the FEBS Journal

Structural and biochemical characterization of the inhibitor complexes of xenotropic murine leukemia virus-related virus protease

Mi Li[1,2], Alla Gustchina[1], Krisztina Matúz[3], Jozsef Tözsér[3], Sirilak Namwong[4], Nathan E. Goldfarb[5], Ben M. Dunn[5] and Alexander Wlodawer[1]

1 Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD, USA
2 Basic Research Program, SAIC-Frederick, MD, USA
3 Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Hungary
4 Department of Biotechnology, Faculty of Science and Technology, Suan Sunandha Rajabhat University, Bangkok, Thailand
5 Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, USA

Keywords
aspartic protease; enzyme inhibition; inhibitor binding; retrovirus

Correspondence
A. Wlodawer, National Cancer Institute, MCL, Bldg 536, Rm 5, Frederick, MD 21702-1201, USA
Fax: +1 301 846 6322
Tel: +1 301 846 5036
E-mail: wlodawer@nih.gov

(Received 5 August 2011, revised 15 September 2011, accepted 21 September 2011)

doi:10.1111/j.1742-4658.2011.08364.x

Interactions between the protease (PR) encoded by the xenotropic murine leukemia virus-related virus and a number of potential inhibitors have been investigated by biochemical and structural techniques. It was observed that several inhibitors used clinically against HIV PR exhibit nanomolar or even subnanomolar values of K_i, depending on the exact experimental conditions. Both TL-3, a universal inhibitor of retroviral PRs, and some inhibitors originally shown to inhibit plasmepsins were also quite potent, whereas inhibition by pepstatin A was considerably weaker. Crystal structures of the complexes of xenotropic murine leukemia virus-related virus PR with TL-3, amprenavir and pepstatin A were solved at high resolution and compared with the structures of complexes of these inhibitors with other retropepsins. Whereas TL-3 and amprenavir bound in a predictable manner, spanning the substrate-binding site of the enzyme, two molecules of pepstatin A bound simultaneously in an unprecedented manner, leaving the catalytic water molecule in place.

Structured digital abstract

- XMRV PR and XMRV PR bind by x-ray crystallography (View interaction)

Introduction

Several links between infection by a retrovirus and human pathology have been identified to date, the best known being the finding that HIV is responsible for the development of AIDS [1,2]. However, other potential links between retroviral infections and human disease are less well characterized. The presence of xenotropic murine leukemia virus-related virus (XMRV) was recently reported in tissues of patients suffering from two vastly different diseases. XMRV was found in prostate cancer cells [3], and in cells isolated from patients suffering from chronic fatigue syndrome [4,5]. These reports led to considerable controversy over whether the presence of the virus is indeed linked to these two or any other diseases [6–11], with the most recent results indicating that XMRV may have been created by passaging human tumors in mice [12]. Nevertheless, structural studies of XMVR proteins were initiated almost immediately after a possible connection of the virus with disease was first postulated. Almost all proteins encoded in the XMRV

Abbreviations
MMLV, Moloney murine leukemia virus; MMLVGagΔ2, recombinant Moloney murine leukemia virus Gag fragment; PDB, Protein Data Bank; PR, protease; XMRV, xenotropic murine leukemia virus-related virus.

genome have been cloned and expressed [13], and the structures of two of them, protease (PR) and RNase H, have been solved [Protein Data Bank (PDB) IDs: 3NR6 and 3P1G, respectively]. The structure of XMRV PR was solved for the apoenzyme only [14], and although some novel topological features were present, especially at the termini, the environment of the active site was found to be similar to what had previously been seen in other retroviral aspartic PRs (retropepsins). On the other hand, only limited susceptibility of XMRV to PR inhibitors developed against HIV-1 PR has previously been reported [15]. To further elucidate these findings, we commenced studies of the enzymatic activity, inhibition, and structural features of the inhibitor complexes of XMRV PR.

Table 1. Inhibition of XMRV PR by approved anti-HIV drugs (top six compounds) and by other protease inhibitors designed against retroviral PRs and malarial aspartic PRs. ND, not determined.

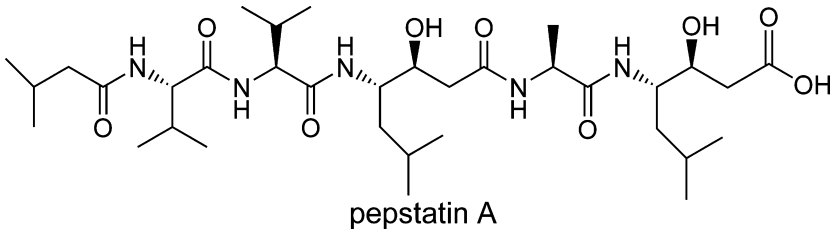
Inhibitor	K_i[a] (nM)	K_i[b] (nM)
Amprenavir	0.2	10 ± 2.0
Atazanavir	1.8	22 ± 1.4
Darunavir	ND	15 ± 0.7
Tipranavir	ND	27 ± 2.1
Lopinavir	ND	32 ± 3.5
Ritonavir	ND	36 ± 4.2
TL-3	102	ND
Pepstatin A	1442	ND
DMP 323	ND	18 ± 1.4
132830[c]	ND	97 ± 8.5
129463[c]	ND	63 ± 2.8
14[d]	ND	26 ± 3.5
2[d]	ND	202 ± 36.1
1[d]	ND	329 ± 21.2
3[d]	ND	333 ± 35

[a] K_i determined by assay with oligopeptide substrate RSLLY↓–PALTP, using an HPLC-based method at 37 °C. [b] K_i determined by assay with chromogenic substrate KARVnL↓NphEAnLG at 25 °C. [c] Compounds 132830 and 129463 were identified by virtual screening of the XMRV PR active site against the library of the Developmental Therapeutics Program of NCI/NIH. [d] Compound numbers from table 1 of [51].

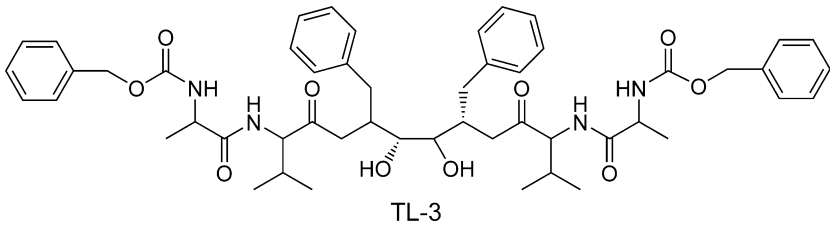
The enzymatic properties of XMRV PR were investigated with substrates specific for the enzyme from the closely related Moloney murine leukemia virus (MMLV), as well as substrates developed for HIV PR. In the absence of inhibitors specifically targeting XMRV PR, we investigated a number of compounds that have been previously shown to be broadly specific against a wide range of retroviral PRs. One such inhibitor is pepstatin A, characterized almost 40 years ago as a generic inhibitor of aspartic PRs [16], and over 20 years ago as an inhibitor of HIV PR [17,18]. Another inhibitor used in our studies was TL-3 [19,20], initially developed to inhibit FIV PR, but subsequently shown to also be a relatively potent inhibitor of other retroviral PRs. As some Food and Drug Administration-approved inhibitors used clinically against HIV show activity against PRs from other retroviruses, interactions of a number of them with XMRV PR were characterized biochemically, and the crystal structure of a complex with amprenavir [21], shown here to be the most potent among them against XMRV PR, was solved. These structural and biochemical data are discussed below.

Results and Discussion

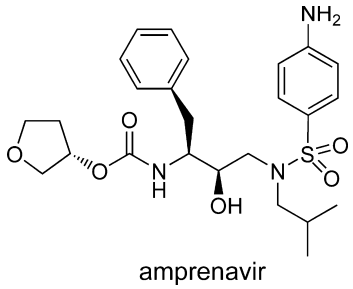
Inhibition of XMRV PR by a number of compounds known to inhibit aspartic PRs was evaluated in this study (Table 1), and crystal structures of the complexes with three of them were solved. The inhibitors that we succeeded in cocrystallizing were pepstatin A



Scheme 1. Structure of pepstatin A.



Scheme 2. Structure of TL-3.



Scheme 3. Structure of amprenavir.

(Scheme 1) [16], TL-3 (Scheme 2) [19], and amprenavir (Scheme 3) [22]. Pepstatin A inhibits a variety of aspartic PRs [23–25], TL-3 has been shown to be active against several retropepsins [26,27], and amprenavir is a potent anti-HIV drug [28]. Crystals of the inhibitor complexes successfully crystallized in this work could only be grown under conditions different from those used to grow the crystals of the apoenzyme [14], and are not isomorphous with them. However, crystals of all three inhibitor complexes are isomorphous in the space group $P2_12_12_1$ and contain an enzyme dimer in the asymmetric unit. The structures of the three complexes were solved and refined at high resolution, 1.4 Å for the complex with TL-3, 1.5 Å for the complex with pepstatin A, and 1.75 Å for the complex with amprenavir, with acceptable quality of the final models (Table 2).

As expected, the flaps (β-hairpins that cover the active sites in retropepsins) that were partially disordered in the apoenzyme were fully ordered in the inhibitor complexes (Fig. 1A). Surprisingly, although the crystals of the inhibitor complexes diffracted better than the crystals of the apoenzyme, the visible parts of the termini of enzyme molecules were less complete. Ten residues at the N-terminus of each protomer in all inhibitor complexes and three residues at the C-terminus in the crystals of the complex with pepstatin A were found to be disordered (Fig. 1B). Superpositions with the program ALIGN [29] of the dimers of the complexes with pepstatin, TL-3 and amprenavir onto the dimer of the apoenzyme yielded rmsd values of 1.75, 1.81 and 1.86 Å for 205, 202 and 207 superimposed Cα pairs, respectively. Such deviations are unusually large for comparison of the structures of the same protein. When only monomers A for the corresponding pairs as above were compared, the deviations were smaller (1.36, 1.11 and 1.35 Å for 106, 100 and 103 target pairs, respectively), indicating that the two protomers must have moved relative to each other. A comparison of the positions of the monomers in the dimers of the inhibitor complexes with the apoenzyme revealed the rotation of each monomer in the former

Table 2. Data collection and structure refinement.

	XMRV PR–TL-3	XMRV PR–pepstatin A	XMRV PR–amprenavir
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Molecules/a.u. Unit cell, $a = b = c$ (Å); $\alpha = \beta = \gamma$ (°)	46. 6, 65.5, 69.8; 90	46.4, 65.46, 69.8; 90	46.6, 65.1, 69.2; 90
Resolution (Å)[a]	50.0–1.40 (1.45–1.40)	50.0–1.50 (1.55–1.50)	30.0–1.75 (1.81–1.75)
R_{merge}[b]	5.0 (43.8)	8.2 (31.1)	8.0 (64.0)
No. of reflections (measured/unique)	309 877/42 500	122 809/34 189	125 158/21 646
$<I/\sigma I>$	42.72 (2.77)	15.85 (2.38)	18.94 (2.1)
Completeness (%)	99.2 (95.5)	98.4 (92.7)	99.2 (96.2)
Redundancy	7.3 (4.1)	3.6 (2.5)	5.8 (4.6)
Refinement			
Resolution (Å)	47.78–1.40	14.84–1.50	27.77–1.75
No. of reflections (refinement/R_{free})	41 334/1095	32 901/1073	20 457/1096
R/R_{free}[c]	0.176/0.201	0.174/0.196	0.188/0.234
No. of atoms			
Protein	1805	1724	1753
Ligand/ion	85	76	35
Water	261	166	204
Rmsd from ideal			
Bond lengths (Å)	0.016	0.013	0.012
Bond angles (°)	1.75	1.6	1.5
PDB code	3SLZ	3SM1	3SM2

[a] The highest-resolution shell is shown in parentheses. [b] $R_{merge} = \sum_h\sum_i|I_i - \langle I\rangle|/\sum_h\sum_i I_i$, where I_i is the observed intensity of the ith measurement of reflection h, and $\langle I\rangle$ is the average intensity of that reflection obtained from multiple observations. [c] $R = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively, calculated for all data. R_{free} was as defined in [52].

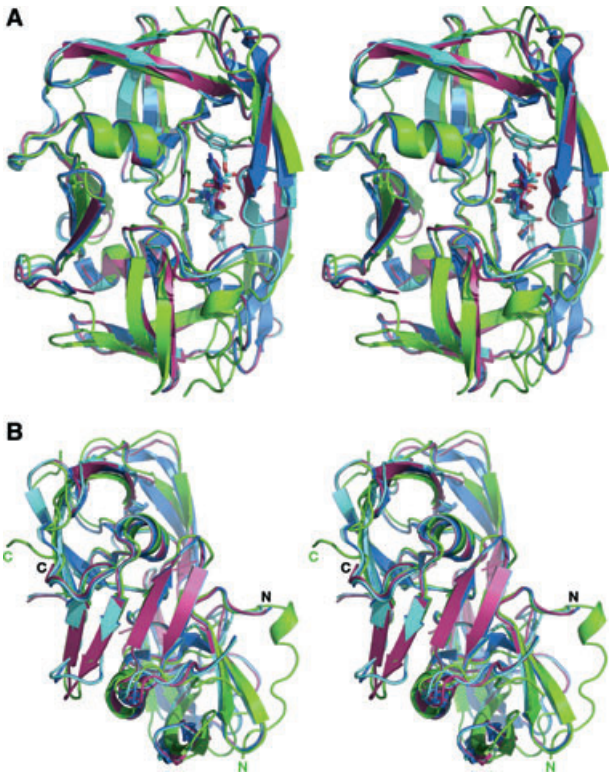


Fig. 1. Two views of the superimposed apoenzyme (green) and three inhibitor complexes of XMRV PR with TL-3 (cyan), pepstatin A (blue), and amprenavir (dark pink). (A) The classic view, revealing the closed conformation of the flaps in the inhibitor complexes, with the inhibitors shown as sticks. (B) An orientation showing the dimer interface and the extended termini in the apoenzyme, with the inhibitors removed. The N-termini of monomers A and C-termini of monomers B are labeled (black for the inhibitor complexes, and green for the apoenzyme).

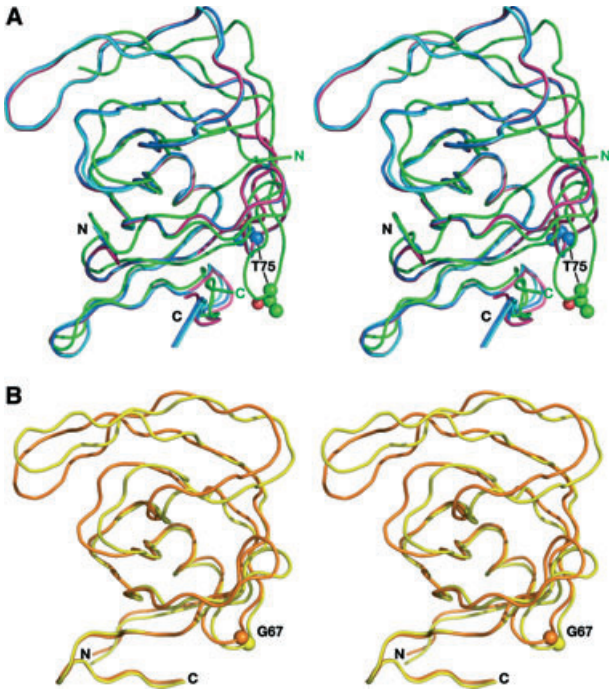


Fig. 2. Superposition of the monomers of retropepsins. (A) Ribbon representation of the protein chains in the inhibitor complexes and apoenzyme of XMRV PR, colored as in Fig. 1. The termini are marked as in Fig. 1B, and Thr75, the residue with the largest shift between the structures, is marked. (B) Superposition of apo-HIV-1 PR (yellow: 3HVP) and amprenavir complex of HIV-1 PR (orange: 3NU3) in a view corresponding to (A). Gly67 (marked) is a structural equivalent of Thr75 in XMRV PR.

towards the ligand by 6.8°, thus narrowing the cleft between the domains. A combination of global rotation and local adjustments results in significant movement (rmsd of more than 1 Å) of $\sim 2/3$ of a molecule, leading to shifts as large as 6.7 Å for Cα atoms of Thr75 (Fig. 2A). This residue is located at the tip of the β-hairpin (74–75), which is part of a structural element, an 'elbow' [30], that comprises two loops, the flap at one end and the following loop with Thr75 at the tip. Motions of the corresponding two loops in other retropepsins upon ligand binding have also been shown to be correlated [31]. Another residue exhibiting a very considerable shift is Val124, located near the C-terminus of the molecule (Fig. 2A), which moves by up to 5.5 Å in the amprenavir complex. When the shifts between the inhibited and the apo states of XMRV and HIV-1 PRs (Fig. 2) are compared, a much larger movement of the 'elbow' is seen in the former than in the latter. The resulting open conformation of an elbow in the apo-XMRV PR is stabilized by the intramolecular interactions between loop 74–75 in with the much longer C-terminal fragment that is not present in HIV-1 PR (Fig. 2).

The C-terminal strand in XMRV PR makes a distorted β-turn that involves residues 116–120, but the nascent third strand contributing to half of the dimer interface within each monomer, corresponding topologically to its counterpart in Ddi1 retroviral PR [14,32], is redirected in XMRV PR at Leu122 away from its neighboring strand (Fig. 3). The length of the visible part of that strand and the directions of the termini of the polypeptide chain differ slightly between different structures, but the hydrogen bonding pattern did not change upon ligand binding.

The inhibitor molecule in the complex with the C2-symmetric inhibitor TL-3 is bound to the protease dimer in a canonical extended conformation. The electron density corresponding to the inhibitor is very clear (Fig. 4A), and TL-3 appears to bind principally in one direction, with only a small contribution (20%) of binding in the opposite direction. Although the inhibitor is symmetric, its mode of binding to the

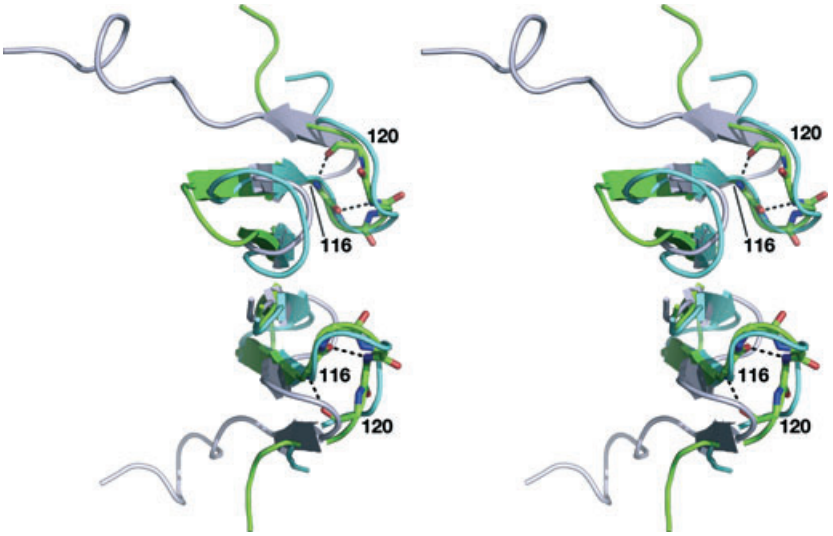


Fig. 3. Superposition of the dimer interfaces in apo-XMRV PR (green) and inhibited XMRV PR (cyan) with Ddi1 (gray). The main chain of the fragment that forms a distorted β-turn in each monomer of apo-XMRV PR is shown as sticks, with the hydrogen bonds dashed.

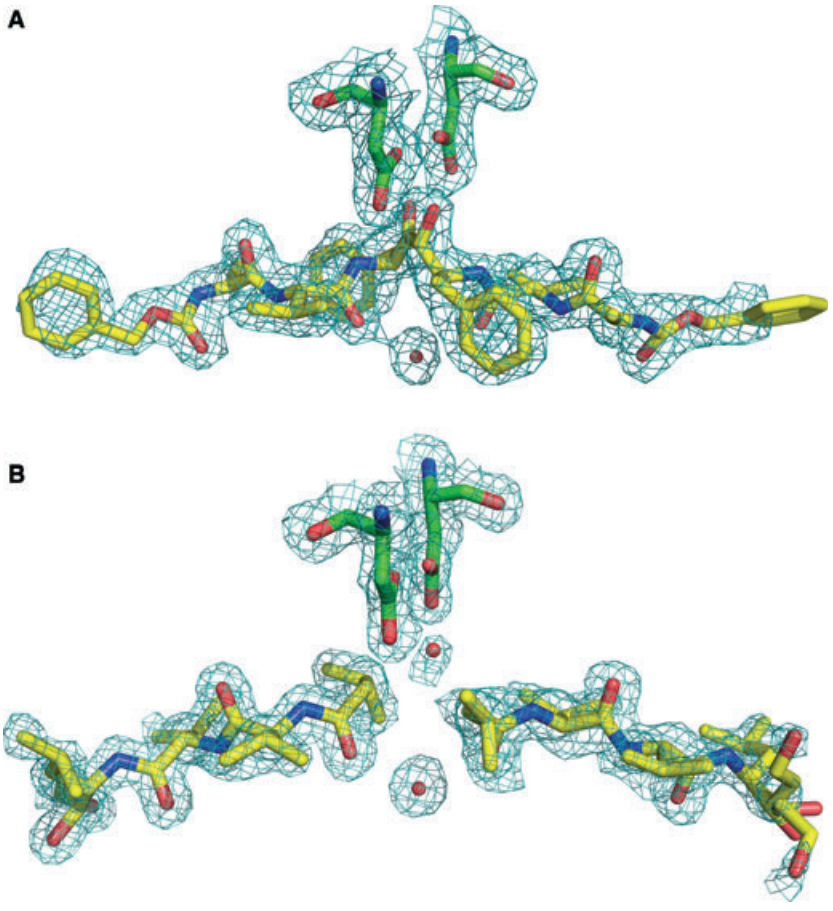


Fig. 4. Electron density maps for the inhibitors and the active site residues. (A) $2F_o - F_c$ electron density map for TL-3. (B) $2F_o - F_c$ map for the two molecules of pepstatin A. Both maps were contoured at a 1.0σ level.

enzyme is not, so the directionality of the inhibitor is determined by the fact that only one of two hydroxyls in its central core occupies the position of a nucleophilic water between the two catalytic aspartates. The second hydroxyl interacts with only one of the two catalytic aspartates in the dimer, which is itself not fully symmetric, as both molecules are crystallographically independent. Amprenavir binds in a canonical manner and in a single orientation, although at only 75% occupancy. The mode of binding of pepstatin A,

however, is very unusual and has not previously been seen in any of its complexes with aspartic PRs. It is clear that, instead of a single pepstatin A molecule binding to the enzyme, two molecules bind on the two sides of the catalytic aspartates, with the 'N-terminal' isovaleryl groups being very clearly visible in the proximity of the aspartates (Fig. 4B). Only four amino acids of pepstatin A are ordered in either molecule, with the C-terminal parts of the inhibitors disordered and not seen in the electron density. The catalytic water molecule that is usually present between the two aspartates in the structures of uninhibited aspartic PRs is also present (Fig. 4B), although that site is most likely occupied only partially, as indicated by comparatively weak electron density. On the other hand, the water molecule that mediates the contacts between the inhibitor and the flaps of the enzyme (named Wat301 in the first structure of an inhibitor complex of HIV-1 PR [33]) is fully occupied in both inhibitor complexes (Fig. 4B).

Although pepstatin A binds to the enzyme in a very different way from the other inhibitors, with two molecules of pepstatin A binding to the XMRV PR dimer rather than a single molecule of TL-3 or amprenavir, the interactions between the inhibitors and the enzyme are remarkably similar (Tables 3–6; Fig. 5). As the TL-3 inhibitor is symmetric and the two molecules of pepstatin A also bind in a symmetric manner, definition of the primed and unprimed binding sites on the enzyme [34] is arbitrary, and refers only to the interactions with the A and B protein molecules in the asymmetric unit of the crystal (the latter marked with a prime in Tables 3–6). The phenyl side chain of TL-3 that occupies the S1 pocket is much larger than the corresponding N-terminal isovaleryl group of pepstatin A, leading to a large shift in the location of Pro89 and Tyr90. Tyr90 also moves in the S1′ pocket, although Pro89 occupies almost the same position in the complexes with both inhibitors. Both inhibitors contain almost exactly superimposable valines at subsites P2/P2′. The P3/P3′ alanines of TL-3 occupy the same area in the very open subsites S3/S3′, whereas subsites S4/S4′, which are also quite open, are filled by the side chains of statine in pepstatin A and benzyl carbamate in TL-3. The main chain of pepstatin A continues away from the protein, and only a few further atoms are still visible.

With the peptide chains of TL-3 and pepstatin running in opposite directions, their peptide carbonyl groups still superimpose well, but their amide nitrogens are shifted by $\sim$ 1.5 Å (Fig. 5). Nevertheless, the peptides of both inhibitors are hydrogen bonded to the same groups on the enzyme, with the length of the hydrogen bonds being very similar (Tables 3–6).

As TL-3 binds to XMRV PR in a manner that is very typical for inhibitor binding to other retropepsins,

Table 4. Interactions between XMRV PR and the inhibitors. Hydrogen bonds with TL-3.

TL-3	XMRV PR, waters	Distance (Å)
O8	N Gln55	2.90
O9	OE1 Gln36	3.28
N4	OE1 Gln36, Wat53	3.11, 3.39
O4	N Gln36	2.88
N2	O Gln55	2.95
O2	Wat3	2.80
N1	O Gly34	2.92
O1	OD1 Asp32′,	2.96
O51	OD1 Asp32, Asp32′	2.54, 2.76
O51	OD2 Asp32, Asp32′	3.07, 3.33
N51	O Gly34′	2.95
O52	Wat3	2.70
N52	GLn55′	2.86
O54	N Gln36′, Wat16	2.99, 3.24
N54	OE1 Gln36′	3.11
O58	N Gln55′	2.82
O59	OE1 Gln36′	3.18

Table 3. Interactions between XMRV PR and the inhibitors. Hydrophobic interactions within 4.5 Å.

Subsite	TL-3	Pepstatin A	Amprenavir
P4	Gln36, His37, Ala52, Trp65	Gln36, His37, Ala52, Trp53, Gln55, Trp65, Leu83	
P3	Gln36, Tyr90′	Gln36, Gln55, Val54, Tyr90′	
P2	Ala35, Val39, Val54	Gly34, Ala35, Gln55, His37, Val54, Gln36, Ala57′	Ala35, Gln36, His37, Val39, Val54, Leu92
P1	Ala57, Gly56, Pro80′, Tyr90′	Asp32, Gly34, Gln55, Gly56, Ala57, Leu30′, Asp32′, Leu92′, Cys88′, Pro89′	Val54, Gly34, Gln55, Gly56, Ala57, Cys88′, Pro89′, Tyr90′
P1′	Gln55′, Gly56′, Ala57′, Cys88, Pro89, Tyr90, Leu92	Asp32, Cys88, Pro89, Leu92, Gly56′, Ala57′	Leu30, Asp32, Cys88, Pro89, Tyr90, Leu92, Gly34′, Gln55′, Gly56′, Ala57′
P2′	Gly34′, Ala35′,Val39′, Val54′, Gln55′, Leu83′, Leu92′	Ala57, Gly34′, Ala35′, Gln36′, His37′, Gln55′, Gly56′, Ala57′, Leu92′	Ala57, Ala35′, Gln36′, His37′, Val39′, Val54′, Gln55′, Leu83′, Leu92′
P3′	Gln36′, Gln55′, Tyr90	Pro89, Tyr90, Gln36′, Trp53′, Val54′, Gln55′	
P4′	Gln36′, Gln37′, Ala52′, Trp65′	Gln36′, His37′, Ala52′, Trp53′, Trp65′, Leu83′	

Table 5. Interactions between XMRV PR and the inhibitors. Hydrogen bonds with pepstatin A (the two pepstatin molecules are labeled M and J).

Pepstatin A	XMRV PR, waters	Distance (Å)
O Iva1 (M)	Wat160	2.72
N Val2 (M)	O Gly34	3.30
O Val2 (M)	N Gln36	2.88
N Val3 (M)	O Gln55	2.99
O Val3 (M)	N Gln55	2.86
O Sta4 (M)	OE1 Gln36	2.94
OH Sta4 (M)	O Trp53	3.24
N Sta4 (M)	OE1 Gln36	2.96
O Iva1 (J)	Wat160	2.69
N Val2 (J)	O Gly34′	3.04
O Val2 (J)	N Gln36′, Wat120	2.88, 3.12
N Val3 (J)	O Gln55′	2.87
O Val3 (J)	N Gln55′	2.82
OH Sta4 (J)	O Trp53′	2.68
N Sta4 (J)	OE1 Gln36′	2.91

Table 6. Interactions between XMRV PR and the inhibitors. Hydrogen bonds with amprenavir.

Amprenavir	XMRV PR, waters	Distance (Å)
O5, O4	Wat10	2.54, 3.42
N3	O His37′	3.24
O3	OD1, OD2 Asp32	2.66, 3.21
O3	OD1, OD2 Asp32′	2.80, 2.89
N1	O GLy34	3.11
O2	Wat10	2.84
O1	Wat36	3.02
O6	N His37	3.45

its conformation closely resembles the conformation of the same inhibitor bound to feline immunodeficiency virus PR and HIV-1 PR (Fig. 6). This is particularly true for side chains P1/P1′ and P2/P2′, whereas larger deviations can be seen for groups occupying subsites S3/S3′ of the enzymes. The conformation of amprenavir bound to XMRV PR is almost identical to the principal conformation of this inhibitor reported in an atomic-resolution structure of its complex with HIV-1 PR (PDB code: 3NU3 [35]). Previous modeling studies of the mode of binding to human T-lymphotropic virus 1 PR of approved drugs directed against HIV-1 PR also suggested a reasonable fit of amprenavir to the former enzyme [36].

The unusual mode of binding of pepstatin A to XMRV PR demonstrates that mechanism-based inhibitors of aspartic PRs could function even without direct interactions between the transition state isosteres and the catalytic residues. Known inhibitors usually contain such an isostere in the form of hydroxyethylene, phosphinate, difluoroketone, etc., with a hydroxyl group substituting for the catalytic water molecule. However, even though such an isostere is present in the statine of pepstatin A, its hydroxyl is not located in the expected position. It appears that the maintenance of the specific hydrogen bonding pattern along the chain and filling of the substrate-binding subsites by the side chains of an inhibitor (Fig. 5) provides sufficient binding energy to block the active site and to inhibit the enzyme. Noncanonical modes of binding of aspartic PR inhibitors have been noted in the past, e.g. in inhibitor complexes of plasmepsins I and II [37,38] and several inhibitors of histo-aspartic PR [39]. However, no similarities with the binding mode of pepstatin A described here have been reported. It remains to be shown whether this mode of binding is specific for XMRV PR, or whether it is also present in other aspartic PRs.

Inhibition constants for a number of different known inhibitors of retropepsins were determined in

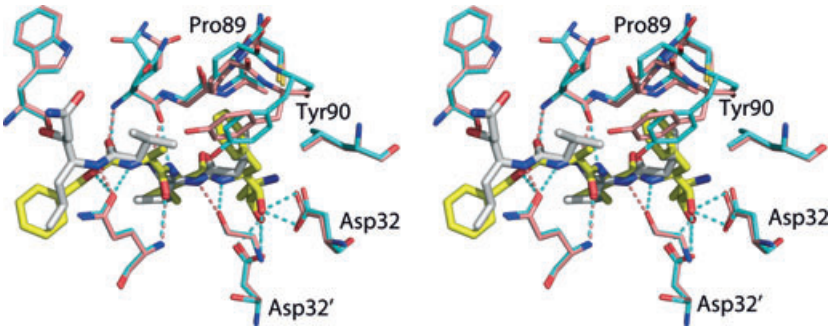


Fig. 5. Interactions between the inhibitor molecules and XMRV PR. The 'N-terminal' half of TL-3 and molecule M of pepstatin A are shown, together with the surrounding residues of XMRV PR, thus delineating the P1–P4 sites of the inhibitors. It should be noted that this is an arbitrary assignment, as the structures are symmetric, and rely only on the interactions with the crystallographically distinct protomers A and B (the latter marked with primes in Table 2) of the enzyme. Carbon atoms of the protein side chains are blue for the TL-3 complex and pink for the pepstatin A complex, and the inhibitors are yellow and gray, respectively.

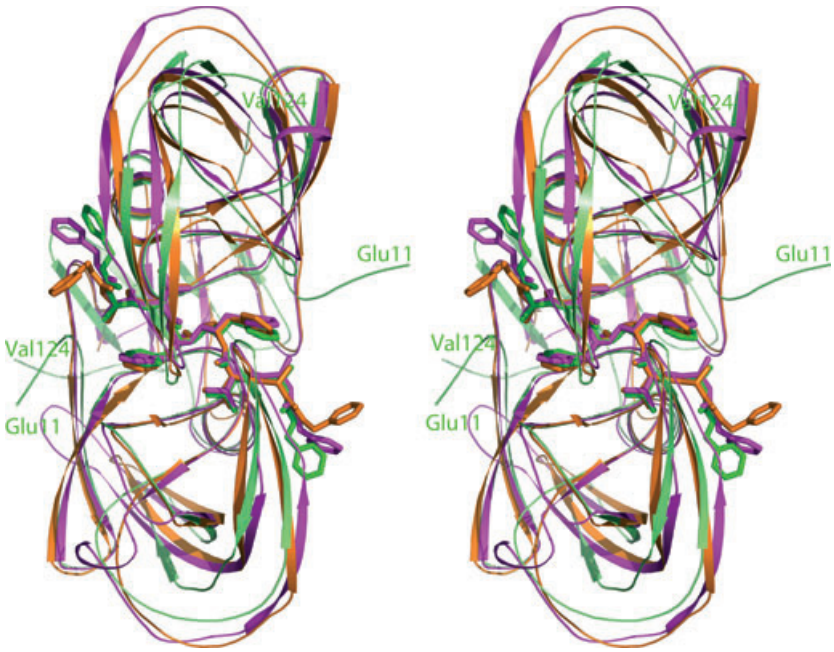


Fig. 6. Superposition of the crystal structures of the complexes of TL-3 with XMRV PR (green), HIV-1 PR (orange), and FIV PR (magenta).

two parallel sets of experiments. In one approach, *in vitro* inhibition constants were determined for selected retroviral protease inhibitors with an HPLC-based method (Table 1). The oligopeptide substrate utilized for this assay, RSLLY↓PALTP (where the arrow indicates the point of cleavage), demonstrated the following kinetic parameters: K_m = 0.216 ± 0.027 mM, k_{cat} = 0.55 ± 0.04 s^{-1}, and k_{cat}/K_m = 2.55 ± 0.37 $mM^{-1}{\cdot}s^{-1}$. The specificity constant (k_{cat}/K_m) was very similar to that determined for MMLV PR (2.74 ± 0.32 $mM^{-1}{\cdot}s^{-1}$), and in the same range as values determined for MMLV Gag cleavage site-mimicking peptides with MMLV PR [40]. The K_i values suggested that amprenavir is the most potent inhibitor of XMRV PR among those tested (Table 1), and it was utilized for active site titration of the enzyme. Comparison of the protein amount determined by the Bradford method with the active amount of enzyme suggested that 12% of the pure protein was active. The much less efficient inhibition by pepstatin A suggests that it does not function as a transition state analog inhibitor of XMRV PR, which is in good agreement with the unusual binding of the inhibitor to the enzyme (Fig. 4B).

The activity of XMRV PR was also tested with a recombinant MMLV Gag fragment (MMLVGagΔ2) that contains the p12/CA, CA/NC and NC/PR cleavage sites [40]. Cleavage of MMLVGagΔ2 with XMRV PR provided the characteristic CA (31 kDa) and Δp12-CA (34 kDa) fragments, which are also seen after cleavage by MMLV PR [40]. Amprenavir appeared to be a substantially more potent inhibitor of

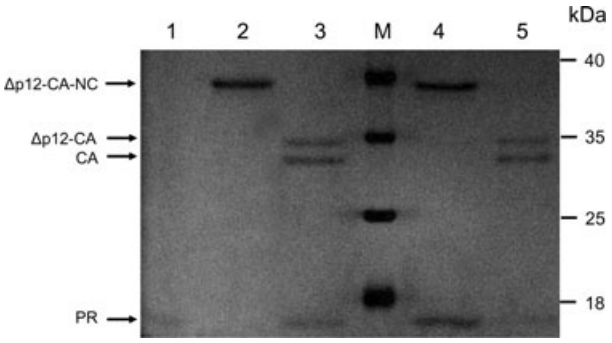


Fig. 7. Cleavage of MMLVGagΔ2 by XMRV PR. XMRV PR (30 nM) and MMLVGagΔ2 were incubated for 1 h alone (lanes 1 and 2) or together in the absence of any inhibitor (lane 3), in the presence of amprenavir (3.3 μM, lane 4), or in the presence of TL-3 (1 mM, lane 5). Reactions were stopped by the addition of loading buffer and subjected to SDS/PAGE followed by Coomassie staining. Molecular masses of the protein markers (lane M) are indicated. Arrows indicate the uncleaved recombinant protein (Δp12-CA-NC) and its fragments.

MMLVGagΔ2 cleavage than TL-3 (Fig. 7). As the protein cleavage was performed under conditions of substantially lower ionic strength than the peptide processing, the higher potency of amprenavir against XMRV PR appears to be independent of the ionic strength.

In the second kinetic approach, inhibition constants were determined at 25 °C with a continuous spectrophotometric assay that was originally developed [41] for analyses of HIV-1 PR using KARVnL↓NphEAnLG,

where nL is norLeucine, Nph is *p*-nitrophenylalanine, and the arrow indicates the point of cleavage. With this particular substrate, the following kinetic parameters were determined: $K_m = 0.019 \pm 0.0035$ mM, $k_{cat} = 9.8 \pm 1.7\ s^{-1}$, and $k_{cat}/K_m = 520 \pm 130\ mM^{-1} \cdot s^{-1}$. These values are all similar to those determined for HIV-1 PR with the same substrate. Again, amprenavir was the strongest inhibitor in these assays (Table 1), but several more compounds showing inhibition in the 10–40 nM range of K_i were identified (Table 1). Two of these compounds were discovered by *in silico* screening against the XMRV PR apoenzyme. The higher values of K_i determined in this assay were probably attributable to differences in pH (5.6 versus 5.0), ionic strength (2 M NaCl versus 1 M NaCl), and temperature (37 °C versus 25 °C).

The results of structural and biochemical studies of XMRV PR indicate that, despite significant differences in the topology of the enzyme as compared with other retropepsins, its interactions with the well-characterized inhibitors of this class of enzymes are similar, in both structural and kinetic terms, with the exception of the unusual binding mode of pepstatin A. The detailed description of the substrate-binding pockets presented here may assist in the development of inhibitors that are more specific for this subfamily, which includes XMRV, MMLV, and similar retroviruses.

Experimental procedures

Crystallization of XMRV PR

XMRV PR was expressed and purified following previously described procedures [13,14]. Recombinant XMRV PR engineered with an N-terminal noncleavable His_6 purification tag was expressed in *Escherichia coli* and purified on a nickel column. The resulting polypeptide consisted of 132 amino acids (initial Met, His_6, and the complete 125-residue PR). Before addition of the inhibitors for crystallization, the PR sample buffer was exchanged for 20 mM sodium citrate (pH 5.5), also including 0.2 M NaCl, and was concentrated to 6 $mg \cdot mL^{-1}$. The inhibitors were added at an XMRV PR (monomer)/inhibitor molar ratio of 4 : 1 for TL-3 and pepstatin, and at a 1 : 1 ratio for amprenavir. All crystallizations were carried out with the hanging drop vapor diffusion method. Each drop contained 4 μL of the complex sample mixed with 2 μL of well solution, and was equilibrated with 500 μL of the latter. The conditions yielding the crystals of the TL-3 complex were 3.5 M sodium formate (pH 5.5), whereas crystals of the pepstatin A complex grew at pH 7.0, and crystals of the amprenavir complex grew at pH 4.75. The crystals grew slowly, taking over a month to reach a size of $0.1 \times 0.1 \times 0.15$ mm for the TL-3 complex, $0.05 \times 0.05 \times 0.2$ mm for the pepstatin A complex, and $0.1 \times 0.2 \times 0.1$ mm for the amprenavir complex.

Diffraction data for the TL-3, pepstatin A and amprenavir complexes extending to 1.4-, 1.5- and 1.75-Å resolution, respectively, were collected with one crystal of each complex. Data were measured on the beamline 22-ID of the Southeast Regional Collaborative Access Team, located at the Advanced Photon Source, Argonne National Laboratory, with a MAR300CCD detector. Crystals were cryoprotected before rapid freezing, and diffraction intensities were measured at 100 K. Diffraction data for the TL-3 complex were collected in two passes: 50–1.4 Å with exposure of 3 s per degree and 50–2.4 Å with exposure of 2 s per deg. Diffraction data for the pepstatin A and amprenavir complexes were measured in a single pass at 2 s per degree. Data were indexed, integrated and scaled with the HKL2000 package [42]. Despite the differences in crystallization conditions, the crystals of all three complexes were isomorphous in the orthorhombic space group $P2_12_12_1$ (Table 2). The structures were solved by molecular replacement with PHASER [43], with a monomer of XMRV PR (PDB ID: 3NR6) as a search model. The structures were refined with REFMAC5 [44], with the final parameters listed in Table 2. Figures were prepared with PYMOL [45].

Determination of the inhibition constants

XMRV PR was diluted with 20 mM Pipes (pH 7.0), containing 100 mM NaCl, 10% glycerol, and 0.5% NP40. The activity of XMRV PR was measured with an HPLC-based assay, as described previously for MMLV PR [40]. Briefly, the PR assays were initiated by mixing 5 μL of PR, 10 μL of 2 × incubation buffer (0.5 M potassium phosphate buffer, pH 5.6, containing 10% glycerol, 10 mM dithiothreitol, and 4 M NaCl), and 5 μL of 0.12–0.8 mM RSLLY↓PALTP, a P3 Leu-substituted peptide mimicking the MA/p12 cleavage site of MMLV PR (RSSLY↓PALTP). Inhibitors were assayed by using 4.8 μL of substrate and 0.2 μL of inhibitor in dimethylsulfoxide or dimethylsulfoxide alone. The reaction mixture was incubated at 37 °C for 1 h, and the reaction was stopped by the addition of 180 μL of 1% trifluoroacetic acid. The enzyme concentration in the assay was selected to cause < 20% substrate hydrolysis. Separation of cleavage products with reversed-phase chromatography was performed as described previously [46]. Cleavage products were identified by retention time as compared with previous runs performed with MMLV and HIV-1 PRs, and the amounts of cleavage product were determined on the basis of integration value–peptide amount correlation determined by amino acid analysis for HIV-1 PR-mediated cleavage. The K_i values were obtained from the IC_{50} values determined from the inhibitor dose–response curves, with the equation $K_i = (IC_{50} - [E]/2)/(1 + [S]/K_m)$, where [E]

and [S] are the PR and substrate concentrations, respectively [47]. The exact amount of active PR in the preparations used for kinetic measurements was determined by active center titration with amprenavir, using the HPLC method. Kinetic parameters were determined by fitting the data to the Michaelis–Menten equation with ENZYME KINETICS MODULE 1.1 of SIGMAPLOT 8.0 (Systat Software).

In an alternative approach, inhibition constants were also determined with a continuous spectrophotometric assay. XMRV PR activity was assayed kinetically in 250 mM sodium acetate, 200 mM imidazole, and 1 M NaCl (pH 5.0), at 25 °C, with the chromogenic substrate KARVnL↓NphEAnLG (Nph = *p*-nitrophenylalanine) [48]. Confirmation of cleavage at the position indicated by the arrow was obtained by observing the shift in absorbance maximum from 280 to 272 nm, as previously reported for HIV-1 PR cleavage of the same peptide [49].

Reactions were not carried out at 37 °C, owing to autoproteolysis, which was prevented by performing the reactions at lower temperature. Cleavage of the substrate was monitored with a Cary 50 Bio Varian spectrophotometer equipped with an 18-cell multitransport system. For determination of K_i, the enzyme was preincubated with inhibitor for 5 min at 25 °C. Reactions were initiated by the addition of 50 μM chromogenic substrate, and the initial rates of substrate hydrolysis were monitored over a range of inhibitor concentrations at 25 °C. The dimethylsulfoxide concentration for all reactions was 2%. K_i values were calculated by fitting initial velocities to the Dixon equation [50], with ENZYME KINETICS MODULE 1.1 of SIGMAPLOT 10.0 (Systat Software). K_i values for all inhibitors were measured under the same conditions.

Cleavage of recombinant MMLV Gag fragment with XMRV PR

MMLVGagΔ2 (3.7 μM) was incubated in 75 mM phosphate buffer (pH 5.6) and 0.5 mM EDTA for 1 h at 37 °C in the absence of XMRV PR, or with XMRV PR (30 nM) in the absence and presence of amprenavir (3.3 μM) or TL-3 (1 mM). Reactions were stopped by the addition of loading buffer and subjected to SDS/PAGE, followed by staining with Coomassie Brillant Blue. A protein ladder (Fermentas) was used to determine the molecular masses of protein fragments.

Acknowledgements

The help of B. Farkas with kinetic and inhibition measurements is greatly appreciated. We acknowledge the use of beamline 22-ID of the Southeast Regional Collaborative Access Team, located at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. The work of K. Matúz and J. Tözsér was supported by the TÁMOP 4.2.1./B-09/1/KONV-2010-0007 project and by the Hungarian Science and Research Fund (OTKA K68288). This work was also supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract HHSN261200800001E, and by grant R37 AI28571 from NIAID to B. M. Dunn. N. E. Goldfarb was supported by a Ruth L. Kirschstein National Research Service Award 5T32 CA009126-33 Training Grant in Cancer Biology. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, and nor does the mention of trade names, commercial products or organizations imply endorsement by the US Government.

References

1 Broder S & Gallo RC (1984) A pathogenic retrovirus (HTLV-III) linked to AIDS. *N Engl J Med* **311**, 1292–1297.
2 Montagnier L (1985) Lymphadenopathy-associated virus: from molecular biology to pathogenicity. *Ann Intern Med* **103**, 689–693.
3 Schlaberg R, Choe DJ, Brown KR, Thaker HM & Singh IR (2009) XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors. *Proc Natl Acad Sci USA* **106**, 16351–16356.
4 Lombardi VC, Ruscetti FW, Das GJ, Pfost MA, Hagen KS, Peterson DL, Ruscetti SK, Bagni RK, Petrow-Sadowski C, Gold B *et al.* (2009) Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. *Science* **326**, 585–589.
5 Mikovits JA, Lombardi VC, Pfost MA, Hagen KS & Ruscetti FW (2010) Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. *Virulence* **1**, 386–390.
6 Groom HC, Boucherit VC, Makinson K, Randal E, Baptista S, Hagan S, Gow JW, Mattes FM, Breuer J, Kerr JR *et al.* (2010) Absence of xenotropic murine leukaemia virus-related virus in UK patients with chronic fatigue syndrome. *Retrovirology* **7**, 10.
7 Hohn O, Strohschein K, Brandt AU, Seeher S, Klein S, Kurth R, Paul F, Meisel C, Scheibenbogen C & Bannert N (2010) No evidence for XMRV in German CFS and MS patients with fatigue despite the ability of the virus to infect human blood cells in vitro. *PLoS ONE* **5**, e15632.

8 Hue S, Gray ER, Gall A, Katzourakis A, Tan CP, Houldcroft CJ, McLaren S, Pillay D, Futreal A, Garson JA *et al.* (2010) Disease-associated XMRV sequences are consistent with laboratory contamination. *Retrovirology* **7**, 111.
9 Smith RA (2010) Contamination of clinical specimens with MLV-encoding nucleic acids: implications for XMRV and other candidate human retroviruses. *Retrovirology* **7**, 112.
10 Cohen J (2011) Retrovirology. More negative data for link between mouse virus and human disease. *Science* **331**, 1253–1254.
11 Knox K, Carrigan D, Simmons G, Teque F, Zhou Y, Hackett J Jr, Qiu X, Luk KC, Schochetman G, Knox A *et al.* (2011) No evidence of murine-like gammaretroviruses in CFS patients previously identified as XMRV-infected. *Science* **333**, 94–97.
12 Paprotka T, Delviks-Frankenberry KA, Cingoz O, Martinez A, Kung HJ, Tepper CG, Hu WS, Fivash MJ Jr, Coffin JM & Pathak VK (2011) Recombinant origin of the retrovirus XMRV. *Science* **333**, 97–101.
13 Gillette WK, Esposito D, Taylor TE, Hopkins RF, Bagni RK & Hartley JL (2010) Purify first: rapid expression and purification of proteins from XMRV. *Protein Expr Purif* **76**, 238–247.
14 Li M, DiMaio F, Zhou D, Gustchina A, Lubkowski J, Dauter Z, Baker D & Wlodawer A (2011) Crystal structure of XMRV protease differs from the structures of other retropepsins. *Nat Struct Mol Biol* **18**, 227–229.
15 Smith RA, Gottlieb GS & Miller AD (2010) Susceptibility of the human retrovirus XMRV to antiretroviral inhibitors. *Retrovirology* **7**, 70.
16 Umezawa H, Aoyagi T, Morishima H, Matsuzaki M & Hamada M (1970) Pepstatin, a new pepsin inhibitor produced by Actinomycetes. *J Antibiot (Tokyo)* **23**, 259–262.
17 Seelmeier S, Schmidt H, Turk V & von der Helm K (1988) Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. *Proc Natl Acad Sci USA* **85**, 6612–6616.
18 Fitzgerald PMD, McKeever BM, VanMiddlesworth JF, Springer JP, Heimbach JC, Leu C-T, Herber WK, Dixon RAF & Darke PL (1990) Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0 Å resolution. *J Biol Chem* **265**, 14209–14219.
19 Slee DH, Laslo KL, Elder JH, Ollmann IR, Gustchina A, Kervinen J, Zdanov A, Wlodawer A & Wong C-H (1995) Selectivity in the inhibition of HIV and FIV protease: inhibitory and mechanistic studies of pyrrolidine-containing α-keto amide and hydroxyethylamine core structures. *J Am Chem Soc* **117**, 11867–11878.
20 Wlodawer A, Gustchina A, Reshetnikova L, Lubkowski J, Zdanov A, Hui KY, Angleton EL, Farmerie WG, Goodenow MM, Bhatt D *et al.* (1995) Structure of an inhibitor complex of the proteinase from feline immunodeficiency virus. *Nat Struct Biol* **2**, 480–488.
21 Fung HB, Kirschenbaum HL & Hameed R (2000) Amprenavir: a new human immunodeficiency virus type 1 protease inhibitor. *Clin Ther* **22**, 549–572.
22 Kim EE, Baker CT, Dwyer MD, Murcko MA, Rao BG, Tung RD & Navia MA (1995) Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J Am Chem Soc* **117**, 1181–1182.
23 James MN, Sielecki A, Salituro F, Rich DH & Hofmann T (1982) Conformational flexibility in the active sites of aspartyl proteinases revealed by a pepstatin fragment binding to penicillopepsin. *Proc Natl Acad Sci USA* **79**, 6137–6141.
24 Bott R, Subramanian E & Davies DR (1982) Three-dimensional structure of the complex of the Rhizopus chinensis carboxyl proteinase and pepstatin at 2.5-Å resolution. *Biochemistry* **21**, 6956–6962.
25 Bailey D, Cooper JB, Veerapandian B, Blundell TL, Atrash B, Jones DM & Szelke M (1993) X-ray-crystallographic studies of complexes of pepstatin A and a statine-containing human renin inhibitor with endothiapepsin. *Biochem J* **289**, 363–371.
26 Li M, Morris GM, Lee T, Laco GS, Wong C-H, Olson AJ, Elder JH, Wlodawer A & Gustchina A (2000) Structural studies of FIV and HIV-1 proteases complexed with an efficient inhibitor of FIV protease. *Proteins Struct Funct Genet* **38**, 29–40.
27 Heaslet H, Lin YC, Tam K, Torbett BE, Elder JH & Stout CD (2007) Crystal structure of an FIV/HIV chimeric protease complexed with the broad-based inhibitor, TL-3. *Retrovirology* **4**, 1.
28 Noble S & Goa KL (2000) Amprenavir: a review of its clinical potential in patients with HIV infection. *Drugs* **60**, 1383–1410.
29 Cohen GE (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J Appl Crystallogr* **30**, 1160–1161.
30 Perryman AL, Lin JH & McCammon JA (2006) Restrained molecular dynamics simulations of HIV-1 protease: the first step in validating a new target for drug design. *Biopolymers* **82**, 272–284.
31 Perryman AL, Zhang Q, Soutter HH, Rosenfeld R, McRee DE, Olson AJ, Elder JE & Stout CD (2010) Fragment-based screen against HIV protease. *Chem Biol Drug Des* **75**, 257–268.
32 Sirkis R, Gerst JE & Fass D (2006) Ddi1, a eukaryotic protein with the retroviral protease fold. *J Mol Biol* **364**, 376–387.
33 Wlodawer A, Miller M, Jaskólski M, Sathyanarayana BK, Baldwin E, Weber IT, Selk LM, Clawson L, Schneider J & Kent SBH (1989) Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **245**, 616–621.

34 Schechter I & Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157–162.
35 Shen CH, Wang YF, Kovalevsky AY, Harrison RW & Weber IT (2010) Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobic clusters. *FEBS J* **277**, 3699–3714.
36 Li M, Laco GS, Jaskólski M, Rozycki J, Alexandratos J, Wlodawer A & Gustchina A (2005) Crystal structure of human T-cell leukemia virus protease, a novel target for anti-cancer drug design. *Proc Natl Acad Sci USA* **102**, 18322–18337.
37 Bhaumik P, Horimoto Y, Xiao H, Miura T, Hidaka K, Kiso Y, Wlodawer A, Yada RY & Gustchina A (2011) Crystal structures of the free and inhibited forms of plasmepsin I (PMI) from *Plasmodium falciparum*. *J Struct Biol* **175**, 73–84.
38 Prade L, Jones AF, Boss C, Richard-Bildstein S, Meyer S, Binkert C & Bur D (2005) X-ray structure of plasmepsin II complexed with a potent achiral inhibitor. *J Biol Chem* **280**, 23837–23843.
39 Bhaumik P, Gustchina A & Wlodawer A (2011) Structural studies of vacuolar plasmepsins. *Biochim Biophys Acta* In press, doi:10.1016/j.bbapap.2011.04.008
40 Feher A, Boross P, Sperka T, Miklossy G, Kadas J, Bagossi P, Oroszlan S, Weber IT & Tozser J (2006) Characterization of the murine leukemia virus protease and its comparison with the human immunodeficiency virus type 1 protease. *J Gen Virol* **87**, 1321–1330.
41 Richards AD, Phylip LH, Farmerie WG, Scarborough PE, Alvarez A, Dunn BM, Hirel PH, Konvalinka J, Strop P & Pavlickova L (1990) Sensitive, soluble chromogenic substrates for HIV-1 proteinase. *J Biol Chem* **265**, 7733–7736.
42 Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
43 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) *Phaser* crystallographic software. *J Appl Crystallogr* **40**, 658–674.
44 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr* **67**, 355–367.
45 DeLano WL (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA.
46 Tözsér J, Gustchina A, Weber IT, Blaha I, Wondrak EM & Oroszlan S (1991) Studies on the role of the S4 substrate binding site of HIV proteinases. *FEBS Lett* **279**, 356–360.
47 Maibaum J & Rich DH (1988) Inhibition of porcine pepsin by two substrate analogues containing statine. The effect of histidine at the P2 subsite on the inhibition of aspartic proteinases. *J Med Chem* **31**, 625–629.
48 Coman RM, Robbins AH, Fernandez MA, Gilliland CT, Sochet AA, Goodenow MM, McKenna R & Dunn BM (2008) The contribution of naturally occurring polymorphisms in altering the biochemical and structural characteristics of HIV-1 subtype C protease. *Biochemistry* **47**, 731–743.
49 Dunn BM, Scarborough PE, Davenport R & Swietnicki W (1994) Analysis of Proteinase Specificity by Studies of Peptide Substrates: The Use of UV and Fluorescence Spectroscopy to Quantitate Rates of Enzymatic Cleavage. In: *Methods in Molecular Biology*, Vol. 36: Peptide Analysis Protocols (Dunn BM & Pennington MW eds.), pp. 225–243, Humana Press, Totowa.
50 Dixon M (1972) The graphical determination of Km and Ki. *Biochem J* **129**, 197–202.
51 Janka L, Clemente J, Vaiana N, Sparatore A, Romeo S & Dunn BM (2008) Targeting the plasmepsin 4 orthologs of Plasmodium sp. with 'double drug' inhibitors. *Protein Pept Lett* **15**, 868–873.
52 Brünger AT (1992) The free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.

