeptide) while P3 had a type I' β turn at G-D-Pro-GR. Nascent helices are observed in the C-terminal end of these peptides. Fig. 12 shows the stereoplot of folded type II β turn involving residues 13-20 of P1.

Gupta *et al.* also studied[129-131] several peptide sequences from the V3 loop and observed that inspite of the sequence variation, the conserved structure is located inside the loop. Cyclization of P5 results in an N-terminal loop between residues 2 to 9, stabilization of type II β turn at GPGR and formation of two turns at residues 23-26 and 30-33. This incipient helix, found in aqeuous solution, shows propensity of a helix in less polar solvents. The structural rigidity and a desirable structure makes the cyclic peptide a better ligand for monoclonal antibodies than the linear form. Peptide P7, P8 and P9 also show very similar structural features with a GPG(R/K/Q) crest in the centre of the neutralizing domain, two extended regions flanking the central crest and a helical segment in the C-terminal region. These results lead us to study[132] much smaller peptides P10, P11

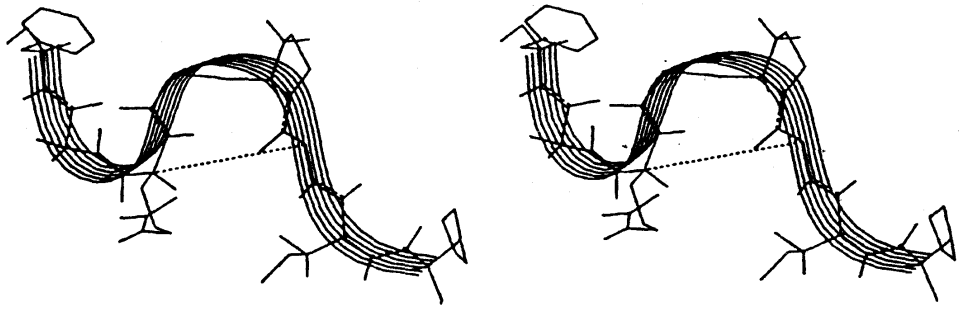

Fig. 12 Stereoplot of folded type II β conformation of P1 (amino acids 13-20) (adapted with permission from C Griesinger *et al. Eur J Biochem* **237** (1996) 188 Copyright © FEBS 1996)

and P12 and look for some incipient structure in them. Even these small peptides showed propensity of a turn in GPGR in DMSO solution. While P10 and P12 seem to have both type I and type II β turns, P11 showed propensity of only type II β turn. In presence of an appropriate receptor these peptides may take the required turn conformation and can be antigenic. The resin bound conformation of GPGRAF was also found to be folded[133].

Using NMR, Zvi *et al.* have also attempted to map the binding of an antibody and a specific antibody binding epitope from an HIV-1 III B peptide[134]. More explicit and detailed results were, however, obtained from the X-ray study of a complex of peptide RP142 with Fab fragment of a neutralizing antibody[135]. It was found that the tip of V3 loop with GPGRAFY, shows a multiple β-turn conformation. GPGR has a type II β turn followed by type III and type I β turn for GRAF and RAFY residues respectively. Some of the NMR results on peptides from V3 loop with β turn at GPGR and a "nascent helix" conformations are consistent with these X-ray studies.

Despite the conserved PND sequence, the hypervariability of V3 loop region possibly leads to altered conformation and thus the HIV antibody recognition of different V3 peptides is controlled by the primary as well as secondary structure of the antigen. It therefore appears that the vaccine development to kill HIV requires fresh line of thinking and some novel approaches will have to be devised. A renewed effort is being made in this direction, involving cytotoxic T lymphocytes (CTL)[119], which play an important role in the body's natural defence, as they search out and destroy virus infected cells. Other approaches involving genetically engineered viruses, specially the vesicular stomatitis virus (VSV) with CD4 and CCR5 receptors at its surface, have shown great promise[137].

Structure of HIV Integrase

^{1}H NMR spectroscopy has also been used to study the structure of the fragments of integrase and other structural proteins of the HIV. HIV integrase is responsible for the insertion of a DNA copy of the viral genes into host DNA, an essential step in the replicative cycle of HIV. HIV integrase comprises three functional domains. The central core domain, consisting the catalytic site, is flanked by a small N-terminal zinc binding domain and C-terminal DNA binding domain. While the catalytic core domain can carry out a simple polynucleotidyl transfer, termed disintegration, all three domains are required for the 3' processing and DNA transfer activities, that accomplish integration of the viral genome. Thus even the N- and C-terminal domains are potential useful targets for rational drug design for AIDS. NMR studies have been carried out on the N-terminal and C-terminal domains[138,139]. The C-terminal domain (residues 220-270) of HIV-1 integrase is a dimer in solution. Each subunit consists of a five-stranded β-barrel with a topology very similar to SH3 domains, which are found in numerous proteins[140] involved in signal transduction. The N-terminal zinc binding domain (residues 1-55) of integrase contains a His_2Cys_2 motif that is conserved in all retroviral integrases. It binds one equivalent of zinc and is necessary for full integrase activity. The isolated domain is unstructured in the absence of zinc, but folds into a dimer with high α-helical content in the presence of zinc. Each monomer has four helices with zinc

tetrahedrally coordinated to His 12, His 16, Cys 40 and Cys 43. This fragment exists in two interconverting conformational states, differing in the coordination of two histidine side chains to zinc. The monomer fold of the N-terminal domain is remarkably similar to that of a member of helical DNA binding protein, containing a helix-turn-helix motif. We would not discuss these results in further detail as they have not yet been exploited in AIDS therapy. Similarly we have not included in this chapter the NMR studies on other structural proteins of HIV.

Outlook

There is considerable evidence that paradigm shift is occurring now in the area of drug discovery[141]. Presently medicinal chemists are not only generating new structures, they are also designing more compounds with higher biological activity, higher selectivity, higher bioavailability and at the same time with lesser toxicity. NMR spectroscopy in combination with X-ray crystallography and appropriate theoretical methods for structural elucidation, constitute the best method for the development of structure activity relationship in drug design. Combinatorial chemistry[142-145] and high throughput screening have been revolutionizing the drug discovery programmes for the past few years. Combinatorial chemistry is already supplementing the historical approach of medicinal chemistry, starting from natural products leads. Modern and fast screening methods yield oligomers and analogues with high receptor-binding affinities. Interestingly such screening can often be accomplished without precise information about the receptor. The combinatorial libraries result in large number of molecules, invariably obtained through solid-phase synthesis. Molecules bound to resins can exhibit biological activity and immunogenic properties. A major challenge has been to find out what these libraries contain. ^{1}H and ^{13}C NMR methods[146-150], using magic angle sample spinning (MAS) of organic molecules covalently attached to beads, are being developed. These allow non-destructive monitoring of solid-phase reactions and should play an important role in future developments in the pharmaceutical applications. Pursch *et al.*[151] have used macrobeads, with particle size 400μm, based on resins with long flexible tentacles and tethers. Synthesis was carried out on a single macrobead. Each step in the synthesis was followed by ^{1}H MAS NMR using a single bead, which makes the characterization of the building blocks of molecules in the polymeric support quite convenient and straightforward.

Recently Fesik *et al.* suggested an NMR based structure activity method (SAR by NMR)[152,153]. This method involves screening of libraries of small molecules to identify, optimize and link components that bind to proximal binding sites of ^{15}N labelled protein. The compounds were screened by monitoring perturbation of the amide chemical shifts of the labelled protein, upon the addition of protein ligands. Using this approach and combinatorial synthesis, Kessler[154] has identified two weakly FKBP binding ligands with association constants K_d=0.1 mM and K_d=2 mM. Linking these two ligands with various $-(CH_2)_n-$ groups resulted in the recognition of a molecule with great affinity (K_d=19 nM) with FKBP. It seems likely that, as combinatorial chemistry is applied more widely in drug discovery, the rules for molecular recognition will become more transparent making structure-based drug design as the method of choice. This will enable finding leads, and developing them into drugs two to three times faster than has been possible. Such studies may also have important applications in the development of vaccines.

We have presented a view of NMR application to some interesting and contemporary problems in the area of immunology. An attempt has been made to show how important is the knowledge of three dimensional structure of drugs, receptors and drug-receptor complexes, in designing an effective and efficacious drug.

Acknowledgements

The authors are thankful to Dr T K Chakraborty and Dr B Jagannadh for several helpful discussions and to Mrs S Reddy for the help in the preparation of the manuscript. The author (AVSS) is thankful to UGC for the Senior Research Fellowship.

References

1 R R Ernst, G Bodenhausen and A Wokaun *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* Clarendon Press Oxford (1987)

2 K Wüthrich, *NMR of Protein and Nucleic Acids* John Wiley and Sons New York (1986)

3 D J Craik *NMR in Drug Design* CRC Press Boca Raton (1995)

4 R E Handschumacher and I M Armitage *NMR Methods for Elucidating Macromolecule-Ligand Interactions: An Approach to Drug Design* Pergamon Press Oxford (1990)

5 D J Craik, K J Nielson and K A Higgins *Annu Rep NMR Spectroc* **32** (1996) 143

6 (a) S W Fesik *J Med Chem* **34** (1991) 2937; (b) S W Fesik, E R P Zuiderweg, E T Olejniczak and R T Gampe Jr *Biochem Pharmacol* **40** (1990) 161

7 (a) M K Rosen and S L Schreiber *Angew Chem Int Ed Eng* **31** (1992) 384; (b) W Braun, J Kallen, V Mikol, M D Walkinshaw and K Wuthrich *FASEB J* **9** (1995) 63

8 (a) G W Vuister, R Boelens and R Kaptein *J Magn Reson* **80** (1988) 176: (b) H Oschkinat, C Griesinger, P J Kraulis, O W Sorensen, R R Ernst, A M Gronenborn and G M Clore *Nature* **332** (1988) 374

9 L E Kay, G M Clore, A Bax and A M Gronenborn *Science* **249** (1990) 411

10 M M McEvroy, A F delaCruz and F H Dahlquvist *Nature Struc Bio* **4** (1997) 9

11 G M Crippen and T F Havel *Distance Geometry and Molecular Conformation* John Wiley and Sons New York (1988)

12 (a) W F van Gunsteren and H J C Berendsen *Angew Chem Int Ed Engl* **29** (1990) 992; (b) A T Brunger and M Karplus *Acc Chem Res* **24** (1991) 54

13 J F Borel *Immunology* **31** (1976) 631

14 R E Handschumacher, M W Harding, J Rice, R J Drugge and D W Speicher *Science* **226** (1984) 544

15 H Tanaka, A Kuroda, H Marusawa, H Hatanaka, T Kino, T Goto, M Hashimoto and T Taga *J Am Chem Soc* **109** (1987) 5031

16 J J Siekierka, S H Y Hung, M Poe, C S Lin and N H Sigal *Nature* **341** (1989) 755

17 M W Harding, A Galat, D E Uehling and S L Schreiber *Nature* **341** (1989) 758

18 A J Findlay and L Radics *Can J Chem* **58** (1980) 579

19 F X Schmid *Annu Rev Biophys Biomol Struct* **22** (1993) 123

20 M Kruse, M Brunke, A Escher, A A Szalay, M Troschug and R Zimmermann *J Biol Chem* **270** (1995) 2588

21 M Thali *Mol Med Today* **1** (1995) 287

22 F J Dumont, M J Staruch, S L Koprak, M R Melino and N H Sigal *J Immunol* **144** (1990) 251

23 B E Bierer, P K Somers, T J Wandless, S J Burakoff, S L Schreiber *Science* **250** (1990) 556

24 B E Bierer, P S Mattila, R F Standaert, L A Herzenberg, S J Burakoff, G Crabtree and S L Schreiber *Proc Natl Acad Sci USA* **87** (1990) 9231

25 C B Klee, T H Crouch, M H Krinks *Proc Natl Acad Sci USA* **76** (1979) 6270

26 J Liu, J D Farmer Jr, W S Lane, J Friedman, I Weissman and S L Schreiber *Cell* **66** (1991) 807

27 H Kessler, H-R Loosli, H Oschkinat in *'Peptides 1984' Proc 18th Eur Pept Symp* (Ed U Ragnarsson Almqvist and Wiksell: Int Stockholm (1984) 65

28 H Kessler, H-R Loosli, H Oschkinat *Helv Chim Acta* **68** (1985) 661

29 H-R Loosli, H Kessler, H Oschkinat, H-P Weber, T J Petcher and A Widmer *Helv Cheim Acta* **68** (1985) 682

30 H Kessler, H Oschkinat, H-R Loosli in *Two Dimensional NMR Spectroscopy* (Eds W R Croasman and R M K Carlson) Verlag Chemie Weinheim (1987) p 259

31 H Kessler, C Griesinger and K Wagner *J Am Chem Soc* **109** (1987) 6927

32 (a) J Lautz, H Kessler, J M Blaney, R M Scheek, W F van Gunsteren *Int J Peptide Protein Res* **33** (1989) 281; (b) J Lautz, H Kessler, R Kaptein, W F van Gunsteren *J Comput Aided Mol Design* **1** (1987) 219; (c) J Lautz, H Kessler, H P Weber, R M Wenger, W F van Gunsteren *Biopolymers* **29** (1990) 1669

33 H Kessler, M Kock, T Wein and M Gehrke *Helv Chim Acta* **73** (1990) 1818

34 H Kessler, M Gehrke, J Lautz, M Kock, D Seebach and A Thaler *Biochem Pharmacol* **40** (1990) 169

35 J Kallen, C Spitzfaden, M G M Zurini, G Wider, H Widmer, K Wüthrich and M D Walkinshaw *Nature* **353** (1991) 276

36 K Wüthrich, C Spitzfaden, K Memmert, H Widmer and G Wider *FEBS Lett* **285** (1991) 237; (b) C Spitzfaden, W Braun, G Wider, H Widmer and K Wuthrich *J Biomol NMR* **4** (1994) 463

37 R T Clubb, S B Ferguson, C T Walsh and G Wagner *Biochemistry* **33** (1994) 2761

38 M Ottiger, O Zerbe, P Guntert and K Wuthrich *J Mol Biol*, **272** (1997) 64

39 L Sawyer *Nature* **327** (1987) 659

40 M E Newcomer, T A Jones, J Aqvist, J Sundelin, U Eriksson, I Rask and P A Peterson , *EMBO J* **3** (1984) 1451

41 R Huber, M Schneider, O Epp, I Mayr, A Messerschmidt, J Pflugrath and H Kayser *J Mol Biol* **195** (1987) 423

42 H M Holden, W R Rypniewski, J H Law and I Rayment *EMBO J* **6** (1987) 1565

43 M Z Papiz, L Sawyer, E E Eliopoulos, A C T North, J B C Findlay, R Sivaprasadarao, T A Jones, M E Newcomer and P J Kraulis *Nature* **324** (1986) 383

44 S W Fesik, R T Gampe Jr , T F Holzman, D A Egan, R Edalji, J R Luly, R Simmer, R Hefrich, V Kishore and D H Rich *Science* **250** (1990) 1406

45 H Ke, L D Zydowsky, J Liu and C T Walsh *Proc Natl Acad Sci U S A* **88** (1991) 9483

46 Y Theriault, T M Logan, R Meadows, L Yu, E T Olejniczak, T F Holzman, R Simmer and S W Fesik *Nature* **361** (1993) 88

47 C Weber, G Wider, B von Freyberg, R Traber, W Braun, H Widmer and K Wüthrich *Biochem* **30** (1991) 6563

48 S W Fesik, R T Gampe Jr, H L Eaton, G Gemmecker, E T Olejniczak, P Neri, T F Holzman, D A Egan, R Edalji, R Simmer, R Helfrich, J Hochlowski and M Jackson *Biochem* **30** (1991) 6574

49 V L Hsu and I M Armitage *Biochem* **31** (1992) 12778

50 R Freeman, T H Mareci and G A Morris *J Magn Reson* **42** (1981) 341

51 N H Sigal, F Dumont, P Durette, J J Siekierka, L Peterson, D H Rich, B E Dunlap, M J Staruch, M R Melino, S L Koprak, D Williams, B Witzel, J M Pisano *J Expt Med* **173** (1991) 619

52 V F J Quesniaux, M H Schreier, R M Wenger, P C Hiestand, M W Harding and M H V van Regenmortel *Eur J Immun* **17** (1987) 1359

53 J P Griffith, J L Kim, E E Kim, M D Sintchak, J A

Thomson, M J Fitzgibbon, M A Fleming, P R Caron, K Hsiao and M A Navia *Cell* **82** (1995) 507

54 P J Belshaw and S L Schreiber *J Am Chem Soc* **119** (1997) 1805

55 (a) M K Rosen, S W Michnick, M Karplus and S L Schreiber *Biochem* **30** (1991) 4774; (b) S W Michnick, M K Rosen, T J Wandless, M Karplus and S J Schreiber *Science* **251** (1991) 836

56 J A Moore, D A Peattie, M J Fitzgibbon and J A Thomson *Nature* **351** (1991) 248

57 T J Wandless, S W Michnick, M K Rosen, M Karplus and S L Schreiber *J Am Chem Soc* **113** (1991) 2239

58 (a) R P Meadows, D G Nettesheim, R X Xu, E T Olejniczak, A M Petros, T F Holzman, J Severin, E Gubbins, H Smith and S W Fesik *Biochem* **32** (1993) 754. (b) R X Xu, D Netlesheim, E T Olejniczak, R Meadows, G Gemmecker and S W Fesik *Biopolymers* **33** (1993) 535

59 T K Chakraborty, H P Weber and K C Nicalaou *Chem Biol* **2** (1955) 157

60 M K Rosen, R F Standaert, A Galat, M Nakatsuka and S L Schreiber *Science* **248** (1990) 863

61 P Karuso, H Kessler and D F Mierke *J Am Chem Soc* **112** (1990) 9434

62 D F Mierke, P Schnieder, P Karuso and H Kessler *Helv Chim Acta* **74** (1991) 1027

63 G D van Duyne, R F Standaert, P A Karplus, S L Schreiber and J Clardy *Science* **252** (1991) 839

64 G D van Duyne, R F Standaert, S L Schreiber and J Clardy, *J Am Chem Soc* **113** (1991) 7433

65 G Fischer, H Bang, C Mech *Biomed Biochim Acta* **43** (1984) 1101

66 (a) M W Albers, C T Walsh and S L Schreiber *J Org Chem* **55** (1990) 4984. (b) J L Kofron, P Kusmic, V Kishore, E Colon-Bonilla and D H Rich *Biochem* **30** (1991) 6127

67 K Wüthrich, B V Freyberg, C Weber, G Wider, R Traber, H Widmer and W Braun *Science* (1991) 953

68 W L Jorgensen *Science* (1991) 954

69 R M Wenger, J France, G Bovermann, L Walliser, A Widmer and H Widmer *FEBS Lett* **340** (1994) 255

70 A M Petos, J R Luly, H Liang and S W Fesik *J Am Chem Soc* **115** (1993) 9920

71 C A Lepre, D A Pearlman, J-W Cheng, M T DeCenzo, D L Livingston and J M Moore *Biochem* **33** (1994) 13571

72 A I Denesyuk, M Vihinen, J Lundell, V P Zav'yalov and T Korpela *Biochem Biophys Res Commun* **192** (1993) 912

73 J Luban, K L Bossolt, E K Franke, G V Kalpana and S P Goff *Cell* **73** (1993) 1067

74 (a) E K Frank, H E Yuan and J Luban *Nature* **372** (1994) 359 (b) M Thali, A Bukovsky, E Kondo, B Rosenwirth, C T Walsh, J Sodaroski and H G Gottlinger *Nature* **372** (1994) 363

75 M A McCornack, L T Kakalis, C Casesla, R E Handschumacher and I M Armitage *FEBS Lett* **414** (1997) 84

76 A Steinkasserer, R Harrison, A Billich, F Hammerschmid, G Werner, B Wolff, P Peichl, G Palfi, W Schnitzel, E Mlynar and B Rosen *J Virol* **69** (1995) 814

77 E De Clercq *Antiviral Res* **12** (1989) 1

78 A S Perelson, A U Neumann, M Markowitz, J M Leonard and D D Ho *Science* **271** (1976) 1582

79 Special Report: AIDS *Science* **272** (1996) 1876 (b) R J Wurtman *Nature Medicine* **3** (1997) 714

80 D Hazuda and L Kuo *Nature Medicine* **3** (1997) 836

81 G V T Swapna, B Jagannadh, M K Gurjar and A C Kunwar *Biochem Biophys Res Commun* **164** (1989) 1086

82 H Bazin, J Chattopadhyaya, R Datema, A-C Erickson, G Gilljam, N G Johansson, J Hansen, R Koshida, K Moelling, B Oberg, G Ramaud, G Stening, L Vrang, B Wahren and J C Wu *Biochem Pharmacol* **38** (1989) 109

83 H Rosemeyer, G Toth, B Golankiewicz, Z Kazimierczuk, W Bourgeois, U Kretschmer, H P Muth and F Sheela *J Org Chem* **55** (1990) 5784

84 G R Painter, J P Shockco and C W Andrews *Adv Mol Model* **2** (1990) 135

85 (a) B Jagannadh, D V Reddy and A C Kunwar *Biochem Biophys Res Commun* **179** (1991) 386; (b) A C Kunwar, B Jagannadh and D V Reddy *Magnetic Resonance Current Trends* (Ed. C L Khetrapal and G Govil) Narosa Publishing House, New Delhi (1991) 171

86 N Hicks, O W Howarth and D W Hutchinson *Carbohydrate Res* **216** (1991) 1

87 (a) J Plavec, L H Koole, A Sandstrom and J Chattopadhyaya *Tetrahedr* **47** (1991) 7363; (b) J Plavec, L H Koole, J Chattopadhyaya *J Biochem Biophys Methods* **25** (1992) 253

88 M K Gurjar, A C Kunwar, D V Reddy, A Islam, S V S Lalitha, B Jagannadh and A V Rama Rao *Tetrahedron* **49** (1993) 4373

89 D O Cicero, G Barbato and R Bazzo *Tetrahedron* **51** (1995) 10303

90 J J Barchi Jr, L S Jeong, M A Siddiqui and V E Marquez *J Biochem Biophys Methods* **34** (1997) 11

91 W Saenger *Principles of Nucleic Acid Structure*, Springer New York (1984)

92 C Altona and M Sundaralingam *J Am Chem Soc* **94** (1972) 8205

93 G I Birnbaum and D Shugar *Topics in Nucleic Acis Structure: Part 3* (Ed S Neidle) Macmillan London (1987) 1

94 P Van Roey, G M Salern, C K Chu and R F Schinazi *Proc Natl Acad Sci USA* **86** (1989) 3929

95 (a) D V Reddy, B Jagannadh and A C Kunwar *Proc Indian Acad Sci* **106** (1994) 1481; (b) D V Reddy, B Jagannadh and A C Kunwar, *J Biochem Biophys Methods* **31** (1996) 13

96 J P.Albrand, B Birdsall, J Feeney, G C K Roberts and A S V Burgen *Int J Biol Macromol* **1** (1979) 37

97 G R Painter, A E Aulabaugh, C W Andrews *Biochem Biophys Res Commun* **191** (1993) 1166

98 J Cohen *Science* **277** (1997) 32

99 M A Navia, P M D Fitzerald, B M Mckeever, C T Leu, J C Heimbach, W K Herber, I S Sigal, P L Drake and J P Springer *Nature* **337** (1989) 615

100 A Wlodawer and J W Erickson *Annu Rev Biochem* **62** (1993) 543

101 P Y S Lam, P K Jadhav, P K Eyermann, C N Hodge, Y Ru, L T Bacheler, J L Meek, M J Otto, M M Rayner,

Y N Wong, C H Chang, P C Weber, D A Jackson, T R Sharpe and S Erickson-Viitanen *Science* **263** (1994) 380
102 T Yamazaki, L K Nicholson, D A Torchia, S J Stahl, J D Kaufman, P T Wingfield, P J Domaille and S Campbell-Burk *Eur J Biochem* **219** (1994) 707
103 T Yamazaki, A P Hinck, Y –X Wang, L K Nicholson, D A Torchia, P Wingfield, S J Stahl, J D Kaufman, C H Chang, P J Domaille and P Y S Lam *Protein Science* **5** (1996) 495
104 S Grzesiek, A Bax, L K Nicholson, T Yamazaki, P Wingfield, S J Stahl, C J Eyermann, D A Torchia, C N Hodge, P Y S Lam, P K Jadhav and C H Chang *J Am Chem Soc* **116** (1994) 1581
105 M J M Burgering, R Boelens, M Caffrey, J N Breg and R Kaptein *FEBS Lett* **330** (1993) 105
106 P J M Folkers, R H A Folmer, R N H Konings and C W Hilbers *J Am Chem Soc* **115** (1993) 3798
107 T Yamazaki, L K Nicholson, P Wingfield, S J Stahl, J D Kaufman, C J Eyermann, C N Hodge, P Y S Lam, Y Ru, P K Jadhav, C Chong-Hwan, P C Weber and D A Torchia *J Am Chem Soc* **116** (1994) 10791
108 L K Nicholson, T Yamazaki, D A Torchia, S Grzesiek, A Bax, J D Kaufman, S J Stahl, P Wingfield, P Y S Lam, P K Jadhav, C N Hodge, P J Domaille, C Chong-Hwan *Nature Struct Biol* **2** (1995) 274
109 J R Rose, R Salto and C S Craik *J Biol Chem* **268** (1993) 11939
110 R Smith, I M Brereton, R Y Chai and S B H Kent *Nature Struct Biol* **3** (1996) 946
111 Y X Wang, D J Freedberg, T Yamazaki, P T Wingfield, S J Stahl, J D Kaufman, Y Kiso and D A Torchia *Biochemistry* **35** (1996) 9945
112 A M Mildner, D J Rothrock, J W Leone, C J Bannow, J M Lull, I M Reardon, J L Sarcich, W J Howe, C S C Tomich, C W Smith, R L Heinrikson and A G Tomasselli *Biochemistry* **33** (1994) 9405
113 P Poignard, P J Klasse and Q J Sattentau *Immunology Today* **17** (1996) 239
114 J R Rusche, K Javaherian, C McDanal, J Petro, D L Lynn, R Grimaila, A Langlois, R C Gallo, L O Arthur, P J Fischinger, D P Bolognesi, S D Putney and T J Metthews *Proc Natl Acad Sci USA* **85** (1988) 3198
115 T J Palker, M E Clark, A J Langlois, T J Matthews, K J Weinhold, R R Randall, D P Bolognesi and B F Haynes *Proc Natl Acad Sci USA* **85** (1988) 1932
116 J Goudsmit, C Debouck, R H Meloen, L Smit, M Bakker, M Asher, A V Wolff, C J Gibbs Jr and D C Gajdusek *Proc Natl Acad Sci USA* **85** (1988) 4478
117 S Matsushita, M Robert-Guroff, J Rusche, A Koito, T Hattori, H Hoshino, K Javaherian, K Takatsuki and S D Putney *J Virol* **62** (1988) 2107
118 G J LaRosa, J P Davide, K Weinhold, J A Waterbury, A T Profy, J A Lewis, A J Langlois, G R Dreesman, R N Boswell, P Shadduck, L H Holley, M Karplus, D P Bolognesi, T J Matthews, E A Emini and S D Putney *Science* **249** (1990) 932; *Science* **251** (1991) 811
119 K Javaherian, A J Langlois, G J LaRosa, A T Profy, J R Rusche, D P Bolognesi, W C Herlihy, S D Putney and T J Matthews *Science* **250** (1990) 1590
120 T Hattori, A Koito, K Takatsuki, H Kido and N Katunuma *FEBS Lett* **248** (1989) 48
121 D R Burton *Proc Natl Acad Sci USA* **94** (1997) 10018
122 P E Stephens, G Clements, G T Yarranton, J Moore *Nature* **343** (1990) 219
123 S J O'Brien and M Dean *Sci Amer* **277** (3) (1997) 28
124 J Cohen *Science* **275** (1997) 1261
125 K Chandrasekhar, A T Profy and H J Dyson *Biochem* **30** (1991) 9187
126 A Zvi, R Hiller and J Anglister *Biochem* **31** (1992) 6972
127 (a) R de Lorimier, M A Moody, B F Haynes and L D Spicer *Biochem* **33** (1994) 2055; (b) H M Vu, R de Lorimier, M A Moody, B F Haynes and L D Spicer *Biochem* **35** (1996) 5158
128 R L M Markert, H Ruppach, S Gehring, U Dietrich, D F Mierke, M Kock, H Rubsamen-Waigmann and C Griesinger *Eur J Biochem* **237** (1996) 188
129 P Catasti, J D Fontenot, E M Bradbury and G Gupta *Biol Chem* **270** (1995) 2224
130 G Gupta, G M Anantharamaiah, D R Scott, J H Eldrige and G Myer *J Biomol Struct & Dyn* **11** (1993) 345
131 P Catasti, E M Bradbury and G Gupta *J Biol Chem* **271** (1996) 8236
132 A V Sarma, T V Raju and A C Kunwar *J Biochem Biophys Methods* **34** (1997) 83
133 R Jelinek, A P Valente, K G Valentine and S J Opella, *J Magn Reson* **125** (1997) 185
134 A Zvi, I Kustanovich, D Feigelson, R Levy, M Eisenstein, S Matsushita, P Richaletsecorel, M Regenmortel and J Anglister *Eur J Biochem* **229** (1995) 178
135 J B Ghira, E A Stura, R L Stanfield, A T Profy and I A Wilson *Science* **264** (1994) 82
136 J Cohen *Science* **276** (1997) 1177
137 P J Lodi, J A Ernst, J Kuszewski, A B Kickman, A Engleman, R Craigie, G M Clore and A M Gronenborn *Biochem* **34** (1993) 9826
138 A P Eijkelenboom, R A Lutzke, R Boelens, R H A Plasterk, R Kaptein and K Hard *Nature Struct Biol* **2** (1995) 807
139 M Cai, R Zheng, M Caffrey, R Craigie, G M Clore, A M Gronenborn *Nature Struct Biol* **4** (1997) 567
140 J Kuriyan and D Cowburn *Curr Opin Struct Biol* **3** (1993) 828
141 *Intelligent Drug Design Nature (supplement)* **384b** (1996)
142 S P A Fodor, J L Read, M C Pirneng, L Stryer, A T Lu and D Solas *Science* **251** (1991) 767
143 K S Lam, S E Salman, E M Hersh, V J Hruby, W M Kazmierski and R J Knapp *Nature* **354** (1991) 82
144 R A Houghten, C Pinilla, S E Blondelle, J R Appel, C T Dooley and J H Cuervo *Nature* **354** (1991) 84
145 S H Dewitt, J S Kiely, C J Stankovic, M C Schroeder, M R Cody and M R Pavia *Proc Natl Acad Sci USA* **90** (1993) 6909
146 G C Look, C P Holmes, J P Chinn and M A Gallop *J Org Chem* **59** (1994) 7588
147 W L Fitch, G Detre, C P Holmes, J N Shoolery and P A Keifer *J Org Chem* **59** (1994) 7055
148 R C Anderson, M A Jarema, M J Shapiro, J P Stokes and M Ziliox *J Org Chem* **60** (1995) 2650
149 P A Keifer, L Baetusis, D M Rice, A A Tymiak and J N Shoolery *J Magn Reson* **A119** (1996) 65

150 S K Sarkar, R S Garigipati, J L Adams and P A Keifer *J Am Chem Soc* **118** (1996) 2305

151 M Pursch, G Schlotterbeck, L-H Tseng, K Albert and W Rapp *Angew Chem Int Ed Engl* **35** (1996) 2867

152 S B Suker, P J Hajduk, R P Meadows and S W Pesik *Science* **274** (1996) 1531

153 P J Hajduk, G Sheppard, D G Nettesheim, E T Olejniczak, S B Shuker, R P Meadows, D H Steinnsan, G M Carrera Jr , P A Marcotte, J Severin, K Walter, H Smith, E Gubbins, R Simmer, T F Holtzman, D W Morgan, S K Davidsen, J B Summers and S W Fesik *J Am Chem Soc* **119** (1997) 5818

154 H Kessler *Angew Chem Int Ed Engl* **36** (1997) 829