

A peptide inhibitor of HIV-1 protease using α,β -dehydro residues : A structure based computer model

Mohd. Imran Siddiqui, Sarika Kataria, Vivek Ahuja and Gita Subba Rao*

Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India

Accepted 3 October 2000

HIV-1 encodes an aspartic protease, an enzyme crucial to viral maturation and infectivity. It is responsible for the cleavage of various protein precursors into viral proteins. Inhibition of this enzyme prevents the formation of mature, infective viral particles and therefore, it is a potential target for therapeutic intervention following infection. Several drugs that inhibit the action of this enzyme have been discovered. These include peptidomimetic inhibitors such as ABT-538 and saquinavir, and structure based inhibitors such as indinavir and nelfinavir. Several of these have been tested in human clinical trials and have demonstrated significant reduction in viral load. However, most of them have been found to be of limited clinical utility because of their poor pharmacological properties and also because the viral protease becomes rapidly resistant to these drugs on account of mutations in the enzyme. One way to overcome these limitations is to design an inhibitor that interacts mainly with the conserved residues of HIV-1 protease. By a rational drug design approach based on the high resolution X-ray crystal structure of the HIV-1 protease with – MVT 101 (a substrate based inhibitor) and the specific design principles of peptides containing dehydro-Alanine (Δ Ala) derived from our earlier studies, we have designed a tetrapeptide with the sequence: NH_2 -Thr- Δ Ala- Δ Ala-Gln-COOH. Energy minimization and molecular modelling of the interaction of the designed tetrapeptide with the inhibitor binding site indicate that the inhibitor is in an extended conformation and makes excessive contacts with the viral enzyme at the interface between the protein subunits. The designed inhibitor has 33% of its interaction with the conserved region of HIV-1 protease which is of the same order as that of MVT 101 with the enzyme.

Introduction

The late stages in the replication of HIV include crucial virus-specific secondary processing of certain proteins by an enzyme HIV-1 protease. The activity of the protease is essential for the proper assembly and maturation of the fully infectious virions of HIV-1¹ as well as other retroviruses. Thus HIV-1 protease has become an important target for the rational design of antiviral agents for AIDS. HIV-1 protease functions as a symmetric homodimer compounds consisting of two identical 99-amino acids subunits which combine to form one active site with the signature sequence of Asp-Thr-Gly. The search for HIV protease inhibitors was launched after it was ascertained that HIV protease is required for viral infectivity¹. This search was facilitated by the vast knowledge of other aspartyl protease (renin inhibitors) the cloning and purification of the HIV protease, the elucidation of its three-dimensional(3-D) structure, and the development of rapid enzyme assays for screening potential inhibitors. In the past

few years there has been a virtual explosion of new X-ray crystal structures from numerous laboratories aimed at characterization, at an atomic level, of the structures of the HIV-1 protease-inhibitor complexes².

Several inhibitors of HIV-1 protease have been discovered to date. The identification of HIV-1 protease cleavage sites proved useful in designing the appropriate inhibitors. Several peptidomimetic HIV-1 protease inhibitors like Ro 31-8959³, U-81749⁴, KNI-227⁵, A-74704⁶, A-77003⁷ and L-700,414⁸ have been designed. Structure-based ligand design has been quite successful in predicting inhibitors of HIV-1 protease. Indinavir sulfate and nelfinavir mesylate were designed with the help of X-ray crystallography and molecular mechanics calculations.

The principal issue for the development of protease inhibitors seems to be the capacity of protease to tolerate single residue (point) mutations that renders it's inhibitors ineffective yet do not significantly affect the activity of the enzyme⁹. Analysis of data shows that mutations that produce inhibitor resistant proteases are concentrated in regions of the enzyme that have been shown to form subsites of the substrate binding pocket. Key amino acids whose change produces resistant phenotypes are at position 82 and

* Author for correspondence.

Tel: 91-11-6594816; Fax: 91-11-6862663;

Email: gitarao@midinst.ernet.in

84 in the S_1 and S_1' subsites, at position 32 in S_2 and S_2' subsite, at position 8, 46, 47 and 48 in S_4 and S_4' subsites. Most of these mutating residues have been shown to be critical for determining retroviral protease specificity¹⁰.

It is, therefore, desirable to design an inhibitor which has a large number of interactions with the conserved residues in the binding pocket of protease so as to minimize the loss of specificity of interactions on account of point mutations. The starting point and basis for the design of the inhibitor is the X-ray crystal structure of a complex between a substrate based peptide inhibitor with the sequence: N-Ac-Thr-Ile-Nle- ψ [CH₂-NH]-Nle-Gln-Arg-amide(Nle,Norleucine), also called MVT 101, with chemically synthesised HIV-1 protease determined at 2.3Å resolution¹¹ (PDB accession number 4HVP). The inhibitor binds to the protease in an extended conformation. The geometry of the binding pocket suggests that the required peptide inhibitor should have an extended conformation. It has been shown in our earlier work¹²⁻¹⁴ that peptides containing some α,β -unsaturated (dehydro) alanine residues do have such a conformation.

Dehydro amino acids play an important role in the inverse protein folding problem, i.e., the problem of predicting the amino acid sequence that would fold into a required 3-D structure. They occur naturally in a variety of peptide antibiotics¹⁵⁻¹⁸ and in some

proteins¹⁹. The steric effects caused by dehydro residue influence the backbone conformation in a specific manner. Therefore such residues can be used to generate specific peptide structures²⁰. X-ray crystallography^{12,21-23}, NMR spectroscopy²⁴ and theoretical calculations^{21, 25-27} on peptides containing dehydro-Ala (Δ Ala) have shown that the dehydro residue adopts an extended conformation. It also induces steric effects on the neighbouring saturated residues and forces them to adopt a constrained conformation¹². Thus, peptides containing dehydro residues are highly suitable for designing specific structures. In addition to their occurring naturally, there is an evidence to suggest that peptides containing dehydro amino acids are more resistant to enzymatic degradation^{28,29}.

Using the design principles generated with Δ Ala we had, in our earlier work, designed an inhibitor for bovine trypsin which had a structure similar to that of the binding loop region of bovine pancreatic trypsin inhibitor (BPTI). We have also designed an inhibitor of HIV-1 reverse transcriptase³⁰, which is much more resistant to mutations as compared to the existing drugs.

Materials and Methods

Although the minimal length substrate for HIV-1 protease was observed to be a heptapeptide³¹, subsequent studies showed that tri and tetrapeptide

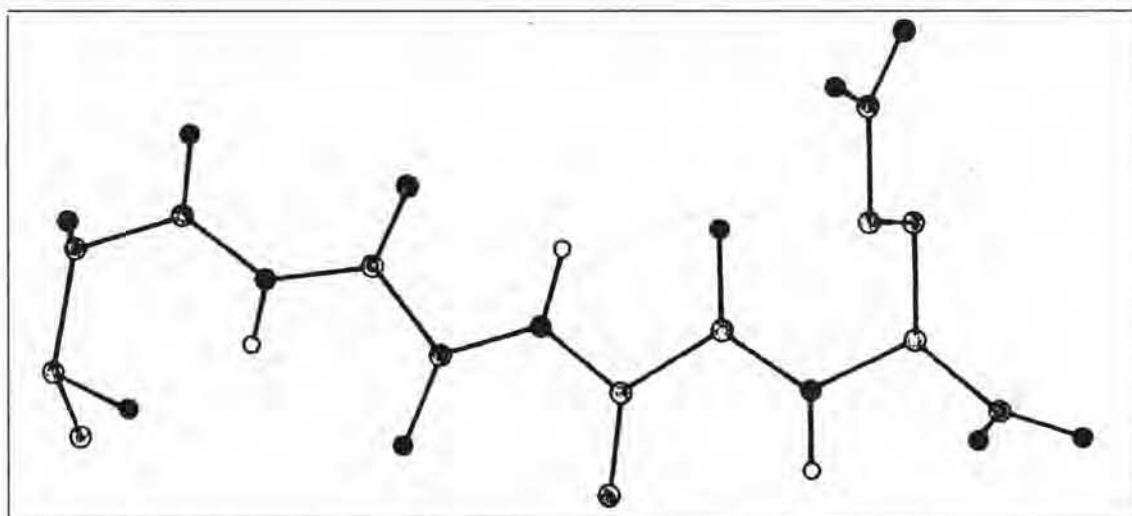


Fig. 1—Theoretically calculated minimum energy conformation structure of NH₂-Dehydro-Ala-Dehydro-Ala-Gln-COOH

Table 1—Minimum energy conformation of the designed HIV-1 protease inhibitor with the sequence: NH₂-Thr- Δ Ala- Δ Ala-Gln-COOH

ϕ_1	ψ_1	χ_1	ϕ_2	ψ_2	ϕ_3	ψ_3	ϕ_4	ψ_4	χ_4
157°	-74.7	26.7	180	150	180	150	156.6	-51.3	53.1

mimetics can also be potent inhibitors³ as the strength of binding is not a function of the total number of contacts or the total number of subsite interactions². A minimal inhibitor model consists of a core structure containing four side chains corresponding to residues P_2 through P_2' . The protease side chains comprising the pockets S_1 and S_1' are mostly hydrophobic. Thus, the inhibitor should have a hydrophobic moiety at P_1 and P_1' . The S_2 and S_2' pockets are also hydrophobic but both hydrophobic and hydrophilic residues can occupy these sites. The positional preference is greatest for P_2' which is occupied almost exclusively by glutamine or glutamic acid and for P_1 which is almost invariably occupied by hydrophobic residues.

The X-ray crystal structure of HIV-1 protease complexed with MVT 101¹¹ was used as the starting point. The binding site was carefully studied to find out prospective shapes that might fit in the binding site.

Design of the peptide

Theoretical calculations done on peptide sequences containing Δ Ala¹²⁻¹⁴ indicated that dehydro alanine is

always found in an extended conformation with $\phi = 180^\circ$ and $\psi = 150^\circ$. Moreover, the neighbouring saturated residues are constrained to adopt a conformation in which $\phi = 140^\circ$ and $\psi = -40^\circ$. Therefore, a peptide sequence with suitably placed dehydro-alanine residues can have an extended conformation.

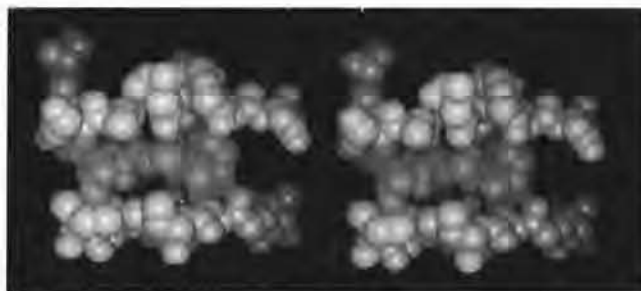


Fig. 2—Representation of final docked position of the designed inhibitor in yellow in drug binding cavity of HIV-1 protease formed by residues shown in pink

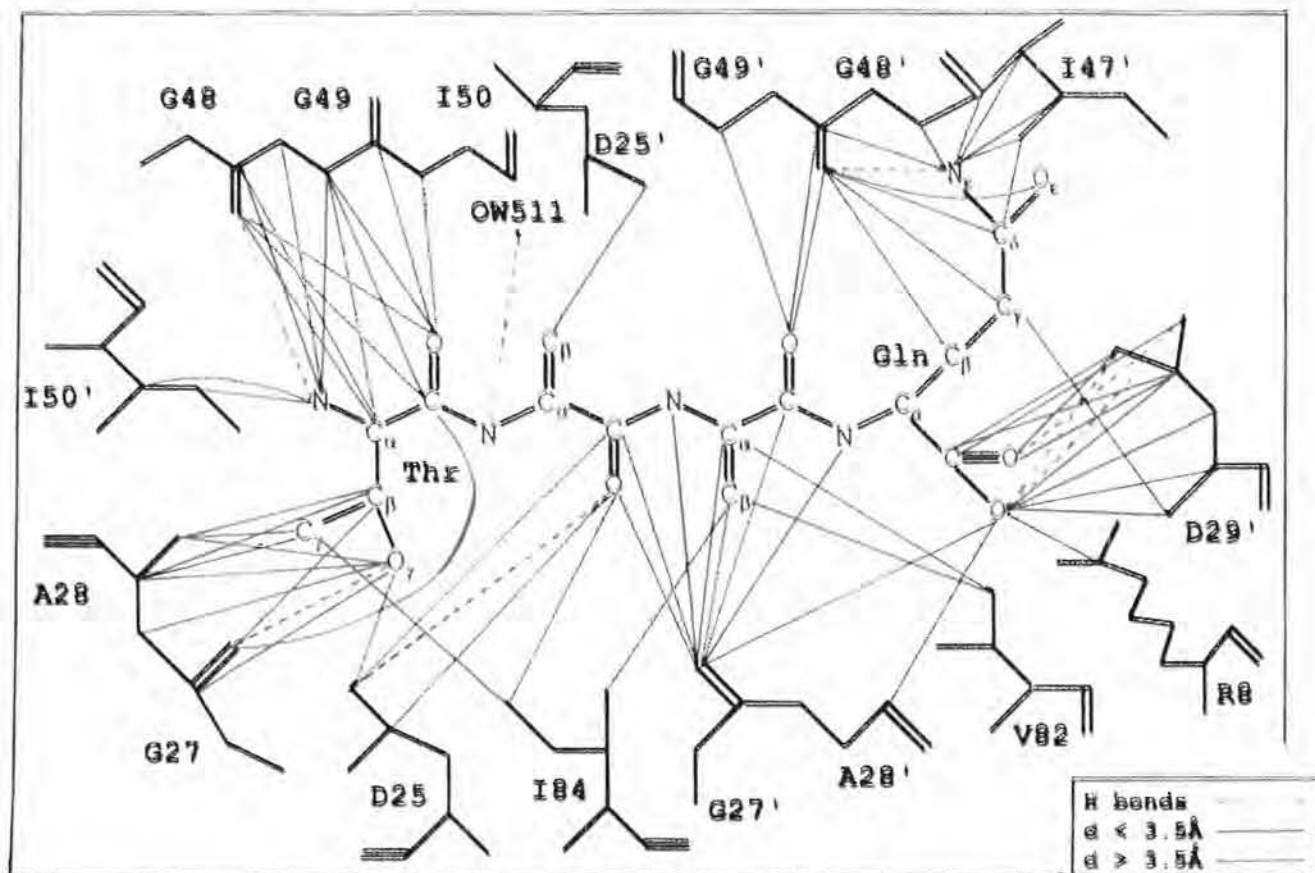


Fig. 3—Contacts between drug binding pocket of HIV-1 protease and inhibitor

Table 2—Hydrogen bond lengths and van der Waals distances between the inhibitor and HIV-1 protease

Hydrogen bond distances between inhibitor and protease			van der Waals contacts between inhibitors and main chain atoms of protease		
Inhibitor atom	Protease atom	Distances (Å)			
Thr ¹¹ (N)	Gly ⁴⁸ (O)	2.5	Gln ⁴¹ (N)	Gly ²⁷ (O)	3.5
Thr ¹¹ (Oγ1)	Gly ²⁷ (O)	2.5	Gln ⁴¹ (Cβ)	Gly ⁴⁸ (O)	3.6
ΔAla ²¹ (N)	Wat ⁵¹¹	3.0	Gln ⁴¹ (Cγ)	Gly ⁴⁸ (O)	3.8
ΔAla ²¹ (O)	Asp ²⁵ (Oδ2)	2.9	Gln ⁴¹ (Cγ)	Asp ²⁰ (N)	4.0
Gln ⁴¹ (Ne2)	Gly ⁴⁸ (O)	2.6	Gln ⁴¹ (Cδ)	Gly ⁴⁸ (O)	3.3
Gln ⁴¹ (O)	Asp ²⁹ (Oδ2)	2.7	Gln ⁴¹ (Oε1)	Gly ⁴⁸ (O)	4.2
C-Ter-O	Asp ²⁹ (Oδ2)	2.9	Gln ⁴¹ (Ne2)	Ile ⁴⁷ (Cα)	4.2
C-Ter-O	Asp ²⁹ (Oδ1)	2.6	Gln ⁴¹ (Ne2)	Ile ⁴⁷ (Cβ)	3.4
van der Waals contacts between inhibitors and main chain atoms of protease			Gln ⁴¹ (Ne2)	Ile ⁴⁷ (C)	4.2
Thr ¹¹ (N)	Gly ⁴⁸ (C)	3.3	Gln ⁴¹ (Ne2)	Gly ⁴⁸ (N)	3.5
Thr ¹¹ (N)	Gly ⁴⁹ (N)	3.6	Gln ⁴¹ (Ne2)	Gly ⁴⁸ (C)	3.7
Thr ¹¹ (N)	Gly ⁴⁹ (Cα)	3.4	C-Ter-O	Asp ²⁹ (Cβ)	3.4
Thr ¹¹ (N)	Ile ⁵⁰ (Cβ)	3.8	C-Ter-O	Asp ²⁹ (Cα)	3.8
Thr ¹¹ (Cα)	Gly ⁴⁸ (O)	3.1	C-Ter-O	Asp ²⁹ (N)	3.3
Thr ¹¹ (Cα)	Gly ⁴⁸ (C)	4.0	C-Ter-O	Ala ²⁸ (C)	4.1
Thr ¹¹ (Cα)	Gly ⁴⁹ (Cα)	3.7	C-Ter-O	Gly ²⁷ (O)	3.9
Thr ¹¹ (Cβ)	Ala ²⁸ (Cα)	3.5	van der Waals contacts between inhibitor and side chains of mutable residues		
Thr ¹¹ (Cβ)	Ala ²⁸ (Cβ)	3.4	Thr ¹¹ (Cγ2)	Ile ⁸⁴ (Cδ1)	3.2
Thr ¹¹ (Cβ)	Gly ²⁷ (O)	3.5	ΔAla ²¹ (O)	Ile ⁸⁴ (Cδ1)	3.9
Thr ¹¹ (Oγ1)	Ala ²⁸ (Cα)	3.1	ΔAla ³¹ (Cα)	V ⁸² (Cγ1)	4.2
Thr ¹¹ (Oγ1)	Ala ²⁸ (Cβ)	3.1	ΔAla ³¹ (Cβ)	V ⁸² (Cγ1)	3.0
Thr ¹¹ (Oγ1)	Ala ²⁸ (N)	3.5	ΔAla ³¹ (Cβ)	Ile ⁸⁴ (Cγ2)	4.1
Thr ¹¹ (Oγ1)	Gly ²⁷ (C)	3.2	Gln ⁴¹ (Cδ)	Ile ⁴⁷ (Cγ2)	4.1
Thr ¹¹ (Cγ2)	Ala ²⁸ (Cα)	3.6	Gln ⁴¹ (Ne2)	Ile ⁴⁷ (Cγ2)	4.2
Thr ¹¹ (Cγ2)	Ala ²⁸ (Cβ)	2.8	C-Ter-O	R ⁸ (Cz)	4.2
Thr ¹¹ (C)	Gly ⁴⁸ (O)	4.0	van der Waals contacts between inhibitor and side chains of other residues		
Thr ¹¹ (C)	Gly ⁴⁹ (Cα)	3.5	Thr ¹¹ (N)	Ile ⁵⁰ (Cγ1)	3.4
Thr ¹¹ (C)	Gly ²⁷ (O)	2.6	Thr ¹¹ (Oγ1)	Asp ²⁵ (Oδ2)	4.0
Thr ¹¹ (O)	Gly ⁴⁸ (O)	4.2	ΔAla ²¹ (Cβ)	Asp ²⁵ (Oδ1)	3.5
Thr ¹¹ (O)	Gly ⁴⁹ (Cα)	3.1	ΔAla ²¹ (C)	Asp ²⁵ (Oδ2)	4.0
Thr ¹¹ (O)	Gly ⁴⁹ (C)	3.8	ΔAla ²¹ (O)	Asp ²⁵ (Cγ)	4.0
Thr ¹¹ (O)	Ile ⁵⁰ (N)	3.9	Gln ⁴¹ (Ne2)	Ile ⁵⁰ (Cδ1)	3.9
ΔAla ²¹ (C)	Gly ²⁷ (O)	4.0	Gln ⁴¹ (C)	Asp ²⁹ (Oδ2)	3.0
ΔAla ²¹ (O)	Gly ²⁷ (O)	4.2	Gln ⁴¹ (C)	Asp ²⁹ (Cγ)	3.3
ΔAla ³¹ (N)	Gly ²⁷ (O)	3.5	Gln ⁴¹ (C)	Asp ²⁹ (Oδ1)	3.4
ΔAla ³¹ (Cα)	Gly ²⁷ (O)	3.1	Gln ⁴¹ (O)	Asp ²⁹ (Cγ)	3.3
ΔAla ³¹ (Cβ)	Gly ²⁷ (O)	4.0	Gln ⁴¹ (O)	Asp ²⁹ (Oδ1)	3.4
ΔAla ³¹ (C)	Gly ²⁷ (O)	3.6	C-Ter-O	Asp ²⁹ (Cγ)	2.7
ΔAla ³¹ (O)	Gly ⁴⁹ (Cα)	3.5			
ΔAla ³¹ (O)	Gly ⁴⁸ (C)	4.1			
ΔAla ³¹ (O)	Gly ⁴⁸ (O)	3.2			

Table 3—Inhibitor /HIV PR contacts in subsites S_2/S_2' and S_1/S_1'

Inhibitor residue	Subsite	Enzyme residue	No. of contacts
Thr	S_2	D25, G27, A28, G48, I50, I84, I50'	27
Gln	S_2'	R8, I50, G27', A28', D29', I47', G48'	30
Δ Ala	S_1	D25, I84, D25', G27'	8
Δ Ala'	S_1'	V82, I84, G27', G48', G49'	10

Based on the above consideration, a peptide was designed with the following sequence :

$\text{NH}_2\text{-Thr-}\Delta\text{Ala-}\Delta\text{Ala-Gln-COOH}$

A "model" structure was assigned to the peptide on the basis of the folding rules for peptides containing Δ Ala. This structure was used as a starting point for further energy minimization using the Global optimization method of Subba Rao *et al.*³², with modified torsional potentials for Δ Ala, the details of which are described in our earlier papers^{13,14}.

Molecular modeling

The modeling of the interaction of the designed inhibitor with the drug-binding region of HIV-1 protease was carried out on a Silicon Graphics INDIGO workstation, using Biopolymer and Docking modules of the BIOSYM software package. A manual docking of the inhibitor was done with the binding pocket of protease as determined by the X-ray crystal structure¹¹, after extraction of MVT 101 molecule. All steric clashes were removed by turning on the 'bump monitor'.

Results and Discussion

The minimum energy conformation of the inhibitor (Fig. 1 and Table 1) indicates that the inhibitor is in an extended conformation, as has been shown for other known inhibitors.

The inhibitor molecule makes extensive contacts with the enzyme at the interface between the protein subunits as can be seen in space filling model in Fig. 2. The inhibitor binds to the enzyme in an asymmetric fashion, with only the carboxy terminal amino acid partially exposed to the surrounding solvent.

Extensive van der Waals contacts of the inhibitor with the enzyme residues define the binding pockets of the complex (Fig. 3 and Table 2). Side chains P_2 and P_2' are located in well defined hydrophobic pockets (Table 3). There are a total of 83 interactions of 4.2Å or less. Out of these, 33% are with residues

which are sites of resistance mutations, which is of the same order as those of MVT 101 with the enzyme.

Analysis of the subsite composition reveals that some residues contact the inhibitor at more than one subsite position (Table 3). Residues G27, A28, G48, G49, I50, I50' are unique to subsite S_2 , while residues A28', D29' and I47' are unique to S_2' . The preferences G27', A28' and D29' for S_2' ; G27', G48' and G49' for S_1' ; D25, G27, G48, G49 and I50 for S_2 which have been observed for many existing inhibitors are also seen here, while they are absent in MVT 101.

Most of the current substrate analog inhibitors are those in which the scissile $P_1\text{-}P_1'$ amide bond has been replaced by a non-hydrolyzable isostere with tetrahedral geometry. In contrast with the inhibitor designed by us, we have used another variant of normal saturated residues — dehydro residues, which apart from bestowing a unique conformation on the peptide, are also resistant to enzymatic degradation.

Acknowledgement

We thank Prof. T.P. Singh for making available to us the Silicon Graphics workstation. One of the authors, GSR thanks the National AIDS Control Organization, Ministry of Health and Family Welfare for financial support, while MIS thanks the Council of Scientific and Industrial Research for the grant of a Senior Research Fellowship.

References

- Kohl NE, Emini EA, Schleif WA, Davis LJ, Heimbach JC, Dixon RAF, Scolnick EM & Sigal IS (1988) *Proc Nat Acad Sci USA* 85, 4686-4690
- Wlodawer A & Erickson JW (1993) *Annu Rev Biochem* 62, 543-585
- Roberts NA, Martin JA, Kitchington D, Broadhurst AV, Craig JC, Duncan IB, Calpin SA & Handa BK (1990) *Science* 248, 358-361
- McQuade TJ, Tomaselli AG, Liu L, Karacostas V, Moss B, Sawyer TK, Heinrichson RL & Orson FM (1990) *Science* 247, 4543-456
- Kageyama S, Mimoto T, Murakawa Y, Nomiza M, Ford H, Shirasaka T, Gulnik S, Erickson J, Takada K, Hayashi H, Broader S, Kiso Y & Mitsuya H (1993) *Antimicrob Agents Chemother* 37, 810-817
- Erickson J, Neidhart DJ, Van Drie J, Kempf DJ, Wang XC, Norbeck DW, Plattner JJ, Rittenhouse JW, Turon M, Wideburg N, Kohlbrenner WE, Simmers R, Helfrich R, Paul DA & Knigge M, (1990) *Science* 249, 527-533
- Kempf DJ, Marsh KC, Paul DA, Knigge MF, Norbeck DW, Kohlbrenner WE, Codacavi L, Vasavanonda S, Bryant P, Wang XC, Wideburg NE, Clement JJ, Plattner JJ & Erickson J, (1991) *Antimicrob Agents Chemother* 35, 2209-2214
- Bone R, Vacca JP, Anderson PS & Holloway MK. *J Am Chem Soc* 113, 9382-9384

- 9 Moore ML & Huffman WT (1993) *Perspective in Drug Discovery and Design* 1, 85-108
- 10 Leis J & Cameron CE (1994) *Curr Opin Biotech* 5, 403
- 11 Miller M, Schneider J, Sathyanarayan BK, Toth MV, Marshall GR, Clawson L, Selk L, Kent BH & Wlodawer A (1989), *Science* 246, 1149- 1152
- 12 Padmanabhan B, Dey S, Khandelwal B, Subba Rao G & Singh TP (1992) *Biopolymers* 32, 1271-1276
- 13 Bhatnagar S, Subba Rao G & Singh TP, (1995) *Biosystems* 34, 143- 148
- 14 Subba Rao G, Siddiqui MI & Arora S (1996) *J Biomol Struct Dynam* 14, 21-24
- 15 Gross LM & Morell JL (1967) *J Am Chem Soc* 89, 2791-2792
- 16 Gross LM & Morell JL & Craig CL (1964) *Proc Natl Acad Sci USA* 62, 953-961
- 17 Abraham EP & Newton GCF (1961), *Biochem J*, 79, 377-393
- 18 Allgaier H, Jung G, Werner RG, Schneider U & Zaehner H (1986) *Eur J Biochem* 160, 9-22
- 19 Givot IL & Abeles RH (1970) *J Biol Chem* 245, 3271-3273
- 20 Singh TP, Narula P & Patel HC (1990) *Acta Cryst B*46, 539-545
- 21 Ajo D, Granozzi G, Tondello E, Del Pra & Zanotti J (1979) *J Chem Soc Perkin Trans 2*, 927-929
- 22 Palmer DE, Pattaroni K, Nunami K, Chaddha RK, Goodman M, Wakamiya T, Fukase K, Horimoto S, Katisawa M, Fujita H, Kubo A & Shiba T (1992) *J Am Chem Soc* 114, 5634-5642
- 23 Pascard C, Ducruix A, Lunel J & Prange T (1977) *J Am Chem Soc* 99, 6418-6423
- 24 Gupta A & Chauhan VS (1990), *Biopolymers* 30, 395-403
- 25 Ajo D, Granozzi G & Tondello E (1977) *J Mol Struct* 41, 131-137
- 26 Ajo D, Granozzi G, Tondello E & Del Pra *Biopolymers* (1980) 19, 469-475
- 27 Ajo D, Casarin N & Granozzi G (1982) *J Mol Struct* 86, 297-300
- 28 Fisher GH, Breyer P, Ryan JW, Chauhan VS & Stammer CH (1981) *Arch Biochem Biophys* 211, 269-275
- 29 Brady SF, Cochran DW, Ruth FN, Holly FW, Bennett CD, Palevada WJ, Curley PE, Arison BH, Saperstein R & Veber DF (1984) *Int J Peptide Protein Res* 23, 212-222
- 30 Subba Rao G, Kataria S & Siddiqui MI (1998) *J Biomol Struct Dynam* 16, 347-354
- 31 Moore ML, Bryan WM, Fakhoury SA, Maguard VW & Hoffman WF (1989) *Biochem Biophys Res Commun* 159, 420
- 32 Subba Rao G, Tyagi RS & Mishra RK (1981) *J Theor Biol* 90, 377-389