HIV-1 Protease crystallography at BARC

M. V. Hosur AND Vishal Prashar

Abstract | Human Immunodeficiency Virus (HIV) is a complex retrovirus, and is the causative agent of Aquired Immuno Deficiency Syndrome (AIDS), a disease condition that is proving fatal to millions of Indians. Since a HIV/AIDS pandemic can destroy the very fabric of nations, devastate the most productive members of the societies, and orphan children, there is an urgent need to rein-in this disease. Crystallography can play a major role in this endeavour by providing accurate structural data on the chemical interactions between a chosen target and the designed inhibitor. Because of its essential role in the maturation process of HIV, the viral protease (HIV-1 PR) is an important target for inhibitor design. Many of the effective anti-HIV drugs in the market are inhibitors of HIV-1 PR, optimized for active-site binding through high-resolution crystal structures. However, occurence of drug resistant mutations in the protease necessitates continuous development of newer inhibitors through novel strategies. Inhibitors based on the mechanism of the enzyme are less likely to be overcome by point mutations, and thus may represent a novel class of inhibitors. Although hundreds of structures of active HIV-1 protease inhibitor complexes are reported, there are no reports of active enzyme substrate complexes. We have produced, for the first time, atomic-level snap-shots of enzyme-substrate complexes trapped at different stages of the cleavage reaction. These crystal structures may provide insight into the molecular mechanism of HIV-1 PR. For example, the formation of a low barrier hydrogen bond between catalytic aspartates had not been envisaged in earlier mechanistic proposals. Similarly, the hydrogen bond between cleavage products suggests protonation of scissile nitrogen atom by the hydroxyl group of the tetrahedral intermediate, in the reaction mechanism. The structures also open up the possibility that the enzyme uses different mechanisms to cleave substrates containing either a proline residue or a non-proline residue at the cleavage site. Finally, the structures provide a template for the design of mechanism-based inhibitors.

1. Introduction

Human Immunodeficiency Virus (HIV), discovered 25 years ago and established to be the causative agent of Aquired Immune Deficiency Syndrome (AIDS), has already claimed 25 million human lives. It is further estimated that 40 million people are already infected world wide underlining the magnitude of havoc this virus is causing, especially for the developing countries, where 90% of the infected people live. The number of Indians infected is estimated to be about 2.5 million. What is very alarming is that this virus strikes people during their most productive age, thereby affecting the economies of their families and even of the state. Viral infection is diagnosed through ELISA and Western Blot techniques by detecting antibodies to

Solid State Physics Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085
hosur@barc.gov.in

the virus in the blood sample. However, antibodies are generally not detectable until 3–6 months following infection. After exposure to HIV, some people have a flu-like illness that lasts between a week to a month, and these people suffer from: fever and sweats, headache, enlarged lymph nodes, fatigue, weight loss, persistent skin rashes or yeast infections, and short-term memory loss [1–3]. These symptoms are all those caused generally by a decrease in the white blood T cells (lymphocytes) carrying the CD4 receptor molecule. HIV infection causes depletion of CD4+ lymphocytes at a rate of 2 billion per day. Fortunately, the virus is very labile, and the half-life of the free virus is only six to eight hours. There is no known cure against HIV/AIDS. Inhibitors of the three viral enzymes are being developed as drugs against AIDS by controling replication of the virus inside target cells.

We, at the Bhabha Atomic Research Centre, Trombay, Mumbai, have focused on one of the three enzymes, viz. HIV protease (HIV PR). We are using the technique of single-crystal X-ray crystallography to understand the enzyme mechanism, and also for purposes of structure-based drug design [4–12]. We have provided, for the first time, structural snapshots of the enzyme '*in action*' [7–9]. Drug-resistant mutants of HIV-1 PR have been prepared in the laboratory using methods of genetic engineering. Three dimensional structures of these mutants complexed with resistance causing drugs have been determined in an effort to understand, at an atomic level, the phenomenon of drug resistance [10–12]. Such knowledge is the driving force behind development of new drugs against AIDS. This review, however, is limited to results on active protease true substrate complexes.

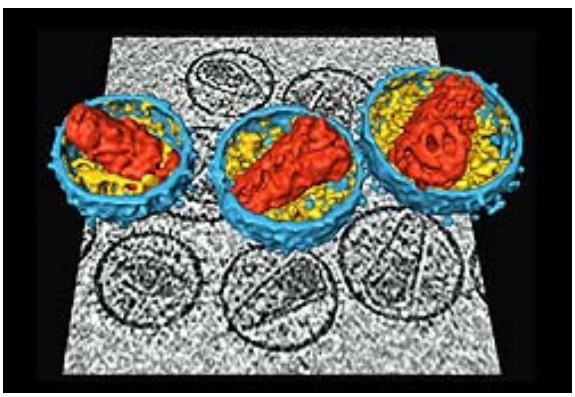
2. HIV: Morphology and molecular biology

HIV is a complex virus consisting of a membrane surrounding a capsid (greek for shell) which encloses two copies of viral RNA chromosome, and a number of other viral proteins. The size and shape of the virus change during its life cycle, and Figure 1 shows the first three dimensional view of an infectious HIV, obtained using the technique of single particle reconstruction of cryo-EM images recorded at various tilt angles [13]. The membraneous bilayer is embedded with viral proteins (gp120, gp40), which recognize receptors on target cells. The capsid is assembled from the viral capsid protein (p24), while the matrix protein provides the connection between the envelope and the capsid.

HIV is a retrovirus and belongs to the family of lentiviruses. There are two types of HIV, HIV-1 and HIV-2, the vast majority type being HIV-1. HIV-1 is more virulent compared to HIV-2. Within type 1, there are three groups of the virus, M, O and N depending on the antigenic properties. In the most prevalent group, M, there are sub-types, A, B, C, D, F, G, H, J and K. Different sub-types are concentrated in different parts of the world; A in Central and Western Africa, B in Europe, Japan, Australia and America, C in South Africa, India and Nepal. The genome of HIV-1 is a short RNA molecule of length 9200 kbases, which codes for structural and regulatory proteins. There are no introns in the viral genome. Arrangement of the genes is shown in Figure 2.

The genome is composed of 9 genes encoding 3 structural (GAG-derived MATRIX, CAPSID and Nucleo Capsid), 2 envelope (gp120 and gp41), and 6 regulatory proteins in addition to 3 enzymes

Figure 1: Cryo-EM image of infectious HIV. The oblong shaped capsid(red) containing the chromosome is surrounded by cellular membrane(blue), with matrix protein (yellow) occupying the region in between.



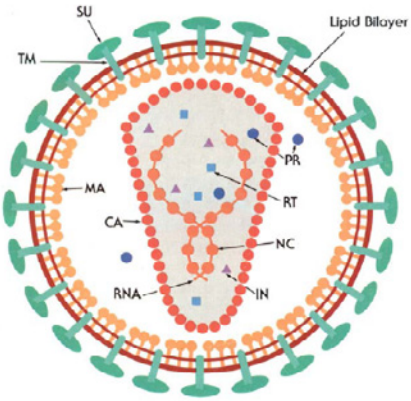


Figure 2: HIV-1 genome and arrangement of genes. No introns are present in the genome.

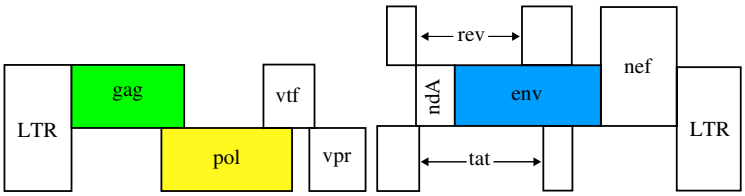
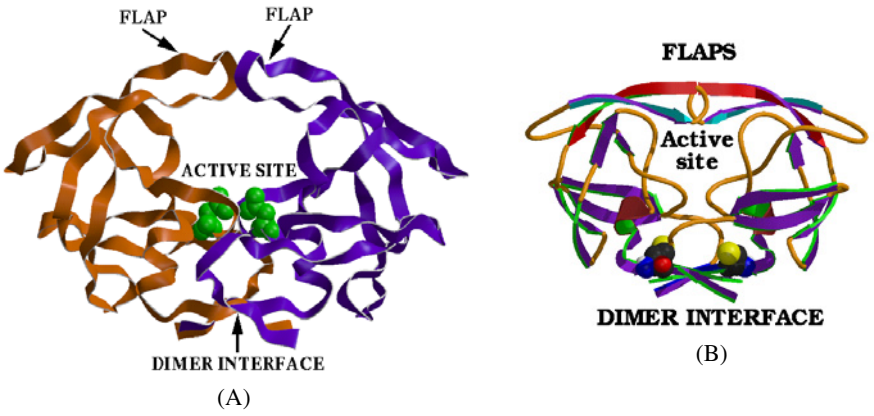


Figure 3: The HIV-1 PR homo-dimer. (A) open-flap conformation. Each sub-unit is drawn as a ribbon of single colour. The FLAP domain, the 4-stranded beta sheet dimer interface and the active site are indicated. The side chains of catalytic aspartic acids (one from each sub-unit) drawn in green. (B) closed-flap conformation.



(Reverse Transcriptase, INtegrase and PRotease). The structural proteins form part of the capsid, while the envelope proteins are utilized in the recognition of the target cell. The three enzymes are essential for the replication of the virus to produce infectious progeny. Regulatory genes contain information required for the production of proteins that control HIV's ability to infect a cell, reproduce and cause disease. The specific role of each of the regulatory protein is identified. It is estimated that, the virus also uses almost 250 host proteins during its replication cycle. The whole genome is transcribed and then the full-length RNA is spliced appropriately to produce messenger RNA molecules corresponding to each protein. Differential splicing of the viral genome controls, in a temporal fashion, the expression of structural and regulatory genes. Replication cycle is similar to, but under much tighter control than, that for other retroviruses. In HIV, the **RNA** replication is extremely rapid, but without any proof-reading mechanism. Therefore mutations accumulate at a rate 10^4–10^6 times that observed in DNA counterparts. As a consequence, the progeny of a single virus can differ greatly in antigenic configuration (gp120 and gp40 proteins) from the parent. This antigenic variation makes development of a vaccine extremely difficult in the case of HIV. Inhibitors of the three viral enzymes essential for viral replication and infectivity are being developed as drugs against AIDS. The success of HAART (Highly Active Anti-Retroviral Therapy) is largely due to presence of inhibitors of the PR enzyme in the medicine cocktail.

3. HIV PR

HIV PR plays a critical role in the life cycle of HIV [14–16]. The polyproteins that are produced in the course of the viral replication process have to be cleaved at nine different sites for a successful replication. The amino acid sequences at these cleavage sites are given in Table 1, where '$*$' denotes the peptide bond cleaved. The first residues toward the N- and C-terminal sides of the scissile bond are designated P1 and P1' respectively.

It may be seen that almost all cleavage sites are hydrophobic in character, have widely varying

Table 1: Amino acid sequences at the cleavage sites of HIV –1 PR.

Sr. No.	Peptide sequences	Cleavage domain (k_{cat}/K_M)
Cleavage sites in GAG:		
1.	SQNY*PIVQ	MA-CA (17980)
2.	ARVL*AEAM	CA-p2 (720)
3.	ATIM*MQRG	p2-NC (27970)
4.	QANF*LGKI	NC-p1 (NA)
5.	PGNF*LQST	p1-p6 (370)
Cleavage sites in POL:		
6.	SFNF*PQIT	TF-PR (113, 400)
7.	TLNF*PISP	PR-RT (8930)
8.	AETF*YVDG	RT-RH (11, 400)
9.	RKIL*FLDG	RH-IN (845)

k_{cat}/K_M values [17], and three of them contain a proline residue in the P1' position. HIV PR is active over a pH range of 3–6.5, the optimum pH being 5.5. The functional enzyme contains two identical polypeptides non-covalently associated into a dimer. Each monomer has four structural elements: two distinct hairpin loops, a wide loop containing the catalytic aspartic acid and an alpha helix. The crystal structure established HIV-1 PR to be an aspartyl protease with each sub-unit contributing one aspartic acid to the active site located at the interface of the two sub-units (Figure 3A). The two sub-units are held together through: a four stranded anti-parallel beta sheet, the conserved complex scaffold of hydrogen bonds called the "fireman's grip", and several hydrophobic interactions between residues forming loops in each subunit. The beta hair pin described as the 'flap' forms a flexible domain of each sub-unit. The flap becomes ordered and folds onto the active site to form what is described as a 'closed flap' conformation, when bound by a ligand in the active site cavity (Figure 3B). It has been shown that flap closure and specific hydrogen bonds between flap residues and the substrate are essential for catalysis [18]. In the three dimensional structure, the C-terminus of one sub-unit is about 5 Å away from the N-terminus of the second sub-unit. These two termini have been linked covalently thorough oligopeptides to produce what are known as tethered HIV-1 PRs [19]. The activity and substrate specificity of tethered HIV-1 PRs are shown to be identical to those of the native dimeric enzyme. Site directed mutagenesis studies using the tethered HIV-1 PR have established

Figure 4: General acid-base mechanism proposed for peptide bond cleavage by HIV-1 PR. Note absence of hydrogen bonding between inner oxygen atoms of catalytic aspartates.

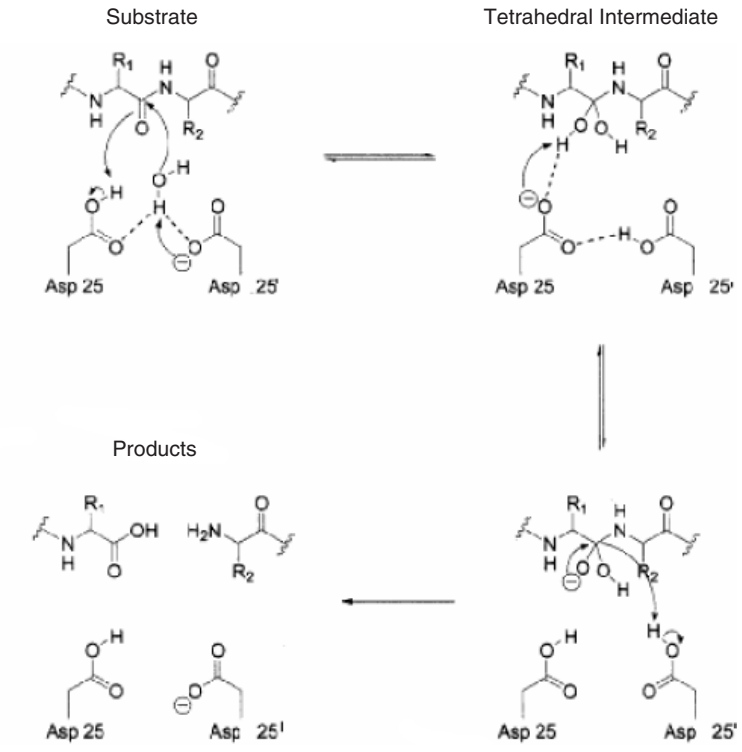
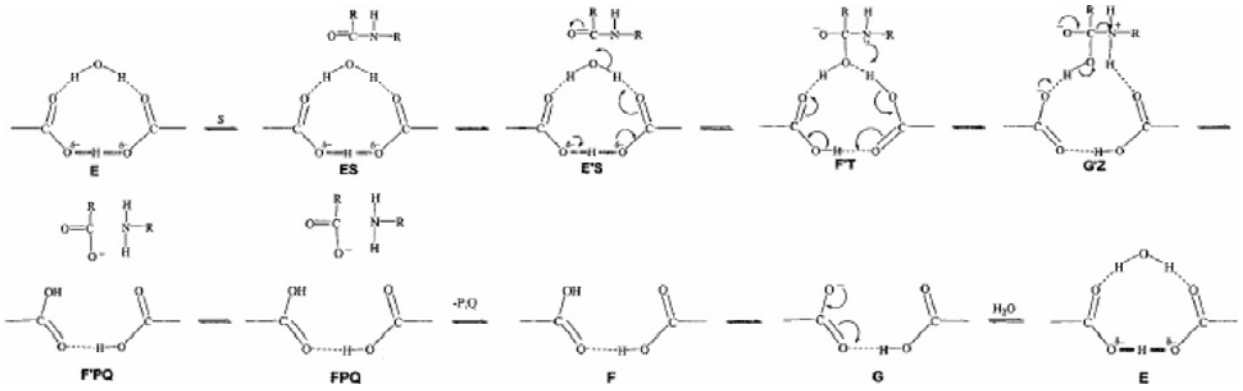


Figure 5: Two step kinetic iso-mechanism proposed by Northrop for peptide bond cleavage by HIV PR. Note the changing character of the hydrogen bond between inner oxygen atoms of catalytic aspartates.



that both aspartic acid residues are required for activity. Because of the vital role played by HIV PR in the viral life cycle, many efforts are being made to develop varieties of inhibitors of this enzyme as drugs against AIDS [20,21]. Since HIV PR is functional only as a homodimer, besides blocking the active site, prevention of dimerisation is pursued as another way of inhibiting this enzyme [22]. However, there has not been significant success so far in this approach. In fact all the drugs currently administered as part of HAART are active site binders, and have been developed using the crystallographic technology of structure-based drug-design.

4. Drug resistance

These drugs have been helping enormously AIDS patients in improving the quality of their life. However, it has been observed that patients very quickly develop resistance to treatment by these drugs, and this effect is traced to specific mutations in the HIV PR. Table 2 gives a list of mutations that have been found to be associated with different administered drugs [23]. The mutations given in bold are the primary or major mutations conferring resistance, while others are secondary or compensatory mutations. While some mutations provide resistance across several drugs, there are others that are specific to particular drugs. Interestingly, these drug-resistant mutations, which drastically reduce (almost to 5%) the proteolysis efficiency, occur both in the active site and non-active site regions of the enzyme. There is an urgent need to generate three dimensional structural data to gain insight into this phenomenon of drug-resistance to enable design of next generation drugs. It is generally believed that drugs that interfere with some step in the mechanism of the enzyme are less prone to be overcome by the enzyme. For this reason, establishing the mechanism has been another field of intense AIDS research [24–31].

5. Reaction mechanism

The mechanisms proposed for peptide bond cleavage by HIV PR are based on: (1) analysis of kinetic and isotope data from in-vitro studies on hydrolyses of oligopeptide substrates, and (2) structural data on protease inhibitor complexes. The mechanisms are all general acid-base mechanisms in which the nucleophile is a water molecule attacking the scissile carbonyl carbon atom (Figure 4).

There are many variants of this basic mechanism, the variations being in the pattern and nature of hydrogen bonding interactions and proton transfers involving the substrate and the catalytic aspartates. An important step in this class of mechanisms is the protonation of the scissile nitrogen atom just prior to peptide-bond breakage. A vast majority of proposals invoke protonation by the outer oxygen atom of a catalytic aspartate. In some others, the protonation is suggested to be by the hydroxyl group of the tetrahedral reaction intermediate. More recently, Northrop has proposed an iso-mechanism which is claimed to resolve for the first time, all uncertainties about pH-profile, transpeptidation activity and anomalous isotope effects in aspartyl proteases. (Figure 5).

In this mechanism, the inner oxygen atoms of catalytic aspartates are hydrogen bonded to one another at all stages of the reaction in sharp contrast to earlier proposals. The character of this hydrogen bond, however, varies, being a low barrier hydrogen bond at some stage and a normal hydrogen bond at some other stages of the reaction. This mechanism also invokes protonation of scissile nitrogen by catalytic aspartates.

We evaluate below these mechanisms in the light of our results on structures of enzyme substrate complexes.

Table 2: Table 2. List of mutations in HIV PR associated with drug resistance in AIDS patients.

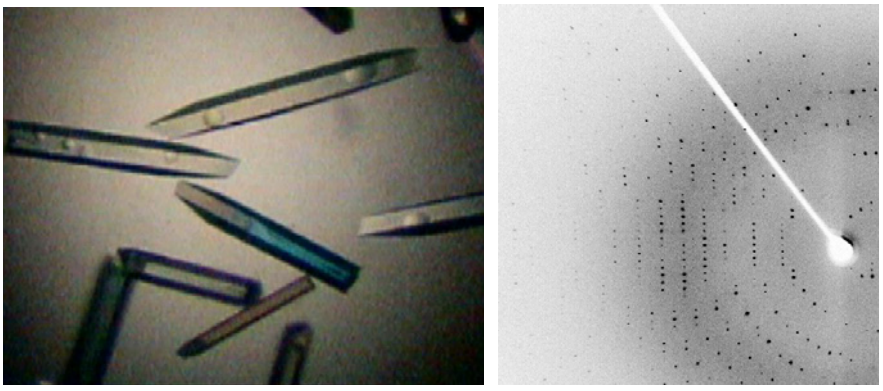
Drug	Mutations (wild type, position, mutant residues)
Atazanavir +/– ritonavir[18]	L10 I/F/V/C; G16 E; K20 R/M/I/T/V; L24 I; V32 I; L33 I/F/V; E34 Q; M36 I/L/V; M46 I/L; G48 V; I**50** L; F53 L/Y; I54 L/V/M/T/A; D60 E; I62 V; I64 L/M/V; A71 V/I/T/L; G73 C/S/T/A; V82 A/T/F/I; I**84** V; I85 V; N**88** S; L90 M; I93 L/M
osamprenavir/ ritonavir	L10 F/I/R/V; V32 I; M46 I/L; I47 V; I**50** V; I54 L/V/M; G73 S; L76 V; V82 A/F/S/T; I**84** V; L90 M
Darunavir/ ritonavir[19]	V11 I; V32 I; L33 F; I47 V; I**50** V; I**54** M/L; G73 S; L**76** V; I**84** V; L89 V
Indinavir/ ritonavir[20]	L10 I/R/V; K20 M/R; L24 I; V32 I; M36 I; M**46** I/L; I54 V; A71 V/T; G73 S/A; L76 V; V77 I; V**82** A/F/T; I**84** V; L90 M
Lopinavir/ ritonavir[21]	L10 F/I/R/V; K20 M/R; L24 I; V**32** I; L33 F; M46 I/L; I**47** V/A; I50 V; F53 L; I54 V/L/A/M/T/S; L63 P; A71 V/T; G73 S; L76 V; V**82** A/F/T/S; I84 V; L90 M
Nelfinavir[20,22]	L10 F/I; D**30** N; M36 I; M46 I/L; A71 V/T; V77 I; V82 A/F/T/S; I84 V; N88 D/S; L**90** M
Saquinavir/ ritonavir	L10 I/R/V; L24 I; G**48** V; I54 V/L; I62 V; A71 V/T; G73 S; V77 I; V82 A/F/T/S; I84 V; L**90** M
Tipranavir/ ritonavir[23]	L10 V; I13 V; K20 M/R; L**33** F; E35 G; M36 I; K43 T; M46 L; I47 V; I54 A/M/V; Q58 E; H69 K; T74 P; V**82** L/T; N83 D; I**84** V; L90 M

6. Experimental procedures and results

Purified tethered HIV-1 PR was prepared as described before [4,32]. Drug-resistant mutations have been introduced using standard methods of genetic engineering. The mutations have been confirmed by sequencing the cDNA. Single crystals (Figure 6A) have been grown at room temperature either through manual crystallization set-ups in hanging drops or through sitting drop setups prepared by the crystallization robot. X-ray diffraction data has been collected by the oscillation method (Figure 6B), both on the in-house RAXIS-IIC image plate diffractometer and on protein crystallography beamlines at synchrotrons. Oscillation film processing, structure refinement and electron density interpretation have been carried out using widely distributed application software suites, DENZO[33], XDS[34], CNS[35], REFMAC[36], PHENIX[37], CCP4[38], and O[39].

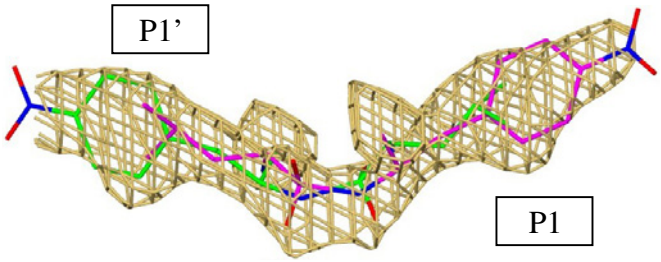
As already stated, mechanism proposals are based on solution studies and also on crystal structures of complexes between HIV PR and inhibitors or substrate analogs. One however needs structural data on complexes between unmodified substrates and active enzymes. The first attempts to provide such data were not totally successful [40]. In these attempts, crystals of the complex were prepared by using the method of co-crystallisation, in the hope that crystal formation preceded substrate cleavage and product release. However, in the crystals obtained, the situation was found to be quite different. The substrate was cleaved before formation of crystals, and the enzyme crystallized as complex with product peptides. HIV-1 PR was

Figure 6: (A) Single crystals of HIV-1 PR. (B) X-ray diffraction photograph for crystals of HIV-1 PR heptapeptide complex recorded using the in-house RAXIS-IIC detector mounted on a RU300 rotating anode X-ray generator fitted with multilayer mirrors.



(A) (B)

Figure 7: OMIT Electron density (yellow) in the active site for the substrate tetrahedral reaction intermediate. Only two residues on either side of the scissile peptide bond are shown. The substrate binds in two 2-fold related orientations (magenta and green).



found to be complexed with only the N-terminal (P) product peptide, while SIV PR was found to be complexed with a C-terminal (Q) product peptide. No structures with both products bound simultaneously were obtained. In contrast, we have attempted to prepare the complex by the soaking method [6], thanks to the closed-flap (Figure 3B) conformation of unliganded HIV-1 protease in the crystals we have obtained [4] under certain conditions. The complete substrate molecules would diffuse to the active site through the solvent channels of the crystal, and the reaction will take place in the active sites of the enzyme molecules packed in the crystal [7–9]. We have varied the soaking conditions to trap the enzyme-substrate complex in different stages of the reaction. HIV-1 PR is one of very few enzymes which cleave next to a proline residue. For purposes of a comparative study we have selected, for soaking, two types of substrate peptides; one containing a proline residue at the P1' position, and the other containing a non-proline residue at P1' position. Depending on the chemical conditions during soaking, the substrate oligopeptide has been trapped either as a tetrahedral intermediate (Figure 7) or as product peptides (Figures 8 and 9).

7. Implications toward mechanism of HIV PR

(a) The nucleophile

The hydrolysis of the peptide bond by aspartyl proteases is through a nucleophilic attack. There has been a debate as to whether the exact nucleophile attacking the scissile carbonyl carbon atom is a protein atom or a water molecule. The observation of transpeptidation activity in HIV-1 PR suggested that the attack was directly by the side chain of a catalytic aspartate. Many theoretical calculations also still leave open the possibility of a direct attack by an aspartate [28,29]. The observation of no covalent attachment between the aspartates and the substrate in forming the tetrahedral intermediate

Figure 8: OMIT Electron density (blue) in the active site for the peptides produced in situ by cleavage. The complete substrate peptide is shown. The substrate binds in a single orientation. The substrate peptide sequence is AETF*YVDGAA.

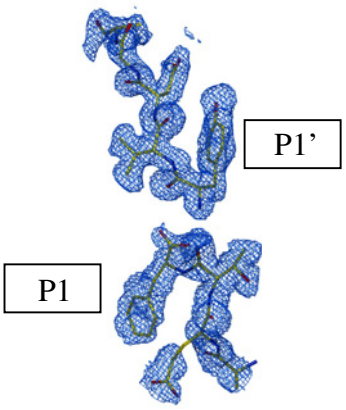
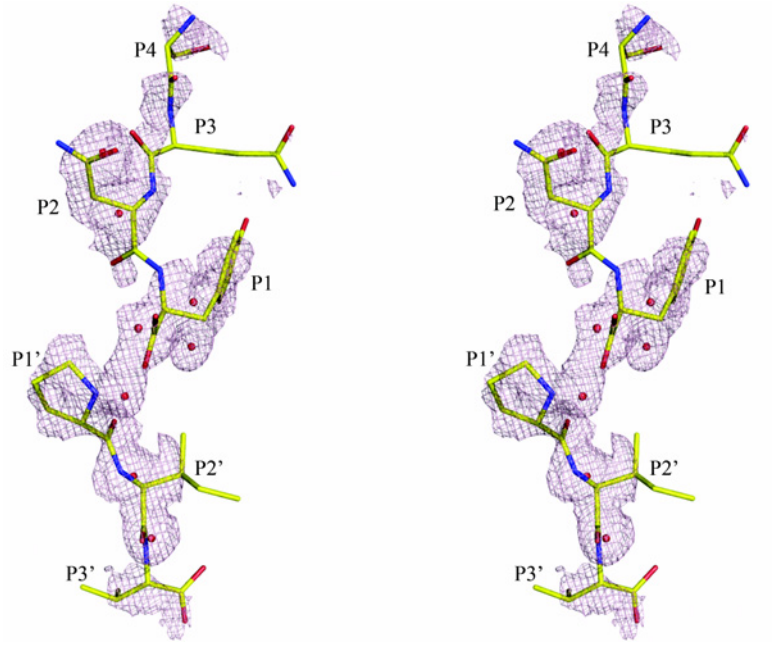
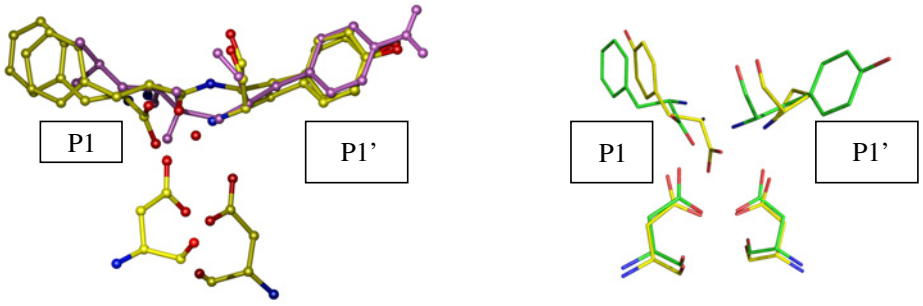


Figure 9: Stereo diagram of the electron density (magenta) in the active site for the peptides produced in situ by cleavage. The complete substrate peptide is shown. The substrate binds in a single orientation and has a occupancy of 0.73. The substrate peptide sequence is SQNY*PIV. Red dots are partially occupied water molecules.



in our studies [7] clearly proves that the aspartates do not directly attack the scissile carbon atom, and that the nucleophile is a water molecule. The tetrahedral intermediate structure also identifies, from among the different water molecules in the active site, the water molecule that performs the attack. In this structure [7], the 'flap water' is present while the water molecule coordinating the aspartates in unliganded protease is missing. It is reasonable to assume then that the water molecule coordinating the aspartates is 'used-up' while performing the nucleophilic attack on the in-coming substrate. This interpretation is also consistent with suggestions based on kinetic and biochemical experiments [24–27]. The tetrahedral intermediate structure [7] thus provides the first direct evidence that the

Figure 10: (A) Superposition of substrate at different stages of the cleavage reaction: binding (yellow), tetrahedral intermediate (magenta) and product after cleavage (atomic colour), (B) Structural superposition of proline-containing heptapeptide complex (yellow) and the non-proline containing decapeptide complex (green). The substrate is cleaved in both structures. The positions of the products in the heptapeptide complex are laterally shifted by about 1.6 Å compared to the decapeptide complex. Only P1 and P1′ residues are drawn.



(A) (B)

nucleophile is the water molecule coordinating the catalytic aspartates.

(b) Atomic rearrangement on the reaction pathway

Figure 10A shows a structural superposition of the 'substrates' in the enzyme active site, at three different stages in the reaction pathway: (1) soon after substrate binding [41], (2) after nucleophilic attack leading to formation of tetrahedral intermediate [7] and (3) just when the products begin to separate [8]. The position and orientation of scissile N atom is very different in stages 1 and 2. While the amide hydrogen atom is pointing away from the catalytic aspartates in stage 1, it is pointing toward aspartates in stage 2. It is clear that, after binding, the substrate undergoes significant conformational change by rotation about the C–N bond of the tetrahedral intermediate. The N atom in the tetrahedral intermediate is in a position to hydrogen bond with catalytic aspartate. It is also interesting that during product separation process (stage 3), only the N-terminal product moves. This may be because the C-terminal product is anchored in place through the hydrogen bond between the scissile N atom and the outer oxygen atom of the catalytic aspartate. The situation appears to be different when the scissile N atom belongs to a proline residue.

(c) Protonation of scissile nitrogen atom

Figure 10B gives a structural superposition of the in-situ products obtained by cleavage of the two types of substrates by HIV-1 PR [8,9]. It may be seen that both products derived from the proline-containing substrate, SQNYPIV, have moved to the right in relation to the aspartates. The lateral shift of the scissile nitrogen atom is about 1.6 Å. This difference in position is to avoid short contacts due to the presence of proline ring next to the cleavage site in the heptapeptide substrate. As a consequence of this difference in positioning, the outer oxygen atom of catalytic aspartate does not form any hydrogen bond with the scissile nitrogen atom. If one makes the reasonable assumption that protonation happens through hydrogen bonds, then in the case of proline containing substrate, scissile nitrogen atom is not protonated by the aspartate. This would then require the enzyme to use different mechanisms to cleave the two types of substrates. Alternatively, since it is found that the hydrogen bond between scissile nitrogen atom and one of the newly generated carboxyl oxygens is common to complexes of both types, this may be the hydrogen bond that acts as a conduit in the protonation of scissile N atom.

(d) Inter-aspartate hydrogen bond

In the peptide bond cleavage reaction, the C–N separation changes from that of a covalent bond at the start of the reaction to that of van der Waal separation at the end, with intermediate values representing intermediate stages of the reaction. Table 3 gives interatomic distances between inner oxygens of catalytic aspartates and the corresponding scissile C–N distances in the three substrate complexes that we have solved. These distances are suggestive of a correlation between strength of the inter-aspartate hydrogen bond and the stage of the cleavage reaction. The C…N separation of 3.5 Å in the heptapeptide complex is normal van der Waal separation and the steric repulsion between the scissile C and N atoms

Table 3: Correlation between inter-aspartate hydrogen bond length and separation between C and N atoms of the scissile peptide bond.

Tethered HIV-1 PR complexed with	Hydrogen bond length (D25OD1 — D1025OD1)	Separation between atoms C and N of the Scissile peptide bond	Reference
Tetrahedral intermediate	3.0 Å	1.65 Å	[7]
Cleaved substrate	2.3 Å	2.7 Å	[8]
Cleaved substrate	2.8 Å	3.5 Å	[9]

is expected to be less. The corresponding inter-aspartate hydrogen bond is relatively long and weak. Similarly, in the tetrahedral reaction intermediate before bond-breakage, where the scissile covalent bond is only stretched and not yet broken, the steric repulsion is again less and the inter-aspartate hydrogen bond is again relatively long and weak (row 1 in Table 3). On the other hand when the C…N separation is significantly shorter than van der Waal separation (row 2 in Table 3), the inter-aspartate hydrogen bond is very short and strong. Since these three structures are closely related and were obtained using almost identical experimental methodology, the correlation mentioned above may be significant. It is therefore suggested that in the reaction mechanism of HIV-1 protease, the inter-aspartate hydrogen bond modulates to make up for inter-product steric repulsion at different stages of the cleavage reaction. Modulation of hydrogen bond strengths during enzyme catalysis is widely prevalent, but these hydrogen bonds are normally between functional groups from the enzyme and the reactants [41]. But what is suggested here is modulation of a hydrogen bond between functional groups of the enzyme alone. This suggestion is different also from the inter-aspartate hydrogen bond modulation envisaged in the two stage kinetic *iso*-mechanism proposed by Northrop [26]. It is however desirable to have higher resolution structures to put this suggestion on a firmer footing.

8. Conclusions

Three dimensional structures of active tethered HIV-1 PR complexed with three separate unmodified 'substrate' oligopeptides, one of which contains a proline residue at P1' position, have been elucidated, for the first time, using X-ray crystallography. In the three structures, the substrates have been processed to different stages of the cleavage reaction. The tetrahedral intermediate structure provides direct evidence for nucleophilic attack by a water molecule. The attacking water molecule is also identified as the one coordinating catalytic aspartates in the free enzyme structure. The product-complex structures give a snap shot of the atomic rearrangements that occur within the active site as the bond cleavage reaction progresses. The hydrogen bonding interactions observed at the catalytic center are not consistent with any of the mechanisms proposed earlier. New possibilities for the mechanism are suggested. The inter-aspartate hydrogen bond is suggested to play an important role in overcoming steric repulsion during the cleavage process.

Acknowledgements

We thank the National Facility for Macromolecular Crystallography, SSPD, BARC, and Dr. Jean-Luc Ferrer, IBS/ESRF, Grenoble, for the X-ray diffraction and biochemistry equipment. The authors acknowledge collaborations with Dr. Smita Mahale, National Institute for Research in Reproduction, Parel, Mumbai, and also with Prof. Sujatha Bhat, Department of Chemistry, IIT-Bombay, Mumbai. We thank Drs. R. Chidambaram, K. K. Kannan, Mukesh Kumar, and Mr. Amit Das and Mr. Subhash Bihani, for scientific discussions, and Mr. S. R. Jadhav for technical help.

Received 03 July 2008; revised 02 September 2008.

References

1. UNAIDS, World Health Organization. 2007. AIDS epidemic update. Available from: http://data.unaids.org/pub/EPISlides/2007/2007_epiupdate_en.pdf
2. Kaiser Family Foundation. 2007. The multi-sectoral impact of the HIV/AIDS epidemic–a primer. Available from: www.kff.org/hivaids/upload/7661.pdf
3. UNICEF. The state of the world's children 2007. Available from: www.unicef.org/publications/index_36587.html
4. Pillai, B., Kannan, K. K. & Hosur, M. V. (2001) *Proteins* **43**: 57–64.
5. Mukesh Kumar and M. V. Hosur**,** (2003) Eur. J. Biochem., 270, 1231–1239.
6. Pillai, B., Bhat, S. V., Kannan, K. K. & Hosur, M. V. (2004) *Acta Crystallogr. D* **60**: 594–596.
7. Kumar, M., Prashar, V., Mahale, S. & Hosur, M. V. (2005) *Biochem. J.* **389**: 365–371
8. Das, Amit, Prashar Vishal, Mahale Smita, Serre Laurence, Ferrer Jean-Luc and Hosur M. V. (2006) *Proc. Natl. Acad. Sci. USA* **103,** 18464–18469.
9. Bihani et al. (2008) PROTEINS: Structure, Function and Bioinformatics, In Press.
10. Vishal Prashar and M. V.Hosur, (2004), Biochem. Biophys. Res. Commun., **323**, 1229–1235.
11. Amit Das, Dharmaraj Rao and M. V. Hosur, (2007) Protein and Peptide Letters, **14**, 565–568.
12. Subhash Bihani, J. L. Ferrer and M. V. Hosur (2008), To be published.
13. John A. G. Briggs, Kay Gru¨newald, Ba¨rbel Glass, Friedrich Fo¨rster, Hans-Georg Kra¨usslich, and Stephen D. Fuller (2006) Structure **14**, 15–20.

14. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M. & Sigal, I. S. (1988) *Proc. Natl. Acad. Sci. USA* **85**: 4686–4690.
15. Swanstrom, R., and Wills, J.W. (1997). Synthesis, assembly, and processing of viral proteins. In Retroviruses, J. M. Coffin, S. H. Hughes, and H. E. Varmus, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 263–334.
16. Heagarty, M. C. (2003) AIDS: The Battle Rages On, Issues in science and technology online: http://www.nap.edu/issues/19.4/heagarty.html#top.
17. Hilary B. Schock, Victor M. Garsky, and Lawrence C. Kuo, (1996), J. Biol. Chem. 31957–31963.
18. Manuel Baca and S. B. H. Kent(1993), Proc. Natl. Acad. Sci. (USA) **90**, 11638–11642.
19. Cheng, Y. S. E., Yin, F. H., Foundling, S., Blomstrom, D. & Kettner, C. A. (1990) *Proc. Natl. Acad. Sci. USA* **87**: 9660–9664.
20. Wolfenden, R. (1999) *Bioorg. Med. Chem.* **7**: 647–652.
21. Wlodawer, A. & Vondrasek, J. (1998) *Annu. Rev. Biophys. Biomol. Struct.* **27**: 249–284; Vondrasek, J. & Wlodawer, A. (2002) *Proteins Struct. Funct. Genet.* **49**: 429–431.
22. Pavlina Rezacova, Julien Lescar, Jiri Brynda, Milan Fabry, Magda Horejsi, Juraj Sedlacek, and Graham A. Bentley(2001). Structure, **9**, 887–895.
23. http://www.iasusa.org/resistance_mutations/index.html
24. Hyland, L. J., Tomaszek Jr., T. A. & Meek, T. D. (1991) *Biochemistry* **30**: 8454–8463.
25. Jaskolski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Hendrickson, R. L., Schneider, J., Kent, S. B. H. & Wlodawer, A. (1991*) Biochemistry* **30**: 1600–1609.
26. Northrop, D.B. (2001) *Acc. Chem. Res.* **34**: 790–797.
27. Dunn, B. M. (2002) *Chem. Rev.* **102**: 4431–4458.
28. Trylska, J., Grochowski, P. & McCammon, J.A. (2004) *Protein Science* **13**: 513–528.
29. Brik A., Wong, C.-H. (2003) *Org. Biomol. Chem.* **1**, 5–14.
30. Kovalevsky, A. V., Chumanevich, A. A., Liu, F., Louis, J. M. and Weber I. T. (2007) *Biochemistry*, **46**, 14854–14864.
31. Rodriguez, E.J., Angeles, T.S., and Meek, T.D. (1993). *Biochemistry* **32**, 12380–12385.
32. Kumar, M., Kannan, K. K., Hosur, M. V., Bhavesh, N. S., Chatterjee, A., Mittal, R. & Hosur, R. V. (2002) *Biochem. Biophys. Res. Commun.* **294**: 395–401.
33. Z. Otwinowski and W. Minor, "Processing of X-ray Diffraction Data Collected in Oscillation Mode", *Methods in Enzymology*, Volume **276**: Macromolecular Crystallography, part A, p.307–326, 1997,C.W. Carter, Jr. & R. M. Sweet, Eds., AcademicPress (New York).
34. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**: 795–800.
35. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et.al.* (1998) *Acta Crystallogr. D* **54:** 905–921.
36. Murshudov, G. N., Vagin, A. A., Dodson, E. J. (1997) *Acta Crystallogr. D* **53**: 240–255.
37. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R.,McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Crystallogr. D* 58, 1948–1954. (http://www.phenix-online.org).
38. COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4, (1994). "The CCP4 Suite: Programs for Protein Crystallography". *Acta Crystallogr. D* 50, 760–763
39. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47:** 110–119.
40. Rose, R. B., Craik, C. S., Douglas, N. L. & Stroud, R. M. (1996) *Biochemistry* **35**: 12933–12944.
41. Jeyabalan, M. P., Nalivaika, E. & Schiffer, C. A. (2002) *Structure (London)* **10**: 369–381.
42. Cleland, W. W., Frey, P. A. & Gerlt, J.A. (1998) *J. Biol. Chem.* **273**: 25529–25532.

M. V. Hosur has obtained his Ph.D. from Department of Physics, IISc., Bangalore. He was awarded Martin Forster gold medal for his Ph.D. thesis. He has done Post-Doctoral research at Purdue University in the laboratory of Prof. Jack Johnson. He has been Visiting Scientist at National Cancer Institute, Frederick, Maryland USA, and Werner-Grenn Fellow, Lund University, Sweden. Presently, He is Head, Protein Crystallography Section in Solid State Physics Division, Bhabha Atomic Research Centre, Mumbai. His research work is in the area of structure-function studies and drug design, primarily using the technique of X-ray Crystallography.

Vishal Prashar completed his M.Sc (Major: Plant Breeding; Minor: Biotechnology) in 1999 from Punjab Agricultural University, Ludhiana. After completing the orientation course in Biology and Radiobiology in 2000 from Bhabha Atomic Research Centre training school, he Joined Solid State Physics Division at Bhabha Atomic Research Centre. Presently, he is working as Scientific Officer (D) in Protein Crystallography Section of Solid State Physics Division at Bhabha Atomic Research Centre, Mumbai.