J|A|C|S ARTICLES

Published on Web 04/16/2010

X-ray Snapshot of HIV-1 Protease in Action: Observation of Tetrahedral Intermediate and Short Ionic Hydrogen Bond SIHB with Catalytic Aspartate

Amit Das,[†] Smita Mahale,[‡] Vishal Prashar,[†] Subhash Bihani,[†] J.-L. Ferrer,[§] and M. V. Hosur*[,†]

Protein Crystallography Section, Solid State Physics Division, Bhabha Atomic Research Centre, Trombay, Mumbai-400085, India, National Institute for Research in Reproductive Health, Parel, Mumbai-400074, India, and LCCP/GSY, Institute de Biologie Structurale, J.-P. Ebel CEA-CNRS-UJF, 41, rue Jules Horowitz, F-38027 Grenoble Cedex 1, France

Received January 1, 2010; E-mail: hosur@barc.gov.in

Abstract: Structural snapshots of each step in the catalytic cycle would help development of inhibitors of human immunodeficiency virus type 1 protease (HIV-1 PR) as effective drugs against HIV/AIDS. We report here one snapshot obtained by determining the structure of enzyme−substrate complex under conditions where the catalytic activity of the enzyme is greatly reduced. The 1.76 Å crystal structure shows the oligopeptide substrate, AETFYVDGAA, converted *in situ* into a *gem*-diol tetrahedral intermediate (TI). The *gem*-diol intermediate is neutral and one of the hydroxyl oxygens forms a very short hydrogen bond (2.2 Å) with the anionic aspartate of the catalytic dyad, which is monoprotonated. Further, there is no hydrogen atom on the outer oxygen of the neutral aspartate. These two observations provide direct evidence that, in the reaction mechanism, hydrogen bonding between catalytic aspartate and scissile carbonyl oxygen facilitates water attack on the scissile carbon atom. Comparison with the structural snapshot of the biproduct complex involving the same substrate reveals the reorganization of the hydrogen bonds at the catalytic center as the enzymatic reaction progresses toward completion. Accumulation of TI in the crystals provides direct evidence that collapse of TI is the rate-limiting step of hydrolysis.

Introduction

Human immunodeficiency virus type 1 protease (HIV-1 PR) is a homodimeric aspartyl protease in which each monomer contributes one catalytic aspartic acid. The enzyme cleaves the viral polyproteins at eight different sites in the maturation process of the virus and is therefore an important target for structure-based drug design.[1−4] Majority crystal structures of enzyme complexes available in the literature are of two types: (1) inactive HIV-1 PR complexed with substrate peptides and (2) active enzyme complexed with noncleavable inhibitors. These complexes provide only approximate information about (1) structures of Michaelis complex and tetrahedral reaction intermediates and (2) protonation states of catalytic aspartates. Further, additional inputs such as the mechanism of catalysis and interactions of the active enzyme with natural peptide substrates rather than with analogues, are required to deal with challenges posed by drug-resistant mutations. Both theoretical and experimental techniques have been used to structurally characterize the substrate/enzyme complexes. A number of computer simulations employing chemical theories of varying degrees of sophistication have revealed[5−9] that the preferred route to substrate hydrolysis is through a noncovalently bound tetrahedral intermediate. Presence of a tetrahedral intermediate has also been suggested by isotope exchange studies. Earlier attempts to prepare crystals of active-enzyme/substrate-oligopeptide complexes by cocrystallization were not successful in trapping the whole peptide.[10] We have successfully prepared crystalline complexes by employing the soaking method, thanks to our discovery of closed-flap conformation of unliganded tethered HIV-1 PR (TD).[11−13] In addition, the soaking method opens up the possibility of exploring if one can structurally map different stages of the cleavage reaction by suitably modifying

[†] Bhabha Atomic Research Centre.
[‡] National Institute for Research in Reproductive Health.
[§] LCCP/GSY, Institute de Biologie Structurale.

(1) Debouck, C.; Gorniak, J. G.; Strickler, J. E.; Meek, T. D.; Metcalf, B. W.; Rosenberg, M. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8903–8906.
(2) Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A.; Scolnick, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686–4690.
(3) Wolfenden, R. *Bioorg. Med. Chem.* **1999**, *7*, 647–652.
(4) Wlodawer, A.; Vondrasek, J. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249–284.
(5) Venturini, A.; Lopez-Ortiz, F.; Alvarez, J. M.; Gonzalez, J. *J. Am. Chem. Soc.* **1998**, *120*, 1110–1111.
(6) Okimoto, N.; Tsukui, T.; Hata, M.; Hoshino, T.; Tsuda, M. *J. Am. Chem. Soc.* **1999**, *121*, 7349–7354.
(7) Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1996**, *118*, 3946–3950.
(8) Bjelic, S.; Aqvist, J. *Biochemistry* **2006**, *45*, 7709–7723.
(9) Park, H.; Suh, J.; Lee, S. *J. Am. Chem. Soc.* **2000**, *122*, 3901–3908.
(10) Rose, R. B.; Craik, C. S.; Douglas, N. L.; Stroud, R. M. *Biochemistry* **1996**, *35*, 12933–12944.
(11) Pillai, B.; Kannan, K. K.; Hosur, M. V. *Proteins* **2001**, *43*, 57–64.
(12) Das, A.; Prashar, V.; Mahale, S.; Serre, L.; Ferrer, J.-L.; Hosur, M. V. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (49), 18464–18469.
(13) Bihani, S.; Das, A.; Prashar, V.; Ferrer, J.-L.; Hosur, M. V. *Proteins: Struct., Funct., Bioinf.* **2009**, *74* (3), 594–602.

 10.1021/ja100002b

either the chemical conditions of the soaking solution and/or the duration of the soak. The catalytic activity of HIV-1 PR is known to be pH dependent, with the activity decreasing drastically below pH 3.0.[14] We have therefore soaked native crystals of TD in substrate peptide solutions of pHs ranging from 2.5 to 9.0, and for periods varying from 6 to 48 h before collecting X-ray diffraction data under cryoconditions. Here we report observation of the substrate decapeptide AETFYVDGAA trapped as a tetrahedral intermediate (TI) in the active site of TD, at pH 2.5. The amino acid sequence of the decapeptide substrate used corresponds to the RT-RH junction of the viral Gag-Pol polyprotein. Thus, the structure represents the first X-ray snapshot of HIV-1 PR in action on a natural substrate. Another interesting feature of the TI/TD complex is the presence of a very short hydrogen bond (2.2 Å) between anionic catalytic aspartate D1025 and the diol-oxygen O1, which was originally the scissile carbonyl oxygen. The results provide structural insights into the catalytic mechanism of hydrolysis by HIV-1 PR.

Methods

Protein Crystallization and Data Collection. Production and crystallization of TD has been reported earlier.[15−18] The 10-residue peptide of amino acid sequence AETFYVDGAA, which acts as a substrate for HIV-1 PR was synthesized at the National Institute for Research in Reproductive Health, Parel, Mumbai, India, using an Applied BioSystem Peptide Synthesizer. The peptide was dissolved in reservoir buffer (0.1 M sodium dihydrogen phosphate and 0.2 M sodium citrate) at pH 2.5 to prepare a 5 mM solution for soaking. Soaking experiment was performed as reported earlier.[17] The soaking time was systematically varied from 6 to 48 h with a view to trap the substrate at various stages of the cleavage reaction. At the end of soaking at room temperature, the crystal was equilibrated for 1 min in the low pH cryoprotectant solution (25% glycerol and 75% reservoir buffer) before flash freezing in a stream of nitrogen gas cooled to liquid nitrogen temperature. X-ray diffraction data were recorded for each one of the soaked crystals using the FIP (ESRF, Grenoble, France) beamline.[19] Electron density in the active site, calculated by difference Fourier syntheses, was monitored to gauge the degree of presence of the soaked substrate in the crystals. Crystals that were soaked for more than 36 h either did not diffract or crumbled to pieces during equilibration with cryoprotectant solution. There was weak difference density in the enzyme active site for crystals soaked for a period of 6−12 h. It was found that a soak-time of 24 h consistently resulted in good difference density for the substrate in the enzyme active site. High-resolution diffraction data set on the chosen crystal was then collected on the ID23-2 beamline at ESRF (Grenoble, France). A total of 120 contiguous diffraction images, each for an oscillation range of 1.0° and exposure time of 0.8 s, were recorded. The oscillation images were indexed, integrated, and scaled to 1.76 Å resolution using the computer program *XDS*.[20] Crystal and intensity data statistics are given in Table 1.

Table 1. Crystal and Intensity Data Statistics[a]

crystal parameters	TD−TI complex
unit cell (Å)	$a = b = 62.56$, $c = 81.86$
wavelength (Å)	0.97945
resolution range (Å)	50−1.76 (1.81−1.76)
space group	$P6_1$
number of reflections measured	128747
number of unique reflections	17226
completeness (%)	95.3 (88.5)
R-merge (%)	7.2 (46.5)
$I/\sigma(I)$	21.76 (2.98)

[a] The numbers in parentheses indicate the value in the highest resolution shell.

Refinement. The structure was refined in the *Crystallography and NMR System* (CNS) using standard simulated annealing (SA) protocols and the amplitude-based maximum likelihood target function.[21,22] A total of 5% of randomly selected reflections were set aside for cross validation.[23] All reflections in the resolution range 50−1.76 Å were included in the refinement. Since the crystals of the present complex are almost isomorphous to those of the unliganded protein structure (PDB entry 1LV1),[24] protein coordinates extracted from the structure 1LV1 were used as the starting model for refinement. In the initial stages of SA refinement and during calculation of SA omit maps, the model was heated to a temperature of 2000 °C, and then annealed at a cooling rate of 25 °C per iteration. Initial anisotropic B-factor and bulk solvent corrections were applied. The relative weighting between geometric and X-ray terms in the target function was determined automatically in *CNS*. When the substrate was being treated as a tetrahedral reaction intermediate, the scissile PHE residue at the P1 position was modified by conversion of the carbonyl carbon atom into a *gem*-diol carbon atom. *CNS* parameter and topology files for this modified PHE residue, designated as PHD, were generated manually. The force constants in the parameter file for bonds and angles involving the *gem*-diol carbon atom were set to comparatively low values so that the actual geometry is not tightly restrained, but is dictated by the diffraction terms. Water molecules were added manually by examining environment around electron densities that were present in both mFo-DFc and 2mFo-DFc maps. Composite omit maps were calculated by leaving out 3% of the amino acid residues at a time. In the final stages of refinement *PHENIX*[25] was used for TLS[26,27] refinement on the protein model which brought the crystallographic residuals down by ∼2%. The entire model building was carried out using the software *O*[28] and the structural

(14) (a) Hyland, L. J.; Tomaszek, T. A.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454–8463. (b) Hyland, L. J.; Tomaszek, T. A.; Roberts, G. D.; Carr, S. A.; Magaard, V. W.; Bryan, H. L.; Fakhoury, S. A.; Moore, M. L.; Minnich, M. D.; Culp, J. S.; DesJarlias, R. L.; Meek, T. D. *Biochemistry* **1991**, *30*, 8441–8453.

(15) Cheng, Y-S. E.; Yin, F. H.; Foundling, S.; Blomstrom, D.; Kettner, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9660–9664.

(16) Kumar, M.; Kannan, K. K.; Hosur, M. V.; Bhavesh, N. S.; Chatterjee, A.; Mittal, R.; Hosur, R. V. *Biochem. Biophys. Res. Commun.* **2002**, *294*, 395–401.

(17) Pillai, B.; Bhat, S. V.; Kannan, K. K.; Hosur, M. V. *Acta Crystallogr.* **2004**, *D60*, 594–596.

(18) Kumar, M.; Prashar, V.; Mahale, S.; Hosur, M. V. *Biochem. J.* **2005**, *389*, 365–371.

(19) Roth, M.; Carpentier, P.; Kaikati, O.; Joly, J.; Charrault, P.; Pirocchi, M.; Kahn, R.; Fanchon, E.; Jacquamet, L.; Borel, F.; Bertoni, A.; Israel-Gouy, P.; Ferrer, J.-L. *Acta Crystallogr.* **2002**, *D58*, 805–814.

(20) Kabsch, W. *J. Appl. Crystallogr.* **1993**, *26*, 795–800.

(21) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gross, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr.* **1998**, *D54*, 905–921.

(22) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gross, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Pannu, N. S.; Read , R. J.; Rice, L. M.; Simonson, T. *International Tables for Crystallography*; Rossmann, M. G.; Arnold, E., Eds.; Crystallography of Biological Macromolecules, Vol. F; Kluwer Academic Publishers: Dordrecht, the Netherlands, 2001; pp 710−716.

(23) Brünger, A. T. *Nature* **1992**, *355*, 472–474.

(24) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(25) 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. *Acta Crystallogr.* **2002**, *D58*, 1948–1954.

(26) Howlin, B.; Butler, S. A.; Moss, D. S.; Harris, G. W.; Driessen, H. P. C. *J. Appl. Crystallogr.* **1993**, *26*, 622–624.

(27) Schomaker, V.; Trueblood, K. N. *Acta Crystallogr.* **1968**, *B24*, 63.

(28) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr.* **1991**, *A47*, 110–119.

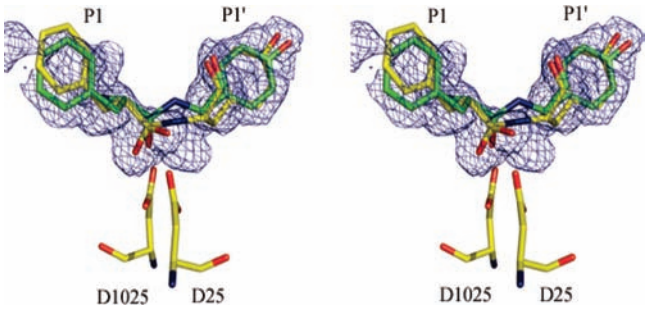

Figure 1. Stereoview of SA omit map contoured at 2.0σ, when the substrate model is not included in map calculation. The refined tetrahedral intermediate, TI (yellow), and the regular peptide (green) models are shown for comparison. Note that the scissile nitrogen atom of the regular peptide model is not inside the electron density. Only the major orientation is shown for the sake of clarity.

superpositions and figures were made using *Pymol*.[29] When comparing different complex structures, only protein atoms were used in the calculation of the superposition matrices.

Results

The Substrate in the Active Site. Connected positive maxima in the 2mFo-DFc map calculated in the enzyme active-site region indicated presence of the substrate peptide AETFYVDGAA. Since the enzyme crystallized is an active enzyme, in principle three possibilities exist for the scissile peptide bond of the soaked substrate: (1) normal *trans*-peptide bond, (2) hydrated peptide bond in which the scissile carbon is converted into a tetrahedral carbon (TI), and (3) cleaved peptide bond in which the scissile carbonyl is converted into a carboxyl group. There was continuous difference electron density in the region of scissile peptide bond, and therefore possibility of a cleaved peptide was ruled out. In view of the pseudo-two-fold-symmetry of HIV-1 PR, the substrate, for both options 1 and 2 above, was modeled in two two-fold related orientations, each with an occupancy of 0.5. Figure 1 is the stereo picture of SA omit map for the P1−P1′ region overlaid with refined P1 and P1′ residues of regular peptide and TI models. It is seen from Figure 1 that there is no density along the C−N bond linking P1 and P1′ residues in the regular peptide model. Further, continuous density near the scissile peptide atoms is unaccounted for in the regular peptide model. On the other hand, the C−N bond in the TI model is well inside the omit density. Real-space correlation coefficients calculated using CNS over the substrate residues, modeled either as a TI or as a regular peptide, are listed in Table 2. The correlation coefficient averaged over either all the six residues or over only the two residues, P1 and P1′, is higher for the TI model. The molecular motif refined further was therefore a complex between the TD and the TI in two orientations. Occupancy refinement of the substrate converged to values of 0.51 and 0.32 for the two orientations. The spherical shape of isolated positive difference density calculated after including the substrate suggested the presence of water molecules rather than large buffer ions. Water molecules have been placed in the remaining 17% of active sites where the enzyme is in apo form. Table 3 gives details of the refinement statistics. In all molecular superpositions only the major orientation is shown in stick representation.

(29) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: San Carlos, CA, 2002; http://www.pymol.org.
(30) Kovalevsky, A. Y.; Chumanevich, A. A.; Liu, F.; Louis, J. M.; Weber, I. T. *Biochemistry* **2001**, *46*, 14854–14864.

Table 2. Real-Space Correlation Coefficient (CC) for Residues in the Two Models of the Substrate

residue no.	tetrahedral intermediate CC	regular peptide CC
2 (P3)	0.680	0.694
3 (P2)	0.783	0.789
4 (P1)	0.842	0.752
5 (P1′)	0.831	0.824
6 (P2′)	0.766	0.760
7 (P3′)	0.854	0.849
4 and 5	0.837	0.788
all residues	0.793	0.778

Table 3. Refinement Statistics

Ligand Model Refined: Tetrahedral Intermediate	
resolution range (Å)	50−1.76
R_{work} (%)	22.28
R_{free} (%)	25.00
number of protein atoms	1515
number of solvent atoms	194
number of ligand atoms	56
RMS deviation of bond lengths (Å)	0.012
RMS deviation of bond angles (deg)	1.99
Average *B* Factor (Å^2)	
protein atoms	31.5
substrate atoms	56.2
water molecules	52.5

Figure 2 is a stereo overlay of two SA omit maps, one calculated by omitting the two oxygen atoms bound to the scissile carbon, and the other calculated by omitting the four Oδ atoms of the two catalytic aspartates. These maps unambiguously confirm the refined positions of the respective atoms. Thus, the species present in the active-site cavity is a TI, consistent with the omit electron density maps and the crystallographic refinement data. Further, the interactions deduced from these coordinates are reliable.

Figure 3 shows molecular fit in the SA omit map of the major occupancy TI model for residues P2−P2′. The electron density beyond P2 and P2′ is weak, suggesting disorder of these substrate residues. The interatomic distance between the scissile N and C atoms is 1.44 Å, while the bond angles around the scissile carbon atom are as follows: Cα−C−*N* = 122.8°, O2−C−O1 = 113.2°, O2−C−*N* = 99.7°, O1−C−*N* = 94.0°, Cα−C−O2 = 113.3°, and Cα−C−O1 = 112.1°. The tetrahedron about the scissile carbon is thus slightly distorted as has

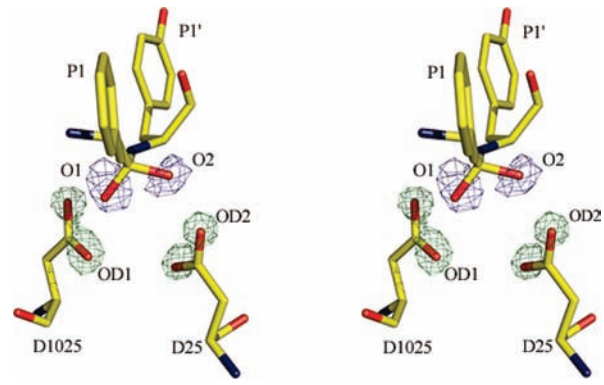

Figure 2. Stereoview of SA omit map (blue) contoured at 2.4σ when the two oxygen atoms in the major-occupancy-TI are not included in the calculation of structure factors. Also shown is the stereoview of SA omit map (green) contoured at 5.2σ when the four oxygen atoms of the catalytic aspartates are omitted from structure factor calculations.

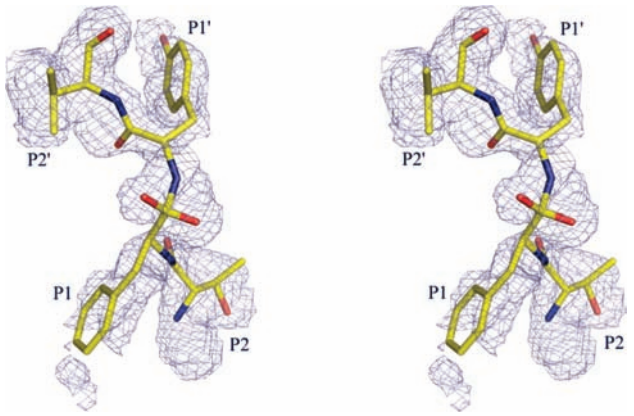

Figure 3. Stereoview of SA omit map contoured at 2.0σ along with the refined TI model shown in the major orientation.

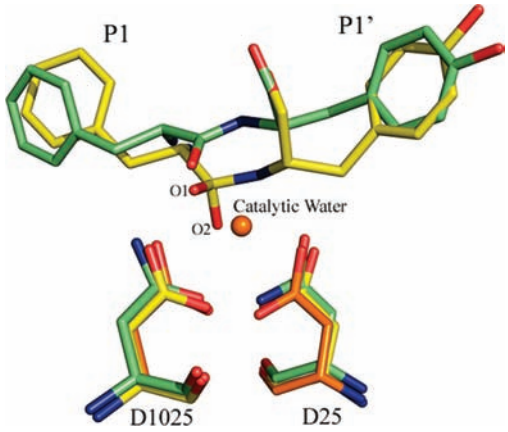

Figure 4. Comparison of catalytic center from three superposed structures: (1) inactive HIV-1 PR/regular peptide complex [PDB ID: 1KJG] (light green), (2) present structure (yellow), and (3) unliganded HIV-1 PR [PDB ID: 2G69] (orange). Note the proximity of O2 to the putative catalytic water in unliganded structure, and of O1 to the peptidyl oxygen in the inactive complex.

been observed in other TI structures.[30,31] The torsion angle around the C−N bond is 106° compared to the value of 180° in the normal *trans*-peptide.

To explore which of the two oxygens O1 and O2 in the TI is derived from the attacking water molecule, we have compared the present structure with the two structures recently published by Torbeev et al.[31] These two structures are complexes of a noncleavable keto-methylene inhibitor with active and inactive HIV-1 PRs. While in the complex with active HIV-1 PR the inhibitor is converted into a TI, in the inactive complex the inhibitor is bound unchanged. The additional oxygen atom bound to scissile carbon in the active complex is inferred to be the water-derived hydroxyl oxygen in the TI. This hydroxyl oxygen superposes onto the O2 oxygen atom of the present structure to within 0.7 Å. Further, O2 is also closer to the putative catalytic water molecule hydrogen bonding to active-site aspartates in unliganded HIV-1 PR (Figure 4). The O1 atom is in close proximity to the position of the scissile carbonyl oxygen (0.7 Å) when the present structure is superposed on the structure [PDB ID: 1KJG][32] of the complex between inactive HIV-1 PR and the oligopeptide corresponding to the RT-RH

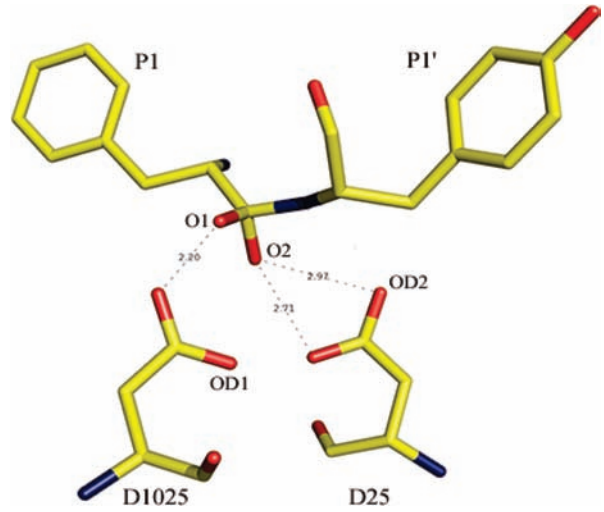

Figure 5. Hydrogen bonds between aspartates and *gem*-diol of the tetrahedral intermediate at the catalytic center. This hydrogen-bonding pattern has not been observed before.

cleavage site (Figure 4). The O1 atom in the present structure is therefore likely to be derived from the scissile carbonyl oxygen atom of the substrate. Such identification is crucial from the point of view of understanding the molecular mechanism of HIV-1 PR.

Hydrogen Bonds in the Active Site. Figures 5 and 6 show the hydrogen-bonding interactions observed in the active-site cavity of the present structure. The TI is involved in nine hydrogen bonds shorter than 3.1 Å, to protein atoms, either directly or through water molecules (Figure 6). The TI interacts with the carboxyl oxygens of catalytic aspartates in an asymmetric way: O2 forms two hydrogen bonds with both oxygen atoms of D25, while O1 forms only a single hydrogen bond with the outer oxygen atom of D1025. The hydrogen bond with D1025 is very short (2.2 Å). The scissile nitrogen atom is at a distance of (3.2 Å) from OD2 atom of D25. The inner oxygen atoms of the catalytic aspartates are 3.1 Å apart (Figure 5).

Figure 7 is a comparison of the active-site region of the present structure with the biproduct complex[12] showing the atomic rearrangement that accompanies collapse of TI into the cleavage products. There is a change in the conformation of P1′ residue as evidenced by a substantially different position for the Cβ atom in the TI. There is a new hydrogen bond between the carboxyl group of P3′ ASP and G1048 carbonyl oxygen atom (d = 2.74 Å) in the low pH structure. This hydrogen bond requires protonation of the the P3′ ASP residue in the substrate (Figure 6). Shifts in the positions of enzyme residues have been very small, of the order of 0.2−0.5 Å (Figure 7). The most significant change has been in the side chains of catalytic aspartates. The virtual dihedral angle OD2 (D1025)−OD1 (D1025)−OD1(D25)−OD2(D25), which is a measure of the coplanarity of the two carboxyl groups of catalytic aspartates, has changed from 67° in the TI/TD complex to the more coplanar value of 23° in the product complex. These conformational changes have led to alterations in hydrogen-bond interactions from catalytic aspartates. For example, in the TI/TD complex, the inner oxygen atoms do not form any hydrogen bond, whereas they form a low-barrier hydrogen bond (LBHB) in the product complex. Similarly, both *gem*-

(31) Torbeev, V. Y.; Mandal, K.; Terechko, V. A.; Kent, S. B. H. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4554–4557.

(32) Jeyabalan, M. P.; Nalivaika, E.; Schiffer, C. A. *Structure* **2002**, *10*, 369–381.

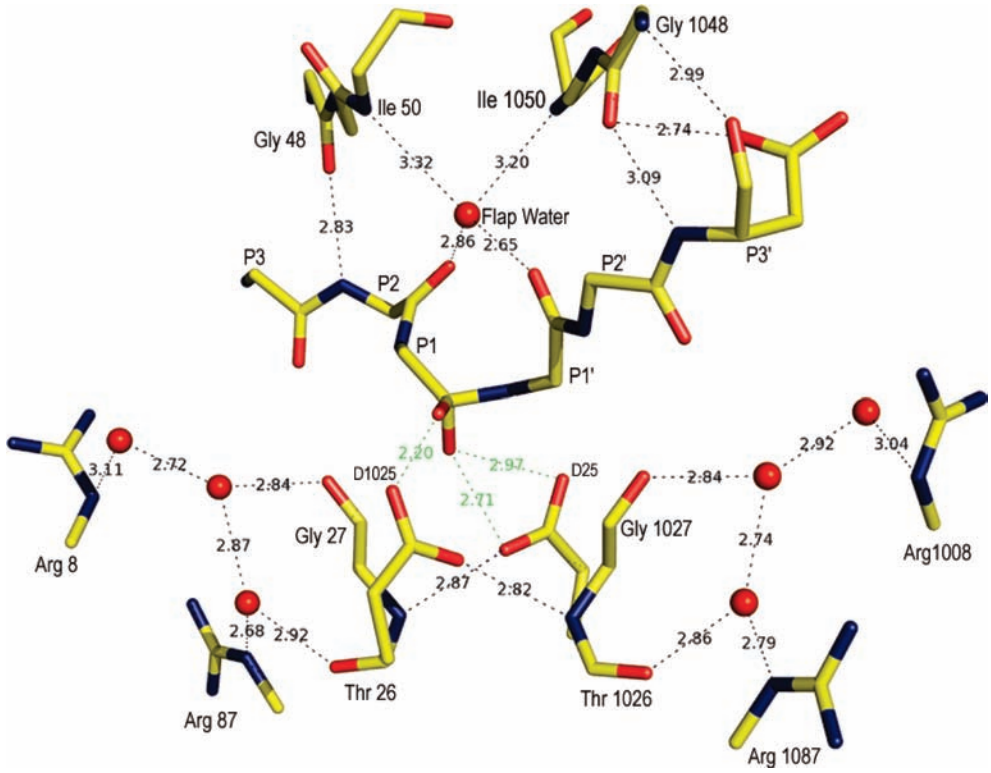

Figure 6. Hydrogen-bonding network in the active-site cavity. The interactions of the *gem*-diol hydroxyls with the catalytic aspartates are shown in green. Only main-chain atoms are shown for clarity, except for side-chain atoms of P3′, D25/1025, R8/1008, and R87/1087.

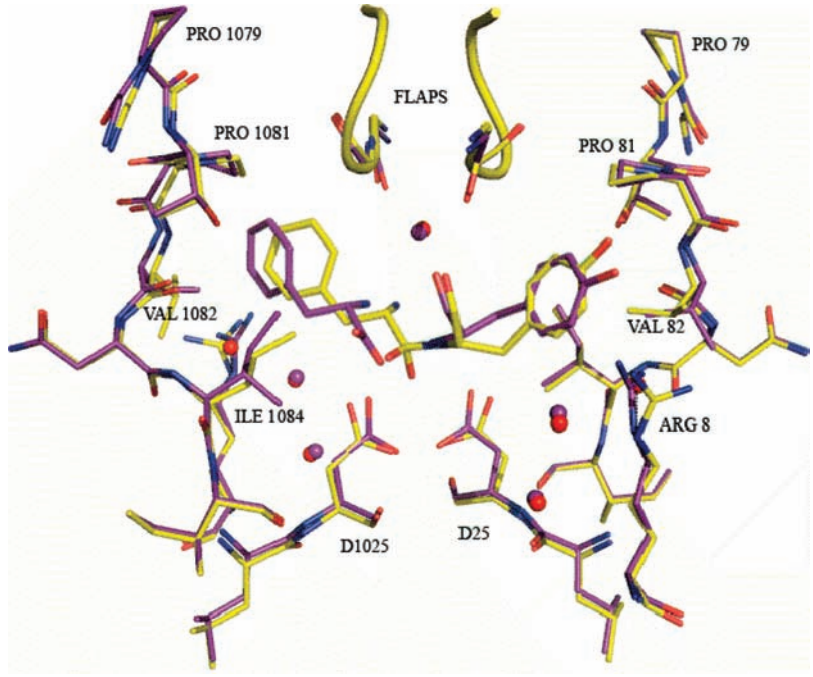

Figure 7. Superposition of the present structure (yellow) with that of the product complex [PDB ID: 2NPH] (purple).

diol oxygen atoms form hydrogen bonds to catalytic aspartates in the TI/TD complex, whereas one carboxyl oxygen atom in the product complex does not make any hydrogen bonds with catalytic aspartates. It is also clear that in going from the TI stage to the product stage, it is the P1 residue that shows variation in its position, while the P1′ residue has not altered its position (Figure 7), despite there being, in the TI/TD complex, no hydrogen bond to anchor in place the scissile N atom of P1′.

Discussion

The Nature of TI. The catalytic mechanism of HIV-1 PR is proposed to consist of two parts: (1) reversible formation of a TI and (2) collapse of the TI to products on protonation of

the scissile N atom.[14,33−36] In some proposals, the two processes happen simultaneously, resulting in a single-step cleavage that leaves no scope for observation of a TI.[34] The TI can contain either a neutral, zwitterionic, or anionic *gem*-diol, and occurrence of all these types has been postulated in different mechanistic proposals.[35,36] The TI has not been directly observed and characterized in experiments carried out in solution, using fast substrates. On the basis of theoretical calculations of energetics of peptide bond hydrolysis, it has even been suggested that the lifetime of TIs are too short to be detected experimentally.[37] However, some success in crystallographic observation of reaction intermediates has been realized by using millisecond Laue crystallography.[38,39] Our strategy of low pH and varying soaking times has allowed us to accumulate the reaction intermediate in the crystals and to analyze the intermediate by conventional protein crystallographic methods. The bond lengths and angles around the tetrahedral *gem*-diol carbon are very comparable to standard values, indicating no significant strain in the intermediate. The atom O1 forms a very definite short hydrogen bond of length 2.2 Å with D1025 OD2. This hydrogen bond establishes O1 to be protonated and D1025 OD2 to be negatively charged, because a common feature among short OHO hydrogen bonds between atoms of dissimilar pK_a values is localization of negative charge on the atom having the lower pK_a.[40−43] Since the pK_a of alcoholic O1 is higher than that of carboxyl oxygen, the negative charge must be located on D1025 OD2, and the hydrogen bond O1···D1025 OD2 is thus an asymmetric short ionic hydrogen bond (SIHB). To decide on the protonation state of O2 we analyze the linearity of hydrogen bonds formed, when the hydrogen atom is placed in a standard hybrid orbital of the donor atom. The orbital hybridization of the donor atom is determined by the geometry of its binding to other non-hydrogen atoms, whose positions are known accurately. Such an approach has been shown to give accurate predictions for positions of polar hydrogen atoms in protein structures.[44,45] The relevant angles for the hydrogen bond O2−(D25 OD2) (Figure 6) are (Cγ D25)−(D25 OD2)−O2 = 89° and C−O2−(D25 OD2) = 117°. Since the latter is closer to the tetrahedral value and the former is too acute, a hydrogen atom on O2 results in a more linear O2−(D25 OD2) hydrogen bond leading to greater stability. This conclusion is also consistent with the fact that O2 is the oxygen derived from the attacking water molecule. Thus, in the TI formed just after nucleophilic attack by a water molecule, the *gem*-diol is neutral. This inference is in agreement with recent QM/MM calculations on the nature of the reaction intermediate in HIV-1 PR catalysis.[46] The neutral *gem*-diol is also consistent with the fact that, in the present structure as well as in any reported structure of HIV-1 PR, there is no oxyanion-binding pocket.

Protonation State of the Aspartates. The identification of the protonation states of the catalytic aspartates in HIV-1 PR is of great interest both in understanding the reaction mechanism and in guiding the design of drugs against HIV/AIDS. It has been realized that the protonation state can be different at different steps in the enzyme's catalytic cycle. It has also been recognized that the nature of the ligand bound in the active site is an important determinant of the protonation state of the aspartates. For this reason, protonation states derived from enzyme complexes with true substrates rather than with nonscissile substrate isosteres are likely to be much more relevant. It is well-known that the effective pK_a values of titratable groups in proteins are influenced by the microenvironment, electrostatic and hydrogen bonding interactions and solvent accessibility.[47] In a computational study on HIV-1 PR substrate oligopeptide complex, the pK_a value of one of the aspartates is estimated to be 2.3, while that of the other was much higher, when all solvent molecules and detailed atomic charges are explicitly included in the calculations.[48] Since the pH used in the present study is higher than 2.3, one aspartate is likely to be deprotonated in the present structure. Based on pH rate profile studies for four oligopeptide substrates and two competitive inhibitors Hyland et al. conclude that the substrates bind only to a form of HIV-1 PR in which one of the two catalytic aspartyl residues is protonated.[14] Since there are strong hydrogen bonds between TI and catalytic aspartates in the present complex, we believe that the aspartic dyad here is in a monoprotonated form. A similar correlation between monoprotonation of an aspartic dyad and its strong hydrogen-bonding interactions with inhibitor molecules has been reported.[49]

Support for monoprotonated dyad at pH 2.5 is also provided by experimental studies involving ^{13}C NMR spectroscopy on HIV-1 PR complexed with inhibitor molecules Pepstatin A[50] and KNI-272.[51] The authors of these studies conclude that, in the complex throughout the range pH 2.5−6.2, one aspartate is protonated and the second is deprotonated. Since, as explained before, D1025 is ionized, D25 has to be protonated to give a monoprotonated aspartic dyad. To infer which of the two oxygens of D25 is protonated, we again follow the procedure described above. There are two hydrogen bonds from D25: O2−(D25 OD2) and O2−(D25 OD1). In the first hydrogen bond O2 is the donor atom as inferred above, and therefore

(33) Northrop, D. B. *Acc. Chem. Res.* **2001**, *34*, 790–797.
(34) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrickson, R. L.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Biochemistry* **1991**, *30*, 1600–1609.
(35) Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003**, *1*, 5–14.
(36) Dunn, B. M. *Chem. Rev.* **2002**, *102*, 4431–4458.
(37) Jacques, F. *Eur. J. Biochem.* **1983**, *135*, 339–341.
(38) Hajdu, J.; Machin, P. A.; Campbell, J. W.; Greenhough, T. J.; Clifton, I. J.; Zurek, S.; Gover, S.; Johnson, L. N.; Elder, M. *Nature* **1987**, *329*, 178–181.
(39) Singer, P. T.; Smalas, A.; Carty, R. P.; Mangel, W. F.; Sweet, R. M. *Science* **1993**, *259* (5095), 669–673.
(40) Katz, B. A.; Spencer, J. R.; Elrod, K.; Luong, C.; Mackman, R. L.; Rice, M.; Sprengeler, P. A.; Allen, D.; Janc, J. *J. Am. Chem. Soc.* **2002**, *124*, 11657–11668.
(41) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, 1997.
(42) Perrin, C. L.; Nielson, J. B. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511–544.
(43) Yamaguchi, S.; Kamikubo, H.; Kurihara, K.; Kuroki, R.; Niimura, N.; Shimizu, N.; Yamazaki, Y.; Kataoka, M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 440–444.
(44) Li, Y.; Roy, A.; Zhang, Y. *PLoS ONE* **2009**, *4*, e6701; 10.1371/journal.pone.0006701.
(45) Hooft, R. W. W.; Sander, C.; Vriend, G. *Proteins: Struct., Funct., Genet.* **1996**, *26*, 363–376.
(46) Carnevale, V.; Raugei, S.; Piana, S.; Carloni, P. *Comput. Phys. Commun.* **2008**, *179*, 120–123.
(47) Harris, T. K.; Turner, G. J. IUBMB. *Life* **2002**, *53*, 85−98.
(48) Trylska, J.; Bała, P.; Geller, M.; Grochowski, P. *Biophys. J.* **2002**, *83*, 794–807.
(49) Piana, S.; Sebastiani, D.; Carloni, P.; Parinello, M. *J. Am. Chem. Soc.* **2001**, *123*, 8730–8737.
(50) Smith, R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. H. *Nat. Struct. Biol.* **1996**, *3* (11), 946–950.
(51) Wang, Y. X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry* **1996**, *35*, 9945–9950.

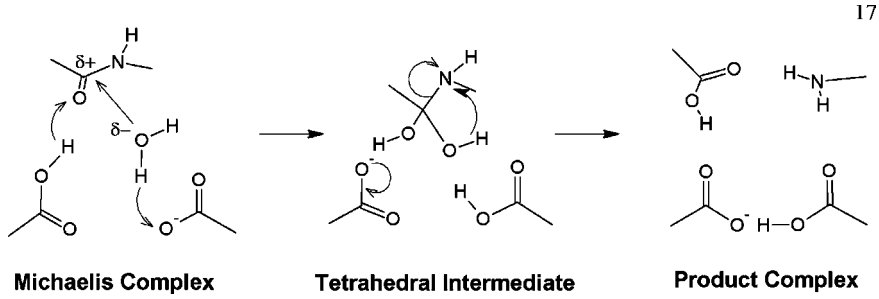

Figure 8. Schematic diagram of the reaction mechanism based on the present and earlier structure [PDB ID: 2NPH] involving the same oligopeptide substrate.

(D25 OD2) is an acceptor. In the second hydrogen bond, we examine the angles (Cγ D25)−(D25 OD1)−O2 = 101° and C−O2−(D25 OD1) = 145°. The former angle is closer to the average value of 112.2° for the C−O−H angle in hydrogen-bonded carboxyl groups determined by neutron diffraction.[52] So placing a proton on (D25 OD1) would lead to a more linear OD1−O2 hydrogen bond. This deduction is also consistent with the fact that O2 is already a donor in the hydrogen bond with (D25 OD2). Therefore we conclude that in the monoprotonated catalytic dyad the proton is bonded to (D25 OD1). Thus, there is no proton on the outer oxygen atom of D25 at this stage of the reaction, and this fact has important implications toward the reaction mechanism. The importance of knowing the protonation state in the context of drug design has been underscored by the recent work of Yu et al.,[53] who have developed a novel method of refining X-ray structures wherein, in the energy function used during refinement, the atoms in the active site are modeled by quantum mechanics while other atoms are represented by molecular mechanics. This treatment, besides fitting atomic structures into experimental electron density, enables very accurate energy rankings of structures having different protonation states of the aspartates. Their calculations imply that, if an incorrect protonation state were to be used, substantial error would be introduced into predicting the binding affinities of different lead compounds.

Reaction Mechanism. The catalytic mechanism of HIV-1 PR involves a TI flanked on either side by two transition states, TS1 and TS2. The activation barrier for TS2 is calculated to be higher than that for TS1.[54] Formation of TI is preceded by two reaction steps: (1) polarization of the scissile carbon atom and (2) activation of nucleophilic water. There is no consensus on the enzyme−substrate interactions that represent these steps. While some invoke, for the polarization of the scissile carbon, a hydrogen bond from scissile N atom to aspartate,[7,34] others invoke a hydrogen bond from the scissile carbonyl oxygen with aspartate.[14] The SIHB between O1 and (D1025 OD2) and the absence of any hydrogen on (D25 OD2) suggest that hydrogen-bonding to the scissile carbonyl oxygen is one of the key interactions in the Michaelis complex prior to TI formation. This hydrogen bond from D1025 polarizes the carbonyl group, while in the same Michaelis complex, a hydrogen bond to the inner oxygen of anionic D25 activates the nucleophilic water molecule. Nucleophilic attack on the scissile carbonyl carbon atom leads to formation of TI, concomitantly with conversion of the hydrogen bond between catalytic aspartate and the scissile carbonyl into an SIHB. Such conversions of normal hydrogen bonds into SIHBs or LBHBs can be important components of enzyme catalysis.[42,55]

The resulting TI is a neutral *gem*-diol, which hydrogen bonds with catalytic aspartates in a unique manner (Figure 5) not observed before. The success of our strategy in trapping the TI also implies directly that the rate-limiting step in peptide bond hydrolysis by HIV-1 PR is the collapse of TI. The collapse of TI is triggered by protonation of the scissile N atom. Given the protonation state of aspartates derived here, the proton for transfer to the scissile N atom has to be provided by the *gem*-diol hydroxyl (Figure 8). It is conceivable that this proton transfer may be facilitated through conformational changes in the enzyme and TI, as the next step in the reaction pathway, which would steer the N atom lone pair to receive the proton.[56] This TS2 activation barrier may be higher at lower pH. The TI collapses into the product complex with the *gem*-diol converting into a planar carboxyl group. The intermolecular SIHB from *gem*-diol is disrupted, and in the product complex a new intra-enzyme LBHB is observed between the inner oxygen atoms of D25 and D1025 (Figure 8).[12]

Conclusion

The structure reported here is a snapshot of the peptide bond cleavage reaction catalyzed by the enzyme HIV-1 PR. The natural cleavage-site oligopeptide substrate has been trapped *in situ* as a *gem*-diol TI by preparing the enzyme substrate complex at a pH value of 2.5. At this stage of the reaction, the aspartic dyad is monoprotonated, with the proton residing on the inner oxygen atom of D25. Binding of the *gem*-diol to catalytic aspartates is asymmetric, with O1 binding to anionic D1025 through one hydrogen bond and O2 binding to neutral D25 through two hydrogen bonds. The hydrogen bond to D1025 is a SIHB, and this observation indicates that hydrogen bonding between D1025 OD2 and scissile carbonyl oxygen atom is an important interaction in the reaction mechanism. The inner oxygens of the catalytic aspartates are too far apart to form any inter-aspartate hydrogen bond. The structure shows that the catalytic aspartates cannot protonate the scissile N atom in the TI. Comparison with an X-ray snapshot of the biproduct complex with the same oligopeptide substrate reveals the exact reorganization of hydrogen bonds at the catalytic center to satisfy different energy requirements

(52) Ichikawa, M. *Acta Crystallogr.* **1979**, *B35*, 1300–1301.
(53) Yu, N.; Hayik, S. A.; Wang, B.; Liao, N.; Reynolds, C. H., Jr.; Merz, K. M. *J. Chem. Theory Comput.* **2006**, *2* (4), 1057–1069.
(54) Piana, S.; Bucher, D.; Carloni, P.; Rothlisberger, U. *J. Phys. Chem.* **2004**, *108*, 11139–11149.
(55) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem.* **1998**, *273*, 25529–25532.
(56) Mesecar, A. D.; Stoddard, B. L.; Koshland, D. E., Jr. *Science* **1997**, *277*, 202–206.

during collapse of TI into products. The inferred protonation scheme and interactions at the active site provide inputs to *in silico* compound screening and also for design of more effective and tighter binding inhibitors.

The refined atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB 3MIM).

Acknowledgment. This work was carried out under an MOU signed between BARC and CEA for cooperation in the field of Life Sciences. We thank the National Facility for Macromolecular Crystallography, SSPD, BARC, for the X-ray diffraction and biochemistry equipment. M.V.H. thanks Drs. R. Chidambaram, K. K. Kannan, and Mukesh Kumar for scientific discussions and S. R. Jadhav for technical help.

JA100002B