

Structural Characterization of B and non-B Subtypes of HIV-Protease: Insights into the Natural Susceptibility to Drug Resistance Development

Mario Sanches¹, Sandra Krauchenco¹, Nadia H. Martins¹
Alla Gustchina², Alexander Wlodawer² and Igor Polikarpov^{1*}

¹Grupo de Cristalografia
Instituto de Física de São Carlos
Universidade de São Paulo
Av. Trabalhador Saocarlense
400, CEP 13560-970
São Carlos, SP, Brazil

²Protein Structure Section
Macromolecular
Crystallography Laboratory
National Cancer Institute
at Frederick, Frederick
MD 21702, USA

Although a majority of HIV-1 infections in Brazil are caused by the subtype B virus (also prevalent in the United States and Western Europe), viral subtypes F and C are also found very frequently. Genomic differences between the subtypes give rise to sequence variations in the encoded proteins, including the HIV-1 protease. The current anti-HIV drugs have been developed primarily against subtype B and the effects arising from the combination of drug-resistance mutations with the naturally existing polymorphisms in non-B HIV-1 subtypes are only beginning to be elucidated. To gain more insights into the structure and function of different variants of HIV proteases, we have determined a 2.1 Å structure of the native subtype F HIV-1 protease (PR) in complex with the protease inhibitor TL-3. We have also solved crystal structures of two multi-drug resistant mutant HIV PRs in complex with TL-3, from subtype B (*Bmut*) carrying the primary mutations V82A and L90M, and from subtype F (*Fmut*) carrying the primary mutation V82A plus the secondary mutation M36I, at 1.75 Å and 2.8 Å resolution, respectively. The proteases *Bmut*, *Fwt* and *Fmut* exhibit sevenfold, threefold, and 54-fold resistance to TL-3, respectively. In addition, the structure of subtype B wild type HIV-PR in complex with TL-3 has been redetermined in space group P6₁, consistent with the other three structures. Our results show that the primary mutation V82A causes the known effect of collapsing the S1/S1' pockets that ultimately lead to the reduced inhibitory effect of TL-3. Our results further indicate that two naturally occurring polymorphic substitutions in subtype F and other non-B HIV proteases, M36I and L89M, may lead to early development of drug resistance in patients infected with non-B HIV subtypes.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: HIV-1 protease structures; non-B subtype; drug resistance; subtype F protease; L89M mutation

*Corresponding author

Present address: M. Sanches, Brazilian Synchrotron Light Laboratory (LNLS), Center for Structural Molecular Biology, Caixa Postal 6192, CEP 13084-971, Campinas, SP, Brazil.

Abbreviations used: HIV-PR, human immunodeficiency virus protease; *Bwt*, subtype B wild-type HIV-PR; *Fwt*, subtype F wild-type HIV-PR; *Bmut*, multi-drug resistant subtype B mutant HIV-PR; *Fmut*, drug-resistant subtype F mutant HIV-PR; PI, a commercial protease inhibitor used in AIDS therapy.

E-mail address of the corresponding author:
ipolikarpov@if.sc.usp.br

Introduction

Acquired immunodeficiency syndrome (AIDS) is a complex of symptoms and diseases resulting from infection by the human immunodeficiency virus (HIV). HIV is a member of the Lentivirus genus, which also includes, among others, simian immunodeficiency virus (SIV) and feline immunodeficiency virus (FIV).¹ HIV is characterized by a wide range of viral genetic diversity among the distinct types, groups, and clades. The two major distinct types of HIV, HIV-1 and HIV-2, are distinguished by

their genome organization and phylogenetic relationship. Further analyses of different strains of HIV-1, from diverse geographical origins, show that isolates can be subdivided into groups, subtypes, sub-subtypes, and circulating recombinant forms (CRFs), based on phylogenetic sequence differences. Groups refer to distinctive HIV-1 lineages, M (for Major), O (for Outlier), and N (for New, or Non-M, Non-O); a majority of strains found worldwide belong to group M. Nine subtypes of HIV-1 group M (A–D, F–H, J and K) are currently identified, as well as 14 circulating recombinant forms (CRFs),² the most common of which are CRF01_AE and CRF02_AG.³ The sequences of these subtypes and recombinant forms differ from one another by 25–35% in the *env* gene, and 10–15% in the *pol* gene, which includes the coding regions for protease (PR) and reverse transcriptase (RT). Although the *pol* gene is the most conserved region of HIV-1, there is sufficient diversity to allow phylogenetic subtype identification.⁴

Almost all studies on drug susceptibility of HIV-1 have been performed in developed countries, where subtype B still dominates the epidemic, but on the worldwide scale this is not a predominant HIV subtype.³ The Brazilian epidemic is characterized by the presence of multiple HIV-1 group M subtypes, primarily subtype B and subtype F, but also subtypes C, D, and other circulating recombinant forms.⁵

HIV-1 PR continues to be one of the primary targets of AIDS drug discovery due to its central role in processing of viral polypeptide precursors.⁶ Although inhibitors of HIV-PR slow down the progress of the disease, they do not completely suppress viral replication, and the rapid development of drug resistance decreases their efficacy. More than 87 mutations have already been reported in at least 49 positions within the 99 residue-long HIV-PR; many of them have been identified as potential contributors to resistance toward one or more inhibitors.⁷ These mutations are labeled as primary and secondary, with the primary mutations directly reducing drug susceptibility, whereas the secondary mutations contribute to resistance by reducing drug susceptibility or improving the replicative fitness of isolates with a primary mutation.⁸ The type and location of possible mutations is restricted by the necessity of the virus to produce an active enzyme with sufficient affinity for the substrate, in order to maintain viability. Most primary mutations occur in the active site and, although they usually preserve its charge and polarity, they alter its geometry. Other mutations cause resistance by altering enzyme catalysis, dimer stability, inhibitor binding kinetics, conformational dynamics, or by reshaping the active site through long-range structural perturbations.⁸ Extensive use of anti-retroviral drugs may cause different effects on distinct subtypes, due to additional selective pressure on HIV-PR. The polymorphism of subtypes could modulate anti-retroviral drug susceptibility and the possibility of drug resistance during therapy as well the fitness of HIV-1 variants.⁴

Studies involving the effects of the available HIV-PR inhibitors in persons infected with different subtypes

of HIV-1 are still at the early stages. Therefore, with the aim of expanding the information about non-B subtypes and their relation to drug resistance, we present the crystal structures and kinetic studies for three multi-drug-resistant variants of HIV-PR isolated from Brazilian patients: a wild-type HIV-PR of subtype F (*Fwt*) obtained from a naïve individual carrying polymorphic mutations I15V, E35D, M36I, S37A, R41K, R57K, D60E, Q61N, I62V, L63S, I64L, L89M and two mutant HIV-PR isolated from patients failing intensive anti-AIDS therapy. One of the latter is of subtype B (*Bmut*), with mutations S37A, R41K, K45R, I54V, L63P, A71V, V82A, L90M, and the other of subtype F (*Fmut*) with mutations L10I, I15V, G16E, K20R, E35D, M36I, S37N, P39S, R41K, M46I, G51R, I54V, R57K, D60E, Q61D, K70R, I72V, T74A, V82A, L89M, with respect to the B wild-type (*Bwt*) sequence. An additional Q7K mutation was included in all the constructs to increase stability toward autolysis.⁹ The C₂-symmetric inhibitor TL-3^{10,11} was used for co-crystallization and inhibitory studies. The structures of *Fwt* and *Fmut* HIV-PR reported here are the first for any non-B subtype HIV-PRs available to date.

Results and Discussion

Protease sequences

The sequence alignment of the four HIV proteases studied by us is presented in Figure 1. Positions of the primary and secondary mutations with respect to the *Bwt* sequence are highlighted. The subtype F wild-type HIV-1 PR (*Fwt*) naturally carries the mutations I15V, E35D, M36I, S37A, R41K, R57K, D60E, Q61N, I62V, L63S, I64L, and L89M with respect to the subtype B sequence. The mutant of B subtype HIV-PR (*Bmut*) contains eight mutations (S37A, R41K, K45R, I54V, L63P, A71V, V82A, and L90M) and the mutant of F subtype HIV-PR (*Fmut*) includes 20 mutations (L10I, I15V, G16E, K20R, E35D, M36I, S37N, P39S, R41K, M46I, G51R, I54V, R57K, D60E, Q61D, K70R, I72V, T74A, V82A, and L89M). Finally, the point mutation Q7K was introduced in all four enzymes to prevent autolysis without affecting enzymatic activity.^{9,12}

Crystallographic structure solution

All four structures of HIV-PR presented here were solved with the inhibitor TL-3 (Figure 2(a)) bound in the active site, in order to allow for precise structural comparison and to assess the differences between the B and F subtypes of HIV PR. TL-3 is a symmetrical inhibitor originally developed against FIV PR, which proved to be a good universal inhibitor of most retroviral proteases, with high potency against HIV-PR.^{10,11}

The choice of the space group for the hexagonal crystals of HIV-1 PR has been a matter of some controversy in the past, with no unambiguous

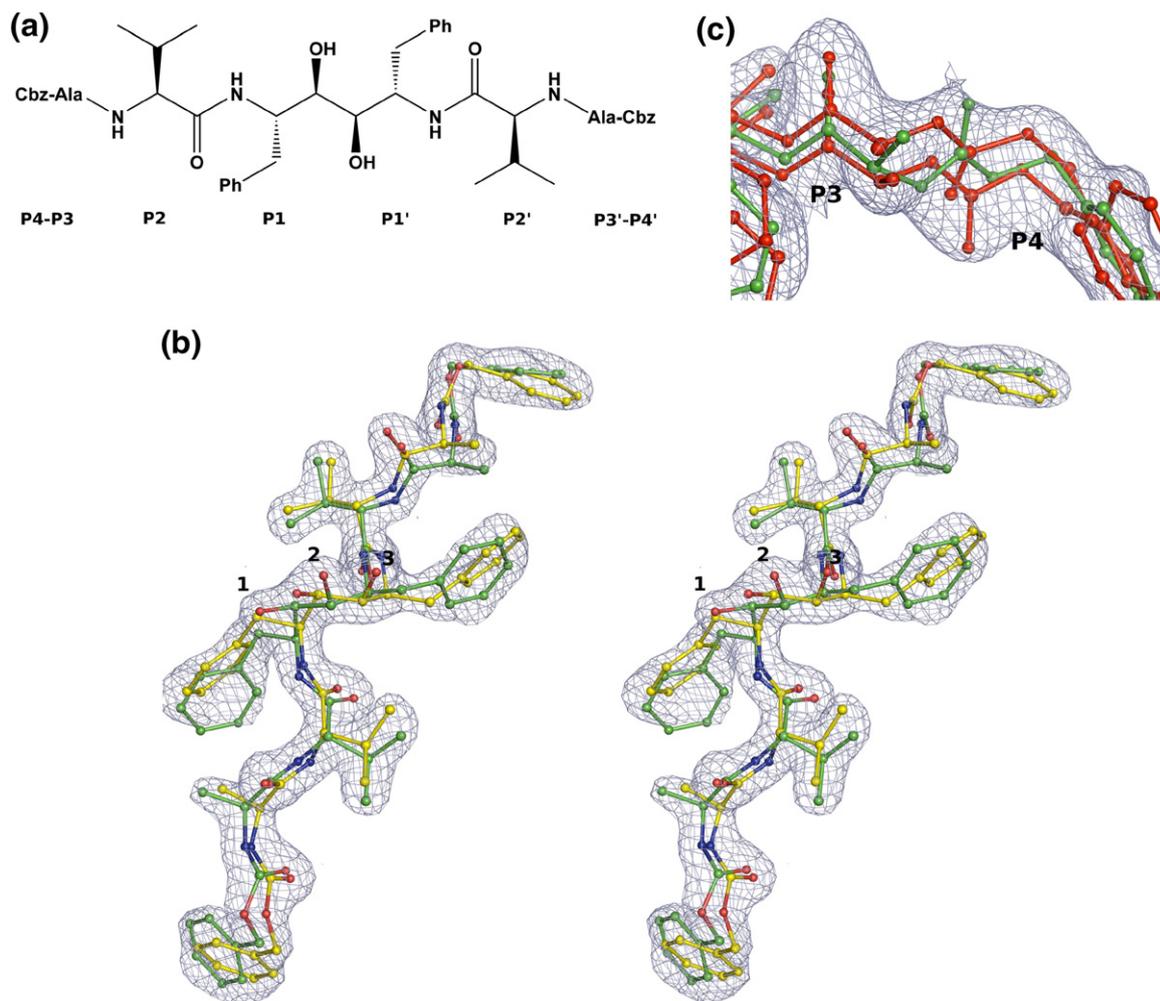


Figure 2. (a) Structural representation of the C_2 -symmetric inhibitor TL-3¹⁰. (b) The final $2F_{\text{obs}}-F_{\text{calc}}$ electron density contoured at 1σ , superposed with the refined model of the TL-3 inhibitor, for the *Bmut* structure. The three lobes of electron density, marked with numbers, are explained by the double conformation of the inhibitor. This central part of the inhibitor, which contains the non-hydrolyzable peptide bond, occupies the active site, with the diols interacting with the catalytic aspartate residues. For clarity, one conformation is colored green and the other one yellow. (c) The final $2F_{\text{obs}}-F_{\text{calc}}$ electron density contoured at 1σ for the P3 and P4 subsites of the inhibitor TL-3 bound to the *Bwt* PR. A model for the TL-3 inhibitor is shown in red (both conformations). The previously published *Bwt* structure in *P6*_{1,22} was superimposed on this model and is shown in green. Apparently, the flip observed in that structure is an artifact.

caused by the refinement of one half of a symmetric inhibitor in a single conformation, not sufficient to model fully its interactions in the HIV-PR active site.

Comparison of the structures

Unless explicitly stated, all the analyses and comparisons performed throughout this article refer to subunit A. As expected, the observed structural differences between the four refined models are very small, with a positional RMSD of 0.43 Å (*Bwt/Bmut*), 0.58 Å (*Bwt/Fwt*), 0.55 Å (*Bwt/Fmut*), 0.44 Å (*Fwt/Bmut*), 0.47 Å (*Fwt/Fmut*), and 0.43 Å (*Bmut/Fmut*) for all 198 C^α atoms of the dimer. The mean temperature factors for each structure are 32.8, 30.1, 36.1 and 47.5 Å², for *Bwt*, *Bmut*, *Fwt* and *Fmut*, respectively. The mean temperature factors for the TL-3 inhibitor in the same

structures are 20.1, 20.7, 24.9 and 21.1 Å², respectively. These low *B* values are typical for well-ordered structures, thus allowing for precise comparison and analysis of these models.

The largest difference among the models is observed for the loop between residues 35 and 41. Superposition of the inhibitors bound to each of the structures shows that they interact with the enzymes in a very similar fashion, with the highest divergence confined to the residue Cbz in the subsite P4/P4', probably due to the lack of interactions between this residue and the S4/S4' pocket of the enzyme.

Differences between HIV-PR of the B and F viral subtypes

None of the polymorphic substitutions that differentiate the wild-type subtype F protease

Table 1. Crystallographic data and statistics for the four HIV-PRs

	Bwt	Bmut	Fwt	Fmut
Cell parameter [$a=b, c$] (Å)	63.13, 83.41	60.93, 82.46	61.43, 80.89	61.25, 82.23
Resolution (Å)	54.72–2.10 (2.21–2.10)	52.78–1.75 (1.84–1.75)	53.23–2.10 (2.21–2.10)	53.07–2.80 (2.95–2.80)
Dataset completeness (%)	99.7 (99.7)	99.6 (99.6)	99.7 (99.7)	99.9 (99.9)
$I/\sigma(I)$	13.6 (2.7)	17.8 (3.4)	13.0 (3.1)	9.8 (2.0)
R_{sym} (%) ^a	7.5 (41.4)	3.7 (34.8)	4.6 (31.8)	8.8 (48.2)
Redundancy	3.7 (3.7)	4.8 (4.8)	2.5 (2.4)	3.2 (3.2)
$R_{\text{cryst}}/R_{\text{free}}$	0.159/0.219	0.183/0.232	0.194/0.265	0.183/0.260
RMSD from ideality				
Bond distance (Å)	0.015	0.015	0.013	0.015
Bond angle (°)	1.57	1.78	1.59	1.69
Number of water molecules in AU	98	189	117	50
RMSD between subunits ^b	0.047	0.058	0.057	0.039
K_i (nM) ^c	3.3±0.9	11±3	24±8	180±29
PDB ID code	2P3B	2P3A	2P3C	2P3D

All structures were solved and refined in the space group $P6_1$. The values in parentheses refer to the highest resolution shell. The inhibition constants (K_i), determined for all four enzymes against TL-3 inhibitor, are also shown.

^a This refers to the R_{merge} value reported by the program SCALA.

^b Root-mean-square deviation between the C^α coordinates of the two equivalent subunits in the asymmetric unit.

^c Inhibition constants (K_i) for the inhibitor TL-3 obtained at 37 °C by measuring the rate of substrate hydrolysis using 5–50 nM protease in 100 mM sodium acetate (pH 4.7), 1 M sodium chloride, 1 mM EDTA, 1 mM DTT, 10% DMSO and 10 μ M substrate plus increasing amounts of inhibitor.

from wild-type subtype B involve residues that create the substrate-binding pockets. In fact, both structures are very similar, with an overall RMSD of only 0.58 Å between the C^α coordinates of equivalent subunits. Nevertheless, significant structural differences are observed in the region between residues 33–42, with a maximum deviation of 2.85 Å between the C^α positions of residue 35 (Figure 3(a)). The residues 33 to 42 are confined to a loop that can be defined as the flap hinge, a region that exhibits extensive rearrangement between the unliganded enzymes and their complexes with inhibitors.¹⁷ Among the differences in the sequence found in this region, only M36I is associated with resistance to inhibitors, and is considered a subtype B secondary mutation (Figure 1) developed along with primary mutations in individuals under intensive treatment, as well as a common substitution in non-B wild-type HIV proteases. Our hypothesis is that the mutation M36I impairs the movement of the hinges of the flaps. The flaps (residues 33–62) extend over the substrate-binding cleft and must be flexible to allow entry and exit of the polypeptide substrates and products, as well as inhibitors.

Careful analysis of the flap hinges of both *Bwt* and *Fwt* PRs shows that the loop containing residue 36 shifts toward the protein core in the *Fwt* structure and that the end of the side-chain in position 36 tends to occupy equivalent positions in both *Bwt* and *Fwt* in order to maintain contacts with the interacting residues (Figure 3(a)). This specific side-chain position and conformation of the residue 36 (Met or Ile) is maintained by van der Waals (VDW) interactions with the side-chains of neighboring residues (Leu33, Leu38, and Lys20). In other words, mutation of a long methionine residue to a shorter isoleucine causes a collapse of the loop,

resulting in retracing of the main chain of the loop containing this mutation, and in its displacement toward the loop 76–83, which forms the S1/S1' pocket. In fact, in *Bwt* PR structure, the loop 34–40 is involved in 27 VDW interactions and three H-bonds with the loop 76–83, compared to 40 VDW and five H-bonds in the *Fwt* PR structure, implying better stabilization of the catalytic pocket loop 76–83 in the latter enzyme. In total, we found 18 H-bonds between the flap hinge (residues 34–40) and the rest of the protein residues in the *Bwt* structure, compared to 25 in *Fwt*. In agreement with our hypothesis, displacement of the flap hinge loop is not observed in the *Bmut* structure, which does not carry the M36I mutation (Figure 3(b)).

The working hypothesis of the stiffening of the flap hinges for F subtype HIV-PR carrying the M36I mutation is also supported by a comparison of the temperature factors (Figure 4). In *Fwt* PR the magnitude of the temperature factors is consistently higher than in *Bwt* PR, an exception made for the residues 33–45. Within this particular region, *Fwt* PR has lower B-factor values as compared with *Bwt*, indicating lower flexibility of the flap hinge region of the former enzyme. An equivalent behavior can be also clearly seen in the published structure of subtype B multi-drug resistant variant containing the M36I mutation complexed with indinavir and ritonavir.¹⁸ This indicates that the differences observed in the flap hinge region of *Fwt* PR are of a general nature and are not restricted to the binding of the particular inhibitor used in this study.

It has been noted that the non-B subtypes HIV-PRs do not usually develop the primary mutation L90M, not even in the individuals under intensive treatment with HIV-PR inhibitors, whereas the L89M

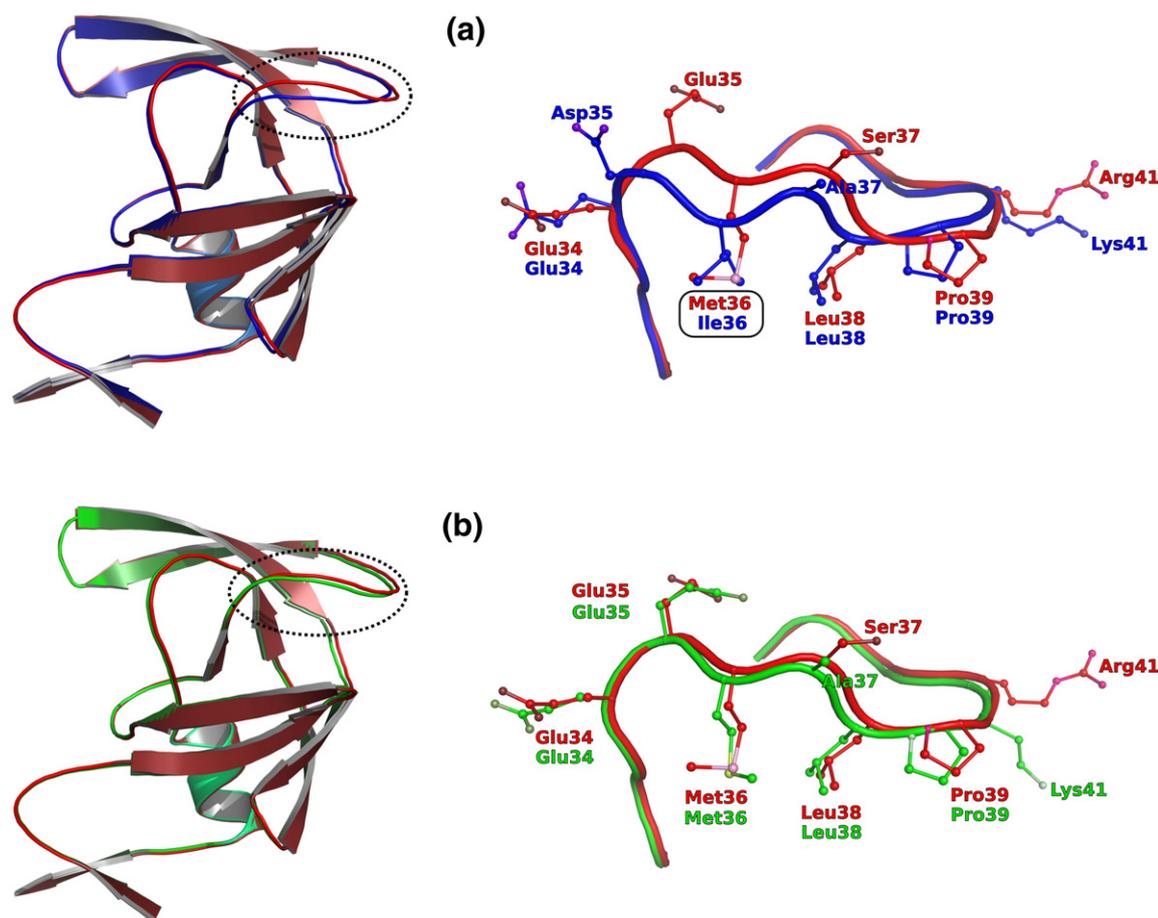


Figure 3. Pairwise comparisons of *Bwt* (red) with *Fwt* (in blue) and with *Bmut* (in green), emphasizing the effects of the M36I mutation. (a) The alignment of *Bwt* and *Fwt* shows a region of high discrepancy (circled) in the otherwise very similar structures, which cannot be seen in the alignment of *Bwt* and *Bmut*. (b) A close-up view of the flap hinges (on the right) reveals that this region contains the substitution M36I in subtype F HIV-PRs. Note that the position occupied by the end of the side-chain of residue 36 is the same in all three structures, forcing the backbone of a few neighboring amino acids in *Fwt* to shift toward the core of the protease, and thus increasing the number of van der Waals contacts between these residues and the rest of the protease. We propose that this displacement may be responsible by the hardening of the flap hinges, which, in turn, leads to higher susceptibility to development of resistance. Apparently the substitution S37A does not play a role in the observed shift, since it appears also in *Bmut*, where the structural difference is not significant.

substitution is very frequent among the non-B subtypes isolated.⁵ The double mutation L89M/L90M is very rare in all subtypes. Here we propose that the L89M substitution mimics the effect of the L90M mutation, which could explain why they are mutually exclusive. Structurally, when the well-characterized mutation L90M occurs, the bulkier side-chain of methionine makes additional van der Waals contacts with the main chain atoms of the active site residues 24 through 26. The increased number of interactions leads to a decrease of the volume of the substrate-binding cavity and compromises the structural flexibility of the S1/S1' pockets, as compared with the wild-type protease. This effect renders resistance to some commercial inhibitors, particularly nelfinavir and saquinavir. This analysis, previously presented for the G48V/L90M double mutant HIV-PR crystallized with saquinavir,¹⁹ can be extended to our *Bmut* structure complexed with TL-3, also containing the L90M mutation (Figure 5(a)).

The structural consequences of the L89M substitution in the PRs that belong to the F subtype can be related to those of the mutation L90M (Figure 5(b)). The longer methionine side-chain at position 89 forces the neighboring Leu90 to shift toward the main chain atoms in the vicinity of the catalytic Asp25, mimicking the effects of the L90M mutation. This displacement increases the number of van der Waals interactions between the side-chain of Leu90 and the catalytic loop (five in *Bwt* to seven in *Fwt*/*Fmut*), thus constraining the S1/S1' binding pockets. The maximum shift observed for the active site residues is 0.3 Å for the carboxyl side-chain of Asp25 in the *Fwt* structure, which is in agreement with the degree of structural rearrangements reported for the mutation L90M.^{19,20}

Apart from M36I and L89M, all other amino acid sequence differences between *Bwt* and *Fwt* HIV-PRs are considered polymorphic. Five of those mutations, D60E, Q61N, I62V, L63S and I64L, are located away from the active site, at the end of a β -strand *b'*

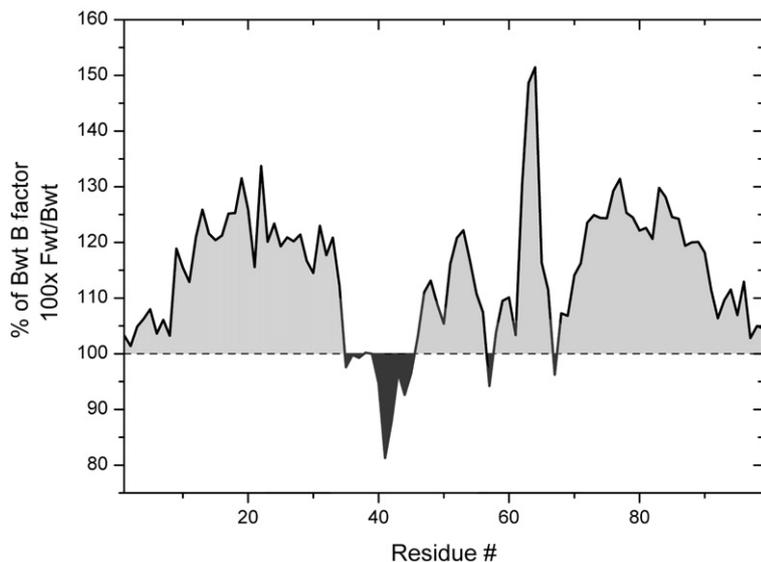


Figure 4. Plot of the *Fwt* mean temperature factors per residue (*Fwt* B factor) relative to the *Bwt* B factor (in %) calculated as $100 \times (\text{Fwt B factor}) / (\text{Bwt B factor})$ against the residue number (chain A). The mean B factors for the *Fwt* structure are consistently higher than those for the *Bwt* (gray shaded) except in a few black shaded areas. Within the region between residues 35 and 45, *Fwt* PR has lower B factor values as compared to *Bwt*, indicating lower flexibility of the flap hinge region of the former enzyme.

(nomenclature according to Wlodawer *et al.*²¹) which forms one side of the flap, while the substitution I15V occurs at the last residue of the β -strand *b*, in direct contact with *b'*. The observed structural differences are responsible for significantly lower inhibition of *Fwt* by TL-3 as compared to *Bwt* (Table 1).

Subtype B and F mutants

Both *Bmut* and *Fmut* contain the primary mutation V82A in the S1/S1' pocket. The effect of this mutation on the arrangement of the TL-3 inhibitor is similar to those previously reported²² for the symmetrical inhibitor A-77003. Structurally, only one of the two hydroxyl groups of the central diol of TL-3 can occupy the middle position between the two catalytic aspartate residues in the active site, forcing inhibitor binding to the protease active site in an asymmetric manner. Such asymmetry forces the remaining hydroxyl to be positioned inside the S1' pocket, as shown in Figure 6. In the wild-type structures, the side-chain of Phe in the P1 position occupies the central part of the pocket and interacts with the side-chain of Val82 through van der Waals interactions. In this arrangement, the P1' Phe ring shifts away from Val82 due to the asymmetric binding of the inhibitor and occupies a crevice between the flap and the loop containing the residue Pro81B (Figure 6). This displacement increases the number of interactions between the P1' Phe side-chain and the S1' pocket when compared to P1/S1 (25 interactions on P1'/S1' against 14 interactions on P1/S1 for the *Bwt* structure). In the structures of *Bmut* and *Fmut* that contain the mutation V82A, the P1' Phe side-chain rotates towards the void created by the mutation. This rearrangement of the inhibitor's side-chain is caused by a collapse of the flap residues and Pro81B, which seals the crevice in which it was inserted in the wild-type structures. All these rearrangements preserve the number of van der Waals interactions observed between P1 and S1 when compared to the *Bwt*

structure. At the S1' side, however, the rotation of the P1' Phe side-chain decreases the number of interactions between P1' and S1' (25/20 VDW interactions in *Bwt/Bmut* and 39/19 VDW interactions in *Fwt/Fmut*), leading to a loosely bound inhibitor and reduced binding affinity. Similar behavior was also observed for the symmetric inhibitor BMS-182193 in an HIV-PR containing the V82A mutation.²³ This indicates that the asymmetry between P1 and P1' is caused by structural collapse of the S1' pocket, and by rearrangements of the inhibitor's side-chain. It is also interesting to note that the rotation of the TL-3 P1 side-chain is one of the largest observed for this mutation. While for A-77003 this difference is so small that it can be barely noticed,²² for TL-3 it amounts to a rotation of 53° between the *Bwt/Bmut* structures and 87° between the *Fwt/Fmut* structures.

The effects of the primary mutation L90M in *Bmut* are equivalent to those described previously,¹⁹ where the bulkier methionine residue shifts the active site loop containing the residues 24–26 towards the S1/S1' pockets, thus decreasing the enzyme flexibility, and ultimately leading to reduced activity.

For two of the secondary mutations in the flap region, M46I (*Fmut* only) and I54V (both *Bmut* and *Fmut*), no significant structural differences could be detected between the native and mutant enzymes. This is consistent with the hypothesis that these two mutations affect inhibitor binding dynamically by altering the flexibility of the flaps, but not the final closed conformation of the enzyme.¹⁸

Consequences of polymorphic substitutions in non-B subtypes

It has been proposed that the presence of secondary mutations in a virus infecting a drug-naïve individual could increase the risk of therapeutic failure.²⁴ It was shown that patients with PR mutation M36I were at

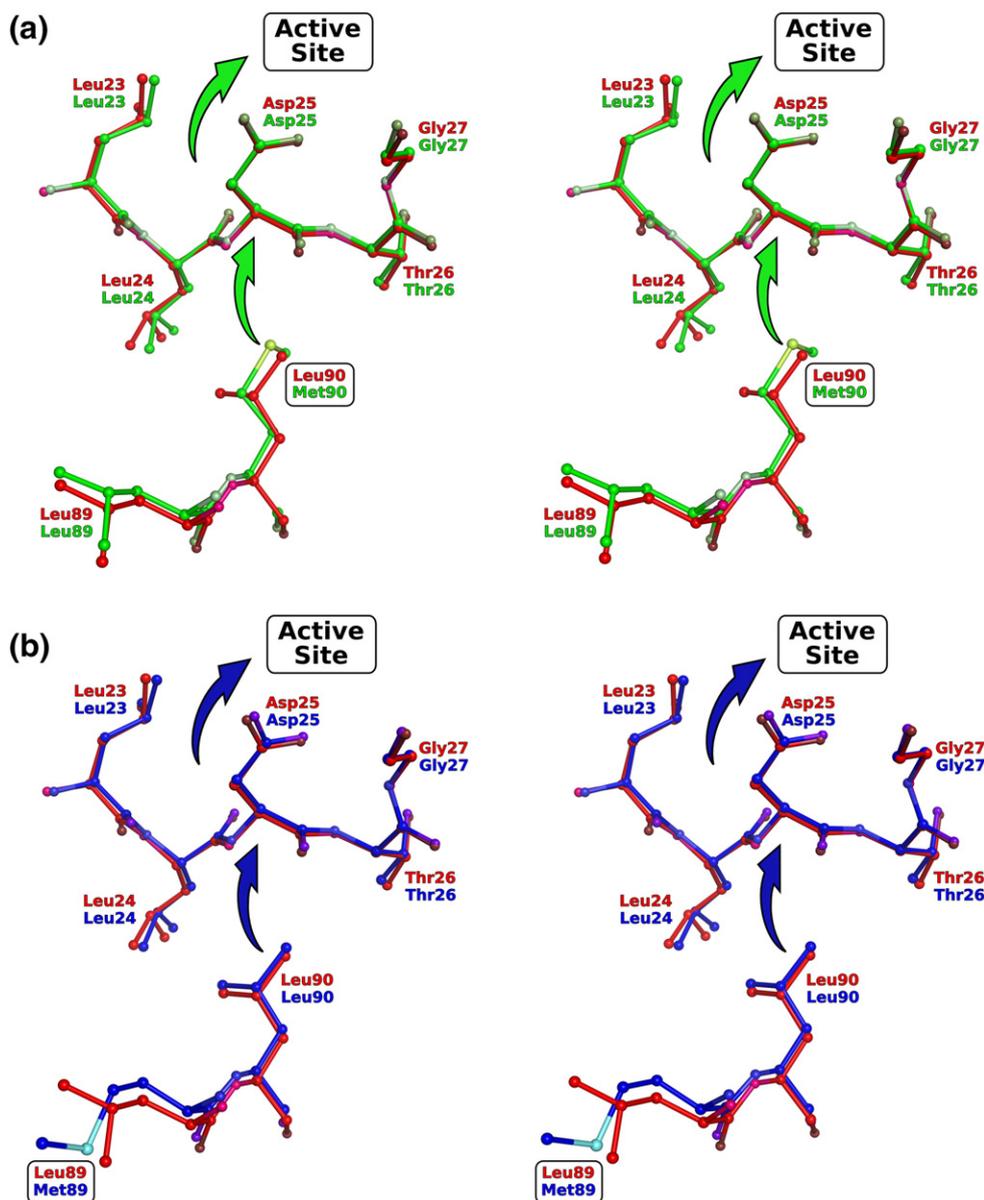


Figure 5. The structures of *Bmut* and *Fwt* PRs aligned on the subtype B wild-type PR (*Bwt*). *Bwt* is colored red, *Bmut* green, and *Fwt* blue. (a) *Bwt* compared to *Bmut*. (b) *Bwt* compared to *Fwt*. It appears that the polymorphic substitution L89M in the subtype F PRs, here represented by the *Fwt* PR, causes the same effect as the mutation L90M induced by anti-retroviral treatment in the subtype B proteases, as shown by *Bmut* PR. Both mutations displace the residue 90 toward the active site loop containing the residue Asp25, thus increasing the amount of van der Waals interactions between residue 90 and the active site loop. This shift slightly compresses the substrate/inhibitor binding cavity, leading to a reduction in volume and flexibility of the active site.

significantly higher risk of having virologic failure at week 24 when compared to patients carrying native protease.²⁴ According to the samples found in the Stanford HIV Database,²⁵ 98% of subtype A, 80% of subtype C, 62% of subtype D, 95% of subtype F, and 100% of subtype G naïve individuals carry a methionine rather than isoleucine at position 36 of the protease, against only 13% of subtype B, which leads to the conclusion that non-B HIV-PR subtypes are more susceptible to develop anti-retroviral resistance. If we consider that the substitution L89M is, at least in part, a structural equivalent of the primary mutation L90M, then this susceptibility is even more

acute, since 99% of subtype A, 85% of subtype C, 69% of subtype F and 98% of subtype G naïve persons carry this substitution as a consequence of natural polymorphism (subtypes B and D proteases show only 1% and 3% of the L89M substitution, respectively). It appears, on the basis of these data, that the subtype B HIV-PR is an exception among the other subtypes.

The polymorphic substitutions L63P and A71V present in *Bmut*, and L10I, K20R and T74A present in *Fmut* occur in persons receiving PIs therapy and, together with other mutations, are associated with resistance to each of the PIs.²⁵ Finally, the proteins

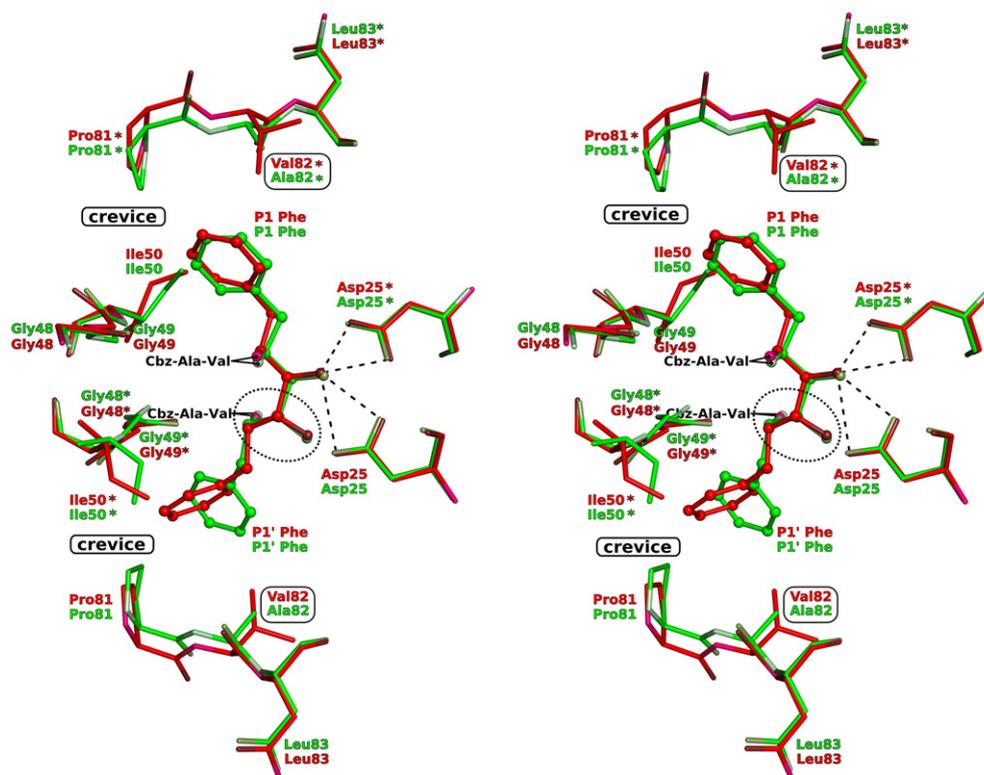


Figure 6. Effects of the mutation V82A on the inhibitor-binding mode. The inhibitor subsite P1/P1' is shown in ball-and-stick representation along with the pocket S1/S1' in stick representation, for the structures *Bwt* (red) and *Bmut* (green). The broken lines represent hydrogen bonds between the active site aspartate residues and the oxygen of the diol in the center of the inhibitor, whereas the asterisks mark residues that belong to chain B. Due to the asymmetric mode of binding of the inhibitor, the S1' pocket, occupied by the P1' phenyl side-chain, accommodates one hydroxyl of the central diol, which is indicated by a dotted circle. Whereas there is no significant modification in the P1 subsite due to V82A mutation (marked with a square), in the P1' a rotation of the Phe ring of the inhibitor is observed. This rotation decreases the number of interactions between the inhibitor's side-chain and the S1' pocket, while the interactions are maintained in the S1 pocket.

discussed here have some other substitutions not found in the Stanford HIV Database: K45R (*Bmut*), S37A (*Bmut* and *Fwt*), R41K (*Bmut*, *Fwt* and *Fmut*), Q61N/L63S/I64L (*Fwt*), I15V/E35D/R57K (*Fwt* and *Fmut*) and G16E/S37N/P39S/G51R/Q61D/K70R/I72V (*Fmut*). These findings call for care in the choice of inhibitors being used for treatment of patients infected with non-B subtypes of HIV.

Inhibitory assays

The inhibition constants (K_i) obtained for the different protease subtypes against the inhibitor TL-3 are summarized in Table 1. The subtype F wild-type protease, as well as the B and F mutant proteases, present different degrees of resistance against TL-3. When compared with *Bwt* PR, the K_i values are about threefold higher for *Bmut* PR, about sevenfold higher for *Fwt* PR, and about 54-fold higher for *Fmut* PR. The presence of the primary and the accessory mutations in *Fwt*, *Fmut* and *Bmut* PRs, alone or together, as well as the described structural differences due to them, sheds light on the structural basis of the decreased affinity of the TL-3 inhibitor to PRs encoded by subtypes B and F of HIV-1. In addition, the presence of the primary mutation V82A

and the accessory flap mutation I54V in *Bmut* and *Fmut* reduces the susceptibility to all PIs, notably including indinavir, ritonavir, and lopinavir.²⁵ If our hypothesis that the substitution L89M is structurally equivalent to the primary mutation L90M is correct, this would lead to intermediate-to-high-level resistance of subtype F HIV-PR to all the commercial protease inhibitors. The L90M mutation is most likely elicited due to administration of nelfinavir and saquinavir in the therapy regimen.²⁶ In order to confirm these points, the affinities of commercial inhibitors to the B and F subtype HIV-PRs described here are currently under investigation in our laboratory. Preliminary results obtained to date clearly demonstrate higher resistance of *Fwt*, *Fmut*, and *Bmut* PRs to the currently used commercial inhibitors (our unpublished data).

Conclusions

Mutations in the substrate cleft cause drug resistance by reducing the binding affinity between the inhibitor and the mutant protease enzyme.²⁷ Mutations elsewhere in the enzyme either compensate for the decreased kinetics of the enzymes with active site

mutations or also cause resistance by altering enzyme catalysis, dimer stability, inhibitor binding kinetics, or by re-shaping the active site through long-range structural perturbations.²⁸ The presence of primary and secondary drug-resistant mutations alone is not sufficient to explain the different degrees of resistance to TL-3 inhibition of each HIV-PR discussed here. In order to understand these differences, the influence of the amino acid polymorphism on the binding of the TL-3 inhibitor must be carefully considered. The results presented here show that the polymorphic substitutions, which need not be in the active site, may amplify the effects of drug-resistant mutations. Therefore, the combined effects of naturally existing polymorphisms and drug-resistant mutations might have important consequences on the viability of current HIV-1 protease inhibitors with pharmaceutical value.

Materials and Methods

Cloning and expression

The cDNA of the whole virus genome from subtypes B mutant and both wild-type and mutant subtype F were obtained from HIV-1 vertically infected seropositive Brazilian children.²⁶ The patients infected with *Bmut* and *Fmut* viruses received treatment consisting of PR inhibitors ritonavir/nelfinavir and reverse transcriptase inhibitors amivudine/stavudine/zidovudine.

The protease coding sequence was amplified by PCR, simultaneously inserting the point mutation Q7K,⁹ and the insert was ligated in the pET11a vector that had been cleaved with the restriction enzymes NdeI and BamHI. *Escherichia coli* BL21(DE3)-RIL (Stratagene) competent cells were transformed with each of the obtained clones and the cells were cultivated at 37 °C in LB medium supplemented with 35 µg/ml chloramphenicol, 100 µg/ml ampicillin, 1% (w/v) glucose and 0.05% antifoam 204 (Sigma). The culture was induced with 1 mM IPTG when the A_{600nm} reached 0.7. After 4 h of induction the cells were harvested by centrifugation.

Inclusion bodies purification and refolding

The protocol for the production of the refolded Brazilian HIV-PRs has been reported.¹⁵ Briefly, the washed inclusion bodies, dissolved in urea-containing buffer, were purified by two successive ion-exchange columns (Q Sepharose and SP Sepharose, respectively). The purified HIV-PR was refolded by dialysis against a urea-free buffer, and concentrated for use in enzymatic assays and crystallization trials.

Crystallization

The protease-inhibitor complexes were prepared by mixing the protease (at ~0.4 mM) with twofold molar excess of TL-3 inhibitor (at 22 mM in DMF/DMSO 1:1, v/v) for 1 h on ice. X-ray diffraction quality crystals were obtained as described.¹⁵ *Bmut* PR crystals appeared after two days and continued to grow for about two weeks (one crystal per drop), in two conditions (0.1 M sodium cacodylate (pH 6.2), 0.32 M ammonium sulfate, 6% (v/v) MPD, 5.1 % (w/v) PEG 3350 (condition 1), and 0.1 M

sodium cacodylate (pH 6.2), 0.32 M ammonium sulfate, 6% MPD (condition 2)). Both *Fwt* and *Fmut* PRs were crystallized using ammonium sulfate as precipitant. A decrease in the crystallization temperature from 291 K to 277 K was used to convert a shower of small rods to just a few small crystals per drop. The crystallization plate was then transferred back to 291 K to allow the crystals to reach their maximum size. It appears that the synergistic effect of a co-precipitant²⁹ is a necessary condition to crystallize the *Bmut* PR, given the fact that no crystals have been obtained using ammonium sulfate as the only precipitant.

Data collection and crystal structure solution

X-ray diffraction data for flash-cooled crystals of *Bmut* (condition 2), *Fwt*, and *Fmut* were collected at the MX1 beamline of the Brazilian Synchrotron Light Laboratory (LNLS)^{30,31} using radiation with 1.43 Å wavelength, and a MarCCD detector. Crystals were frozen in the original crystallization solution containing 20% (v/v) glycerol. Initial diffraction of all protease crystals was significantly improved by successive rounds of annealing, during which the frozen crystals were transferred to the cryo-resolution at room temperature and then were flash-cooled again. Measurements of data for the crystal of *Bwt* PR (Mar345 image plate detector mounted on a Rigaku RU-200 rotating anode) have been reported.¹¹

The new datasets were integrated using MOSFLM 6.2.3³² and scaled with SCALA 3.1.20.³³ The structure of *Bmut* protease was solved by molecular replacement using the structure of wild-type subtype B HIV-PR refined at 1.09 Å (PDB code: 1kzk) as the search model.³⁴ The structure was initially determined with the program AmoRe³⁵ in space group $P6_122$ (although identification of the space group as either $P6_122$ or $P6_1$ was possible, as discussed above). The structures of the other proteases were solved in space groups $P6_122$ and in $P6_1$ using the program PHASER 1.2,³⁶ with the *Bmut* subunit serving as the search model. Positional and temperature factor refinement were performed with CNS 1.1³⁷ and REFMAC 5.1.24.³⁸ Water molecules were automatically inserted with ARP/wARP.³⁹ Real space refinement was carried out by visual inspection of the model using the program O.⁴⁰ TLS parameters were refined in the penultimate cycle of positional refinement, without application of non-crystallographic symmetry. Three TLS groups were defined for each subunit. The last cycle was performed re-establishing the non-crystallographic symmetry, with fixed TLS parameters. When appropriate, double side-chain conformations were modeled. The electron density of the final models for all structures with the exception of *Fmut* clearly indicated chemical modification of the Cys67 amino acid residue. An additional electron density could be satisfactorily modeled as a whole β -mercaptoethanol molecule covalently attached to the cysteine residue (in one subunit of *Bwt* structure and in both subunits of *Bmut* and *Fwt* models). β -Mercaptoethanol was used throughout the whole process from cell lysis to refolding buffers. Three acetate anions, probably carried out from the cation exchange chromatography buffer, were identified on the surface of the *Fwt* structure, one in subunit A and two in subunit B (none of them in the vicinity of the inhibitor molecule).

Enzymatic and inhibitory assays

The catalytic activities of the HIV-1 PRs, as well as the inhibition constants (K_i) for the inhibitor TL-3, were monitored by following the hydrolysis of the fluorogenic

substrate Arg-Glu(EDANS)-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln-Lys(DABCYL)-Arg (Molecular Probes). The proteases were added to a 300 μ l cuvette containing substrate at 37 °C. Final concentrations in the standard assay were: 5–50 nM protease, 10 μ M substrate, 100 mM sodium acetate (pH 4.7), 1 M sodium chloride, 1 mM EDTA, 1 mM DTT, 10% (v/v) DMSO, plus increasing amounts of inhibitor. The fluorescence was monitored by using PC1 Photon Counting Spectrofluorometer (ISS) with excitation and emission wavelengths of 340 nm and 490 nm, respectively.

Different concentrations of EDANS in assay buffer containing the substrate were used for calibration, converting the fluorescence data sampled at regular time intervals into concentrations (nM). The reaction rates ($\text{nM} \times \text{min}^{-1}$) were obtained from the slope of the product concentration progress curves. In cuvettes, the initial rate of wild-type or mutant PR was adjusted to 10–20 $\text{nM} \times \text{min}^{-1}$ by addition of the proper amount of enzyme. The reaction was followed for 8–10 min afterwards.

For each inhibitor concentration, the steady-state rate in the presence of inhibitor (V_i) was divided by the initial rate in the absence of inhibitor (V_0). The resulting ratios V_i/V_0 and the corresponding inhibitor concentrations were used to determine the inhibition constant (K_i) of the inhibitor and the total concentration of active enzyme (Et) by fitting the data to the general equation for competitive tight-binding inhibitors^{41,42}:

$$V_i/V_0 = \frac{([Et] - [It] - K_{i,app}) + \{([Et] - [It] - K_{i,app})^2 + 4[Et]K_{i,app}\}^{1/2}}{2[Et]}$$

where V_i is steady-state rate in the presence of the inhibitor, measured as Δ [product] min^{-1} ; V_0 is initial rate in the absence of the inhibitor, measured as Δ [product] min^{-1} ; [Et] is total concentration of the enzyme (free and bound); [It] is total concentration of the inhibitor (free and bound); $K_{i,app}$ is apparent dissociation constant of the enzyme-inhibitor complex at a given substrate concentration.

Since under our experimental conditions the substrate concentration [S] is lower than the Michaelis-Menten constant (K_m), thus $K_{i,app}$ and K_i are similar because competition of the inhibitor with the substrate is negligible. $K_{i,app}$ corresponds to the IC_{50} , the inhibitor concentration resulting in 50% inhibition of the enzymatic activity under defined experimental conditions.

Protein Data Bank accession codes

The coordinates and structure factors have been submitted to the RCSB Protein Data Bank. The accession codes are 2P3A, 2P3B, 2P3C and 2P3D for the *Bmut*, *Bwt*, *Fwt*, and *Fmut* PRs, respectively.

Acknowledgements

This work was supported in part by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil, *via* grants 99/03387-4, 04/11890-8, 04/12201-1 and 06/00182-8; by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil; and by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2007.03.049](https://doi.org/10.1016/j.jmb.2007.03.049)

References

1. Vaishnav, Y. N. & Wong-Staal, F. (1991). The biochemistry of AIDS. *Annu. Rev. Biochem.* **60**, 577–630.
2. Kuiken, C., Foley, B., Freed, E., Hahn, B., Marx, P., McCutchan, F., *et al.*, (eds). (2002). *HIV sequence Compendium 2002*. Published by Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, LA-UR n 03-3564.
3. Osmanov, S., Pattou, C., Walker, N., Schwardlander, B. & Esparza, J. (2002). Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J. Acquir. Immune Defic. Syndr.* **29**, 184–190.
4. Kantor, R. & Katzenstein, D. (2003). Drug resistance in non-subtype B HIV-1. *J. Clin. Virol.* **29**, 152–159.
5. Caride, E., Hertogs, K., Larder, B., Dehertogh, P., Brindeiro, R., Machado, E. *et al.* (2001). Genotypic and phenotypic evidence of different drug resistance mutation patterns between B and non-B subtype isolates of human immunodeficiency virus type 1 found in Brazilian patients failing HAART. *Virus Genes*, **23**, 193–202.
6. Wlodawer, A. (2002). Rational approach to AIDS drug design through structural biology. *Annu. Rev. Med.* **53**, 595–614.
7. Ohtaka, H. & Freire, E. (2005). Adaptive inhibitors of the HIV-1 protease. *Prog. Biophys. Mol. Biol.* **88**, 193–208.
8. Shafer, R. W. (2002). Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clin. Microbiol. Rev.* **15**, 247–277.
9. Rosé, J. R., Salto, R. & Craik, C. S. (1993). Regulation of autoproteolysis of the HIV-1 and HIV-2 proteases with engineered amino acid substitution. *J. Biol. Chem.* **268**, 11939–11945.
10. Lee, T., Le, V.-D., Lim, D., Lin, Y.-C., Morris, G. M., Wong, A. L. *et al.* (1999). Development of a new type of protease inhibitors, efficacious against FIV and HIV variants. *J. Am. Chem. Soc.* **121**, 1145–1155.
11. Li, M., Morris, G. M., Lee, T., Laco, G. S., Wong, C. H., Olson, J. *et al.* (2000). Structural studies of FIV and HIV-1 proteases complexed with an efficient inhibitor of FIV protease. *Proteins: Struct. Funct. Genet.* **38**, 29–40.
12. Mildner, A. M., Rothrock, D. J., Leone, J. W., Bannow, C. A., Lull, J. M., Reardon, I. M. *et al.* (1994). The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties. *Biochemistry*, **33**, 9405–9413.
13. Murthy, K. H., Winborne, E. L., Minnich, M. D., Culp, J. S. & Debouck, C. (1992). The crystal structures at 2.2- Å resolution of hydroxyethylene-based inhibitors bound to human immunodeficiency virus type 1 protease show that the inhibitors are present in two distinct orientations. *J. Biol. Chem.* **267**, 22770–22778.
14. Wlodawer, A. & Erickson, J. W. (1993). Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* **62**, 543–585.
15. Sanches, M., Martins, N. H., Calazans, A., Brindeiro, R. M., Tanuri, A., Antunes, O. A. C. & Polikarpov, I.

- (2004). Crystallization of a non-B and a B mutant HIV-Protease. *Acta Crystallog. sect. D*, **60**, 1625–1627.
16. Kervinen, J., Lubkowski, J., Zdanov, A., Bhatt, B., Dunn, B. M., Hui, K. Y. *et al.* (1998). Toward a universal inhibitor of retroviral proteases: comparative analysis of the interactions of LP-130 complexed with proteases from HIV-1, FIV, and EIAV. *Protein Sci.* **7**, 2314–2323.
 17. Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L. *et al.* (1989). Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science*, **246**, 1149–1152.
 18. Clemente, J. C., Moose, R. E., Hemrajani, R., Whitford, L. R., Govindasamy, L., Reutzel, R. *et al.* (2004). Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. *Biochemistry*, **43**, 12141–12151.
 19. Hong, L., Zhang, X. C., Hartsuck, J. A. & Tang, J. (2000). Crystal structure of an *in vivo* HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance. *Protein Sci.* **9**, 1898–1904.
 20. Kovalevsky, A. Y., Tie, Y., Liu, F., Boross, P. I., Wang, Y.-F., Leshchenko, S. *et al.* (2006). Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. *J. Med. Chem.* **49**, 1379–1387.
 21. Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T. *et al.* (1989). Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science*, **245**, 616–621.
 22. Baldwin, E. T., Bhat, T. N., Liu, B., Pattabiraman, N. & Erickson, J. W. (1995). Structural basis of drug resistance for the V82A mutant of HIV-1 proteinase. *Nature Struct. Biol.* **2**, 244–249.
 23. Kervinen, J., Thanki, N., Zdanov, A., Tino, J., Barrish, J., Lin, P. F. *et al.* (1996). Structural analysis of the native and drug-resistant HIV-1 proteinases complexed with an aminodiol inhibitor. *Protein Pept. Letters*, **3**, 399–406.
 24. Perno, C. A., Cozzi-Lepri, A., Balotta, C., Forbici, F., Violin, M., Bertoli, A. *et al.* (2001). Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naïve patients treated with protease inhibitor based therapy. *J. Infect. Dis.* **184**, 983–991.
 25. Rhee, S.-Y., Gonzales, M. J., Kantor, R., Betts, B. J., Ravela, J. & Shafer, R. W. (2003). Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucl. Acids Res.* **31**, 298–303.
 26. Brindeiro, P. A., Brindeiro, R. M., Mortensen, C., Hertogs, K., Vroey, V. D., Rubini, N. P. M. *et al.* (2002). Testing genotypic and phenotypic resistance in Human Immunodeficiency Virus type 1 isolates of clade B and other clades from children failing antiretroviral therapy. *J. Clin. Microbiol.* **40**, 4512–4519.
 27. Dunn, B. M., Goodenow, M. M., Gustchina, A. & Wlodawer, A. (2002). Retroviral proteases. *Genome Biol.* **3**, 3006.1–3006.7.
 28. Muzammil, S., Ross, P. & Freire, E. (2003). A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. *Biochemistry*, **42**, 631–638.
 29. Majeed, S., Ofek, G., Belachew, A., Huang, C. C., Zhou, T. & Kwong, P. D. (2003). Enhancing protein crystallization through precipitant synergy. *Structure*, **11**, 1061–1070.
 30. Polikarpov, I., Teplyakov, A. & Oliva, G. (1997). The ultimate wavelength for protein crystallography? *Acta Crystallog. sect. D*, **53**, 734–737.
 31. Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. & Craievich, A. (1998). Set-up and experimental parameters of the protein crystallography beamline at the Brazilian National Synchrotron Laboratory. *J. Synchrotron Rad.* **5**, 72–76.
 32. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, **26**.
 33. Evans, P. R. (1997). In *Proceedings of the CCP4 Study Weekend. Recent Advances in Phasing* (Wilson, K. S., Davies, G., Ashton, A. W. & Bailey, S., eds), pp. 97–102, Daresbury Laboratory, Warrington.
 34. Reiling, K. K., Endres, N. F., Dauber, D. S., Craik, C. S. & Stroud, R. M. (2002). Anisotropic dynamics of the JE-2147-HIV Protease complex: drug resistance and thermodynamic binding mode examined in a 1.09 Å structure. *Biochemistry*, **41**, 4582–4594.
 35. Navaza, J. (1994). AMORE - An automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.
 36. Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallog. sect. D*, **60**, 432–438.
 37. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
 38. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
 39. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463.
 40. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. (1990). O: a macromolecular environment. In *Crystallographic and Modeling Methods in Molecular Design* (Bugg, C. & Ealick, S., eds), pp. 189–195, Springer-Verlag Press, Berlin.
 41. Morrison, J. F. (1969). Kinetics of the reversible inhibition of enzyme catalysed reactions of tight-binding inhibitors. *Biochem. Biophys. Acta*, **185**, 269–285.
 42. Bieth, J. G. (1995). Theoretical and practical aspects of proteinase inhibition kinetics. *Methods Enzymol.* **248**, 59–84.

Edited by R. Huber

(Received 14 December 2006; received in revised form 14 March 2007; accepted 20 March 2007)
Available online 24 March 2007