Protein J (2011) 30:173–183
DOI 10.1007/s10930-011-9316-2

Nine Crystal Structures Determine the Substrate Envelope of the MDR HIV-1 Protease

Zhigang Liu · Yong Wang · Joseph Brunzelle · Iulia A. Kovari · Ladislau C. Kovari

Published online: 12 March 2011

Abstract Under drug selection pressure, emerging mutations render HIV-1 protease drug resistant, leading to the therapy failure in anti-HIV treatment. It is known that nine substrate cleavage site peptides bind to wild type (WT) HIV-1 protease in a conserved pattern. However, how the multidrug-resistant (MDR) HIV-1 protease binds to the substrate cleavage site peptides is yet to be determined. MDR769 HIV-1 protease (resistant mutations at residues 10, 36, 46, 54, 62, 63, 71, 82, 84, and 90) was selected for present study to understand the binding to its natural substrates. MDR769 HIV-1 protease was co-crystallized with nine substrate cleavage site hepta-peptides. Crystallographic studies show that MDR769 HIV-1 protease has an expanded substrate envelope with wide open flaps. Furthermore, ligand binding energy calculations indicate weaker binding in MDR769 HIV-1 protease-substrate complexes. These results help in designing the next generation of HIV-1 protease inhibitors by targeting the MDR HIV-1 protease.

Keywords HIV-1 protease · Multi-drug resistant · Substrate envelope · Crystallography · Drug design

Z. Liu · Y. Wang · I. A. Kovari · L. C. Kovari (✉)
Department of Biochemistry and Molecular Biology, Wayne State University, School of Medicine, 540 E. Canfield Avenue, 4263 Scott Hall, Detroit, MI 48201, USA
e-mail: kovari@med.wayne.edu

J. Brunzelle
Life Sciences Collaborative Access Team, Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Feinberg School of Medicine Chicago, Chicago, IL 60611, USA

Abbreviations

MDR	Multi-drug resistance
HIV	Human immunodeficiency virus
WT	Wild type
HAART	Highly active antiretroviral therapy
PR	Protease
MA	Matrix
CA	Capsid
NC	Nucleocapsid
RT	Reverse transcriptase
IN	Integrase
RH	RNase H
RMSD	Root mean square deviation

1 Introduction

HIV has been a major infectious pathogen since it was discovered in the early 1980s'. In 2009 alone, an estimated 2.6 million new HIV infections occurred worldwide, while about 33.3 million people were living with HIV. In 2009, the global AIDS deaths climbed to 1.8 million (UNAIDS report on the global AIDS epidemic 2010, http://www.unaids.org/globalreport/Global_report.htm).

Highly active antiretroviral therapy (HAART) makes possible the inhibition of in vivo HIV replication to delay the onset of AIDS symptoms [3, 4, 26, 27], while a cure or vaccine of HIV infection is yet unavailable [23, 35]. The usual HAART regimen combines three or more drugs from different classes, such as two nucleoside reverse transcriptase inhibitors (NRTI) and one non-nucleoside reverse transcriptase inhibitor (NNRTI) or protease inhibitor (PI). However, the emerging drug resistance, due to the lack

of proofreading ability of the HIV-1 reverse transcriptase [32, 36, 42], requires the development of new and potent drugs to overcome the resistance problem [7, 8, 11, 12].

HIV-1 protease (HIV-1 PR) is an aspartic protease that cleaves newly synthesized HIV-1 polyproteins at nine major cleavage sites to generate the mature protein components of an infectious HIV-1 virion. The HIV-1 *gag* gene codes for the structural proteins: matrix protein (MA), capsid protein (CA), and nucleocapsid protein (NC), while the *gag-pol* gene encodes both structural proteins (MA, CA, and NC) and enzymes: protease (PR), reverse transcriptase (RT), and integrase (IN). Without effective HIV-1 PR, HIV-1 virion remains uninfectious, hence making HIV-1 protease inhibitors the most potent anti-AIDS drugs and essential therapeutic components of HAART [4, 12].

Crystallographic studies of WT HIV-1 PR complexed with the natural cleavage site peptides uncover a consensus of binding pattern of the peptides in the protease active site cavity [29, 31]. Based on the wealth of structural information of WT HIV-1 protease substrate complexes, new HIV-1 protease inhibitors mimicking the natural substrates are designed. In addition, it is proposed that inhibitors fitting the substrate envelope are less susceptible to drug resistance [1, 5, 6, 25].

Nonetheless, there is no crystallographic information on the MDR HIV-1 protease substrate complexes. The questions remaining to be addressed are: Do the nine substrates bind to MDR HIV-1 protease active site with a similar conformation? What is the relative binding energy of MDR HIV-1 protease substrate complexes compared to WT HIV-1 protease substrate complexes? What does the substrate envelope look like in the MDR HIV-1 protease?

We chose MDR769 HIV-1 protease as our study model to investigate the MDR HIV-1 protease substrate complex structures. The inactive enzyme (D25N) was chosen to trap the uncleaved peptide in the active site cavity. The high resolution crystal structure of MDR769 HIV-1 protease was solved previously by our group, and it showed an expanded active site cavity with mutations at positions 10, 36, 46, 54, 62, 63, 71, 82, 84, 90 [22]. Nine hepta-peptides corresponding to natural cleavage sites P3 to P4′ were chemically synthesized and successfully co-crystallized with the inactive MDR769 HIV-1 protease. Based on the crystal structures refined to 1.6–2.0 Å resolutions, a multi-drug resistant substrate envelop was determined and the envelope was comparable to that of the WT HIV-1 protease despite a certain degree of variation. The MDR HIV-1 protease substrate envelope was expended relative to WT HIV-1 protease substrate envelope. Furthermore, substrate binding energy calculations showed reduced binding affinity of MDR769 HIV-1 protease to the substrates.

Our observations with the MDR769 HIV-1 protease substrate complexes were consistent with the previous conclusions that the active site cavity was expanded in MDR769 HIV-1 protease. These results suggest that a larger HIV-1 protease inhibitor is needed to inhibit MDR variant of HIV-1.

2 Materials and Methods

2.1 Preparation of Substrate Peptides

The nine substrate hepta-peptides shown in Table 1 were purchase from SynBioSci Corporation, Livermore, CA. All the hepta-peptides were purified by HPLC to purity higher than 98%. Peptide powder was dissolved in DMSO to prepare stock solution of 20 mM concentration and the samples were stored at −20 °C.

2.2 Protein Purification and Co-Crystallization

The MDR769 HIV-1 protease A82T was overexpressed by using a T7 promoter expression vector in conjunction with the *E. coli* host, BL21 (DE3). In brief, a fresh *E. coli* transformant of BL21(DE3) with the MDR769 plasmid was cultured in 5 mL LB medium containing 100 μg/mL ampicillin for 7 h, which was used to inoculate a 50 mL LB medium containing 100 μg/mL ampicillin and the culture was grown overnight. Later, two liters of LB medium with 100 μg/mL ampicillin was inoculated with 50 mL seed culture and bacterial growth was monitored by absorbance at A600. After 4–6 h of incubation with shaking at 37 °C, the A600 reached 0.5 and the cells were harvested at 10,000×*g* for 10 min at 4 °C using a Sorvall RC5B Plus centrifuge [40]. The MDR protease was isolated from inclusion bodies by using a series of buffered washes, followed by denaturation in 6 M urea. For

Table 1 Sequences of the nine sites within the HIV-1 Gag and Pol polyproteins that are cleaved by HIV-1 protease

Cleavage site	P3	P2	P1	P1′	P2′	P3′	P4′
MA/CA	Gln	Asn	Tyr	Pro	Ile	Val	Gln
CA/p2	Arg	Val	Leu	Phe	Glu	Ala	Met
p2/NC	Thr	Ile	Met	Met	Gln	Arg	Gly
NC/p1	Gln	Ala	Asn	Phe	Leu	Gly	Lys
p1/p6	Gly	Asn	Phe	Leu	Gln	Ser	Arg
TF/PR	Phe	Asn	Phe	Pro	Gln	Ile	Thr
PR/RT	Leu	Asn	Phe	Pro	Ile	Ser	Pro
RT/RH	Glu	Thr	Phe	Tyr	Val	Asp	Gly
RH/IN	Lys	Val	Leu	Phe	Leu	Asp	Gly

The cleavage sites are named by the proteins released after the site is cleaved

MA matrix, *CA* capsid, *NC* nucleocapsid, *TF* trans frame peptide, *PR* protease, *RT* reverse transcriptase, *RH* RNAse H, *IN* integrase

purification of the unfolded protease, an anion exchange column with pH 8.6 (Q Sepharose, Amersham Biosciences) was used that allowed the protease to pass through, while the *E. coli* contaminants to remained bound to the column. The protease purity was monitored by Coomassie blue-stained SDS–PAGE [16] and the MDR769 HIV-1 protease was purified to greater than 95% homogeneity. The 6 M urea was removed to refold the protease by using a series of dialysis exchanges that were carried out at 4 °C. The first buffer consisted of 0.2 M sodium phosphate monobasic, 0.2 M sodium phosphate dibasic, 1.0 M urea, 0.2% βME, and 10% glycerol. The next step consisted of buffer exchange with 0.2 M sodium phosphate monobasic, 0.2 M sodium phosphate dibasic, 0.2% βME, and 10% glycerol. The last two buffer exchanges were with 10 mM sodium acetate, pH 5.0, 1 mM DTT and 10% glycerol. The protease was concentrated by Amicon to 2 or 3 mg/mL and stored in aliquots at −80 °C.

The substrate hepta-peptides were mixed with 2.5 mg/mL MDR 769 HIV-1 protease and diluted to 0.4 mM final concentration. The mixtures were incubated at 4 °C for 2 h, which were subject to centrifugation to remove the insoluble impurities. The hanging drop vapor diffusion method was used to form the bipyramidal crystals of the MDR 769 protease. Using a grid screen consisting of sodium chloride (0.7–1.4 M) and MES-HEPES buffer (pH 5.5–8.1), the HIV-1 protease substrate complex crystals formed overnight at 22 °C. Routinely, 0.2 mm crystals, in the longest dimension, were obtained after 14 days of incubation. In each well, there were two droplets, containing 1 µL of protease substrate mixture, 1 µL of reservoir solution and 2 µL of protease substrate mixture, 1 µL of reservoir solution, respectively, 07 mL of well solution.

2.3 Data Collection and Crystallographic Refinement

Protease crystals were dipped in 30% glucose for cryo-protection and flash frozen in liquid nitrogen. The diffraction data were collected at 1.00 Å wavelength at the Advanced Photon Source (APS) (LS-CAT 21), Argonne National Laboratory (Argonne, IL). Data were reduced with the program CrystalClear (CrystalClear: An Integrated Program for the Collection and Processing of Area Detector Data, Rigaku Corporation©, 1997–2002). The data statistics are shown in Table 2. In all cases the crystals belong to the same space group $P4_1$. Molecular replacement was performed with Molrep-autoMR in CCP4 with the model previously solved in our lab [20].

Initial refinements were performed without substrate hepta-peptides using Refmac5 [24, 38]. The nine hepta-peptides were built into the difference electron density maps as the refinement proceeded in the COOT program [10]. The electron density maps contained two orientations of bound peptide to the MDR HIV-1 protease. The Gag-Pol peptide binding in two orientations was also reported earlier by Prabu-Jeyabalan et al. [31] for the wild type HIV-1 protease. The MDR HIV-1 protease complexes were modeled and refined with the hepta-peptides in two orientations. The two peptide orientations were related by a pseudo-2-fold symmetry and there was minimal asymmetry for the two orientations of the peptide as was the case earlier for the wild type HIV-1 protease peptide complexes. To carry out the substrate envelope analysis in an analogous way as reported for the HIV-1 wild type complexes, we deleted one hepta-peptide orientation from the model. In all nine crystal structures, the N-terminus of the hepta-peptide was oriented close to monomer A of the AB protease dimer to allow for systematic analysis of the substrate envelope.

Crystallographic waters were added with the ARP/wARP program [17]. The structures were refined in Refmac5 to 1.6–2.0 Å resolution. The final stereo-chemical parameters were checked using PROCHECK [39]. Images were generated in PyMol.

2.4 Analysis

PISA Server was used to calculate the ligand binding energy in both MDR and WT HIV-1 protease substrate complexes [15]. Dissociation constants were calculated based on the formula $\Delta Gint = RTlnKd$, where the ΔG_{int} is the intrinsic binding energy as defined by PISA. Hydrogen bonds with the bound hepta-peptides were analyzed with PISA server. Conserved hydrogen bonds within protease were identified by with the program Ligplot [41]. Temperature factor analysis was performed with the CCP4 program Temperature Factor Analysis. Protease substrate complexes were superimposed based on protease residues 1–99 Cα and the RMSD of Cα was analyzed with the CCP4 program Superpose Molecules [14]. Solvent accessible area was calculated with the CCP4 program Accessible Surface Areas. All structures were visualized using PyMol.

3 Results

3.1 Determination of Crystal Structures of MDR769 HIV-1 Protease-Substrate Complexes and the Calculation of Ligand Binding Energy

We report the crystal structures of nine MDR769 HIV-1 protease substrate complexes. The results support our working hypothesis that the MDR769 HIV-1 protease with wide open flaps binds weaker (20–162,000 times) to the substrates. In addition, the expanded MDR HIV-1 protease substrate envelope requires larger inhibitor to overcome drug resistance.

Table 2 Crystallographic statistics of the nine protease-substrate complexes

Space group	MA/CA $P4_1$	CA/p2 $P4_1$	p2/NC $P4_1$	NC/P1 $P4_1$	p1/p6 $P4_1$	TF/PR $P4_1$	PR/RT $P4_1$	RT/RH $P4_1$	RH/IN $P4_1$
Unit cell dimension (Å)									
a	45.77	45.21	45.77	45.62	45.50	45.54	45.62	45.58	45.02
b	45.77	45.21	45.77	45.62	45.50	45.54	45.62	45.58	45.02
c	102.08	103.4	102.15	102.00	102.10	102.07	102.10	102.37	104.48
Resolution range	41.77–1.7	41.43–1.8	45.4–2.0	27.26–1.4	45.51–1.7	45.6–1.6	45.6–1.65	34.04–1.75	45.0–1.8
Unique reflections[a]	22,986	19,176	14,209	40,904	22,802	27,395	22,519	21,011	19,157
R_{merge} (%) overall (final shell)	0.049[b] (0.140)[c]	0.043 (0.201)	0.097 (0.355)	0.085 (0.458)	0.098 (0.389)	0.088 (0.461)	0.044 (0.228)	0.071 (0.407)	0.055 (0.16)
I/sigma (final shell)	18.9 (8.2)	17.8 (6.2)	9.3 (4.6)	9.3 (3.1)	8.3 (3.2)	8.7 (2.8)	20.9 (6.3)	11.0 (3.5)	19.4 (8.8)
Completeness (%) overall (final shell)	99.7 (100)	99.8 (100)	100 (100)	100 (100)	100 (100)	100 (100)	98.3 (97.2)	100 (100)	99.7 (100)
R_{work} (%)[d]	19.2	19.2	19.6	19.2	18.7	19.5	19.3	18.5	19.1
R_{free} (%)[e]	21.8	24.4	26.2	23.6	22.6	23.2	22.6	23.6	23.8
RMSD									
Bond length (Å)	0.013	0.016	0.015	0.01	0.015	0.011	0.013	0.014	0.015
Bond angle (°)	1.461	1.587	1.476	1.319	1.575	1.31	1.439	1.481	1.599
Average B-factors (Å^2)									
Main chain	15.167	26.121	26.073	16.885	21.097	16.687	18.869	23.85	20.436
Side chain	21.118	31.925	29.653	23.181	27.994	22.471	24.534	30.27	25.302
Peptide[f]	36.687	40.769	47.301	37.760	40.576	36.821	35.549	45.478	33.050
Solvent	30.363	41.809	35.832	32.832	38.521	31.28	33.924	40.326	33.088
No. of waters	285	271	193	299	294	261	247	293	239
Ramachandran plot									
Favorable (%)	90.6	94.4	92.5	91.2	90.7	92.5	93.8	93.8	91.3
Additional (%)	9.4	5.0	6.8	8.1	9.3	7.5	6.2	5.6	8.1
Generous (%)	0	0.6	0.6	0.6	0	0	0	0.6	0.6
Forbidden (%)	0	0	0	0	0	0	0	0	0
PDB code	3OTS	3OUD	3OUC	3OUB	3OUA	3OU4	3OU3	3OTY	3OU1

[a] N_{obs}/N_{unique}

[b] $R_{merge} = \Sigma hkl \Sigma i |I(hkl) - \{I(hkl)|/\Sigma hkl \Sigma i I(hkl)i$, where $I(hkl)_i$ is the intensity of an observation and I(hkl) is the mean value for its unique reflection. Summations cover all reflections

[c] The values after the hyphens indicate the highest resolution shell

[d] $R_{work} = \Sigma hkl|F(hkl)o - \{F(hkl/\Sigma hkl F(hkl)o$

[e] R_{free} was calculated same way as R working, but with the reflections excluded from refinement. The R free set was chosen using default parameters in Refmac5

[f] The peptides were fitted into electron density map with occupancy 0.5

The inactive MDR769 HIV-1 protease is co-crystallized with nine hepta-peptides representing the nine natural substrate cleavage site sequences in the Gag and Gag-Pol polyproteins. The sequences and names of the hepta-peptides are shown in Table 1. The hepta-peptides are named based on proteins released after cleavage: matrix (MA), capsid (CA), nucleocapsid (NC), trans-frame peptide (TF), protease (PR), reverse transcriptase (RT), RNase H (RH), and integrase (IN).

The structures of the complexes are solved with crystallographic statistics shown in Table 2. These complexes crystallize in space group $P4_1$, in which each crystallographic asymmetric unit contains one biologically relevant HIV-1 protease dimer complexed with ligands. The structures are refined to 1.6–2.0 Å.

The ligand binding energy is calculated for all the nine MDR769 HIV-1 protease substrate complexes by using the PISA server [15], and is compared to that of WT HIV-1 protease substrate complexes. Based on the calculated binding energy from crystallographic structure, dissociation constant (K_d) of each substrate cleavage site peptides against MDR and WT HIV-1 proteases is calculated and shown in Table 3. The contribution of hydrogen bonds to the binding free energy and K_d is not considered since there is a high degree of variation in the strength of the various hydrogen bonds based on the distance and angle of the hydrogen bonds.

3.2 The Substrates Bind Weaker to the MDR Protease Relative to the WT Protease, Despite Similar Ligand Conformation as That in WT Protease

All the nine hepta-peptides are bound to the MDR HIV-1 protease active site cavity in an extended conformation. The structure of the MDR HIV-1 protease MA/CA peptide complex, shown in Fig. 1, is representative for all the nine complexes. The hepta-peptides establish extensive contacts with both protease monomers; however, the contacts between monomers are not equally distributed. The asymmetric nature of hepta-peptides breaks the symmetry of protease dimer in complexes by forming more contacts with one monomer than the other. Structural analysis shows that in complexes MA/CA, NC/p1, p1/p6, and TF/PR, RMSD of Cα at residue 79 is significantly different between the two monomers (Liu et al., data not shown). In RH/IN complex the significant difference is present at residues 67 and 68 rather than at residue 79. However, there is no big RMSD variation between monomers in complexes CA/p2, PR/RT and RT/RH (Liu et al., data not shown).

Although the active site cavity of MDR 769 HIV-1 protease is expanded leaving enough space for an alternate conformation, all the nine hepta-peptides are located at the bottom of the active site cavity with flexible amino and carboxy termini. P1 and P1′ reside between Asn 25/25′. Seven out of nine hepta-peptides (except for p1/p6 and RT/RH), point their carbonyl oxygen of P1 (P1 O) in between Asn25/25′, bridging the two "catalytic site" residues. In the p1/p6 complex, the P1 O forms a hydrogen bond with only Asn25 ND2 atom. In contrast, in complex RT/RH, a water molecule bridges the connection between P1 O and Asn25′ ND2.

The substrates bind weaker to the MDR HIV-1 protease relative to the WT HIV-1 protease. Based on the ligand binding energy (ΔG) calculated by PISA server, the binding energy is reduced dramatically (34–78%) in MDR HIV-1 protease substrate complexes compared to that in WT HIV-1 protease substrate complexes (Table 3). Correspondingly, K_d of MDR HIV-1 protease substrate complexes is much higher (20–126,000 folds), which indicates lower binding affinity of the substrate hepta-peptides (Table 3).

3.3 Conserved Substrate Conformation in MDR HIV-1 Protease Substrate Complexes Gives Rise to an Unevenly Expanded Substrate Envelope Compared to That of WT HIV-1 Protease Substrate

Substrate envelope is defined as the overall space occupied by the nine substrate peptides in the superimposed nine complexes based on C_α of the protease dimers. As a result, this substrate envelope demonstrates the maximal flexibility of substrates inside of the activity site cavity of the MDR HIV-1 protease.

The substrate peptides are taken out from the MDR HIV-1 protease active site cavity upon dimer superposition and listed individually in Fig. 2. Moreover, these substrate peptides are shown in superposition mode based on the protease dimer Cα to display the spatial overlay in the MDR HIV-1 protease active site cavity (Fig. 3a). However, when compared to those in WT HIV-1 protease, these cleavage site peptides are more flexible in the active site cavity according to the B factor calculation (data not shown). The flexibility may be partially due to the expanded MDR HIV-1 protease active site cavity. Detailed investigation shows that the enlargement of the substrate envelope is not uniform through the different sites. For example, P1/P1′ is expanded by 12%, while the overall expansion of the substrate envelope is 7%. The comparison of the WT and MDR HIV-1 protease substrate envelope is shown in Fig. 3b–d.

3.4 Altered MDR Protease Binding Pockets Result in Weaker Binding of the Substrates

Detailed investigation of the interaction between protease and nine hepta-peptides reveals altered binding pockets

Table 3 Induced fit, ligand binding energy and K_d calculation, and ligand interaction in both MDR and WT HIV-1 protease substrate peptide complexes

	$MDR769_{A82T}$							Wild type						
	Flap distance (Å)	Dimer interface ($Å^2$)	ΔG dimer	ΔG ligand	K_d ligand	H bonds	Buried surface area	Flap distance	Dimer interface ($Å^2$)	ΔG dimer	ΔG ligand	K_d ligand	H bonds	Buried surface area
Apoprotease	12.2	1,508.5	−17.1	N/A	N/A	N/A	N/A	4.1(3phv)	1,487.7	−18.4	N/A	N/A	N/A	N/A
			11.3											
MA/CA	10.1	1,473.9	−17.1	−2.0	3.4×10^{-2}	10	640	6.0(1KJ4)	1,731.0	−23.9	−8.3	8.2×10^{-7}	12	870
			6.3							6.3				
CA/p2	11.7	1,484.0	−17.1	−2.0	3.4×10^{-2}	13	708	5.9(1F7A)	1,768.0	−23.1	−9.1	2.1×10^{-7}	12	860
			6.6							6.4				
p2/NC	10.4	1,485.5	−17.2	−5.2	1.5×10^{-4}	8	766	6.0(1KJ7)	1,768.8	−23.1	−9.1	2.1×10^{-7}	12	860
			6.4							6.4				
NC/p1	9.9	1,478.9	−16.8	−3.8	1.6×10^{-3}	8	699	6.1(1TSU)	1,724.7	−21.7	−6.7	1.2×10^{-5}	8	778 (6)
			6.3							5.4				
p1/p6	10.0	1,471.1	−16.8	−3.5	2.7×10^{-3}	12	672	6.0(1KJF)	1,680.3	−23.0	−5.3	1.3×10^{-4}	15	856
			6.4							6.4				
TF/PR	9.9	1,480.2	−16.9	−3.0	6.3×10^{-3}	13	719	6.3(2AOJ)	1,752.6	−20.6	−8.8	3.5×10^{-7}	17	919 (redu)
			6.4							11.1				
PR/RT	10.0	1,482.1	−16.9	−3.4	3.2×10^{-3}	9	692	N/A	N/A	N/A	N/A	N/A	N/A	N/A
			6.2											
RT/RH	10.2	1,469.7	−16.7	−2.2	2.4×10^{-2}	11	672	6.0(1KJG)	1,746.0	−21.4	−6.4	2.0×10^{-5}	14	855
			6.4							6.3				
RH/IN	12.3	1,480.7	−17.1	−6.5	1.7×10^{-5}	9	690	6.0(1KJH)	1,748.3	−22.0	−9.8	6.4×10^{-8}	10	912
			6.3							6.5				

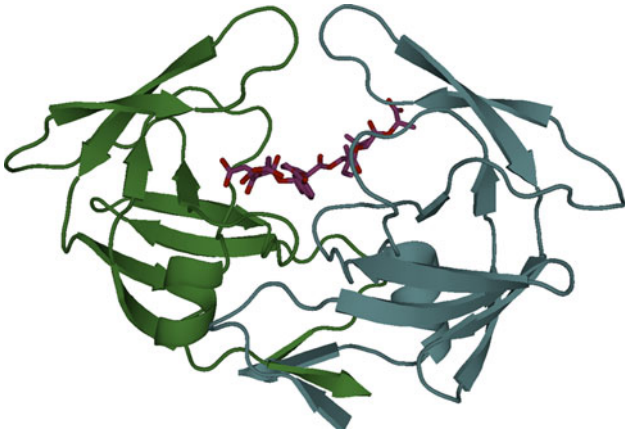

Fig. 1 Structure of the MDR 769 HIV-1 protease MA/CA complex. The two protease monomers are shown in ribbon rendering. The MA/CA substrate peptide is shown by stick model. The substrate shows an extended conformation of the MA/CA peptide binding to the MDR HIV-1 protease dimer, which is representative to all the nine complex structures

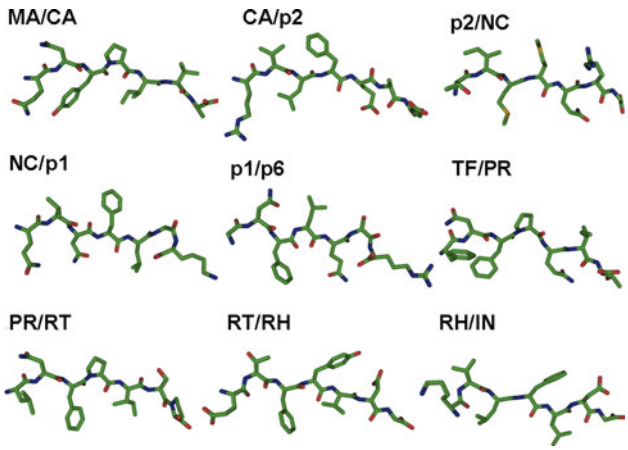

Fig. 2 Orientation of nine substrate peptides in the MDR 769 HIV-1 protease active site cavity. The nine MDR HIV-1 protease substrate peptide complexes are superposed according to the C_α of the protease dimer structure. The substrate peptides are displayed individually after the dimer superposition, showing conserved conformation of these peptides in MDR HIV-1 protease active site cavity

compared to WT HIV-1 protease. Five conserved binding pockets are found in MDR HIV-1 protease, namely S1/S3, S2, S1′/S3′, S2′, and S4′ as shown in Fig. 4. Fewer residues relative to WT HIV-1 protease contribute to the pockets S1/S3, S2, S1′/S3′, S2′, meanwhile a conserved S4′ pocket, missing in WT HIV-1 protease, is found in the MDR HIV-1 protease. The reduced number of amino acid residues lining the binding pocket lead to weaker binding between the active site cavity and the nine substrates. First, the S1/S3 binding pocket involves Gly27′, Asp 29′, Arg8, Thr82, and Asn25 in at least four complexes. Leu23 is involved in NC/p1, MA/CA, and RT/IN complexes, whereas Val84 is involved in NC/p1, RT/RH, and CA/p2 complexes. Side chains of Gly27′, Asp29′, Arg8, Asp25, and Val84 adopt the same conformation in all complexes; whereas side chains of Thr82 and Leu23 have variable conformations. Second, in the S1′/S3′ pocket, Leu23′, Asn25′, Val84′, Arg8, Gly27, and Asp29 are present in at least four complexes; Thr82′ participates in only three complexes p2/NC, RT/TH, and NC/p1. Surprisingly, Gly48 in the flap region is involved in the S1′/S3′ binding pocket in complexes RT/RH, TF/PR, and MA/CA. Except for Thr82′, all residues in S1′/S3′ binding pocket are conserved in the same conformation. Third, side chains from Ala28′, Asp29′ and Asp30′ constitute the binding pocket of the S2 site. Fourth, side chains from Gly27, Ala28, Asp29, and Asp30 constitute the binding pocket at the S2′ site. All residues involved in S2/S2′ binding are highly conserved in conformation, forming a rigid binding pocket. The S2 binding pocket shows fewer contacts relative to that in the S2′ pocket. Finally, the S4′ binding pocket contains two segments: one is flexible and the other is rigid. The flexible segment consists of residues Lys45, Leu46, Ile47 and Gly48. The rigid segment is comprised of Asp29 and Asp30.

3.5 Limited Induced Fit of the MDR Protease Results in Weaker Binding of the Substrates

There is a limited induced fit in the MDR HIV-1 protease when compared to the WT HIV-1 protease substrate complexes, which may be a key reason to the weaker ligand binding. First, flap distance, measured from C_α Ile 50 to C_α Ile 50′, is a good indicator of binding. Analysis of the flap distance shows variation among the nine complexes, ranging from 9.9 to 12.3 Å. When compared to flap distances in WT HIV-1 protease substrate complexes, which range from 5.9 to 6.3 Å [29–31, 37], the MDR HIV-1 substrate complexes are still wide open, i.e. 4–6 Å wider compared to the WT crystal structures. This partial induced fit may contribute to the weaker biding due to fewer contacts between the HIV-1 protease flap area and the substrates. However, when compared to the flap distance in uncomplexed MDR 769 HIV-1 protease (12.2 Å) [22], most complexes show significant reduction in flap distance, except for two complexes: CA/p2 and RH/IN, 11.7 and 12.3 Å, respectively (Table 3). Second, the dimer interface area is also reduced in the MDR HIV-1 protease substrate complexes (1,469.7–1,485.5 $Å^2$) relative to that of WT complexes (1,680.3–1,768.8 $Å^2$) [29–31, 37]. This induced fit in the MDR HIV-1 protease weakens the protease dimer, which may lead to the relative instability of the MDR HIV-1 protease substrate complexes. Surprisingly, the binding of substrate to MDR HIV-1 protease reduces the dimer interface (1,508.5 $Å^2$ in uncomplexed MDR HIV-1 protease), whereas the substrate binding increases the dimer interface area in the WT HIV-1 protease (1,487.7 $Å^2$ in uncomplexed

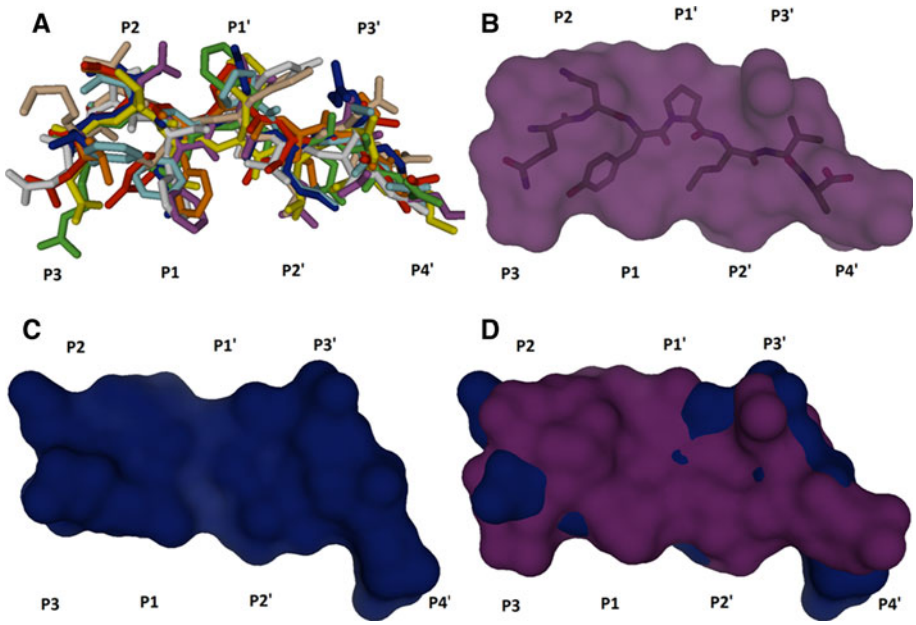

Fig. 3 **Panel A** Superposition of the nine substrate peptides in the MDR HIV-1 protease active site cavity. The superposition shows not only the conserved conformation, but also the conserved position of these substrate peptides in the MDR HIV-1 protease active site cavity. **Panel B** The substrate envelope of the MDR HIV-1 protease. The transparent shading represents the MDR HIV-1 protease substrate envelope. The stick model inside the substrate envelope is the MA/CA substrate peptide which is representative of all the substrate peptides studied here. **Panel C** Substrate envelope of the WT HIV-1 protease. To facilitate direct comparison of the WT and MDR HIV-1 protease complexes, the nano-peptides from the WT studies are truncated to P3 to P4′ hepta-peptides used in this study. **Panel D** Superposition of the WT and MDR HIV-1 protease substrate envelope. The protruding substrate envelope indicates the expanded substrate envelope of the MDR HIV-1 protease

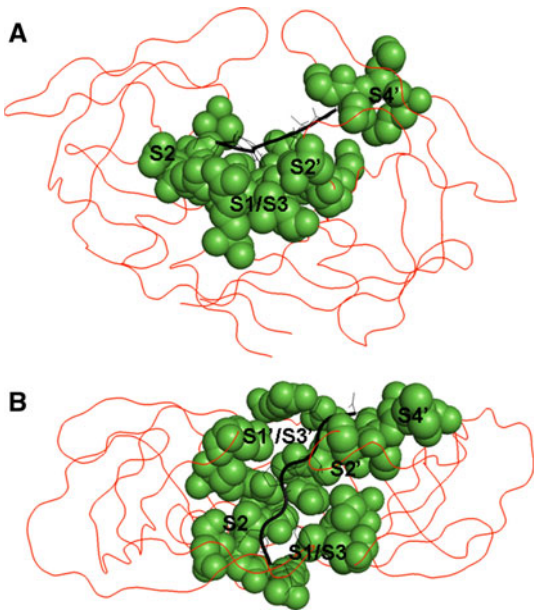

Fig. 4 Substrate binding pocket of the MDR HIV-1 protease. The MDR HIV-1 protease is shown in ribbon rendering. The HIV-1 protease residues involved in peptide binding are shown in space filling model. **Panel A** is the front view of the binding pocket, and **Panel B** is the top view of the binding pocket

WT HIV-1 protease). The factor that ligand binding decreases the stability of MDR HIV-1 protease makes it possible to develop new dimerization inhibitors. Third, one conserved high RMSD pattern is found after the substrates bind to MDR HIV-1 protease. All the nine complexes were superposed with an uncomplexed wide-open MDR769 HIV-1 protease structure, based on C_α. High RMSD regions are located at residues 15 and 16 (loop 2 between β strands 2, 3), residue 21 (β strand 3), residues 40–63 (flap), residues 66–68 (loop between β strands), residues 72 and 73 (loop between β strands), and residues 78–82 (80′ loop interacting with ligands), except for the RH/IN complex which shows high RMSDs at residues 15, 34, 65, and 66. The introduction of these high RMSD regions may contribute to the weaker binding.

3.6 Substrate Conformation Difference Between MDR and WT HIV-1 Protease Complexes

As mentioned before, the overall substrate conformation is conserved between the MDR and WT HIV-1 protease substrate complexes; however, some subtle variation in the substrate conformation makes the MDR HIV-1 protease

substrate complexes less stable (Fig. 5). Details of these substrate conformation differences will be discussed below.

Compared to the side chains, both position and conformation of the substrate backbone are relatively conserved between the WT and MDR HIV-1 protease substrate complexes. The WT and MDR HIV-1 proteases complex with the same substrate peptides are superimposed onto each other according to C_α of the protease dimer. The deviation of corresponding C_α in substrate peptide is shown in Table 4. Except for the substrate peptide in complex RT/IN, the C_α shows more deviation in the C-terminus than in the N-terminus. In addition, the deviation of P1 O is shown in the last column of Table 4 of corresponding substrate peptides from both WT and MDR HIV-1 protease substrate complexes.

In contrast to backbone atoms, the side chain atoms are much more flexible in the MDR complexes. Although no apparent pattern is identified in side chain deviation, it is found that the deviation is larger at both termini.

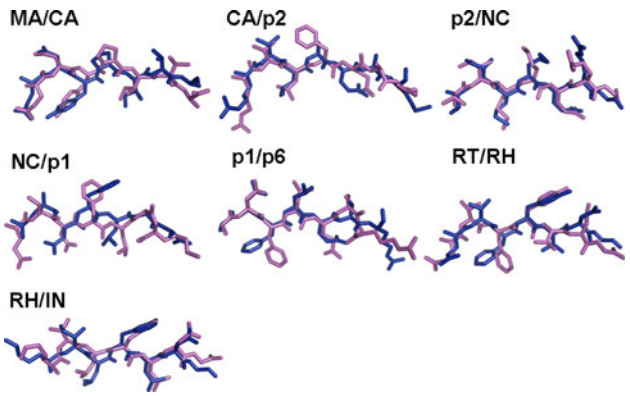

Fig. 5 Overlay of the substrate peptide from the MDR and WT HIV-1 protease substrate complexes based on the C_α superposition of the protease dimer

Table 4 Deviation of substrate peptide Cα and P1 O between MDR and WT HIV-1 protease substrate complexes

	P3	P2	P1	P1′	P2′	P3′	P4′	P1 O
MA/CA	0.9	0.7	1.1	1.0	0.8	1.3	1.9	1.2
CA/p2	0.6	0.7	0.5	1.1	1.2	1.9	1.2	1.4
P2/NC	0.5	0.5	0.6	0.8	0.8	0.7	1.1	1.5
NC/p1	0.9	1.7	0.2	1.8	1.2	0.8	2.2	2.2
P1/p6	0.7	0.4	0.4	0.6	0.4	1.3	1.9	0.8
RT/RH	1.0	1.4	1.1	0.6	0.8	1.4	3.8	3.3
RT/IN	1.0	1.3	1.3	0.7	0.6	0.6	1.1	1.6

Table 5 Distance of P1 O to Asn 25 and Asn 25′ in both MDR and WT HIV-1 protease substrate complexes

	MDR P1O to Asn25/25′		WT P1O to Asn25/25′	
MA/CA	3.2	3.0	4.4	2.7
CA/p2	3.1	2.9	4.4	2.7
P2/NC	3.3	2.6	4.6	2.5
NC/p1	3.3	3.1	5.5	3.2
P1/p6	5.3	2.5	4.8	2.7
RT/RH	5.2	5.0	4.6	2.7
RT/IN	3.2	2.9	4.5	2.7

The relative position of P1 O is totally different from the MDR HIV-1 substrate complexes to the WT HIV-1 protease substrate complexes. The distances between the P1 O atom to Asn25 and Asn25′ in both MDR and WT HIV-1 protease substrate complexes are listed in Table 5. In the MDR HIV-1 protease substrate complexes, the P1 O, Asn25 ND2 and Asn25′ ND2 constitute an isosceles triangle with equal distance between P1 O and Asn25/25′. The only exception to the observation is p1/p6, which is similar to substrates in the WT complexes with unequal distance to Asn25 and Asn25′. The symmetric conformation in the MDR HIV-1 protease substrate complexes may indicate that a symmetric inhibitor might be more suitable to overcome the drug resistance problem.

4 Discussion

4.1 The Hepta-Peptides Bind with Reduced Binding Energy and Increased K_d Values

Although the nine hepta-peptides adopt a similar conformation to bind to MDR HIV-1 protease, the binding energy reveals varied affinities of the hepta-peptides to the protease (Table 3). The binding energy is calculated by using the PISA server [15]. Based on the semi-empirical calculation used in PISA server, the RH/IN complex holds the strongest binding energy between the ligand and protease, although it has the longest inter-flap distance. When compared to the binding energy derived from WT HIV-1 protease substrate complexes, the binding energy for MDR HIV-1 protease substrate complexes is by 34–78%. The K_d value is calculated for both MDR and WT HIV-1 protease substrate complexes, with 20–162,000 fold increase in MDR HIV-1 protease substrate complexes. Among all MDR HIV-1 protease substrate complexes, CA/p2 has the largest decrease in binding energy relative to WT HIV-1 protease CA/p2 complex (22%) and a 162,000 times larger K_d (Table 3).

4.2 The MDR HIV-1 Protease Substrate Envelope Is Enlarged by 12% at P1/P1′ Site

Hepta-peptides bound to MDR769 HIV-1 protease yield a larger substrate envelope relative to that in the WT HIV-1 protease substrate complexes. The volume of substrate envelope determines the size of HIV-1 protease inhibitor required to overcome the drug resistance issue. The substrate envelope is enlarged overall by 7% in MDR HIV-1 protease substrate complexes. The superposition of the MDR with WT HIV-1 protease substrate envelope is shown in Fig. 3d. Further examination shows that the enlargement of MDR HIV-1 protease substrate envelope is not uniform at all binding sites. The MDR HIV-1 protease substrate envelope at P1/P1′ site is enlarged by 12% compared to the P1/P1′ site in the WT HIV-1 protease substrate envelope, while together the P1/P1′ and P2/P2′ increase in surface area by 9.54%. This result suggests that a larger P1/P1′ site is needed to overcome the multi-drug resistance.

4.3 Compensatory Mutations May Restore Enzymatic Activity

With the reduced protease catalytic activity restraining the HIV-1 maturation and replication, HIV-1 virion develops compensatory mutations in and outside the substrate cleavage sites (i.e. mutations at p1/p6 and NC/p1) to restore the protease catalytic activity [2, 9, 30, 33, 43]. Work is currently in progress with the p1/p6and NC/p1 mutant substrate complexes to understand their structural and functional impact on MDR769 HIV-1 protease. New structural information of these two protease-substrate complexes may reveal additional information about the MDR HIV-1 protease substrate envelope and replication of the MDR HIV-1 viron.

4.4 Mutations in CA/p2 Lead to Drug Resistance to the Maturation Inhibitor PA457, Which Is Represented in the Complexes

PA457 (Bevirimat) targeting the CA/p2 cleavage site, blocking the cleavage of CA/p2 by HIV-1 protease, is currently under phase II clinical trials to address the drug resistance problem [13, 18, 19, 21, 34]. It is found that the introduction of single residue mutation to the flanking sequence of CA/p2 cleavage site results in a drug resistant phenotype, which requires no mutations from other parts of the virus, including protease [28, 44, 46]. Among these single mutations, Ala to Phe at P1′ develops drug resistance to PA457 [18, 45]. In our complex structure of CA/p2 with MDR769 HIV-1 protease, at the P1′ position, there is a Phe instead of Ala, and consequently molecular dynamics calculation may elucidate the mechanism of drug resistance to PA457.

4.5 The MDR Protease Forms a Weaker Dimer Compared to the WT Protease Dimer Raising the Possibility of Designing Inhibitors That Disrupt the Dimer Interface

Unlike the WT protease which substrate binding increase dimerization interfaces, MDR HIV-1 protease has a reduced dimerization interface after substrate binding. The dimerization interface area is 16% less in MDR HIV-1 protease substrate complexes than that of WT HIV-1 protease substrate complexes, while the dimerization interface area is almost identical between apo MDR HIV-1 protease and WT HIV-1 protease (Table 3). This partial disrupted dimerization interface in MDR HIV-1 protease substrate complexes raises the possibility of designing inhibitors that disrupt the dimer interface during substrate binding.

4.6 The Catalytic Activity of MDR769 HIV-1 Protease May be Reduced

The catalytic activity may be reduced in MDR 769 HIV-1 protease. Although no direct functional data are available to assess MDR 769 HIV-1 protease catalytic activity, current structural studies, theoretical binding energy calculations, and molecular dynamics simulations indicate the reduced efficiency in peptide cleavage of MDR 769 HIV-1 protease. First, the flap water molecule positioned between Ile50 and Ile50′, bridging the protease flaps and substrate peptides, is missing, and this flap water is crucial in substrate cleavage (Liu et al. unpublished results). Therefore, the MDR HIV-1 protease may adopt other mechanisms to achieve the cleavage of the substrate. As a consequence of the expanded protease active site cavity, both the distance between the two flaps and the distance between flaps and the substrates are considerably increased in the MDR769 HIV-1 protease. As a result, the MDR HIV-1 protease does not bind the substrates tightly in optimal position, hence reducing catalytic cleavage. Second, calculations of substrate binding energies and K_d confirm the results of structural studies of MDR HIV-1 protease substrate complexes. There is a significant decrease of binding energy (34–78% loss) and an increase of K_d (20–162,000 fold gains) in MDR HIV-1 protease substrate complexes. Without tight binding, the substrates flop in and out the active site cavity more frequently, reducing the protease catalytic activity. Third, molecular dynamics calculations with a WT HIV-1 protease substrate complex show a rigid body movement of substrate after a 5 ns simulation, while this movement required for efficient cleavage is missing in

MDR substrate complexes molecular dynamics simulations (Yedidi et al. unpublished results).

Despite all the reasoning pointing to the reduced catalytic activity of MDR HIV-1 protease, the K_{cat} itself may not necessarily be reduced. It is possible that the wide open nature of MDR HIV-1 protease flaps leads to a weaker interaction between the cleavage products and MDR HIV-1 protease. As a result, the release of the cleavage products may be faster in MDR HIV-1 protease substrate complexes compared to that in WT complexes, which may partially compensate for the reduced catalytic activity.

Acknowledgments This research was supported by the National Institutes of Health grant AI65294 and a grant from the American Foundation for AIDS Research (106457-34-RGGN).

References

1. Altman MD, Ali A, Reddy GS, Nalam MN, Anjum SG, Cao H, Chellappan S, Kairys V, Fernandes MX, Gilson MK, Schiffer CA, Rana TM, Tidor B (2008) J Am Chem Soc 130:6099–6113
2. Bally F, Martinez R, Peters S, Sudre P, Telenti A (2000) AIDS Res Hum Retroviruses 16:1209–1213
3. Barbaro G, Lucchini A, Barbarini G (2005) Minerva Cardioangiol 53:153–154
4. Bartlett JA, DeMasi R, Quinn J, Moxham C, Rousseau F (2001) AIDS 15:1369–1377
5. Chellappan S, Kairys V, Fernandes MX, Schiffer C, Gilson MK (2007) Proteins 68:561–567
6. Chellappan S, Kiran Kumar Reddy GS, Ali A, Nalam MN, Anjum SG, Cao H, Kairys V, Fernandes MX, Altman MD, Tidor B, Rana TM, Schiffer CA, Gilson MK (2007) Chem Biol Drug Des 69:298–313
7. Clavel F, Hance AJ (2004) N Engl J Med 350:1023–1035
8. Condra JH, Schleif WA, Blahy OM, Gabryelski LJ, Graham DJ, Quintero JC, Rhodes A, Robbins HL, Roth E, Shivaprakash M (1995) Nature 374:569–571
9. Croteau G, Doyon L, Thibeault D, McKercher G, Pilote L, Lamarre D (1997) J Virol 71:1089–1096
10. Emsley P, Cowtan K (2004) Acta Crystallogr D Biol Crystallogr 60:2126–2132
11. Grabar S, Pradier C, Le Corfec E, Lancar R, Allavena C, Bentata M, Berlureau P, Dupont C, Fabbro-Peray P, Poizot-Martin I, Costagliola D (2000) AIDS 14:141–149
12. Gulick RM, Mellors JW, Havlir D, Eron JJ, Meibohm A, Condra JH, Valentine FT, McMahon D, Gonzalez C, Jonas L, Emini EA, Chodakewitz JA, Isaacs R, Richman DD (2000) Ann Intern Med 133:35–39
13. Holzgrabe U (2004) Pharm Unserer Zeit 33:160
14. Krissinel E, Henrick K (2004) Acta Crystallogr D Biol Crystallogr 60:2256–2268
15. Krissinel E, Henrick K (2007) J Mol Biol 372:774–797
16. Laemmli UK (1970) Nature 227:680–685
17. Lamzin VS, Wilson KS (1993) Acta Crystallogr D Biol Crystallogr 49:129–147
18. Li F, Goila-Gaur R, Salzwedel K, Kilgore NR, Reddick M, Matallana C, Castillo A, Zoumplis D, Martin DE, Orenstein JM, Allaway GP, Freed EO, Wild CT (2003) Proc Natl Acad Sci USA 100:13555–13560
19. Li F, Zoumplis D, Matallana C, Kilgore NR, Reddick M, Yunus AS, Adamson CS, Salzwedel K, Martin DE, Allaway GP, Freed EO, Wild CT (2006) Virology 356:217–224
20. Logsdon BC, Vickrey JF, Martin P, Proteasa G, Koepke JI, Terlecky SR, Wawrzak Z, Winters MA, Merigan TC, Kovari LC (2004) J Virol 78:3123–3132
21. Martin DE, Blum R, Wilton J, Doto J, Galbraith H, Burgess GL, Smith PC, Ballow C (2007) Antimicrob Agents Chemother 51:3063–3066
22. Martin P, Vickrey JF, Proteasa G, Jimenez YL, Wawrzak Z, Winters MA, Merigan TC, Kovari LC (2005) Structure 13:1887–1895
23. McMichael AJ, Hanke T (2003) Nat Med 9:874–880
24. Murshudov GN, Vagin AA, Dodson EJ (1997) Acta Crystallogr D Biol Crystallogr 53:240–255
25. Nalam MN, Ali A, Altman MD, Reddy GS, Chellappan S, Kairys V, Ozen A, Cao H, Gilson MK, Tidor B, Rana TM, Schiffer CA (2010) J Virol 84:5368–5378
26. Natarajan V, Bosche M, Metcalf JA, Ward DJ, Lane HC, Kovacs JA (1999) Lancet 353:119–120
27. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD (1998) N Engl J Med 338:853–860
28. Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH (2004) J Virol 78:8477–8485
29. Prabu-Jeyabalan M, King NM, Nalivaika EA, Heilek-Snyder G, Cammack N, Schiffer CA (2006) Antimicrob Agents Chemother 50:1518–1521
30. Prabu-Jeyabalan M, Nalivaika EA, King NM, Schiffer CA (2004) J Virol 78:12446–12454
31. Prabu-Jeyabalan M, Nalivaika EA, Romano K, Schiffer CA (2006) J Virol 80:3607–3616
32. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA (2000) J Mol Biol 301:1207–1220
33. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA (2002) Structure 10:369–381
34. Roberts JD, Preston BD, Johnston LA, Soni A, Loeb LA, Kunkel TA (1989) Mol Cell Biol 9:469–476
35. Robinson LH, Myers RE, Snowden BW, Tisdale M, Blair ED (2000) AIDS Res Hum Retroviruses 16:1149–1156
36. Schooley RT, Mellors JW (2007) J Infect Dis 195:770–772
37. Sukasem C, Churdboonchart V, Sukeepaisarncharoen W, Piroj W, Inwisai T, Tiensuwan M, Chantratita W (2008) Int J Antimicrob Agents 31:277–281
38. Takeuchi Y, Nagumo T, Hoshino H (1988) J Virol 62:3900–3902
39. Tie Y, Boross PI, Wang YF, Gaddis L, Liu F, Chen X, Tozser J, Harrison RW, Weber IT (2005) FEBS J 272:5265–5277
40. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN (2004) Acta Crystallogr D Biol Crystallogr 60:2184–2195
41. Vaguine AA, Richelle J, Wodak SJ (1999) Acta Crystallogr D Biol Crystallogr 55:191–205
42. Vickrey JF, Logsdon BC, Proteasa G, Palmer S, Winters MA, Merigan TC, Kovari LC (2003) Protein Expr Purif 28:165–172
43. Wallace AC, Laskowski RA, Thornton JM (1995) Protein Eng 8:127–134
44. Weber J, Grosse F (1989) Nucleic Acids Res 17:1379–1393
45. Zhang YM, Imamichi H, Imamichi T, Lane HC, Falloon J, Vasudevachari MB, Salzman NP (1997) J Virol 71:6662–6670
46. Zhou J, Chen CH, Aiken C (2004) Retrovirology 1:15