Crystal Structures of Highly Constrained Substrate and Hydrolysis Products Bound to HIV-1 Protease. Implications for the Catalytic Mechanism†

Joel D. A. Tyndall,*,‡,§ Leonard K. Pattenden,§,|| Robert C. Reid,§ Shu-Hong Hu,§ Dianne Alewood,§ Paul F. Alewood,§ Terry Walsh,|| David P. Fairlie,§ and Jennifer L. Martin§

National School of Pharmacy, University of Otago, P.O. Box 913, Dunedin 9054, New Zealand, Centre for Drug Design and Development, Institute for Molecular Bioscience, University of Queensland, Brisbane QLD 4072, Australia, and Centre for Molecular Biotechnology, Queensland University of Technology, Brisbane QLD 4001, Australia

Received November 22, 2007; Revised Manuscript Received January 17, 2008

ABSTRACT: HIV-1 protease is a key target in treating HIV infection and AIDS, with 10 inhibitors used clinically. Here we used an unusual hexapeptide substrate, containing two macrocyclic tripeptides constrained to mimic a β strand conformation, linked by a scissile peptide bond, to probe the structural mechanism of proteolysis. The substrate has been cocrystallized with catalytically active synthetic HIV-1 protease and an inactive isosteric (D25N) mutant, and three-dimensional structures were determined (1.60 Å). The structure of the inactive HIVPR(D25N)/substrate complex shows an intact substrate molecule in a single orientation that perfectly mimics the binding of conventional peptide ligands of HIVPR. The structure of the active HIVPR/product complex shows two monocyclic hydrolysis products trapped in the active site, revealing two molecules of the N-terminal monocyclic product bound adjacent to one another, one molecule occupying the nonprime site, as expected, and the other monocycle binding in the prime site in the reverse orientation. The results suggest that both hydrolysis products are released from the active site upon cleavage and then rebind to the enzyme. These structures reveal that N-terminal binding of ligands is preferred, that the C-terminal site is more flexible, and that HIVPR can recognize substrate shape rather than just sequence alone. The product complex reveals three carboxylic acids in an almost planar orientation, indicating an unusual hexagonal homodromic complex between three carboxylic acids. The data presented herein regarding orientation of catalytic aspartates support the cleavage mechanism proposed by Northrop. The results imply strategies for design of inhibitors targeting the N-terminal side of the cleavage site or taking advantage of the flexibility in the protease domain that accommodates substrate/inhibitor segments C-terminal to the cleavage site.

Human immunodeficiency virus type 1 protease (HIVPR)[1] is an essential enzyme for viral replication and a major target for therapeutic intervention (*1*–*3*). HIVPR cleaves the precursor viral polyproteins $Pr55^{gag}$ and $Pr160^{gag\text{-}pol}$ to generate both structural proteins such as the matrix, capsid, and nucleocapsid, as well as the viral enzymes reverse transcriptase, RNase H, integrase, and HIVPR itself. When the protease is inactivated, either through mutation of its catalytic aspartic acid residue or through occupancy of its active site by tightly bound inhibitors, viral maturation to an infectious form is halted. So far, ten inhibitors of HIVPR have progressed through to the clinic for the treatment of HIV infections and for preventing/delaying the development of acquired immunodeficiency syndrome (AIDS) (*2*). The HIVPR enzyme is unquestionably one of the most spectacularly successful examples to date for which the principles of structure-based drug design have been validated (*3*–*5*). Although HIVPR inhibitors are routinely used in combination with reverse transcriptase inhibitors, the rapid development of drug resistance continues to be a major problem that compromises complete therapeutic outcomes (*6, 7*). The present study began with the idea that it might be more difficult for HIV to develop resistance to inhibitors that are very close analogues of substrates, preorganized in a preferred conformation for protease binding but flexible enough to accommodate subtle changes within the substrate binding cleft. As a prelude to development of such inhibitors, we examined here the binding of a highly constrained substrate that has an unusually high affinity for HIVPR, with an expectation of gaining structural insights into substrate binding, product dissociation, and perhaps the catalytic mechanism of substrate hydrolysis.

† This work was supported by the Australian National Health and Medical Research Council and the Australian Research Council. J.L.M. is an Australian NH&MRC Senior Research Fellow. D.P.F. is an ARC Professorial and Federation Fellow.

* To whom correspondence should be addressed at the University of Otago. E-mail: joel.tyndall@otago.ac.nz. Phone: +64 3 479 7293. Fax: +64 3 479 7034.

‡ University of Otago.

§ University of Queensland.

|| Queensland University of Technology.

[1] Abbreviations: Aba, L-α-amino-*n*-butyric acid; Abz, 2-aminobenzoic acid; Abz-NF*-6, 2-Abz-Thr-Ile-Nle-Phe(pNO_2)-Gln-Arg-NH_2; HIV, human immunodeficiency virus; HIVPR, chemically synthesized protease from HIV-1 (SF2 isolate) with Cys67 and Cys95 replaced with Aba ([$Aba^{67,95}$]HIVPR); HIVKI, HIVPR with two mutations (Gln7Lys and Leu33Ile) to limit autoproteolysis ([Lys^{7},Ile^{33}, $Aba^{67,95}$]HIVPR); HIVKI(D25N), as for HIVKI with the additional mutation of Asp25Asn; Nle, norleucine.

10.1021/bi7023157 CCC: $40.75
Published on Web 03/01/2008

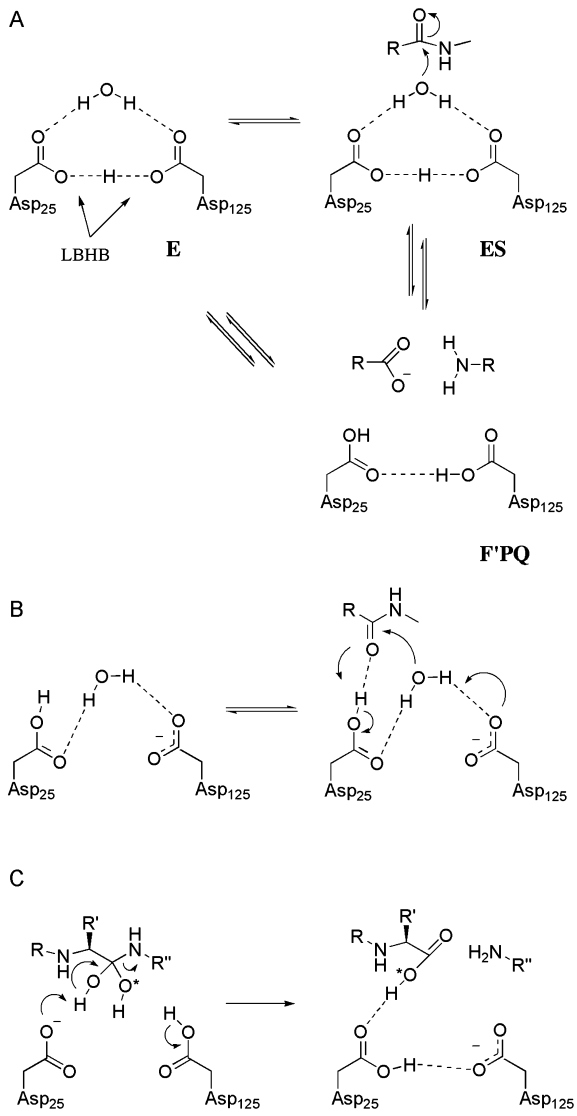

FIGURE 1: Proposed catalytic mechanisms: (A) Starting point (E) of the catalytic mechanism proposed by Northrop (*9*) where a hydrogen is present between the catalytic aspartates in the form of a LBHB and a water of hydrolysis is ideally positioned for nucleophilic attack of the scissile amide bond (ES). Following amide bond cleavage, the F′PQ complex is formed prior to release of the products and regeneration of the active enzyme (E). (B) Mechanism proposed by Veerapandium (*11*) and others (*12, 13*) based on studies of a difluoro ketone inhibitor complexed with endothiapepsin (HIV numbering used). (C) Mechanism proposed by Hyland et al. (*14*) where the product complex formed immediately after hydrolysis shows a hydrogen bond between the two catalytic aspartates.

The catalytic mechanism of HIVPR and other aspartic proteases has been intensively studied over the past 2 decades (*8*). Peptide bond hydrolysis involves the net addition of an active site water molecule, but the precise mechanism of nucleophilic attack and proton shuttling has been controversial. One theory involves a proton permanently located between oxygen (OD) atoms of the two catalytic aspartates (Figure 1A) forming a low-barrier hydrogen bond (LBHB; <2.5 Å between atoms sharing the proton) (*9, 10*). Another hypothesis based on structural information (X-ray and neutron diffraction) suggests that there is no hydrogen atom between the two catalytic aspartate oxygen atoms (Figure 1B) (*11–13*). Others have proposed a mechanism that ends with a hydrogen bond between the two oxygen atoms (Figure 1C) (*14–16*). Elucidation of the mechanism of substrate recognition and cleavage has been assisted by the determination of structures for catalytically inactive HIVPR(D25N) bound to peptide substrates corresponding to the cleavage sites capsid−p2, matrix−capsid, p2−nucleocapsid, p1−p6, reverse transcriptase−RNase H, and RNase H−intergrase (*17, 18*) and complexes with captured products such as Ac-SLNF-OH and PIV-NH_2 (*19*).

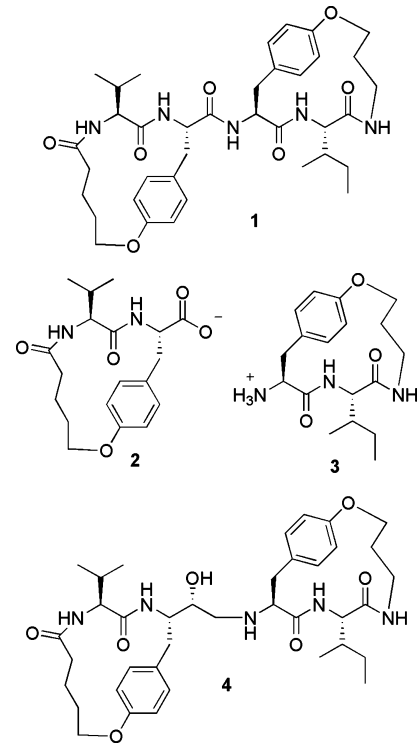

FIGURE 2: Hydrolysis products, **2** (**2a** and **2b**) and **3**, derived from proteolysis of the macrocyclic peptide substrate, **1**. Also shown is the analogous macrocyclic HIVPR inhibitor, **4**.

More recently (*20*), the structure of the *in situ* hydrolysis products AETF-OH and NH_2-YVDGA complexed with active HIVPR, generated after soaking crystals with the substrate AETF*YVGDA, has been determined. This closely represents the F′PQ intermediate (Figure 1A) proposed by Northrop (*9*). This structure implies a LBHB between the catalytic aspartates with a distance of 2.3 Å (*20*). Previous reports of substrate complexes with inactive HIVPR(D25N) enzyme (*17, 18, 21, 22*) as well as macrocyclic substrate analogues with active HIVPR (*23–29*) have implied that the shape of the substrate, rather than a specific sequence alone, can influence HIVPR specificity. In particular, it has been pointed out by Prabu-Jeyabalan and co-workers (*18*) that the volume of substrate residues occupying sites P1−P3 (*30*) forms a toroid-like shape. That shape has also been a feature of constrained macrocyclic peptide inhibitors of HIVPR that have low nanomolar inhibitory potencies against HIV-1 and HIV-2 (*23–25, 27, 31*).

We have developed an unusual bicyclic compound, **1** (Figure 2), consisting of two macrocycles that constrain a hexapeptide backbone into a β strand conformation (*28*), the shape of almost all peptidic ligands that are preferentially recognized by HIVPR and all proteases (*26, 28, 29*). The two monocyclic components of **1** are joined by an amide bond that is the putative cleavage site if **1** were a substrate for HIVPR. In this work compound **1** is a high-affinity

substrate that is cleaved by HIVPR to monocyclic amine **2** and acid **3** products which also have reasonable affinities for HIVPR. The cyclic restraints of **1** essentially preorganize the hexapeptide backbone into a β strand for binding to HIVPR. We anticipated that it was feasible to obtain structures for both this substrate and the specific hydrolysis products complexed with HIVPR as both substrate **1** and putative hydrolysis products **2** and **3** have their backbones locked into the preferred β strand conformation (*28, 32*). We expected that substrate/product binding would be very similar. Thus enzyme residues may not be perturbed by the different induced fits to substrate versus products or by conformational changes to the enzyme that are known to accompany product dissociation from the active site. We therefore expected to gain insights into conformational changes in the enzyme directly related to the catalytic mechanism of hydrolysis and hence learn more about substrate recognition, cleavage, and product dissociation.

In this paper we present a structural study of the binding to and processing of the unusual bismacrocyclic hexapeptide substrate **1** by HIVPR that adopts a highly constrained β strand backbone before and after interacting with enzyme. The crystal structures are reported for substrate **1** bound to catalytically inactive HIVPR(D25N) and to catalytically active HIVPR. Substrate **1** remains intact in the enzyme active site in the former structure but has been hydrolyzed to monocyclic products in the latter structure and reveals two molecules of the proteolytic N-terminal cleavage product **2** bound consecutively within the active site. By comparing these structures with an earlier structure of a related bismacrocyclic transition state analogue, **4**, bound to catalytically active HIVPR (*31*), we are able to reveal the reaction profile for proteolytic hydrolysis of a constrained β strand substrate by HIVPR, enabling comment on the mechanism of enzyme catalysis.

EXPERIMENTAL PROCEDURES

Protein and Substrate Synthesis. Both HIVPR(D25N) and HIVPR were chemically synthesized using solid-phase peptide synthesis (*31, 33, 34*). The bicyclic substrate **1**, monocyclic products **2** and **3**, and bicyclic transition state analogue inhibitor **4** were synthesized as described (*24, 28*).

Competitive Substrate Binding. Synthetic HIVPR was purified from rpHPLC fractions, solubilized in 6 M Gua·HCl (0.05 mg mL^{-1}), and then refolded for 60 min in buffer (20 mM phosphate, pH 7.0, 20% v/v glycerol, 10 mg mL^{-1} BSA). To assess whether bicyclic compound **1** was a substrate for HIVPR, we used a modified assay (*34*) in which the protease/buffer solution was added to **1** and the resulting product mixture was analyzed at different times by rpHPLC in tandem with mass spectrometry, with both putative monocyclic cleavage products **2** and **3** being identified. To obtain some quantitative measure of the binding of **1**–**3** to HIVPR, they have been assessed as competitive inhibitors of the cleavage of a known fluorogenic substrate Abz-NF-6* (*35*). At $t = 0$, the buffer/protease solution (10 μL) was added to the known fluorogenic substrate Abz-NF-6* (50 μM), buffer B (MES, pH 6.5, 37 °C, 100 mM NaCl, 10% v/v glycerol) at 37 °C, and varying concentrations of bicyclic substrate **1** or monocyclic products **2** and **3**, resulting in K_i = 0.8, 8, and 45 μM, respectively, compared with K_i = 3 nM for the bicyclic inhibitor **4** under the same conditions (*28*).

Protein Crystallization. Lyophilized HIVPR(D25N) and HIVPR (2 mg) [HIVPR is [Aba^{67,95}]HIVKI, where cysteines 67 and 95 were replaced with the isosteric analogue L-α-amino-*n*-butyric acid (Aba) and the mutations Q7K and L33I were made to limit autolysis] (*31*) were refolded by dissolving in 50% acetic acid followed by dilution in 25 volumes of buffer C (100 mM acetate, pH 5.5, 5% v/v ethylene glycol, 10% v/v glycerol). Acetic acid was then removed by buffer exchange into buffer C and the protein concentrated in Centricon YM-10 ultrafiltration units to a final concentration of 5 mg mL^{-1}. Protein was mixed with substrate (which had been dissolved in dimethyl sulfoxide to a concentration of 50 mM) to give a 10-fold molar excess of the substrate and then equilibrated for 16 h at 4 °C. Crystals were grown at 20 °C by hanging drop vapor diffusion. Drops of volume 2–5 μL were prepared by mixing equal volumes of the protein/substrate complex and a reservoir containing 100 mM acetate buffer (pH 5.5) and 35–60% saturated ammonium sulfate as the precipitant (*36*). Crystals appeared within 1 week and grew over 4 weeks to a maximum size of 0.4 × 0.1 × 0.1 mm^3.

X-ray Data Collection. Diffraction data were measured from single crystals using a Rigaku R-AXIS IIC imaging plate area detector with Cu Kα radiation (1.5418 Å) generated from a Rigaku RU-200 copper target rotating anode (46 kV, 60 mA) with a 0.00015 in. nickel filter. Cryocrystallographic data were measured by transferring the HIVPR crystal into a solution containing 25% glycerol and 50% ammonium sulfate followed by cryocooling in a nitrogen gas stream (100 K). Diffraction data were integrated and scaled using DENZO and SCALEPACK (*37*) (Table 1).

Structure Determination. The structures were determined by difference Fourier methods, using the structure of the HIVPR complexed with a macrocyclic peptidic inhibitor (PDB code 1CPI) (*23, 31*) as the model with inhibitor and solvent molecules omitted. Structures had 10% of the data set aside for cross-validation using R_{free} (*38*). Model building was performed using O (*39*), and initial models were refined using simulated annealing in X-PLOR 3.851 (8 Å cutoff) (*40, 41*). Poorly ordered residues were positioned using OMIT maps (*42*), where 10–15 residues were omitted at a time. Refinement was completed using positional and individual *B*-factor refinement in CNS v1.1 (*43, 44*). Topology and parameter files for substrate and products were generated using PRODRG (*45*). Water molecules were included where peaks were found in both $F_o - F_c$ (3σ) and $2F_o - F_c$ (1σ) maps and where at least one stereochemically reasonable hydrogen bond could be formed.

Protein residues are listed by monomer chains A and B. The substrate complex had a number of residues modeled with alternate conformations (Arg A14, Gln A18, Lys A45, Ile A47, Lys A55, Val A82, Glu B134, Asn B137, Met B146, Ile B150, Lys B170, Val B182), and portions of several residue side chains have been modeled with half-occupancy (Lys A7, Glu A34, Glu A35, Gln A61). The product complex had the following residues modeled with alternate conformations (Arg A14, Glu A21, Val A82, Ile A84, Glu B121, Glu B134, Met B146, Ile B150, Ile B184). Portions of several residue side chains have been modeled with half-occupancy

Table 1: Statistics for Crystallographic Data Measurement and Structure Refinement

	HIVPR(D25N)/ substrate	HIVPR/ product
Data Collection		
no. of observations	70017	65431
no. of unique reflections	22699	22094
completeness[a] (%)	91.7 (64.1)	90.0 (52.3)
$I/\sigma I$ (top shell)	22.2 (5.9)	20.9 (4.8)
R_{merge}[b] (%)	3.4 (9.3)	3.1 (14.1)
Refinement Statistics		
no. of reflections	22508	21887
R_{factor}[c]	0.179 (0.220)	0.177 (0.272)
R_{free}[d]	0.204 (0.260)	0.205 (0.330)
no. of protein atoms	1553	1534
no. of ligand atoms	49	52
no. of water atoms	206	201
no. of sulfate atoms	4 × 5	3 × 5
rmsd from ideal		
bond lengths (Å)	0.006	0.008
bond angles (deg)	1.41	1.45
dihedral angles (deg)	26.16	16.14
improper angles (deg)	0.96	1.05
Ramachandran statistics		
% in most favored region	94.2	95.5
% in disallowed region	0.0	0.0
coordinate error from Luzatti plot (Å)	0.17–0.19	0.17–0.20
average *B*-factor (Å^2), all atoms	13.63	18.51
average *B*-factor (Å^2), ligand atoms	11.92	23.36

[a] Completeness indicates the number of measured independent reflections divided by the total number of theoretical independent reflections. [b] $R_{merge} = \sum|I_o - I_{av}|/\sum I_{av}$, over all symmetry-related observations. [c] $R = \sum|F_o - F_c|/\sum|F_o|$, over all reflections. [d] R_{free} is calculated as for R_{factor} from 10% of the data excluded from refinement. Values in parentheses are for the top shell of data (1.60−1.61 Å).

(Lys A7, Gln A18, Glu A35, Lys B17, Gln B161, and the C-terminal carboxylate group of **2b**).

Analysis and Alignment. Structures were viewed and analyzed using PyMOL (www.pymol.org). Pairwise superimpositions were carried out using protein Cα atoms, allowing the ligands to be unbiased in the alignment. All superimpositions were conducted with the substrate complex as the reference structure except where substrate comparison was not appropriate. All figures were generated using PyMOL. The quality and geometry of the structures were evaluated by Procheck (*46*). Coordinates and structure factors for HIVPR(D25N) complexed with substrate **1** and HIVPR complexed with products **2a** and **2b** have been deposited with the Protein Data Bank (3BXR and 3BXS, respectively).

RESULTS AND DISCUSSION

*Compound **1** Is a Substrate.* The bicyclic compound **1** is a substrate for HIVPR (K_m = 110 μM), being cleaved into monocyclic amine **2** and acid **3** products that are readily identified by HPLC and mass spectrometry. We had previously tested the bicyclic compound **1** and its linear pseudoanalogue Ac-LVFFIV-NH_2 (K_m = 20 μM) as competitive inhibitors of HIVPR, using a known fluorogenic substrate Abz-NF*-6 (K_m = 26 μM), enabling determination of K_i = 0.8 and 60 μM, respectively (*28*). The higher affinity of the bicyclic substrate **1** over the acyclic hexapeptide analogue Ac-LVFFIV-NH_2 is attributed to the preservation of a protease-recognizing β strand structure in **1** which offers an entropic advantage for binding. The monocyclic cleavage products **2** and **3** are also competitive inhibitors of HIVPR (K_i = 8 and 45 μM, respectively) under the same conditions with the N-terminal product **2** having a 6-fold higher affinity over the C-terminal cycle **3** (*28*). The submicromolar affinity of **1** for HIVPR is to our knowledge the highest affinity of any substrate for this enzyme. The higher affinity of bicyclic hexapeptide **1** over monocyclic tripeptides **2** and **3** can be accounted for by its larger size and greater potential interactions with the enzyme.

A driving interest was to investigate whether we could crystallize compounds **1**–**3** in the active site of HIVPR and thereby allow structural profiling of the catalytic process of substrate cleavage. Two significant problems with previous attempts to structurally analyze substrate processing by HIVPR are that (i) both substrates and products had weak affinity for the enzyme and (ii) substrate and products had different structures that likely required different induced fits with the surrounding enzyme structure. We hoped that the similar shapes and binding modes of the cyclic substrate and cyclic products together with their distinctive β strand structures would translate into similar binding to the HIVPR, without inducing different ligand binding enzyme conformations.

HIVPR/Substrate Complex. Incubation of the constrained macrocyclic substrate **1** with the inactive protease mutant HIVPR(D25N) under conditions optimized for cocrystallization produced crystals that were isomorphous with other previously determined complexes of HIVPR (*31*). Data sets for both complexes [HIVPR(D25N)/substrate and HIVPR/substrate] were measured to a resolution of 1.6 Å using cryocrystallographic techniques.

The enzyme/substrate complex shows overall structural features that are common to other known HIVPR complexes with inhibitors and substrates (*3*). Electron density within the active site of the enzyme shows a single orientation for the bicyclic substrate **1** (Figure 3A). The macrocyclic substrate is well ordered with an average *B*-factor of 11.9, slightly lower than the overall *B*-factor for the enzyme–substrate structure (Table 1). The substrate interacts with the enzyme via numerous contacts, including 11 conserved hydrogen bonds that are also found in other substrate complexes. Four of these interactions are mediated via three conserved water molecules (waters 301, 302, and 303, Figure 3B). The carbonyl oxygen of the scissile amide group forms a hydrogen bond with the nitrogen of Asn B125. Some flexibility is seen in the active site of the enzyme, represented by multiple conformations of the side chains for residues Ile A47, Ile B150, and Val B182 (S3−S1) and Val A82 (S1′−S3′) (*30*).

The complex of inactivated enzyme HIVPR(D25N) with the substrate **1** compares well with other substrate/HIVPR(D25N) complexes, in particular, that of HIVPR with the peptide substrate corresponding to the reverse transcriptase−RNase H cleavage site (AETF↓YVDGA) which closely resembles the sequence of substrate **1** and gives a protein-based pairwise rmsd of 0.27 Å (Figure 3C; 198 Cα atoms) (*18*). Atoms corresponding to the Phe-Phe dipeptide occupying P1 and P1′ and incorporating the scissile amide bond were also compared, giving an atom-based pairwise rmsd of 0.17 Å (Phe-Phe heavy atom comparison, 22 atoms from each molecule). Similarly, the pairwise rmsd of the protein Cα atoms in the HIVPR(D25N) complex with **1** and in the HIVPR complex of the bicyclic inhibitor **4** was 0.26 Å (Figure 3D). The difference in conformation of the N-

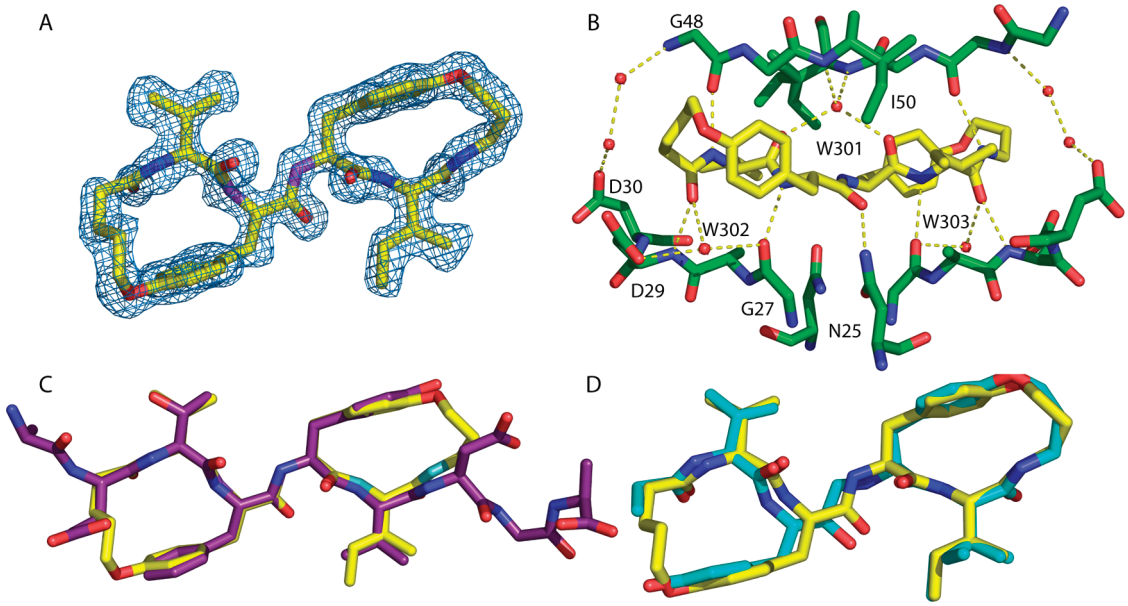

FIGURE 3: (A) Electron density for the substrate, **1**, bound within the active site of HIVPR(D25N). The $2F_o - F_c$ map is contoured at 1.0σ. (B) Hydrogen-bonding interactions between **1** and HIVPR(D25N). (C) Superimposition of bound substrate **1** (yellow carbon atoms) and the substrate from the HIVPR/substrate (AETF↓YVDGA) complex (PDB code 1kjg; purple carbon atoms). (D) Superimposition of bound substrate **1** (yellow carbon atoms) and bound inhibitor **4** (PDB code 1b6p; cyan carbon atoms).

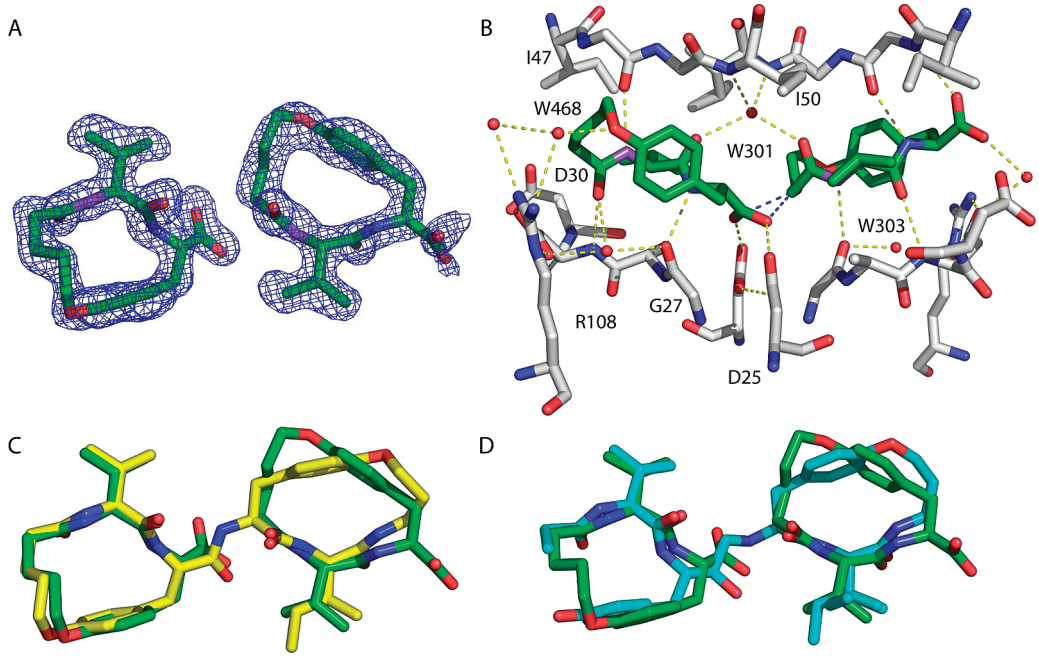

FIGURE 4: (A) Electron density for the products, **2a** (left) and **2b**, bound within the active site of HIVPR. The $2F_o - F_c$ map is contoured at 1.0σ. (B) Yellow dashed lines show hydrogen-bonding interactions between **2a** and **2b** and HIVPR; blue dashed lines illustrate the close contact between the carboxylate oxygens of **2a** and the methylene linker of **2b**. (C) Superimposition of bound substrate **1** (yellow carbon atoms) and bound products (**2a** and **2b**; green carbon atoms). (D) Superimposition of bound products (**2a** and **2b**; green carbon atoms) and bound bicyclic inhibitor **4** (PDB code 1b6p; cyan carbon atoms).

terminal cycle can be attributed to the addition of a carbon atom in the hydroxyethylamine isostere of **4** compared with the macrocyclic peptide substrate **1** as well as the amine−aspartate interaction seen in the complex of **4**.

HIVPR/Product Complex. The HIVPR/product complex was obtained from cocrystallizing the active HIVPR enzyme with the macrocyclic substrate **1**. The enzyme cleaved the substrate during the crystallization experiment, as indicated by electron density in the active site corresponding to the putative product molecules. However, instead of an enzyme complexed with the N-terminal product **2** and C-terminal product **3** (Figure 2), the electron density matched two copies of the N-terminal product **2** bound within the active site: one in the expected nonprime position (termed **2a**) and one in the prime site (P1′−P3′) (termed **2b**, Figure 4) where we had expected to see the C-terminal product **3**. The macrocycle **2b** is located at the P1′ to P3′ sites and can be described as a translation of macrocycle **2a** by approximately 9 Å along the central axis of the active site, followed by a 180° rotation around this axis. This places the tyrosine residue at P3′ and essentially presents an almost continuous peptidic backbone from P3 to P3′ within the entire active site.

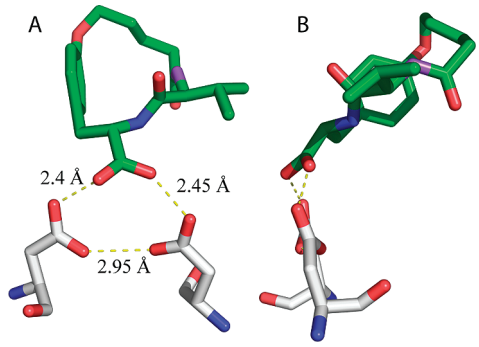

FIGURE 5: (A) Interactions between the catalytic aspartates and the carboxylate of macrocycle **2a**. (B) Side view rotation of 90°.

The electron density for **2a** is well-defined and shows unambiguous evidence that proteolysis has occurred (Figure 4A). The density around the second product, **2b**, clearly illustrates the presence of the N-terminal product molecule **2** and not the expected C-terminal product **3**. Density matching the tyrosinyl moiety of **2b**, in particular, is clearly seen in the S3′ subsite and not in the expected S1′ position if **3** were present. Correspondingly, density in the S1′ subsite matches the four-carbon methylene linker present in **2b**. Density corresponding to valine and not isoleucine can be seen at P2′, confirming the presence of a second moiety of **2**. However, the $2F_o - F_c$ density around the carboxylate is poor, and the three atoms are modeled with 0.5 occupancy.

The two product molecules interact with the enzyme via numerous contacts including 14 hydrogen bonds (Figure 4B). Two key hydrogen bonds between water 301 and the oxygen of the valine residue of **2a** (2.72 Å) and the carbonyl oxygen adjacent to the methylene linker of **2b** (2.63 Å) are present. The products make more hydrogen-bonding interactions (H-bonds) than substrate **1** primarily due to the interactions of the two carboxylate moieties. There is an additional H-bond between water 468 and the tyrosinyl ether oxygen of **2a**, as well as the loss of an H-bond to water 303 from the P2′ carbonyl (compared to the substrate complex). The carboxylate group of **2a** is in close proximity to the α carbon of the methylene linker of macrocycle **2b** (2.84 and 2.69 Å; Figure 4B, blue lines). The free carboxylate of the product **2a** (formed upon substrate cleavage) interacts with both catalytic aspartates via a triangular arrangement of very short hydrogen bonds involving the six oxygen atoms in an almost planar manner (Figure 5). This triangular H-bond interaction provides space for **2b** to be accommodated as well as appearing to be optimal for H-bond interaction with the catalytic aspartates. Some flexibility exists in the active site as indicated by several residues modeled with alternative conformations (Ile B150, S2; Ile B184, S1; Ile A84, S1′; Val A82, S1′-S3′), all of which are sites of resistance mutation which are known to alter inhibitor binding (*7, 47*).

Substrate/Product Comparison. Figure 4C shows the binding conformations of the substrate, **1**, and the two products, **2a** and **2b**. The sites spanning P3−P1 show a high degree of similarity with the greatest deviation occurring at the scissile amide carbon (or carboxylate carbon; 1.2 Å apart). The P1′−P3′ sites show greater deviation between structures. This is clearly due to different orientations of the monocycles occupying these sites. The main chain atoms of P1′ and P2′ are very close, indicating that the enzyme can accommodate a range of amino acids and chemical shapes.

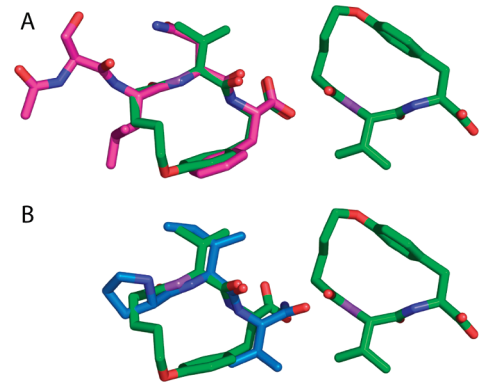

FIGURE 6: Superimposition of product complex (**2a** and **2b**; green carbon atoms) with (A) peptide product Ac-SLNF-COOH (PDB code 1yth) and (B) peptide product PIV-NH_2 (PDB code 1ytg) (*19*).

The residues within the enzyme active site do not differ notably between substrate and product complexes (ignoring disordered residues).

HIVPR/Product Comparison. The presence of duplicate N-terminal hydrolysis products (**2a** and **2b**) in the active site of catalytically active HIVPR, while unexpected, is not without precedent. A substrate (Ac-SLNF↓PIV-amide) previously cocrystallized with active protease also produced a complex between HIVPR and the C-terminal product, with the PIV-amide product positioned in the nonprime, P3−P1 (P) site where one would expect occupation by the N-terminal Ac-SLNF-CO_2H product (*19*). However, the present case is the first example that we know of where the C-terminal product, **3**, is replaced by a second copy of the N-terminal product and where there is also one copy bound in the same position as in the substrate. This occupation of the entire active site of HIVPR may be due to the higher affinity of the enzyme for the N-terminal product, and thus there has been a net exchange of **3** by the second copy of **2**.

The macrocyclic products (**2a** and **2b**) compare well with several other product complexes of HIV and SIV (*19*). Figure 6A illustrates the almost identical binding conformations for the macrocyclic product **2a** and the peptidic product Ac-SLNF-CO_2H. This supports the idea that the N-terminal side has a more stringent requirement for a ligand and that the carboxylate group of **2a** is not influenced by the close proximity of **2b** in our case. Figure 6B shows the C-terminal hydrolysis product PIV-amide positioned in the nonprime side of the active site compared with **2a**. The binding conformation of these two products differs as a result of the hydrogen-bonding interactions between the C-terminal amide (PIV-$CONH_2$) and the catalytic aspartates whereas **2a** forms a more regular triangular interaction (Figure 5) similar to that seen for Ac-SLNF-CO_2H (Figure 6A). It is clear that both N- and C-terminal macrocycles are released from the active site, before recombining with the enzyme since this is the only way that two copies of the N-terminal cycle can coincide in the active site. The observed preference for the recombination product is attributed to the 6-fold higher affinity of **2** over **3**, resulting in only **2** competitively binding following regeneration of the apoenzyme (substrate is at 10-fold excess to enzyme in the crystallization conditions). High concentrations of the substrate could inhibit the enzyme's ability to hydrolyze the peptide via product inhibition.

Bicyclic Inhibitor. The macrocycles **2a** and **2b** bind in a similar fashion when compared to the two monocyclic components of the inhibitor **4** (pairwise rmsd 0.22, 198 Cα atoms; Figures 4D and 2). The most conserved positions are at P2 (Val) and P2′ (Ile, Val). An obvious deviation occurs in the prime side of the active site where the two macrocycles are positioned in opposite orientations (tyrosine of **4** at P1′ and tyrosine of **2b** at P3′). The additional methylene present in the transition state isostere of the inhibitor, as well as the amine−aspartate interaction, forces a rotation of the N-terminal cycle centered around the valine. The residues within the enzyme active site do not differ notably between product and inhibitor complexes (ignoring disordered residues).

The structures presented here, together with the structure for inhibitor **4** in complex with HIVPR, provide an opportunity to consider the conversion from substrate (**1**) via a transition state (similar to **4**) to the product complex (**2a** and **2b**). While not a true transition state isostere, the inhibitor **4** shows only a minor deviation in the position of the N-terminal cycle compared to the substrate **1** (Figure 3D). One may expect the N-terminal cycles (**2a** and **4**) to be more conserved than the C-terminal cycles; however, the charged amino group of **4** interacts with the catalytic aspartates and may influence this movement. The hydroxyl group mimics the position of the carbonyl oxygen prior to nucleophilic attack of the catalytic water (not seen). A superimposition of the substrate complex with the recent unliganded HIVPR structure (PDB code 2g69) shows that the catalytic water is approximately 2.3 Å away from the amide carbonyl carbon and adjacent to the catalytic aspartates (not shown) (*48*). The water is positioned appropriately to nucleophilically attack the scissile amide bond from below the plane of the carbonyl (allowing for a small rotation of the substrate amide). This water molecule is effectively replaced by a transition state isostere in other HIVPR inhibitors, but not in the case of **4** due to the stereochemistry around the tetrahedral carbon. The water molecule is also close (1.0 Å) to the newly formed carboxylate of **2a** when superimposed and would clearly indicate this water as the attacking nucleophile. However, no structure to date has been determined with a substrate analogue as well as the catalytic water so these findings can only be inferred. Overall, both the substrate and the product complexes show their N-terminal cycle to be closer to the catalytic aspartates than that of the inhibitor. However, this shift in the inhibitor ring is well accommodated due to the more flexible bridge between macrocycles.

Catalytic Aspartate Interactions. Figure 5 illustrates the almost planar hydrogen bond network between the catalytic aspartate residues and the newly formed carboxylate group of **2a**. The H-bonds made between the ligand carboxylate and the catalytic aspartates are very short (2.40 and 2.45 Å) and by inference would be termed LBHB (*10*). However, the distance between the two oxygens of the catalytic aspartates is long by comparison (2.95 Å) and would not appear to be classified as a typical LBHB. This triangular formation is similar (but with shorter distances) to that found in the complex with Ac-SLNF-CO_2H (Figure 6A). Without the aid of ultra-high-resolution structural data, positioning of hydrogens can only be inferred based on distance, and there is conjecture as to whether a proton exists between the catalytic aspartates. James et al. have suggested that such

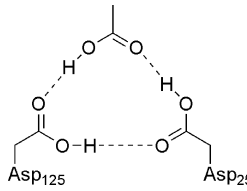

FIGURE 7: Proposed homodromic hydrogen-bonding network between **2a**, Asp25, and Asp125.

configurations could exist without a shared proton given that oxygen atoms of the catalytic aspartates have a van der Waals radius of 1.4 Å, and this is a reasonable nonbonded contact distance (*12*). The distance here (2.95 Å) is longer than in the recent *in situ* product complex (OD1−OD1 distance 2.30 Å), but similar to a (high resolution 1.03 Å) phenylnorstatine inhibitor complex where a hydrogen atom is seen approximately midway between the catalytic aspartates (*49*). If we assume that the product complex contains a proton between the aspartate OD1 atoms, the system would be homodromic and would require the unusual situation of all three carboxylates to be protonated (Figure 7). This state is very similar to the unconventional intermolecular hydrogen-bonding pattern seen for the phenylnorstatine inhibitor (*49*). Mechanistically, this deviates from the model proposed by Northrop et al. where the carboxylate of the transition state F′PQ is unprotonated (Figure 1A). However, it cannot be assumed that **2a** remains in the active site upon amide bond hydrolysis, so we cannot compare this directly as the F′PQ transition state also contains the newly formed amide. If a proton is present between the catalytic aspartate OD1 atoms, this would support the isomechanism proposed by Northrop. An *ab initio* study based on the pepstatin A/HIV-1 complex also suggested that both aspartates are protonated when in complex with the inhibitor with ligand binding stabilizing the diprotonated dyad (*50*), which suggests that a homodromic conformation with **2a** is favorable.

Conclusion. High-resolution crystal structures have been presented for (i) the complex of the catalytically inactive protease mutant HIVPR(D25N) bound to the constrained bismacrocyclic peptide substrate **1**, an analogue of the hexapeptide LVY↓YIL, and (ii) the complex of the catalytically active enzyme HIVPR bound to the same substrate **1** but showing two cleavage products within the active site, albeit two copies of monocyclic product **2** rather than one molecule each of monocycles **2** and **3**.

The complexes between HIVPR and either the constrained bicyclic substrate **1** or products **2a** and **2b** are similar to other HIVPR complexes. However, they show that HIVPR clearly recognizes the shape of the substrate rather than the amino acid sequence alone. This is reinforced by the unexpected orientation of product **2b**, where the tyrosinyl moiety is at position S3′ and not the anticipated S1′. Although the substrate and active HIVPR were coincubated at low temperature (4 °C, 16 h), the crystals were actually grown at 20 °C over 1–4 weeks, so it is no surprise that the substrate is cleaved, and our results suggest that both cleavage products are released and subsequently rebound. The resulting HIVPR/product complex indicates that monocycle **2** binds more tightly than **3**, which is consistent with the inhibition constants. Comparison of the structures of the substrate, bicyclic inhibitor, and product complexes reveals some important similarities and differences that may help to explain

the mechanism of catalysis. A striking feature is the almost identical location of the N-terminal macrocycles in the substrate and product complexes (Figure 4C), but there is a shift of the N-terminal cycle in the inhibitor complex largely due to the long amide bond replacement (hydroxyethylamine transition state isostere) between cycles (Figures 3D and 4D).

This study used an unusual bicyclic compound **1** (*28*), consisting of two macrocycles that constrain a hexapeptide backbone into a β strand conformation known to be preferentially recognized in peptidic ligands by HIVPR and all proteases (*26, 28, 29*). The cyclic restraints in **1** essentially preorganize the peptide backbone for protease binding but leave sufficient flexibility for the compound to "breathe" in response to subtle changes that might be induced in the active site of the protease through mutation. The structures presented here suggest that there is more space for the inhibitor to "breathe" on the C-terminal side, and so it may be possible to take advantage of the flexibility in the protease domain that accommodates substrate/inhibitor segments C-terminal to the putative cleavage site in inhibitor design. By contrast, the N-terminal side has more specific requirements for ligand binding, with the N-terminal side remaining in the same position suggesting that it effectively organizes the surrounding enzyme in a similar conformation that is not dependent upon what is on the C-terminal side. This is supported by our structural work on other inhibitors with an N-terminal cycle but with variable C-terminal appendages (*31*). It may therefore be possible to design inhibitors based on these cyclic structural templates that can closely mimic substrates of HIVPR and yet offer sufficient flexibility to maintain a bound conformation even under pressure from enzyme mutation.

REFERENCES

1. Roberts, N. A., Martin, J. A., Kinchington, D., Broadhurst, A. V., Craig, J. C., Duncan, I. B., Galpin, S. A., Handa, B. K., Kay, J., Krohn, A., et. al. (1990) Rational design of peptide-based HIV proteinase inhibitors. *Science 248*, 358–361.
2. Abbenante, G., and Fairlie, D. P. (2005) Protease Inhibitors in the Clinic. *Med. Chem. 1*, 1–33.
3. Wlodawer, A., and Vondrasek, J. (1998) Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct. 27*, 249–284.
4. Klebe, G. (2000) Recent developments in structure-based drug design. *J. Mol. Med. 78*, 269–281.
5. Ghosh, A. K., Dawson, Z. L., and Mitsuya, H. (2007) Darunavir, a conceptually new HIV-1 protease inhibitor for the treatment of drug-resistant HIV. *Bioorg. Med. Chem. 15*, 7576–7580.
6. Hertogs, K., Bloor, S., Kemp, S. D., Van den Eynde, C., Alcorn, T. M., Pauwels, R., Van Houtte, M., Staszewski, S., Miller, V., and Larder, B. A. (2000) Phenotypic and genotypic analysis of clinical HIV-1 isolates reveals extensive protease inhibitor cross-resistance: a survey of over 6000 samples. *Aids 14*, 1203–1210.
7. Clavel, F., and Hance, A. J. (2004) HIV drug resistance. *N. Engl. J. Med. 350*, 1023–1035.
8. Dunn, B. M. (2002) Structure and Mechanism of the Pepsin-Like Family of Aspartic Peptidases. *Chem. Rev. 102*, 4431–4458.
9. Northrop, D. B. (2001) Follow the protons: a low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Acc. Chem. Res. 34*, 790–797.
10. Cleland, W. W. (2000) Low-barrier hydrogen bonds and enzymatic catalysis. *Arch. Biochem. Biophys. 382*, 1–5.
11. Veerapandian, B., Cooper, J. B., Sali, A., Blundell, T. L., Rosati, R. L., Dominy, B. W., Damon, D. B., and Hoover, D. J. (1992) Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases. *Protein Sci. 1*, 322–328.
12. James, M. N., Sielecki, A. R., Hayakawa, K., and Gelb, M. H. (1992) Crystallographic analysis of transition state mimics bound to penicillopepsin: difluorostatine- and difluorostatone-containing peptides. *Biochemistry 31*, 3872–3886.
13. Coates, L., Erskine, P. T., Wood, S. P., Myles, D. A., and Cooper, J. B. (2001) A neutron Laue diffraction study of endothiapepsin: implications for the aspartic proteinase mechanism. *Biochemistry 40*, 13149–13157.
14. Hyland, L. J., Tomaszek, T. A., Jr., and Meek, T. D. (1991) Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry 30*, 8454–8463.
15. Jaskolski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Heinrikson, R. L., Schneider, J., Kent, S. B., and Wlodawer, A. (1991) Structure at 2.5-Å resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor. *Biochemistry 30*, 1600–1609.
16. Silva, A. M., Cachau, R. E., Sham, H. L., and Erickson, J. W. (1996) Inhibition and catalytic mechanism of HIV-1 aspartic protease. *J. Mol. Biol. 255*, 321–346.
17. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2000) How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease. *J. Mol. Biol. 301*, 1207–1220.
18. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure (Cambridge) 10*, 369–381.
19. Rose, R. B., Craik, C. S., Douglas, N. L., and Stroud, R. M. (1996) Three-dimensional structures of HIV-1 and SIV protease product complexes. *Biochemistry 35*, 12933–12944.
20. Das, A., Prashar, V., Mahale, S., Serre, L., Ferrer, J. L., and Hosur, M. V. (2006) Crystal structure of HIV-1 protease in situ product complex and observation of a low-barrier hydrogen bond between catalytic aspartates. *Proc. Natl. Acad. Sci. U.S.A. 103*, 18464–18469.
21. Wu, J., Adomat, J. M., Ridky, T. W., Louis, J. M., Leis, J., Harrison, R. W., and Weber, I. T. (1998) Structural basis for specificity of retroviral proteases. *Biochemistry 37*, 4518–4526.
22. Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M., and Schiffer, C. A. (2003) Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy. *J. Virol. 77*, 1306–1315.
23. Abbenante, G., March, D. R., Bergman, D. A., Hunt, P. A., Garnham, B., Dancer, R. J., Martin, J. L., and Fairlie, D. P. (1995) Regioselective Structural and Functional Mimicry of Peptides—Design of Hydrolytically-Stable Cyclic Peptidomimetic Inhibitors of Hiv-1 Protease. *J. Am. Chem. Soc. 117*, 10220–10226.
24. Reid, R. C., March, D. R., Dooley, M. J., Bergman, D. A., Abbenante, G., and Fairlie, D. P. (1996) A novel bicyclic enzyme inhibitor as a consensus peptidomimetic for the receptor-bound conformations of 12 peptidic inhibitors of HIV-1 protease. *J. Am. Chem. Soc. 118*, 8511–8517.
25. March, D. R., Abbenante, G., Bergman, D. A., Brinkworth, R. I., Wickramasinghe, W., Begun, J., Martin, J. L., and Fairlie, D. P. (1996) Substrate-based cyclic peptidomimetics of Phe-Ile-Val that inhibit HIV-1 protease using a novel enzyme-binding mode. *J. Am. Chem. Soc. 118*, 3375–3379.
26. Tyndall, J. D. A., and Fairlie, D. P. (1999) Conformational homogeneity in molecular recognition by proteolytic enzymes. *J. Mol. Recognit. 12*, 363–370.
27. Tyndall, J. D. A., Reid, R. C., Tyssen, D. P., Jardine, D. K., Todd, B., Passmore, M., March, D. R., Pattenden, L. K., Bergman, D. A., Alewood, D., Hu, S. H., Alewood, P. F., Birch, C. J., Martin, J. L., and Fairlie, D. P. (2000) Synthesis, stability, antiviral activity, and protease-bound structures of substrate-mimicking constrained macrocyclic inhibitors of HIV-1 protease. *J. Med. Chem. 43*, 3495–4504.
28. Fairlie, D. P., Tyndall, J. D. A., Reid, R. C., Wong, A. K., Abbenante, G., Scanlon, M. J., March, D. R., Bergman, D. A., Chai, C. L., and Burkett, B. A. (2000) Conformational selection of inhibitors and substrates by proteolytic enzymes: implications for drug design and polypeptide processing. *J. Med. Chem. 43*, 1271–1281.
29. Tyndall, J. D., Nall, T., and Fairlie, D. P. (2005) Proteases universally recognize beta strands in their active sites. *Chem. Rev. 105*, 973–1000.
30. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun. 27*, 157–162.

31. Martin, J. L., Begun, J., Schindeler, A., Wickramasinghe, W. A., Alewood, D., Alewood, P. F., Bergman, D. A., Brinkworth, R. I., Abbenante, G., March, D. R., Reid, R. C., and Fairlie, D. P. (1999) Molecular recognition of macrocyclic peptidomimetic inhibitors by HIV-1 protease. *Biochemistry 38*, 7978–7988.
32. Reid, R. C., Pattenden, L. K., Tyndall, J. D., Martin, J. L., Walsh, T., and Fairlie, D. P. (2004) Countering cooperative effects in protease inhibitors using constrained beta-strand-mimicking templates in focused combinatorial libraries. *J. Med. Chem. 47*, 1641–1651.
33. Schneider, J., and Kent, S. B. (1988) Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. *Cell 54*, 363–368.
34. Bergman, D. A., Alewood, D., Alewood, P. F., Andrews, J. L., Brinkworth, R. I., Englebretsen, D. R., and Kent, S. B. H. (1995) Kinetic Properties of HIV-1 Protease Produced by Total Chemical Synthesis with Cysteine Residues Replaced by Isosteric L-Alpha-Amino-N-Butyric Acid. *Lett. Pept. Sci. 2*, 99–107.
35. Toth, M. V., and Marshall, G. R. (1990) A simple, continuous fluorometric assay for HIV protease. *Int. J. Pept. Protein Res. 36*, 544–550.
36. Hui, J. O., Tomasselli, A. G., Reardon, I. M., Lull, J. M., Brunner, D. P., Tomich, C. S., and Heinrikson, R. L. (1993) Large scale purification and refolding of HIV-1 protease from *Escherichia coli* inclusion bodies. *J. Protein Chem. 12*, 323–327.
37. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
38. Brunger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*, 472–475.
39. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A 47* (Part 2), 110–119.
40. Brunger, A. T., Kuriyan, J., and Karplus, M. (1987) Crystallographic R factor refinement by molecular dynamics. *Science 245*, 458–460.
41. Brunger, A. T., Krukowski, A., and Erickson, J. W. (1990) Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. A 46* (Part 7), 585–593.
42. Bhat, T. N., and Cohen, G. H. (1984) OMITMAP: an electron density map suitable for the examination of errors in a macromolecular model. *J. Appl. Crystallogr. 17*, 244–248.
43. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
44. Brunger, 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., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (2001) Crystallography & NMR System, Yale University.
45. Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 1355–1363.
46. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.
47. Dauber, D. S., Ziermann, R., Parkin, N., Maly, D. J., Mahrus, S., Harris, J. L., Ellman, J. A., Petropoulos, C., and Craik, C. S. (2002) Altered substrate specificity of drug-resistant human immunodeficiency virus type 1 protease. *J. Virol. 76*, 1359–1368.
48. Liu, F., Kovalevsky, A. Y., Louis, J. M., Boross, P. I., Wang, Y. F., Harrison, R. W., and Weber, I. T. (2006) Mechanism of drug resistance revealed by the crystal structure of the unliganded HIV-1 protease with F53L mutation. *J. Mol. Biol. 358*, 1191–1199.
49. Brynda, J., Rezacova, P., Fabry, M., Horejsi, M., Stouracova, R., Sedlacek, J., Soucek, M., Hradilek, M., Lepsik, M., and Konvalinka, J. (2004) A phenylnorstatine inhibitor binding to HIV-1 protease: geometry, protonation, and subsite-pocket interactions analyzed at atomic resolution. *J. Med. Chem. 47*, 2030–1036.
50. Piana, S., Sebastiani, D., Carloni, P., and Parrinello, M. (2001) Ab initio molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site. *J. Am. Chem. Soc. 123*, 8730–8737.

BI7023157