reproduction technique using spermatozoa free from HIV-1. *AIDS*. 2006; 20:967–973.
23. Kinai E, Hanabusa H, Kato S. Prediction of the efficacy of antiviral therapy for hepatitis C virus infection by an ultrasensitive RT-PCR assay. *J Med Virol*. 2007;79:1113–1119.
24. Damond F, Loussert-Ajaka I, Apetrei C, et al. Highly sensitive method for amplification of human immunodeficiency virus type 2 DNA. *J Clin Microbiol*. 1998;36:809–811.
25. Tamura K, Dudley J, Nei M, et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007;24: 1596–1599.
26. Lole KS, Bollinger RC, Paranjape RS, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999;73: 152–160.
27. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 2007;7:214.
28. Drummond AJ, Ho SY, Phillips MJ, et al. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 2006;4:e88.
29. Wilgenbusch JC, Swofford D. Inferring evolutionary trees with PAUP*. *Curr Protoc Bioinformatics*. 2003;Chapter 6:Unit 6.4.
30. Rodríguez F, Oliver JL, Marín A, et al. The general stochastic model of nucleotide substitution. *J Theor Biol*. 1990;142:485–501.
31. Lemey P, Pybus OG, Wang B, et al. Tracing the origin and history of the HIV-2 epidemic. *Proc Natl Acad Sci U S A*. 2003;100:6588–6592.
32. Pybus OG, Drummond AJ, Nakano T, et al. The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. *Mol Biol Evol*. 2003;20:381–387.
33. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. In: Kuiken CL, Foley B, Hahn B, et al, eds. *Human Retroviruses and AIDS 1999*. Los Alamos, NM: Los Alamos National Laboratory; 1999:492–505.
34. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science*. 2000;288:55–56.
35. Calef C, Mokili J, O'Connor DH, et al. Numbering positions in SIV relative to SIVMM239. In: Kuiken C, Foley B, Hahn B, et al, eds. *HIV Sequence Compendium 2001*. Los Alamos, NM: Los Alamos National Laboratory; 2001:171–181.
36. Lin G, Bertolotti-Ciarlet A, Haggarty B, et al. Replication-competent variants of human immunodeficiency virus type 2 lacking the V3 loop exhibit resistance to chemokine receptor antagonists. *J Virol*. 2007;81: 9956–9966.
37. Wertheim JO, Worobey M. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. *PLoS Comput Biol*. 2009;5:e1000377.
38. Pieniazek D, Ellenberger D, Janini LM, et al. Predominance of human immunodeficiency virus type 2 subtype B in Abidjan, Ivory Coast. *AIDS Res Hum Retroviruses*. 1999;15:603–608.
39. Zeh C, Pieniazek D, Agwale SM, et al. Nigerian HIV type 2 subtype A and B from heterotypic HIV type 1 and HIV type 2 or monotypic HIV type 2 infections. *AIDS Res Hum Retroviruses*. 2005;21:17–27.
40. Hessol NA, Koblin BA, van Griensven GJ, et al. Progression of human immunodeficiency virus type 1 (HIV-1) infection among homosexual men in hepatitis B vaccine trial cohorts in Amsterdam, New York City, and San Francisco, 1978–1991. *Am J Epidemiol*. 1994;139:1077–1087.
41. Veugelers PJ, Page KA, Tindall B, et al. Determinants of HIV disease progression among homosexual men registered in the Tricontinental Seroconverter Study. *Am J Epidemiol*. 1994;140:747–758.
42. UK Register of HIV Seroconverters Steering Committee. The AIDS incubation period in the UK estimated from a national register of HIV seroconverters. *AIDS*. 1998;12:659–667.
43. Pezzotti P, Galai N, Vlahov D, et al. Direct comparison of time to AIDS and infectious disease death between HIV seroconverter injection drug users in Italy and the United States: results from the ALIVE and ISS studies. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1999;20:275–282.
44. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. *Lancet*. 2000;355:1131–1137.
45. Morgan D, Mahe C, Mayanja B, et al. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? *AIDS*. 2002;16:597–603.
46. Poulsen AG, Aaby P, Larsen O, et al. 9-year HIV-2-associated mortality in an urban community in Bissau, west Africa. *Lancet*. 1997;349:911–914.
47. Berry N, Jaffar S, Schim van der Loeff M, et al. Low level viremia and high CD4% predict normal survival in a cohort of HIV type-2-infected villagers. *AIDS Res Hum Retroviruses*. 2002;18:1167–1173.

JOURNAL OF VIROLOGY, Oct. 2010, p. 9995–10003 Vol. 84, No. 19
0022-538X/10/$12.00 doi:10.1128/JVI.00505-10

The Effect of Clade-Specific Sequence Polymorphisms on HIV-1 Protease Activity and Inhibitor Resistance Pathways▽

Rajintha M. Bandaranayake,[1] Madhavi Kolli,[1] Nancy M. King,[1] Ellen A. Nalivaika,[1] Annie Heroux,[2] Junko Kakizawa,[3] Wataru Sugiura,[3,4] and Celia A. Schiffer[1]*

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, Massachusetts 01605[1]; Biology Department, Brookhaven National Laboratory, Upton, New York 11973-5000[2]; Laboratory of Therapeutic Research and Clinical Science, AIDS Research Center, National Institute of Infectious Diseases, 4-7-1 Gakuenn, Musashimurayama, Tokyo 208-0011, Japan[3]; and Department of Infection and Immunology, Clinical Research Center, Nagoya Medical Center, Nagoya, Japan[4]

Received 6 March 2010/Accepted 14 July 2010

The majority of HIV-1 infections around the world result from non-B clade HIV-1 strains. The CRF01_AE (AE) strain is seen principally in Southeast Asia. AE protease differs by ~10% in amino acid sequence from clade B protease and carries several naturally occurring polymorphisms that are associated with drug resistance in clade B. AE protease has been observed to develop resistance through a nonactive-site N88S mutation in response to nelfinavir (NFV) therapy, whereas clade B protease develops both the active-site mutation D30N and the nonactive-site mutation N88D. Structural and biochemical studies were carried out with wild-type and NFV-resistant clade B and AE protease variants. The relationship between clade-specific sequence variations and pathways to inhibitor resistance was also assessed. AE protease has a lower catalytic turnover rate than clade B protease, and it also has weaker affinity for both NFV and darunavir (DRV). This weaker affinity may lead to the nonactive-site N88S variant in AE, which exhibits significantly decreased affinity for both NFV and DRV. The D30N/N88D mutations in clade B resulted in a significant loss of affinity for NFV and, to a lesser extent, for DRV. A comparison of crystal structures of AE protease shows significant structural rearrangement in the flap hinge region compared with those of clade B protease and suggests insights into the alternative pathways to NFV resistance. In combination, our studies show that sequence polymorphisms within clades can alter protease activity and inhibitor binding and are capable of altering the pathway to inhibitor resistance.

Human immunodeficiency virus type 1 (HIV-1) is classified into three groups (M, N, and O), of which group M is further classified into nine major clades (A, B, C, D, F, G, H, J, and K) and 43 circulating recombinant forms (CRFs) based on viral genomic diversity (32, 37). The majority of HIV-1 infections across the globe result from non-B clade HIV-1 variants; clade B accounts for only ~12% of infections (15). However, the development of currently available anti-HIV therapies has been based on the virology of clade B variants. In recent years, several studies have shown that there are clear differences between clades when it comes to viral transmission and the progression to AIDS, an observation which raises questions about the effectiveness of the currently available anti-HIV therapies against the other clades and CRFs (16–18, 39).

HIV-1 protease has been an important drug target in the global effort to curb the progression from HIV infection to AIDS. However, the accumulation of drug-resistant mutations in the protease gene has been a major drawback in using HIV-1 protease inhibitors. The effects of mutations associated with drug resistance in HIV-1 clade B protease have been studied extensively over the years. For the most part, resistance mutation patterns are very similar in HIV-1 clade B and non-B clade proteases (19). However, several alternative resistance pathways have been observed for non-B clade proteases compared with those of clade B protease (1, 12, 13, 26). Limited data are available on how sequence polymorphisms, some of which are associated with drug resistance in clade B protease, might influence the pathway to drug resistance in non-B clade proteases. Furthermore, very little is understood about how sequence polymorphisms in non-B clade proteases affect protease function and inhibitor binding.

HIV-1 CRF01_AE (AE) was the first CRF to be observed in patient populations and is seen principally in Southeast Asia (2, 10, 25). AE protease differs by ~10% in amino acid sequence from that of clade B protease (Fig. 1A). Interestingly, AE protease develops a different resistance pathway from that of clade B protease to confer resistance to the protease inhibitor nelfinavir (NFV) (1). In patients infected with AE, the protease acquires predominantly the N88S mutation in response to NFV therapy, whereas in patients with clade B infection, the protease acquires the D30N/N88D mutations. The fitness of AE viral strains is thought to be similar to that of HIV-1 group M viral strains (11, 41). However, the effect of AE-specific sequence variations as well as drug resistance substitutions on viral fitness has not been studied extensively.

In the present study, biochemical and biophysical methods were used to determine the effect of sequence polymorphisms in AE protease on enzyme activity and inhibitor binding. Through determination of crystal structures and analysis of changes in hydrogen bonding patterns, a structural rationalization is described for the two different pathways observed for clade B and AE proteases to attain resistance to NFV.

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605. Phone: (508) 856-8008. Fax: (508) 856-6464. E-mail: Celia.Schiffer@umassmed.edu.

▽ Published ahead of print on 21 July 2010.

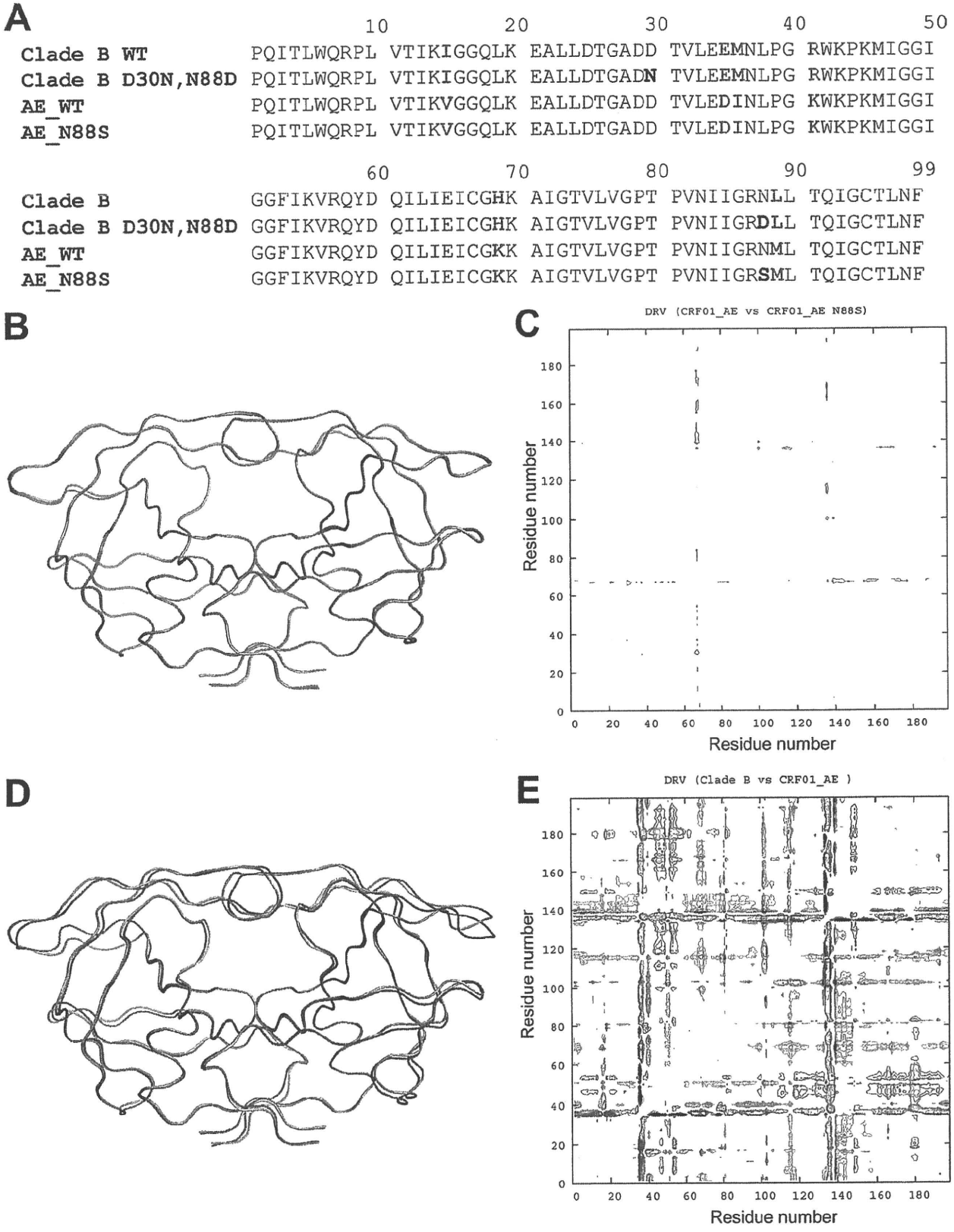

FIG. 1. (A) Amino acid sequence alignments of B-WT and AE-WT and NFV-resistant mutants. Residue positions that differ between clade B and AE are indicated in red. NFV resistance mutations are indicated in blue. (B) Ribbon diagram superposition of $DRV_{AE\text{-}WT}$ (blue) and $DRV_{AE\text{-}N88S}$ (gray). (C) Double-difference plot comparing $DRV_{AE\text{-}WT}$ and $DRV_{AE\text{-}N88S}$. (D) Ribbon diagram superposition of clade $DRV_{AE\text{-}WT}$ (magenta) and clade $DRV_{B\text{-}WT}$ (gray). (E) Double-difference plot comparing $DRV_{AE\text{-}WT}$ and $DRV_{B\text{-}WT}$. The color contours in the double-difference plots indicate distance differences of <1.0 Å (black), 1.0 to 0.5 Å (green), 0.5 to 1.0 Å (blue), and >1.0 Å (red).

MATERIALS AND METHODS

Protease gene construction. The clade B wild-type (B-WT) protease gene was generated as previously described (34). The AE wild-type (AE-WT) protease gene was synthesized in fragments (Integrated DNA Technologies, Coralville, IA), with codons optimized for expression in *Escherichia coli*. The fragments were ligated to form the complete gene, which was then inserted into the pET11a expression vector (Novagen/EMD Chemicals, Gibbstown, NJ). The protease sequence was confirmed by DNA sequencing. The NFV resistance mutations, N88S in AE (AE-N88S) and D30N/N88D in clade B (B-D30N/N88D), were generated by site-directed mutagenesis using a Stratagene QuikChange site-directed mutagenesis kit (Agilent Technologies, La Jolla, CA). Mutagenesis was confirmed by sequencing. The Q7K substitution was introduced to all protease variants to prevent autoproteolysis (38).

Protein expression and purification. The clade B and AE variants were subcloned into the heat-inducible pXC35 expression vector (American Type Culture Collection [ATCC], Manassas, VA) and transformed into *E. coli* TAP-106 cells. Protein overexpression, purification, and refolding were carried out as previously described (20). Protein used for crystallographic studies was further purified with a Pharmacia Superdex 75 fast-performance liquid chromatography column (GE Healthcare, Chalfont St. Giles, United Kingdom) equilibrated with refolding buffer (50 mM sodium acetate [pH 5.5], 10% glycerol, 5% ethylene glycol, and 5 mM dithiothreitol).

Crystallization and structure determination. Protease solutions between 1.0 and 2.0 mg ml^{-1} were equilibrated with a 5-fold molar excess of NFV, darunavir (DRV), and amprenavir (APV) for 1 h on ice. Crystals were grown over a reservoir solution consisting of 126 mM phosphate buffer (pH 6.2), 63 mM sodium citrate, and 18% to 23% ammonium sulfate by the hanging-drop vapor diffusion method. X-ray diffraction data for AE-WT were collected at a BioCARS beamline 14-BM-C at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) at a wavelength tuned to 0.9 Å with a Quantum 315 CCD X-ray detector (Area Detector Systems Corporation, Poway, CA). Diffraction data for AE-N88S were collected by using beamline X29A at the National Synchrotron Light Source (Brookhaven National Laboratory, Upton, NY) at a wavelength tuned to 1.08 Å with a Quantum 315 charge-coupled-device (CCD) X-ray detector (Area Detector Systems Corporation). Data for the B-D30N/N88D variant was collected in-house with an R-Axis IV imaging plate system (Rigaku Corporation, Tokyo, Japan) mounted on a rotating-anode X-ray source (Rigaku Corporation). All data were collected under cryocooled conditions.

The data were indexed, integrated, and scaled using HKL/HKL-2000 software (HKL Research, Charlottesville, VA) (29). Structure determination and refinement were carried out using the CCP4 program suite (4) as previously described (35). The tensor (T), libration (L), and screw (S) parameter files used in TLS refinement were generated using the TLS motion determination server (30). Model building and real-space refinement were carried out with Coot molecular graphics software (8). Structure comparisons were made by superposing the structures using the Cα atoms of the terminal regions (residues 1 to 9 and 86 to 99) from the two monomers. In the case of the AE complexes, which have multiple orientations for the inhibitor, only the orientation common with the clade B structures was used for analysis. Structures were visualized using PyMol molecular graphics software (6).

Double-difference plots were generated for AE and clade B protease structures to graphically visualize structural differences between the clades, as previously described (35). Briefly, distances between all Cα atoms within the dimer were calculated for each complex. A distance difference matrix was then computed for each atom for a given pair of complexes. The distance difference matrix was then plotted as a contour plot using the gnuplot plotting software (44).

Nomenclature. The following nomenclature format will be used to refer to each crystal structure: $\text{inhibitor}_{\text{protease variant}}$. Thus, DRV in complex with AE-WT, clade B-WT, AE-N88S, and AE-D30N/N88D protein are designated $DRV_{AE\ WT}$, $DRV_{B\ WT}$, $DRV_{AE\ N88S}$, and $DRV_{B\ D30N/N88D}$, respectively. Prime notation is used to distinguish the two monomers in the protease dimer. For example, residue 30 from the first monomer would be referred to as Asp30, and the same residue from the second monomer would be referred to as Asp30′.

ITC. Binding affinities and thermodynamic parameters of inhibitor binding to clade B and AE variants were determined by isothermal titration calorimetry (ITC) with a VP isothermal titration calorimeter (MicroCal, LLC, Northampton, MA). The buffer used for all protease and inhibitor solutions consisted of 10 mM sodium acetate (pH 5.0), 2% dimethyl sulfoxide, and 2 mM tris[2-carboxyethyl] phosphine. Binding affinities for all protease variants were obtained by competitive displacement titration using acetyl-pepstatin as the weaker binder. A solution of 30 to 45 μM protease was titrated with 10-μl injections of 200 μM acetyl-pepstatin to saturation. The pepstatin was then displaced by titrating 36 8-μl injections of 200 μM APV or NFV or 41 7-μl injections of 40 μM DRV. Heats of dilution were subtracted from the corresponding heats of reaction to obtain the heat resulting solely from the binding of the ligand to the enzyme. Data were processed and analyzed with the ITC data analysis module (Microcal) for Origin 7 data analysis and graphing software (OriginLab, Northampton, MA). Final results represent the average of at least two measurements.

Measurement of protease activity. Protease activity was assayed by following each variant's ability to hydrolyze the fluorogenic substrate HiLyte Fluor 488-Lys-Ala-Arg-Val-Leu-Ala-Glu-Ala-Met-Ser-Lys (QXL-520) (AnaSpec, Inc., Fremont, CA) that corresponds to the HIV-1 CA-p2 substrate. The CA-p2 cleavage site was used since it is conserved between HIV-1 clades (7). The assay was carried out in a 96-well plate, and the enzymatic reaction was initiated by adding 20 μl of a solution of 100 to 250 nM protease to 80 μl of substrate solution. The buffer used in all reactions consisted of 10 mM sodium acetate (pH 5.0), 2% dimethyl sulfoxide, and 2 mM tris[2-carboxyethyl]phosphine. Final concentrations in each experiment were 0 to 40 μM substrate and 20 to 50 nM protease. Accurate concentrations of properly folded active protease were determined by carrying out ITC experiments for each variant with acetyl-pepstatin as described in the previous section. Fluorogenic response to protease cleavage was monitored at 23°C using a Victor^3 microplate reader (PerkinElmer, Waltham, MA) by exciting the donor molecule at 485 nM and recording emitted light at 535 nM. Data points were acquired every 30 s. The data points in relative fluorescence units (RFU) were converted into concentrations using standard calibration curves generated for HiLyte Fluor 488 at each substrate concentration. In addition to the conversion of RFUs to concentrations, the generation of calibration curves at each substrate concentration allowed us to correct for the inner filter effect (5). Rates of each enzymatic reaction were determined from the linear portion of the data and were fitted against substrate concentrations to determine K_m and catalytic turnover rate (k_{cat}) values using VisualEnzymics enzyme-kinetics software (SoftZymics, Princeton, NJ). Final results for each variant represent the average from at least two experiments.

In order to determine the biochemical fitness of a particular variant in the presence of a given inhibitor, vitality values were calculated using the following equation, based on the vitality function described previously, where K_d is the dissociation constant and k_{cat}/K_m is the catalytic efficiency (14, 43).

$$\text{Vitality} = \frac{[K_d \times (k_{cat}/K_m)]_{\text{mutant}}}{[K_d \times (k_{cat}/K_m)]_{\text{clade B-WT}}}$$

The calculated vitality value for B-WT for a particular inhibitor would be 1.0, and vitality values greater than 1.0 would indicate that a given variant had a selective advantage over the same inhibitor, while values lower than 1.0 would indicate that the variant did not have a selective advantage.

RESULTS

Crystal structures. The AE-WT and NFV-resistant clade B and AE variants were cocrystallized with NFV, DRV, and APV to reveal the structural basis for the altered NFV resistance pathways. In addition, the effects of background polymorphisms in AE-WT on inhibitor binding compared with that of clade B-WT were discerned. Crystals of AE protease in complex with NFV and APV did not diffract to a high resolution; therefore, structural comparisons were carried out for AE and clade B protease in complex with DRV. The structure of $DRV_{B\text{-}WT}$ was solved previously in the laboratory and was used for structural comparisons (Protein Data Bank [PDB] code 1T3R). Both $DRV_{AE\text{-}WT}$ and $DRV_{AE\text{-}N88S}$ crystallized with DRV bound in two orientations in the active site. Crystallographic data and refinement statistics for $DRV_{AE\text{-}WT}$, $DRV_{B\text{-}WT}$, $DRV_{AE\text{-}N88S}$, and $DRV_{B\text{-}D30N/N88D}$ are given in Table 1.

Structural comparisons were carried out for AE and clade B DRV complexes by pairwise structural superposition and double-difference plots (Fig. 1B to E). The $DRV_{AE\text{-}WT}$ and $DRV_{AE\text{-}N88S}$ complexes were structurally similar (Fig. 1B and C). Although the $DRV_{AE\text{-}WT}$ and $DRV_{B\text{-}WT}$ could be superimposed on each other very well (root mean square deviation

TABLE 1. Crystallographic statistics

Parameter[a]	Result for indicated variant			
	$DRV_{B\text{-}WT}$[b]	$DRV_{B\text{-}D30N/N88D}$	$DRV_{AE\text{-}WT}$	$DRV_{AE\text{-}N88S}$
Inhibitor	DRV	DRV	DRV	DRV
Resolution (Å)	1.2	2.15	1.96	1.76
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 6_1	P 6_1
Z	4	4	6	6
Cell dimensions (Å)				
a	54.9	50.9	62.2	61.9
b	57.8	57.7		
c	62.0	61.6	82.7	82.1
Total no. of reflections	302,022	108,838	89,284	79,445
No. of unique reflections	55,056	10,326	12,493	17,277
R_{symm} (%)	3.8	6.7	5.5	4.6
Completeness (%)	95.5	99.6	93.9	97.4
I/σ^I	25.0	9.6	11.2	19.6
R_{work} (%)	14.1	18.1	20.0	19.6
R_{free} (%)	17.9	23.6	25.9	23.9
RMSD				
Bond length (Å)	0.004	0.009	0.009	0.007
Bond angle	1.5	1.9	1.5	1.7
PDB code	1T3R	3LZV	3LZS	3LZU

[a] Z, number of molecules in the unit cell; RMSD, root mean square deviation; $R_{symm} = \Sigma_{hkl}|I_{hkl} - \langle I_{hkl}\rangle|/\Sigma_{hkl}I_{hkl}$; I/σ^I, signal-to-noise ratio; $R_{work} = \Sigma||F_{obs}| - |F_{cal}||/\Sigma|F_{obs}|$; $R_{free} = \Sigma_{test}(|F_{obs}| - |F_{cal}|)^2/\Sigma_{test}|F_{obs}|^2$.
[b] See King et al.(21) and Surleraux et al. (40).

[RMSD] of 0.21 Å), there were clear and significant differences between the variants in the main chain at the flap hinge region (residues 33 to 39) and the protease core region (residues 16 to 22) (Fig. 1D and Fig. 2A to D). These differences were further evident by the presence of significant peaks in the double-difference plot (Fig. 1E). The Ile36 side chain in $DRV_{AE\text{-}WT}$ packs well against the core region through favorable van der Waals interactions and is shorter than the Met36 in $DRV_{B\text{-}WT}$ (Fig. 2C). In addition, the shorter Asp35 in $DRV_{AE\text{-}WT}$ further enhances the packing by being flipped inward against the core, while in $DRV_{B\text{-}WT}$, the longer Glu35 is flipped outward into the solvent and forms a salt bridge with Arg57 (Fig. 2D). The packing of the flap hinge and core regions in $DRV_{AE\text{-}WT}$ is further stabilized by a hydrogen bond between the carbonyl oxygen of Asp35 and Lys20 NZ atom and is not present in $DRV_{B\text{-}WT}$.

The Asp30′ side chain of $DRV_{B\text{-}WT}$ does not directly form a hydrogen bond with DRV but indirectly interacts with the N1 atom of DRV through a water molecule-mediated hydrogen bond network (Fig. 3A). In contrast, the Asp30′ side chain of $DRV_{AE\text{-}WT}$ forms a direct hydrogen bond with the N1 atom of DRV (Fig. 3B). Residue 30 of both NFV-resistant variants also interacts with the N1 atom of DRV through water molecule-mediated hydrogen bonding (Fig. 3C and D). However, in addition to this interaction, Asn30 of $DRV_{B\text{-}D30N/N88D}$ and Asp30 of $DRV_{AE\text{-}N88S}$ are oriented away from the active site, enabling them to form hydrogen bonds with Asp88 and Ser88, respectively. In both cases, the NFV resistance mutations stabilize residue 30 away from the active site via hydrogen bonding.

Binding thermodynamics. To determine the effects of background sequence polymorphisms and NFV resistance mutations on inhibitor binding, the binding thermodynamic parameters of NFV, DRV, and APV binding to WT and resistant AE and clade B variants were determined by isothermal titration calorimetry (Table 2). The AE-WT protease had a 6.9-fold-weaker affinity for NFV and a 2.7-fold-weaker affinity for DRV

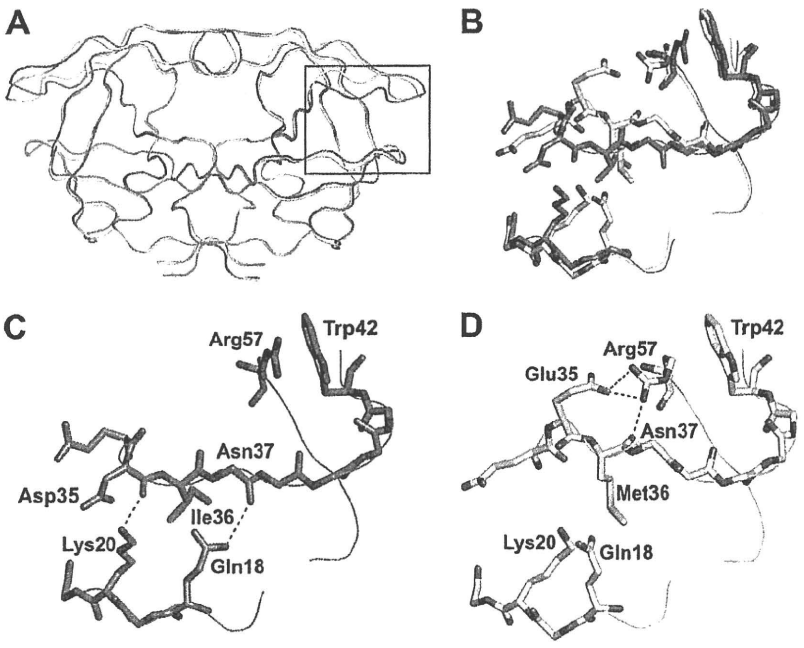

FIG. 2. (A) Ribbon diagram superposition of $DRV_{AE\text{-}WT}$ (blue) and $DRV_{B\text{-}WT}$ (gray). The red box indicates the region of the protease molecule highlighted in panels B to D. (B) Structural rearrangement of the flap hinge and core regions between $DRV_{AE\text{-}WT}$ (blue) and $DRV_{B\text{-}WT}$ (gray). (C) Flap hinge and core regions of $DRV_{AE\text{-}WT}$. (D) Flap hinge and core regions of $DRV_{B\text{-}WT}$ protease. Hydrogen bond interactions are indicated by red dashed lines.

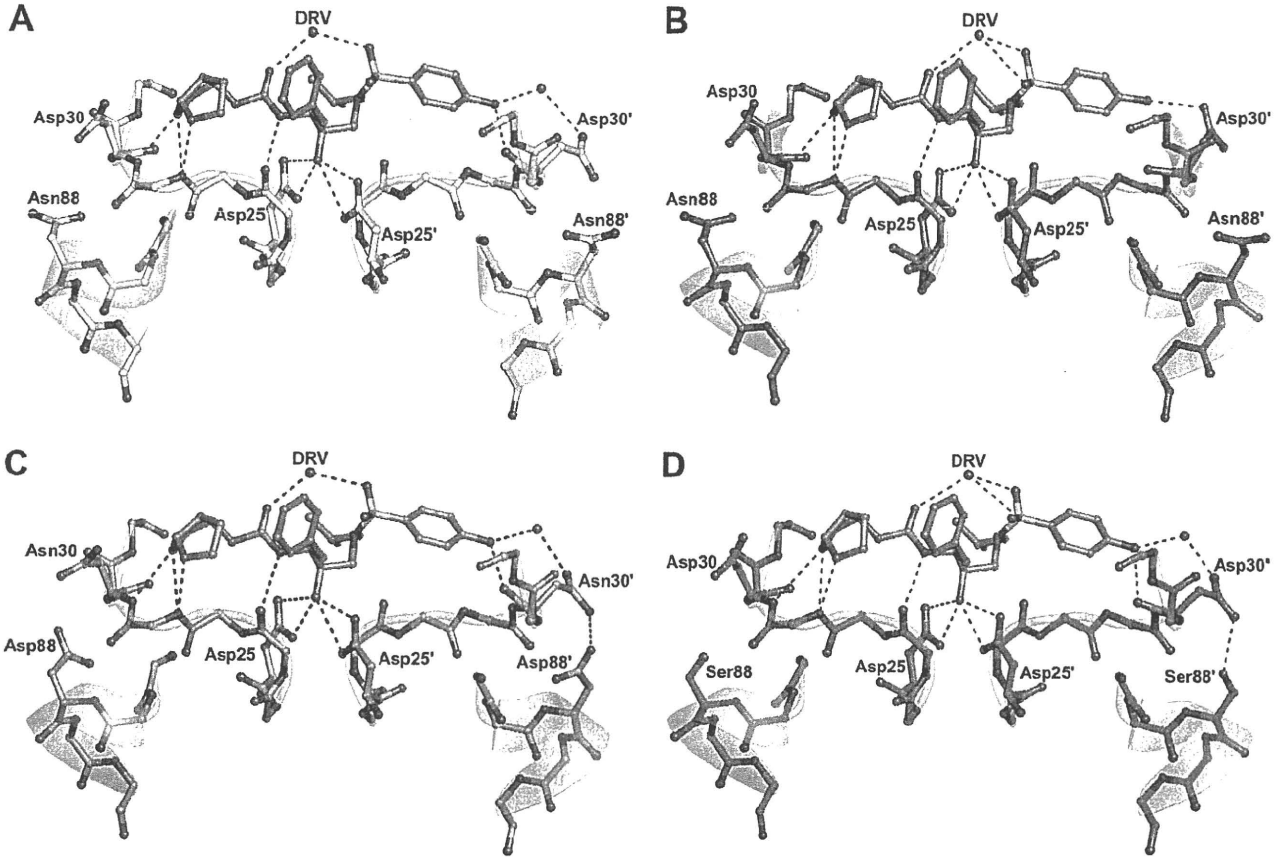

FIG. 3. Protease inhibitor hydrogen bonding interactions. DRV is shown in orange, and hydrogen bond interactions are indicated by red dashed lines. Since the charged states of the Asp25 carboxyl groups and the position of the O18 hydroxyl hydrogen of DRV are not known, all possible hydrogen bond interactions between the Asp25 carboxyl groups and O18 of the DRV molecules are shown. (A) $DRV_{B\text{-}WT}$ (gray). (B) $DRV_{AE\text{-}WT}$ (blue). (C) $DRV_{B\text{-}D30N/N88D}$ (salmon). (D) $DRV_{AE\text{-}N88S}$ (green).

than the affinities of B-WT protease for NFV and DRV, respectively (Table 2). This result indicates that the AE-WT protease has an inherently weaker affinity for NFV and DRV.

No significant differences in the enthalpy of NFV binding were observed among any of the variants. Although the binding of DRV to all protease variants was enthalpically favorable, the enthalpic contributions were reduced with the AE variants (-10.1 kcal mol^{-1} for AE-WT and -5.1 kcal mol^{-1} for AE-

TABLE 2. Binding thermodynamic parameters for NFV, DRV, and APV binding to AE and clade B variants[a]

Inhibitor and protease variant	K_a (M^{-1})	K_d (nM)	K_d ratio	ΔH (kcal mol^{-1})	$\Delta\Delta H$	$-T\Delta S$ (kcal mol^{-1})	$\Delta(-T\Delta S)$	ΔG (kcal mol^{-1})	$\Delta\Delta G$
NFV									
B-WT	$(2.6 \pm 0.5) \times 10^9$	0.39 ± 0.07	1.0	4.4 ± 0.1		-17.0		-12.6 ± 0.1	
B-D30N/N88D	$(1.2 \pm 0.4) \times 10^8$	8.1 ± 2.8	20.7	6.7 ± 0.3	2.3	-17.5	0.5	-10.8 ± 0.2	1.8
AE-WT	$(3.7 \pm 1.0) \times 10^8$	2.7 ± 0.7	6.9	5.0 ± 0.3	0.6	-16.6	0.9	-11.5 ± 0.2	1.1
AE-N88S	$(5.8 \pm 1.2) \times 10^7$	17.2 ± 3.5	44.1	6.2 ± 0.7	1.8	-16.6	0.9	-10.4 ± 0.1	2.2
DRV									
B-WT	$(2.2 \pm 1.1) \times 10^{11}$	0.004 ± 0.002	1.0	-12.1 ± 0.9		3.1		-15.2 ± 0.3	
B-D30N/N88D	$(3.7 \pm 0.7) \times 10^{10}$	0.026 ± 0.005	6.5	-12.5 ± 0.4	-0.4	-1.6	1.5	-14.2 ± 0.1	1.0
AE-WT	$(9.1 \pm 0.3) \times 10^{10}$	0.0109 ± 0.0003	2.7	-10.1 ± 0.5	2.0	-4.6	1.5	-14.7 ± 0.02	0.5
AE-N88S	$(1.1 \pm 0.8) \times 10^{10}$	0.087 ± 0.062	21.8	-5.1 ± 3.6	7.0	-8.4	5.3	-13.5 ± 0.4	1.7
APV									
B-WT	$(2.6 \pm 1.3) \times 10^9$	0.39 ± 0.20	1.0	-7.3 ± 0.9		-5.3		-12.6 ± 0.3	
B-D30N/N88D	$(1.2 \pm 0.2) \times 10^{10}$	0.08 ± 0.01	0.2	-10.2 ± 1.5	-2.9	-3.3	2.0	-13.5 ± 0.09	-0.9
AE-WT	$(3.1 \pm 0.2) \times 10^9$	0.32 ± 0.02	0.8	-5.5 ± 0.3	-1.8	-7.3	-2.0	-12.70 ± 0.03	-0.1
AE-N88S	$(1.3 \pm 0.9) \times 10^{10}$	0.08 ± 0.06	0.2	-5.0 ± 3.6	2.3	-8.6	-3.3	-13.6 ± 0.4	1.0

[a] K_a, association constant; K_d, dissociation constant; H, enthalpy; T, temperature; S, entropy; G, Gibbs free energy.

TABLE 3. Enzyme kinetics parameters for clade B and AE-WT and NFV-resistant variants

Parameter	Result for indicated variant			
	B-WT	B-D30N/N88D	AE-WT	AE-N88S
K_m (μM)	16.7 ± 6.0	35.9 ± 0.1	17.5 ± 4.0	19.0 ± 0.8
k_{cat} (s^{-1})	1.79 ± 0.28	0.13 ± 0.09	0.70 ± 0.08	0.20 ± 0.02
k_{cat}/K_m (s^{-1} μM^{-1})	0.11 ± 0.04	0.004 ± 0.002	0.04 ± 0.01	0.010 ± 0.001

TABLE 4. Vitality values for clade B and AE WT and NFV-resistant variants

Inhibitor	Result for indicated variant		
	B-D30N/N88D	AE-WT	AE-N88S
NFV	0.76	2.52	4.01
DRV	0.24	0.99	1.98
APV	0.01	0.30	0.02

N88S) compared with those for the clade B variants (−12.1 kcal mol^{-1} for B-WT and −12.5 kcal mol^{-1} for B-D30N/N88D). As expected, the NFV-resistant variants showed a significant reduction in binding affinity for NFV compared to that of the wild-type variants. With the AE-N88S variant, the affinity for NFV was reduced 44.1-fold (K_d = 17.2 nM) and was far more significant than the D30N/N88D mutations in clade B protease, which reduced the affinity for NFV 20.7-fold (K_d = 8.1 nM). Similarly, the AE-N88S variant had a 21.8-fold-weaker affinity (K_d = 0.087 nM) for DRV compared to a 6.5-fold-weaker affinity (K_d = 0.026 nM) with the B-D30N/N88D variant. Thus, the single N88S substitution in the AE protease has a profound effect on the binding of NFV and DRV.

In contrast to NFV and DRV, clade-specific sequence differences and NFV resistance mutations had only a minimal effect on the affinities for APV of both AE and clade B protease. Despite this, there were some differences in energy parameters. The binding of APV to the clade B variants appeared to be more enthalpically favorable than that to the AE variants. This was compensated for by an increase in the entropic component to the binding energy for the AE proteases.

Protease activity and vitality. The enzyme-kinetic parameters determined for each clade B and AE variant with the CA-p2 fluorogenic substrate analog are summarized in Table 3. The K_m value for B-D30N/N88D protease (35.9 μM) was 2.1-fold greater than that for B-WT protease (16.7 μM). However, the K_m values for the AE protease variants (17.5 μM for AE-WT and 19.0 μM for AE-N88S) were similar to that of B-WT protease. The turnover rate for B-D30N/N88D protease (k_{cat} = 0.13 s^{-1}) was significantly lower than that of B-WT protease (k_{cat} = 1.79 s^{-1}). Turnover rates for AE-WT (k_{cat} = 0.7 s^{-1}) and AE-N88S (k_{cat} = 0.2 s^{-1}) were 2.5 and 8.5-fold lower, respectively, than that of clade B-WT. The k_{cat}/K_m values, or catalytic efficiency values, for B-D30N/N88D and AE variants were lower than that of B-WT protease. Therefore, the reduction in catalytic efficiency of the B-D30N/N88D protease compared with that of B-WT protease resulted from the combined effects of the K_m and k_{cat} values. However, for the AE variants, the lower turnover rates alone were responsible for the reduced catalytic efficiencies. Overall, these results indicate that the polymorphic sequence differences in AE protease can alter the activity profile of the enzyme compared to results with the clade B protease.

Vitality values were calculated to determine if the protease variants had a selective advantage over NFV, DRV, and APV. AE-WT and AE-N88S protease had calculated vitality values of 2.52 and 4.01 for NFV, respectively, compared with 0.76 for B-WT (Table 4). However, vitality values for DRV were not significantly different from that of B-WT protease. Vitality values for APV were significantly lower for all variants than for B-WT protease. These results indicate that AE-WT may have a selective advantage over NFV compared to B-WT but that the AE variants may not have a significant selective advantage against DRV or APV relative to B-WT.

DISCUSSION

Although the majority of HIV-1 patients are infected with non-B forms of the virus, molecular studies have been carried out predominantly with clade B variants. The AE protease has several polymorphisms that are associated with inhibitor resistance in clade B. AE also shows altered patterns of drug resistance to NFV. We have performed detailed studies to determine the effects of sequence polymorphisms on enzyme structure, activity, and inhibitor binding. These analyses led to a structural rationalization for the altered pathways for drug resistance.

AE-WT protease has an inherently weaker affinity for NFV and DRV than that of B-WT, as is evident from the thermodynamic data (Table 2). The weaker affinity observed for NFV is consistent with previously published data for another AE protease variant (3), as well as for clade A protease (42), which is closely related. The inherent weaker affinity for NFV likely allows the AE protease to gain resistance to NFV through a single nonactive-site substitution, N88S. The clade B protease, in contrast, which has a relatively stronger affinity for NFV, requires a combination of an active-site mutation (D30N) and a nonactive-site mutation (N88D) to gain NFV resistance. The ability of the AE-N88S protease to maintain affinity for substrates is evident from our enzyme kinetics data (Table 4), in which the K_m value for AE-N88S was comparable to that of AE-WT and B-WT protease. The K_m value for clade B-D30N/N88D, on the other hand, was significantly worse than that of the B-WT, likely reflecting the effect of the altered active site.

As an active-site residue, Asp30 plays a key role in substrate recognition by interacting with substrates through side chain-mediated hydrogen bonds with the MA-CA, CA-p2, p1-p6, and p2-NC cleavage sites (36). Therefore, as is evident from our enzyme kinetics data, the D30N/N88D mutations in clade B will likely affect substrate binding and processing. Several studies have observed substrate coevolution in instances in which the protease mutates active-site residues in order to confer inhibitor resistance (22, 23). However, since the AE-N88S protease variant has no active-site mutations, the enzyme retains the ability to effectively recognize substrates while conferring NFV resistance. Therefore, the presence of the N88S substitution in AE protease is unlikely to induce coevolution of

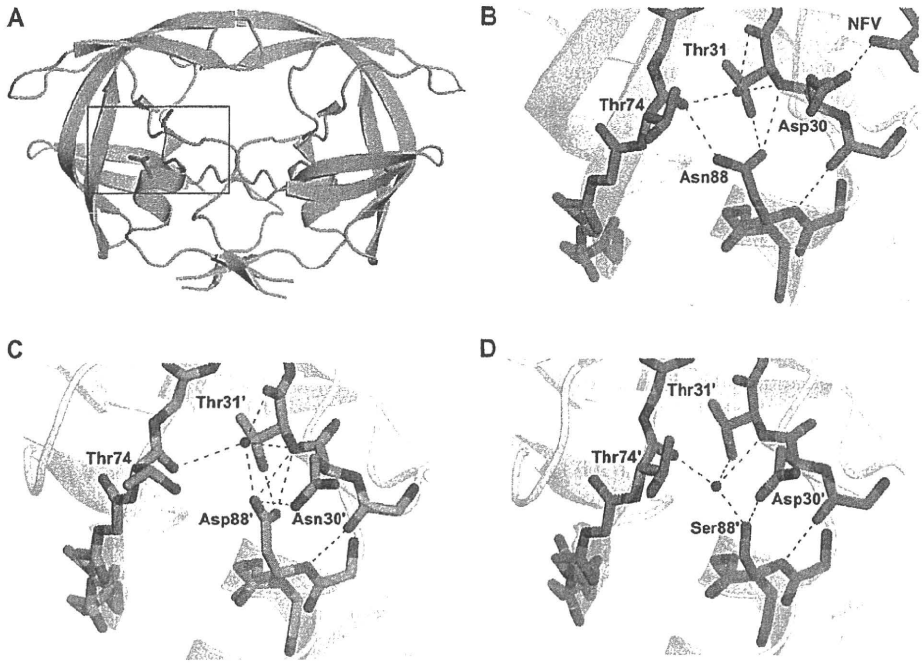

FIG. 4. Hydrogen bond network involving residue 88. (A) Asn88 bridges the terminal helix with Asp30 from the active site and Thr74 from one of the outer beta strands. The red box indicates the region of the protease molecule highlighted in panels B to D. (B) Asn88 in $NFV_{B\text{-}WT}$ (PDB code 3EKX). (C) Asp88 in $DRV_{B\text{-}D30N/N88D}$. (D) Ser88 in $DRV_{N88S\text{-}WT}$. Hydrogen bond interactions are indicated by red dashed lines.

the viral substrates in order to maintain effective enzymatic activity.

Despite having K_m values that were comparable to that of B-WT protease, both AE-WT and AE-N88S had significantly lower catalytic turnover rates (k_{cat}) than that of the B-WT protease (Table 3). As a result, the catalytic efficiency of the AE variants is lower than that of the B-WT protease. The lower turnover rates of the AE variants could be a direct result of the reduced flexibility of the flap hinge (residues 33 to 39) and core regions (residues 16 to 22) of the protein. Molecular dynamics studies have revealed that hydrophobic sliding of the core region facilitates substrate binding through the opening of the active site (9). The unique hydrogen bonds observed between the flap hinge and the core in the AE variants alter movement of the core, thus impacting the ability of the active site to open up for substrate binding and product release. Based on our enzyme kinetics data, this altered flexibility of the flap hinges in the AE variants has little effect on substrate binding but rather affects the catalytic step of the reaction by slowing down product release.

The higher vitality value observed for AE-WT with NFV provides supporting evidence for the reduction in the efficacy of NFV against the AE protease compared with that of clade B (Table 4). This result is consistent with previous vitality calculations for the clade A protease (42). In addition, these results further highlight the idea that background polymorphic sequence variations in the AE protease can affect the potency of NFV. The suboptimal efficacy of NFV against the AE-WT protease likely permits a nonactive-site variant, AE-N88S, to emerge over variants with active-site mutations to effectively confer resistance to NFV.

The impact on other inhibitors, however, is complex. APV and DRV are chemically very closely related compounds, and similar susceptibility and resistance patterns have been observed for these two inhibitors (31). However, this pattern is not evident for this series of resistant variants. Both the N88S mutation in the AE and the D30N/N88D mutations in the clade B proteases result in hypersusceptibility to APV. Similar results have been observed also for a B-N88S protease variant (24, 45). In contrast, the same substitutions in the protease give rise to even greater resistance to DRV. However, since DRV presents a greater genetic barrier to resistance than APV (33), the *in vivo* implications of weaker affinity for DRV in the AE variants are likely negligible. Indeed, our calculated vitality values indicate that DRV maintains its potency against the AE variants despite having a weaker affinity for AE-WT and AE-N88S relative to clade B protease.

A close look at the $NFV_{B\text{-}WT}$ protease complex reveals an important interaction between the Asp30 residue side chain and the inhibitor bound in the active site. (PDB code 3EKX) (Fig. 4A and B). One of the side chain oxygen atoms of Asp30 forms a direct hydrogen bond with the O38 atom of NFV. Our crystal structures of the NFV-resistant variants show that N88S in AE and N88D in clade B have the ability to interact with residue 30 and orient it away from the active site (Fig. 3B and D) and thereby disrupt the interaction between residue 30 and the inhibitor. These structural observations are similar to interpretations made in previous molecular dynamics studies involving NFV-protease complexes (27, 28). Thus, NFV resistance is likely caused in large part due to the loss of this interaction in the NFV-resistant variants.

Overall, mutations that emerge in response to inhibitor ther-

apy need to have a minimal impact on protease structure and activity to maintain the enzyme's function. The D30N substitution, which is associated with NFV resistance, is one of the few drug-resistant mutations that involve a change in charge. The additional substitution of N88D likely helps preserve the net charge on the protein. In AE, resistance to NFV occurs indirectly with the N88S mutation. Likewise, the sole NFV-resistant alteration, N88S, in the AE protease does not change the overall electrostatics. Thus, in both clade B and AE, NFV resistance is attained with no change to the net charge of the enzyme. In the wild-type variants, Asn88 is one of the few internal hydrogen bonding side chains in the core of the protease monomer. The side chain of Asn88 has a key role in the protease structure bridging the terminal helix, with residues 30 and 31 coming from the active site to the backbone of Thr74 in the center of one of the outer beta strands (Fig. 4A and B). With the substitutions of Asp in clade B and Ser in AE for Asn at position 88 in the NFV-resistant protease variants, the hydrogen bonding network is preserved through the coordination of some key water molecules in the core of the protease monomer (Fig. 4B to D). Thus, mutations confer resistance to NFV through a series of interdependent changes that preserve the structural and electrostatic properties of HIV-1 protease.

In conclusion, protease activity and the response to protease inhibitors can be affected by clade-specific sequence differences. Our findings likely extend beyond HIV-1 protease to other drug targets within HIV and underscore the need to consider clade-specific polymorphisms when developing new drugs and formulating treatment plans. Furthermore, drug resistance pathways observed in the context of clade B viruses cannot be assumed to hold true for other HIV-1 clades.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (P01-GM66524) and Tibotec, Inc., to C.A.S. Additionally, this study was supported by a Grant-in-Aid for AIDS research from the Ministry of Health, Labor, and Welfare of Japan (H19-AIDS-007) to W.S.

We thank William Royer, Moses Prabu-Jayabalan, and Madhavi Nalam for helpful discussions and Christina Ng and Brendan Hilbert for assistance with data collection.

We gratefully acknowledge the Mail-In Data Collection Program of the National Synchrotron Light Source, Brookhaven National Laboratory, for collecting X-ray data at the X29A beamline, for which financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the U.S. Department of Energy and from the National Center for Research Resources of the National Institutes of Health. Use of the Advanced Photon Source for X-ray data collection was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract DE-AC02-06CH11357. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under grant RR007707.

The protease inhibitors used in this study were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, NIH.

REFERENCES

1. **Ariyoshi, K., M. Matsuda, H. Miura, S. Tateishi, K. Yamada, and W. Sugiura.** 2003. Patterns of point mutations associated with antiretroviral drug treatment failure in CRF01_AE (subtype E) infection differ from subtype B infection. J. Acquir. Immune Defic. Syndr. **33**:336–342.
2. **Carr, J. K., M. O. Salminen, C. Koch, D. Gotte, A. W. Artenstein, P. A. Hegerich, D. St. Louis, D. S. Burke, and F. E. McCutchan.** 1996. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. J. Virol. **70**:5935–5943.
3. **Clemente, J. C., R. M. Coman, M. M. Thiaville, L. K. Janka, J. A. Jeung, S. Nukoolkarn, L. Govindasamy, M. Agbandje-McKenna, R. McKenna, W. Leelamanit, M. M. Goodenow, and B. M. Dunn.** 2006. Analysis of HIV-1 CRF_01 A/E protease inhibitor resistance: structural determinants for maintaining sensitivity and developing resistance to atazanavir. Biochemistry **45**: 5468–5477.
4. **Collaborative Computational Project, Number 4.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50**:760–763.
5. **Copeland, R. A.** 1996. Enzymes: a practical introduction to structure, mechanism, and data analysis. Wiley-VCH, New York, NY.
6. **DeLano, W. L.** 2002. The PyMol molecular graphics system. DeLano Scientific, San Carlos, CA.
7. **de Oliveira, T., S. Engelbrecht, E. Janse van Rensburg, M. Gordon, K. Bishop, J. zur Megede, S. W. Barnett, and S. Cassol.** 2003. Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? J. Virol. **77**:9422–9430.
8. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**:2126–2132.
9. **Foulkes-Murzycki, J. E., W. R. Scott, and C. A. Schiffer.** 2007. Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease. Structure **15**:225–233.
10. **Gao, F., D. L. Robertson, S. G. Morrison, H. Hui, S. Craig, J. Decker, P. N. Fultz, M. Girard, G. M. Shaw, B. H. Hahn, and P. M. Sharp.** 1996. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. J. Virol. **70**:7013–7029.
11. **Geretti, A. M.** 2006. HIV-1 subtypes: epidemiology and significance for HIV management. Curr. Opin. Infect. Dis. **19**:1–7.
12. **Gomes, P., I. Diogo, M. F. Gonçalves, P. Carvalho, J. Cabanas, M. C. Lobo, and R. Camacho.** 2002. Different pathways to nelfinavir genotypic resistance in HIV-1 subtypes B and G. Conference on Retroviruses and Opportunistic Infections, abstract 46. Conference on Retroviruses and Opportunistic Infections, Seattle, WA.
13. **Grossman, Z., E. E. Paxinos, D. Averbuch, S. Maayan, N. T. Parkin, D. Engelhard, M. Lorber, V. Istomin, Y. Shaked, E. Mendelson, D. Ram, C. J. Petropoulos, and J. M. Schapiro.** 2004. Mutation D30N is not preferentially selected by human immunodeficiency virus type 1 subtype C in the development of resistance to nelfinavir. Antimicrob. Agents Chemother. **48**:2159–2165.
14. **Gulnik, S. V., L. I. Suvorov, B. Liu, B. Yu, B. Anderson, H. Mitsuya, and J. W. Erickson.** 1995. Kinetic characterization and cross-resistance patterns of HIV-1 protease mutants selected under drug pressure. Biochemistry **34**: 9282–9287.
15. **Hemelaar, J., E. Gouws, P. D. Ghys, and S. Osmanov.** 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS **20**:W13–23.
16. **Hu, D. J., A. Buve, J. Baggs, G. van der Groen, and T. J. Dondero.** 1999. What role does HIV-1 subtype play in transmission and pathogenesis?: an epidemiological perspective. AIDS **13**:873–881.
17. **Kaleebu, P., N. French, C. Mahe, D. Yirrell, C. Watera, F. Lyagoba, J. Nakiyingi, A. Rutebemberwa, D. Morgan, J. Weber, C. Gilks, and J. Whitworth.** 2002. Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1-positive persons in Uganda. J. Infect. Dis. **185**:1244–1250.
18. **Kanki, P. J., D. J. Hamel, J. L. Sankale, C. Hsieh, I. Thior, F. Barin, S. A. Woodcock, A. Gueye-Ndiaye, E. Zhang, M. Montano, T. Siby, R. Marlink, I. NDoi, M. E. Essex, and S. MBoup.** 1999. Human immunodeficiency virus type 1 subtypes differ in disease progression. J. Infect. Dis. **179**:68–73.
19. **Kantor, R., D. A. Katzenstein, B. Efron, A. P. Carvalho, B. Wynhoven, P. Cane, J. Clarke, S. Sirivichayakul, M. A. Soares, J. Snoeck, C. Pillay, H. Rudich, R. Rodrigues, A. Holguin, K. Ariyoshi, M. B. Bouzas, P. Cahn, W. Sugiura, V. Soriano, L. F. Brigido, Z. Grossman, L. Morris, A. M. Vandamme, A. Tanuri, P. Phanuphak, J. N. Weber, D. Pillay, P. R. Harrigan, R. Camacho, J. M. Schapiro, and R. W. Shafer.** 2005. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. PLoS Med. **2**:e112.
20. **King, N. M., L. Melnick, M. Prabu-Jeyabalan, E. A. Nalivaika, S.-S. Yang, Y. Gao, X. Nie, C. Zepp, D. L. Heefner, and C. A. Schiffer.** 2002. Lack of synergy for inhibitors targeting a multi-drug-resistant HIV-1 protease. Protein Sci. **11**:418–429.
21. **King, N. M., M. Prabu-Jeyabalan, P. Wigerinck, M.-P. de Béthune, and C. A. Schiffer.** 2004. Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor. J. Virol. **78**:12012–12021.
22. **Kolli, M., S. Lastere, and C. A. Schiffer.** 2006. Co-evolution of nelfinavir-resistant HIV-1 protease and the p1-p6 substrate. Virology **347**:405–409.
23. **Kolli, M., E. Stawiski, C. Chappey, and C. A. Schiffer.** 2009. Human immunodeficiency virus type 1 protease-correlated cleavage site mutations enhance inhibitor resistance. J. Virol. **83**:11027–11042.
24. **Masquelier, B., K. L. Assoumou, D. Descamps, L. Bocket, J. Cottalorda, A. Ruffault, A. G. Marcelin, L. Morand-Joubert, C. Tamalet, C. Charpentier,**

G. Peytavin, Z. Antoun, F. Brun-Vezinet, and D. Costagliola. 2008. Clinically validated mutation scores for HIV-1 resistance to fosamprenavir/ritonavir. J. Antimicrob. Chemother. **61**:1362–1368.

25. **Murphy, E., B. Korber, M. C. Georges-Courbot, B. You, A. Pinter, D. Cook, M. P. Kieny, A. Georges, C. Mathiot, F. Barre-Sinoussi, et al.** 1993. Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the Central African Republic. AIDS Res. Hum. Retroviruses **9**:997–1006.
26. **Nunez, M., C. de Mendoza, L. Valer, E. Casas, S. Lopez-Calvo, A. Castro, B. Roson, D. Podzamczer, A. Rubio, J. Berenguer, and V. Soriano.** 2002. Resistance mutations in HIV-infected patients experiencing early failure with nelfinavir-containing triple combinations. Med. Sci. Monit **8**:CR620–CR623.
27. **Ode, H., S. Matsuyama, M. Hata, T. Hoshino, J. Kakizawa, and W. Sugiura.** 2007. Mechanism of drug resistance due to N88S in CRF01_AE HIV-1 protease, analyzed by molecular dynamics simulations. J. Med. Chem. **50**: 1768–1777.
28. **Ode, H., M. Ota, S. Neya, M. Hata, W. Sugiura, and T. Hoshino.** 2005. Resistant mechanism against nelfinavir of human immunodeficiency virus type 1 proteases. J. Phys. Chem. B **109**:565–574.
29. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276**:307–326.
30. **Painter, J., and E. A. Merritt.** 2006. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr. D Biol. Crystallogr. **62**:439–450.
31. **Parkin, N., E. Stawiski, C. Chappey, and E. Coakley.** 2007. Darunavir/amprenavir cross-resistance in clinical samples submitted for phenotype/genotype combination resistance testing. Conference on Retroviruses and Opportunistic Infections, abstract 607. Conference on Retroviruses and Opportunistic Infections, San Francisco, CA.
32. **Plantier, J. C., M. Leoz, J. E. Dickerson, F. De Oliveira, F. Cordonnier, V. Lemee, F. Damond, D. L. Robertson, and F. Simon.** 2009. A new human immunodeficiency virus derived from gorillas. Nat. Med. **15**:871–872.
33. **Poveda, E., C. de Mendoza, L. Martin-Carbonero, A. Corral, V. Briz, J. Gonzalez-Lahoz, and V. Soriano.** 2007. Prevalence of darunavir resistance mutations in HIV-1-infected patients failing other protease inhibitors. J. Antimicrob. Chemother. **60**:885–888.
34. **Prabu-Jeyabalan, M., E. Nalivaika, N. M. King, and C. A. Schiffer.** 2004. Structural basis for coevolution of the human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease. J. Virol. **78**:12446–12454.
35. **Prabu-Jeyabalan, M., E. A. Nalivaika, K. Romano, and C. A. Schiffer.** 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. J. Virol. **80**:3607–3616.
36. **Prabu-Jeyabalan, M., E. A. Nalivaika, and C. A. Schiffer.** 2002. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. Structure **10**:369–381.
37. **Robertson, D. L., J. P. Anderson, J. A. Bradac, J. K. Carr, B. Foley, R. K. Funkhouser, F. Gao, B. H. Hahn, M. L. Kalish, C. Kuiken, G. H. Learn, T. Leitner, F. McCutchan, S. Osmanov, M. Peeters, D. Pieniazek, M. Salminen, P. M. Sharp, S. Wolinsky, and B. Korber.** 2000. HIV-1 nomenclature proposal. Science **288**:55–56.
38. **Rose, J. R., R. Salto, and C. S. Craik.** 1993. Regulation of autoproteolysis of the HIV-1 and HIV-2 proteases with engineered amino acid substitutions. J. Biol. Chem. **268**:11939–11945.
39. **Spira, S., M. A. Wainberg, H. Loemba, D. Turner, and B. G. Brenner.** 2003. Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance. J. Antimicrob. Chemother. **51**:229–240.
40. **Surleraux, D. L., A. Tahri, W. G. Verschueren, G. M. Pille, H. A. de Kock, T. H. Jonckers, A. Peeters, S. De Meyer, H. Azijn, R. Pauwels, M. P. de Bethune, N. M. King, M. Prabu-Jeyabalan, C. A. Schiffer, and P. B. Wigerinck.** 2005. Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor. J. Med. Chem. **48**:1813–1822.
41. **Tebit, D. M., I. Nankya, E. J. Arts, and Y. Gao.** 2007. HIV diversity, recombination and disease progression: how does fitness "fit" into the puzzle? AIDS Rev. **9**:75–87.
42. **Velazquez-Campoy, A., M. J. Todd, S. Vega, and E. Freire.** 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proc. Natl. Acad. Sci. U. S. A. **98**:6062–6067.
43. **Velazquez-Campoy, A., S. Vega, E. Fleming, U. Bacha, Y. Sayed, H. W. Dirr, and E. Freire.** 2003. Protease inhibition in African subtypes of HIV-1. AIDS Rev. **5**:165–171.
44. **Williams, T., and C. Kelley.** Accessed 28 January 2009. gnuplot, version 4.2. http://www.gnuplot.info/.
45. **Ziermann, R., K. Limoli, K. Das, E. Arnold, C. J. Petropoulos, and N. T. Parkin.** 2000. A mutation in human immunodeficiency virus type 1 protease, N88S, that causes in vitro hypersensitivity to amprenavir. J. Virol. **74**:4414–4419.

5356 J. Med. Chem. 2010, 53, 5356–5360
DOI: 10.1021/jm1003528

Journal of
Medicinal
Chemistry
Brief Article

Peptide HIV-1 Integrase Inhibitors from HIV-1 Gene Products

Shintaro Suzuki,[†,#] Emiko Urano,[‡,#] Chie Hashimoto,[†] Hiroshi Tsutsumi,[†] Toru Nakahara,[†] Tomohiro Tanaka,[†] Yuta Nakanishi,[†] Kasthuraiah Maddali,[§] Yan Han,[‡] Makiko Hamatake,[‡] Kosuke Miyauchi,[‡] Yves Pommier,[§] John A. Beutler,[⊥] Wataru Sugiura,[‡] Hideyoshi Fuji,[∥] Tyuji Hoshino,[∥] Kyoko Itotani,[†] Wataru Nomura,[†] Tetsuo Narumi,[†] Naoki Yamamoto,[‡] Jun A. Komano,[‡] and Hirokazu Tamamura*[,†]

[†]*Department of Medicinal Chemistry, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, 2-3-10 Kandasurugadai, Chiyoda-ku, Tokyo 101-0062, Japan,* [‡]*AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan,* [§]*Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892-4255,* [∥]*Department of Physical Chemistry, Graduate School of Pharmaceutical Sciences, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan, and* [⊥]*Molecular Targets Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, Maryland 21702.* [#]*These authors contributed equally to this work.*

Received March 17, 2010

Anti-HIV peptides with inhibitory activity against HIV-1 integrase (IN) have been found in overlapping peptide libraries derived from HIV-1 gene products. In a strand transfer assay using IN, inhibitory active peptides with certain sequential motifs related to Vpr- and Env-derived peptides were found. The addition of an octa-arginyl group to the inhibitory peptides caused a remarkable inhibition of the strand transfer and 3′-end-processing reactions catalyzed by IN and significant inhibition against HIV replication.

Introduction

Many antiretroviral drugs are currently available to treat human immunodeficiency virus type 1 (HIV-1) infection. Viral enzymes such as reverse transcriptase (RT[a]), protease and integrase (IN), gp41, and coreceptors are the main targets for antiretroviral drugs that are under development. Because of the emergence of viral strains with multidrug resistance (MDR), however, new anti-HIV-1 drugs operating with different inhibitory mechanisms are required. Following the success of raltegravir, IN has emerged as a prime target. IN is an essential enzyme for the stable infection of host cells because it catalyzes the insertion of viral DNA inside the preintegration complex (PIC) into the genome of host cells in two successive reactions, designated as strand transfer and 3′-end-processing. It is assumed that the enzymatic activities of IN have to be negatively regulated in the PIC during its transfer from the cytoplasm to the nucleus. Otherwise, premature activation of IN can lead to the autointegration into the viral DNA itself, resulting in an aborted infection. We speculate that the virus, rather than the host cells, must encode a mechanism to prevent autointegration. The PIC contains in association with the viral nucleic acid, viral proteins such as RT, IN, capsids ($p24^{CA}$ and $p7^{NC}$), matrix ($p17^{MA}$), p6 and Vpr, cellular proteins HMG I (Y), and the barrier to autointegration factor (BAF).[1−4] It is likely that, due to their spatial proximity in the PIC, these proteins physically and functionally interact with each other. For instance, it is already known that RT activity inhibited by Vpr,[5] and that RT and IN inhibit each other.[5−9] Vpr also inhibits IN through its C-terminal domain.[5,10] Because these studies suggest that PIC components regulate each other's function, we have attempted to obtain potent inhibitory lead compounds from a peptide fragment library derived from HIV-1 gene products, an approach which has been successful in finding a peptide IN inhibitor from LEDGF, a cellular IN binding protein.[11]

In this paper, we describe the screening of an overlapping peptide library derived from HIV-1 proteins, the identification of certain peptide motifs with inhibitory activity against HIV-1 IN, and the evaluation of effective inhibition of HIV-1 replication in cells using the identified peptide inhibitors possessing cell membrane permeability.

Results and Discussion

An overlapping peptide library spanning HIV-1 SF2 *Gag*, *Pol*, *Vpr*, *Tat*, *Rev*, *Vpu*, *Env*, and *Nef*, provided by Dr. Iwamoto of the Institute of Medical Science at the University of Tokyo (Supporting Information, SI, Figure 2A), was screened with a strand transfer assay[12] in search of peptide pools with inhibitory activity against HIV-1 IN. The library consists of 658 peptide fragments derived from the HIV-1 gene products. Each peptide is composed of 10−17 amino acid residues with overlapping regions of 1−7 amino acid residues. Sixteen peptide pools containing between 16 and 65 peptides were used for the first screening at the final concentration of 5.0 μM for each peptide (SI Figure 2B). This initial screening gave the results shown in Figure 1. Both Vpr and Env4 pools showed remarkable inhibition of IN strand transfer activity, and consequently a second screening was performed using the individual peptides contained in the Vpr and Env4 pools. A group of consecutive overlapping peptides in the Vpr pool (groups 13−15) and groups 4−6 and 20−21 in the Env4 pool were found to possess IN inhibitory activity (Figure 2). We focused on Vpr15 and Env4-4 peptides because they showed inhibitory activity against IN strand transfer reaction in a dose-dependent manner (Figure 3). The IC_{50} values of Vpr15

*To whom correspondence should be addressed. Phone: +81-3-5280-8036. Fax: +81-3-5280-8039. E-mail: tamamura.mr@tmd.ac.jp.

[a]Abbreviations: HIV, human immunodeficiency virus; IN, integrase; RT, reverse transcriptase; MDR, multidrug resistance; PIC, preintegration complex; BAF, barrier to autointegration factor; R_8, octa-arginyl.

 Published on Web 06/29/2010

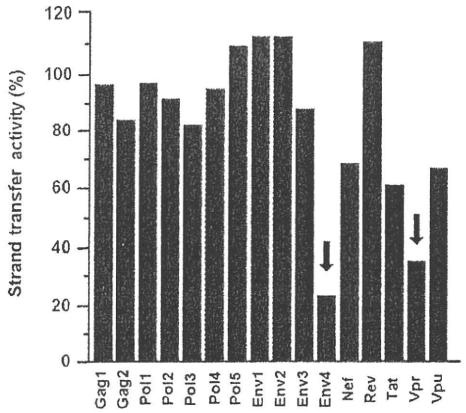

Figure 1. Inhibition of the IN strand transfer activity by peptide pools. Inhibition of the IN strand transfer activity was strongly inhibited by Env4 and Vpr pools (arrows). The *y*-axis represents the IN strand transfer activity relative to the solvent control (DMSO).

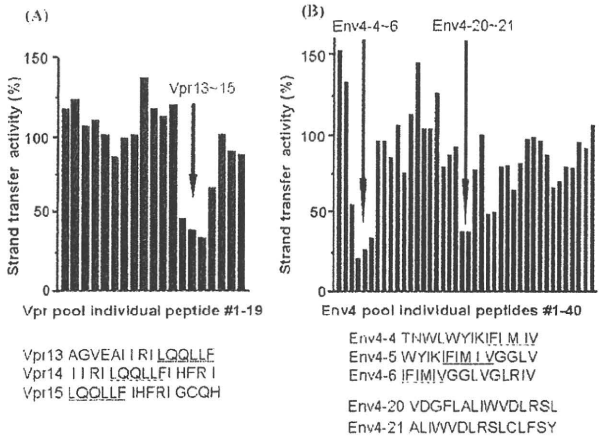

Figure 2. Identification of IN inhibitory peptides in the Vpr (A) and Env4 (B) pools based on the strand transfer activity of IN. The consecutive overlapping peptides display the inhibition of the strand transfer activity of IN (arrows). The *y*-axis represents the IN strand transfer activity relative to the solvent control (DMSO). The concentration of each peptide was 5 μM. The common sequences of individual peptides derived from Vpr and Env4 pools with anti-IN activity are underlined.

and Env4-4 were estimated at 5.5 and 1.9 μM, respectively. These peptides did not show any significant inhibitory activity against HIV-1 RT, suggesting that they might inhibit IN strand transfer reaction selectively.

The overlapping peptides of Vpr13-15 and Env4-4-6 have the vommon hexapeptide sequences LQQLLF and IFIMIV, respectively. The LQQLLF sequence covers positions 64−69 of Vpr, which is a part of the second helix of Vpr. The IFIMIV sequence corresponds to positions 684−689 of gp160, which is a part of the transmembrane domain of TM/gp41. These hexapeptides are thought to be critical to inhibition of IN activity. It was recently reported[5] that similar peptides derived from Vpr inhibit IN with IC_{50} values of 1−16 μM, which is consistent with our data. In this report,[5] the peptide motif was found to be 15 amino acid residues spanning LQQLLF from the overlapping Vpr peptide library. In our study, more precise mapping of inhibitory motif in Vpr peptides was achieved by identifying the shorter effective peptide motif. We focused on the Vpr-derived peptide, LQQLLF (Vpr-1) to develop potent inhibitory peptides. However, the expression of inhibitory activity against IN

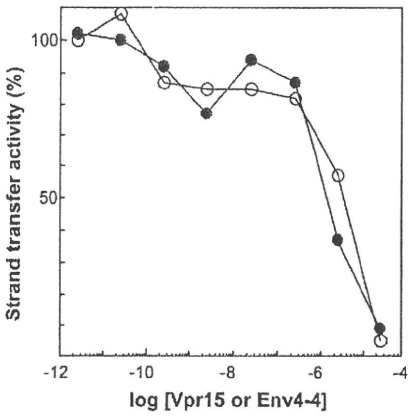

Figure 3. Concentration-dependent inhibition of IN strand transfer activities by Vpr15 (○) and Env4−4 (●) peptides. The *y*-axis represents the IN strand transfer activity relative to the solvent control (DMSO).

in vivo by only hexapeptides might be difficult because these hexapeptides penetrate the plasma membrane very poorly and to achieve antiviral activity, it is essential that they penetrate the cell membrane. To that effect, an octa-arginyl (R_8) group[13] was fused to the Vpr-derived peptides (Table 1). R_8 is a cell membrane permeable motif and its fusion with parent peptides successfully generates bioactive peptides without significant adverse effects or cytotoxicity.[14−18] In addition, the R_8-fusion could increase the solubility of Vpr-derived peptides which have a relatively hydrophobic character.

The inhibitory activity of Vpr-1 and Vpr-1-4 R_8 peptides against IN was evaluated based on the strand transfer and 3′-end-processing reactions in vitro (Table 1).[19,20] Vpr-1 did not show strong inhibition of either IN activity, but the IC_{50} of Vpr-1 R_8 toward the strand transfer reaction of IN was 10-fold lower than that of Vpr-1 lacking the R_8 group. This indicates that the positive charges derived from the R_8 group might enhance the inhibitory activity of the Vpr-1 peptide. Because we were concerned that the strong positive charges close to the LQQLLF motif might interfere with the inhibitory activity, the 6 amino acid sequence (−IHFRIG−) was inserted as a spacer between LQQLLF and R_8 (Vpr-3 R_8). The IHFRIG sequence was used to reconstitute the natural Vpr. The IC_{50} values of Vpr-2 R_8 for the strand transfer and 3′-end-processing activities of IN were 0.70 and 0.83 μM, respectively, while Vpr-3 R_8 showed potent IN inhibitory activities of 4.0 and 8.0 nM against the strand transfer and 3′-end- processing activities, respectively. This result indicates the additional importance of the IHFRIG sequence for inhibitory activities against IN. The increased IN inhibitory activities might be achieved presumably by the synergistic effect of the LQQLLF motif, the IHFRIG sequence, and the R_8 group. Vpr-4 R_8, in which the EAIIRI sequence was attached to further reconstitute the Vpr helix 2, showed inhibitory activities similar to those of Vpr-3 R_8, suggesting that reconstitution of helix 2 of Vpr is not necessary for efficient IN inhibition. Vpr-3 R_8 and Vpr-4 R_8, with $IC_{50} > 0.5$ μM,[21] were less potent inhibitors of RT-associated RNase H activity, indicating that these peptides can selectively inhibit IN. These results suggest that Vpr-derived peptides are novel and distinct from any other IN inhibitors reported to date.

For rapid assessment of the antiviral effect of Vpr-derived peptides, we established an MT-4 Luc system in which MT-4 cells were stably transduced with the firefly luciferase expression cassette by a murine leukemia viral vector (SI Figure 3).

Table 1. Sequences of Vpr-Derived Peptides and Their IC_{50} Values toward the Strand Transfer and 3′-End Processing Reactions of IN

	sequence	IC_{50} (μM) strand transfer	IC_{50} (μM) 3′-end processing
Vpr-1	LQQLLF	68 ± 1.0	> 100
Vpr-1 R_8	Ac-LQQLLF-RRRRRRRR-NH_2	6.1 ± 1.1	> 11
Vpr-2 R_8	Ac-IHFRIG-RRRRRRRR-NH_2	0.70 ± 0.06	0.83 ± 0.07
Vpr-3 R_8	Ac-LQQLLF IHFRIG-RRRRRRRR-NH_2	0.004 ± 0.0001	0.008 ± 0.001
Vpr-4 R_8	Ac-EAIIRI LQQLLF IHFRIG-RRRRRRRR-NH_2	0.005 ± 0.002	0.006 ± 0.006

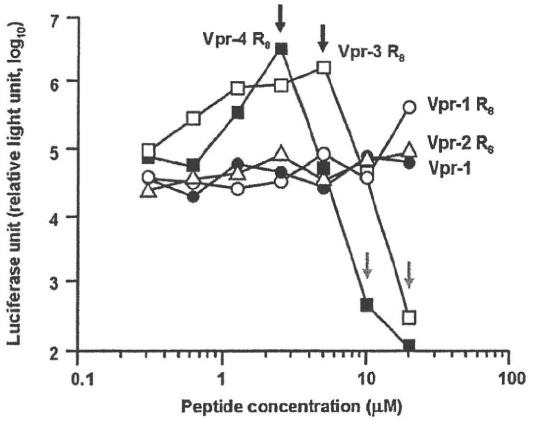

Figure 4. Luciferase signals in MT-4 Luc cells infected with HIV-1 in the presence of various concentrations of Vpr-derived peptides: Vpr-1 (●), Vpr-1 R_8 (○), Vpr-2 R_8 (△), Vpr-3 R_8 (□), Vpr-4 R_8 (■).

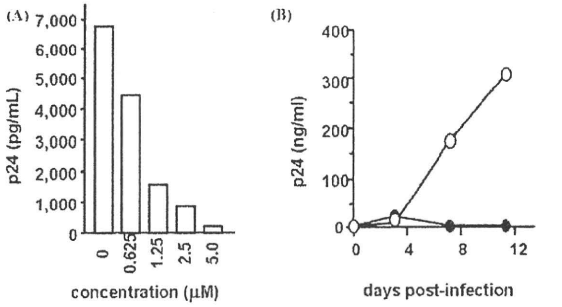

Figure 5. (A) The inhibition of HIV-$1_{JR\text{-}CSF}$ replication in NP2-CD4-CCR5 cells in the presence of various concentrations of Vpr-3 R_8. (B) The replication kinetics of HIV-1_{HXB2} in MT-4 cells in the presence of Vpr-3 R_8 (●). The concentration of Vpr-3 R_8 was fixed at 0.5 μM. Absence of Vpr-3 R_8 (○).

MT-4 Luc cells constitutively express high levels of luciferase which are significantly reduced by HIV-1 infection due to their high susceptibility to cell death upon HIV-1 infection. Protection of MT-4 Luc cells from HIV-1-induced cell death maintains the luciferase signals at high levels. In addition, the cytotoxicity of Vpr-derived peptides can be evaluated by a decrease of luciferase signals in these MT-4 Luc systems. Vpr-2 R_8, which is a weak IN inhibitor, showed no significant anti-HIV-1 activity below concentrations of 20 μM, suggesting that its moderate IC_{50} level in vitro is not sufficient to suppress HIV-1 replication in tissue culture and that the R_8 group is not significantly cytotoxic (Figure 4). Vpr-1 did not show any inhibitory effects against HIV-1 replication; however, Vpr-1 R_8 displayed a weak antiviral effect at a concentration of 20 μM and both Vpr-3 R_8 and Vpr-4 R_8 showed significant inhibitory effects against HIV-1 replication. The R_8 peptide did not show significant anti-HIV activity (IC_{50} > 50 μM, data not shown). These results suggest that the addition of the R_8 group enables Vpr-derived peptides to enter the cytoplasm and access IN, with the result that HIV-1 replication could be effectively inhibited.

Because Vpr-3 R_8 was less cytotoxic than Vpr-4 R_8, the inhibitory activities of Vpr-3 R_8 were further investigated. Two replication assay systems, R5-tropic HIV-$1_{JR\text{-}CSF}$ on NP2-CD4-CCR5 cells and X4-tropic HIV-1_{HXB2} on MT-4 cells, were utilized. NP2-CD4-CCR5 cells were infected with HIV-$1_{JR\text{-}CSF}$ in the presence of various concentrations of Vpr-3 R_8. On day 4 postinfection, the culture supernatant was collected and the concentration of viral p24 antigen was measured by an ELISA assay. The p24 levels decreased in a dose-dependent manner with increasing the concentration of Vpr-3 R_8; 50% inhibition of p24 expression was obtained with approximately 0.8 μM of Vpr-3 R_8 (Figure 5A). This concentration was approximately 10-fold lower than the concentration of Vpr-3 R_8 known to be cytotoxic (Figure 4). Second, MT-4 cells were infected with HIV-1_{HXB2} and the replication kinetics was monitored in the presence of 0.5 μM Vpr-3 R_8. The degree of replication of HIV-1_{HXB2} was quite low in the presence of Vpr-3 R_8, while replication of HIV-1_{HXB2} was robust in the absence of Vpr-3 R_8 (Figure 5B), suggesting that Vpr-3 R_8 strongly suppresses the replication of HIV-1 in cells. To examine whether the HIV-1 replication was blocked through the inhibition of IN activity, quantitative real-time PCR was performed. If IN is inhibited, the efficiency of viral genome integration should be decreased while the reverse transcription of viral genome should not be affected. Accordingly, NP2-CD4-CXCR4 cells were infected with HIV-1_{HXB2} in the presence or absence of 0.5 μM Vpr-3 R_8. Genomic DNA was extracted on day 2 postinfection, and the viral DNA was quantified at the various steps of viral entry phase. The level of "strong stop DNA", representing the total genome of infected virus in Vpr-3 R_8-treated cells, was similar (139.7%) to that in DMSO-treated control cells and the level of viral DNA generated at the late stage of reverse transcription in Vpr-3 R_8-treated cells was slightly decreased (84.4%) compared to control cells. This small decline can probably be attributed to the weak anti-RNase H activity of Vpr-3 R_8. On the other hand, a drastic decrease of Alu-LTR products was observed in Vpr-3 R_8-treated cells (15.8%), indicating an inhibition of integrated viral genome. Concomitantly, the double LTR products, representing the end-joined viral genome catalyzed by host cellular enzymes, were increased by a factor of 8 (779.8%). These results strongly suggest that Vpr-3 R_8 blocks viral infection by inhibiting IN activity in cells, consistent with our in vitro observations. Judging by these results, Vpr-derived peptides with the R_8 group are potent IN inhibitors that suppress HIV-1 replication in vivo.

Finally, in silico molecular docking simulations of Vpr-derived peptides and HIV-1 IN were performed. The Vpr-derived peptides are located in the second helix of Vpr and were thus considered to have an α-helical conformation.[22] Docking simulations of three peptides (Vpr13, Vpr14, and Vpr15), using the predicted structure of the HIV-1 IN dimer as a template,[23] were performed by GOLD software to investigate the binding mode of the peptides, the binding affinity of

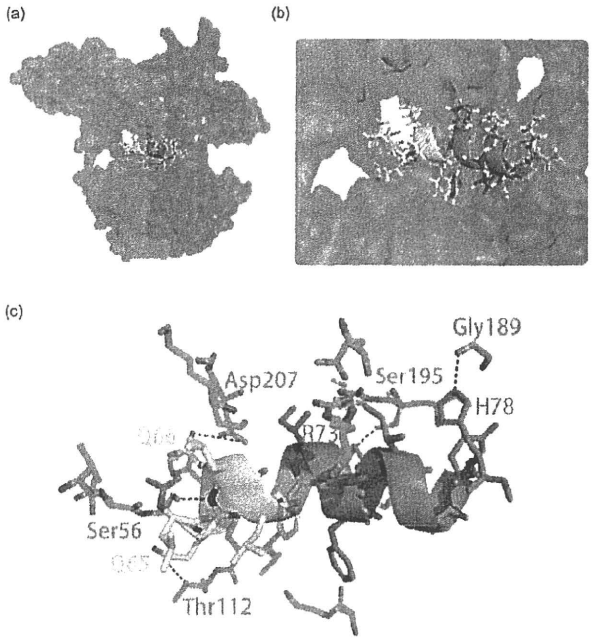

Figure 6. Predicted binding mode of Vpr15 to HIV-1 IN by GOLD. An overall view of (a) the complex obtained by docking Vpr15 with the HIV-1 IN dimer and (b) the closer view of the complex. The predicted structure of full-length HIV-1 IN was used as a template. Each HIV-1 IN monomer was shown as green or cyan surface. The docked Vpr15 is shown as a cartoon. The yellow-colored region is the LQQLLF motif. The GOLD score representing the docking complementarity is 69.83, indicating the high binding affinity between Vpr15 and IN. The hydrogen-bond interactions between HIV-1 IN and Vpr15 were presented by LIGPLOT software shown as blue dotted line (c).

the peptides being evaluated by GOLD Fitness score. The predicted binding mode of Vpr15 to IN is shown in Figure 6. Our results predict that the three Vpr-derived peptides interact with the cleft between the amino-terminal domain and the core domain of HIV-1 IN. This region is distinct from the nucleic acid interacting surfaces, indicating that the Vpr-derived peptides inhibit IN function in an allosteric manner. A previous report provided a model in which a Vpr peptide was bound to IN in a manner similar with our model[5] and, interestingly, the peptides were bound to IN with an exterior surface of Vpr. This earlier report that the full-length Vpr inhibits IN[10] strongly supports the predicted binding mode of Vpr15. Five hydrogen-bond interactions between HIV-1 IN and Vpr15 were identified by LIGPLOT analysis,[24] which invoked the following IN-Vpr amino acids: IN Thr112-Vpr Gln65, IN Ser56-Vpr Gln66, IN Asp207-Vpr Gln66, IN Ser195-Vpr Arg73, and IN Gly189-Vpr His78. The numbering of Vpr amino acids is based on the Vpr full-length coordinate, Figure 6. Additional hydrophobic contacts between IN and Vpr15 were found in which the following IN-Vpr amino acid pairs are involved: IN Lys211-Vpr Gln66, IN Pro109-Vpr Phe69, IN Arg262-Vpr His71, and IN Arg187-Vpr Gln77. These data indicate that the Gln65, Gln66, and Phe69 residues in Vpr-derived peptides play a major role in the interaction between IN and Vpr-derived peptides.

Conclusions

In summary, two peptide motifs, LQQLLF from Vpr and IFIMIV from Env4, possessing inhibitory activity against HIV-1 IN, were identified through the screening of overlapping peptide library derived from HIV-1 gene products. We initially speculate that HIV encodes a mechanism to prevent autointegration in the PIC because integration activity must be regulated until the virus infects cells. This speculation is supported by the finding that IN inhibitors exist in the viral PIC components. Vpr-derived peptides with the R_8 group showed remarkable inhibitory activities against the strand transfer and 3′-end-processing reactions catalyzed by HIV-1 IN in vitro. In addition, Vpr-3 R_8 and Vpr-4 R_8 were shown to inhibit HIV-1 replication with submicromolar IC_{50} values in cells using the MT-4 Luc cell system. In the quantitative analysis of p24 antigen, 50% inhibition of HIV-$1_{JR\text{-}CSF}$ replication was caused by approximately 0.8 μM of Vpr-3 R_8, and the replication of HIV-1_{HXB2} was extensively suppressed in the long term by Vpr-3 R_8 at 0.5 μM concentrations. Our finding suggest that these peptides could serve as lead compounds for novel IN inhibitors. Amino acid residues critical to the interaction of Vpr-derived peptides with IN were identified by our in silico molecular docking simulations, and suggests that more potent peptides[25] or peptidomimetic IN inhibitors represent a novel avenue for future small molecule inhibitors of IN and HIV integration.

Experimental Section

Peptide Synthesis. Vpr-derived peptides containing the R_8 group were synthesized by stepwise elongation techniques of Fmoc-protected amino acids on NovaSyn TGR resin. Coupling reactions were performed using 5.0 equiv of Fmoc-protected amino acid, 5.0 equiv of diisopropylcarbodiimide, and 5.0 equiv of 1-hydroxybenzotriazole monohydrate. Cleavage of peptides from resin and side chain deprotection were carried out with 10 mL of TFA in the presence of 0.25 mL of *m*-cresol, 0.75 mL of thioanisole, 0.75 mL of 1,2-ethanedithiol, and 0.1 mL of water as scavenger by stirring for 1.5 h. After filtration of the deprotected peptides, the filtrate was concentrated under reduced pressure, and crude peptides were precipitated in cooled diethylether. All crude peptides were purified by RP-HPLC and identified by MALDI-TOF MS. Purities of all final compounds were confirmed (>95% purity) by analytical HPLC. Detailed data are provided in SI.

Enzyme Assays. The strand transfer assay for the first screening was performed as described previously.[12] The IN strand transfer and 3′-end-processing assays for peptide motif characterizations were performed as described previously.[19,20] RNase H activity was measured as described by Beutler et al.[21]

Replication Assays. For HIV-1 replication assays, 1×10^5 cells were incubated at room temperature for 30 min with an HIV-1 containing culture supernatant (ca. 0.2−50 ng p24) and then washed and incubated. Culture supernatants were collected at different time points, and then the cells were passaged if necessary. Levels of p24 antigen were measured using a Retro TEK p24 antigen ELISA kit, according to the manufacture's protocol. Signals were detected using an ELx808 microplate photometer.

For MT-4 Luc assays, MT-4 Luc cells (1×10^3 cells) grown in 96-well plates were infected with HIV-1_{XHB2} (ca. 0.2−10 ng p24) in the presence of varying concentrations of Vpr-3 R_8. At 6−7 d postinfection, cells were lysed and luciferase activity was measured using the Steady-Glo assay kits according to the manufacture's protocol. Chemiluminescence was detected with a Veritas luminometer.

Acknowledgment. We thank Prof. A. Iwamoto's group of the Institute of Medical Science at the University of Tokyo for the peptide libraries and Dr. M. Nicklaus from NCI/NIH for providing the modeled structure of full-length HIV-1 IN. T.T. is supported by JSPS research fellowships for young scientists.

This work was supported in part by Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and Health and Labor Sciences Research Grants from Japanese Ministry of Health, Labor, and Welfare. K.M. and Y.P. are supported by the Intramural Program of the National Cancer Institute, Center for Cancer Research.

Supporting Information Available: Additional experimental procedures including MS data and figures; HPLC charts of final compounds, explanation for HIV-1 genes and the peptide pools, and illustration of MT-4 Luc system. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Bukrinsky, M. I.; Haggerty, S.; Dempsey, M. P.; Sharova, N.; Adzhubei, A.; Spitz, L.; Lewis, P.; Goldfarb, D.; Emerman, M.; Stevenson, M. A nuclear-localization signal within HIV-1 matrix protein that governs infection of nondividing cells. *Nature* **1993**, *365*, 666–669.

(2) Miller, M. D.; Farnet, C. M.; Bushman, F. D. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J. Virol.* **1997**, *71*, 5382–5390.

(3) Farnet, C. M.; Bushman, F. D. HIV-1 cDNA integration: Requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell* **1997**, *88*, 483–492.

(4) Chen, H.; Engelman, A. The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15270–15274.

(5) Gleenberg, I. O.; Herschhorn, A.; Hizi, A. Inhibition of the activities of reverse transcriptase and integrase of human immunodeficiency virus type-1 by peptides derived from the homologous viral protein R (Vpr). *J. Mol. Biol.* **2007**, *369*, 1230–1243.

(6) Gleenberg, I. O.; Avidan, O.; Goldgur, Y.; Herschhorn, A.; Hizi, A. Peptides derived from the reverse transcriptase of human immunodeficiency virus type 1 as novel inhibitors of the viral integrase. *J. Biol. Chem.* **2005**, *280*, 21987–21996.

(7) Hehl, E. A.; Joshi, P.; Kalpana, G. V.; Prasad, V. R. Interaction between human immunodeficiency virus type 1 reverse transcriptase and integrase proteins. *J. Virol.* **2004**, *78*, 5056–5067.

(8) Tasara, T.; Maga, G.; Hottiger, M. O.; Hubscher, U. HIV-1 reverse transcriptase and integrase enzymes physically interact and inhibit each other. *FEBS Lett.* **2001**, *507*, 39–44.

(9) Gleenberg, I. O.; Herschhorn, A.; Goldgur, Y.; Hizi, A. Inhibition of human immunodeficiency virus type-1 reverse transcriptase by a novel peptide derived from the viral integrase. *Arch. Biochem. Biophys.* **2007**, *458*, 202–212.

(10) Bischerour, J.; Taue, P.; Leh, H.; De Rocquigny, H.; Roques, B.; Mouscadet, J. F. The (52–96) C-Terminal domain of Vpr stimulates HIV-1 IN-mediated homologous strand transfer of mini-viral DNA. *Nucleic Acids Res.* **2003**, *31*, 2694–2702.

(11) Hayouka, Z.; Rosenbluh, J.; Levin, A.; Loya, S.; Lebendiker, M.; Veprintsev, D.; Kotler, M.; Hizi, A.; Loyter, A.; Friedler, A. Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8316–8312.

(12) Yan, H.; Mizutani, T. C.; Nomura, N.; Tanaka, T.; Kitamura, Y.; Miura, H.; Nishizawa, M.; Tatsumi, M.; Yamamoto, N.; Sugiura, W. A novel small molecular weight compound with a carbazole structure that demonstrates potent human immunodeficiency virus type-1 integrase inhibitory activity. *Antivir. Chem. Chemother.* **2005**, *16*, 363–373.

(13) Suzuki, T.; Futaki, S.; Niwa, M.; Tanaka, S.; Ueda, K.; Sugiura, Y. Possible existence of common internalization mechanisms among arginine-rich peptides. *J. Biol. Chem.* **2002**, *277*, 2437–2443.

(14) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13003–13008.

(15) Matsushita, M.; Tomizawa, K.; Moriwaki, A.; Li, S. T.; Terada, H.; Matsui, H. A high-efficiency protein transduction system demonstrating the role of PKA in long-lasting long-term potentiation. *J. Neurosci.* **2001**, *21*, 6000–6007.

(16) Takenobu, T.; Tomizawa, K.; Matsushita, M.; Li, S. T.; Moriwaki, A.; Lu, Y. F.; Matsui, H. Development of p53 protein transduction therapy using membrane-permeable peptides and the application to oral cancer cells. *Mol. Cancer Ther.* **2002**, *1*, 1043–1049.

(17) Wu, H. Y.; Tomizawa, K.; Matsushita, M.; Lu, Y. F.; Li, S. T.; Matsui, H. Poly-arginine-fused calpastatin peptide, a living cell membrane-permeable and specific inhibitor for calpain. *Neurosci. Res.* **2003**, *47*, 131–135.

(18) Rothbard, J. B.; Garlington, S.; Lin, Q.; Kirschberg, T.; Kreider, E.; McGrane, P. L.; Wender, P. A.; Khavari, P. A. Conjugation of arginine oligomers to cyclosporin A facilitates topical delivery and inhibition of inflammation. *Nature Med.* **2000**, *6*, 1253–1257.

(19) Marchand, C.; Zhang, X.; Pais, G. C. G.; Cowansage, K.; Neamati, N.; Burke, T. R., Jr.; Pommier, Y. Structural determinants for HIV-1 integrase inhibition by beta-diketo acids. *J. Biol. Chem.* **2002**, *277*, 12596–12603.

(20) Semenova, E. A.; Johnson, A. A.; Marchand, C.; Davis, D. A.; Tarchoan, R.; Pommier, Y. Preferential inhibition of the magnesium-dependent strand transfer reaction of HIV-1 integrase by alpha-hydroxytropolones. *Mol. Pharmacol.* **2006**, *69*, 1454–1460.

(21) Parniak, M. A.; Min, K. L.; Budihas, S. R.; Le Grice, S. F. J.; Beutler, J. A. A fluorescence-based high-throughput screening assay for inhibitors of HIV-1 reverse transcriptase associated ribonuclease H activity. *Anal. Biochem.* **2003**, *322*, 33–39.

(22) Morellet, N.; Bouaziz, S.; Petitjean, P.; Roques, B. P. NMR structure of the HIV-1 regulatory protein Vpr. *J. Mol. Biol.* **2003**, *327*, 215–227.

(23) Karki, R. G.; Tang, Y.; Burke, T. R., Jr.; Nicklaus, M. C. Model of full-length HIV-1 integrase complexed with viral DNA as template for anti-HIV drug design. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 739–760.

(24) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT—a program to generate schematic diagrams of protein ligand interactions. *Protein Eng.* **1995**, *8*, 127–134.

(25) Li, H.-Y.; Zawahir, Z.; Song, L.-D.; Long, Y.-Q.; Neamati, N. Sequence-based design and discovery of peptide inhibitors of HIV-1 integrase: insight into the binding mode of the enzyme. *J. Med. Chem.* **2006**, *49*, 4477–4486.

Antiviral Research 87 (2010) 22–29

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Drug-resistant mutation patterns in CRF01_AE cases that failed d4T + 3TC + nevirapine fixed-dosed, combination treatment: Follow-up study from the Lampang cohort

Siriphan Saeng-aroon[a], Naho Tsuchiya[b], Wattana Auwanit[a], Panasda Isarangura Na Ayuthaya[a], Panita Pathipvanich[c], Pathom Sawanpanyalert[a], Archawin Rojanawiwat[a], Mari Kannagi[d], Koya Ariyoshi[b], Wataru Sugiura[e,f,g,*]

[a] National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand
[b] Global COE Program/Department of Clinical Medicine, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan
[c] Lampang Hospital, Lampang, Thailand
[d] Department of Immunotherapeutics, Tokyo Medical and Dental University, Tokyo, Japan
[e] AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan
[f] Department of Infection and Immunology, Clinical Research Center, Nagoya Medical Center, Nagoya, Japan
[g] Department of AIDS Research, Nagoya Graduate School of Medicine, Nagoya, Japan

ARTICLE INFO

Article history:
Received 11 November 2009
Received in revised form 25 March 2010
Accepted 1 April 2010

Keywords:
CRF01_AE
Polymorphism
Drug resistance
Connection domain
RNase H
GPOvir

ABSTRACT

HIV/AIDS patients are treated in Thailand's national antiretroviral treatment (ART) program with a generic combination tablet of stavudine, lamivudine, and nevirapine (GPOvir). To determine GPOvir-resistant mutations, HIV-1 sequences of 59 GPOvir-failure cases from the Lampang cohort were compared with sequences from 76 randomly selected ART-naïve cases. The GPOvir-failure cases had not only known stavudine-, lamivudine- and nevirapine-resistant mutations, but also V118I, G196E, and H221Y. Among the 59 GPOvir-failure cases, 29 were ART-naïve prior to GPOvir (naïve group), and 30 had previous ART (exposed group). To clarify the effect of previous ART in drug-resistant acquisition pathways, naïve and exposed groups were compared. The exposed group had predominantly thymidine analogue-related mutations, whereas the naïve group had a higher prevalence of Q151M and K103N mutations. M184V lamivudine resistance was most frequent in both naïve and exposed groups. To identify which mutations in CRF01_AE pol were polymorphisms, the connection and RNase domains were also analyzed. CRF01_AE-specific polymorphisms were found in 19 residues, and GPOvir-failure cases had significantly higher frequency of N348I, E399D, P537S, and I542M. Our results expand identification of mutations in CRF01_AE pol that are polymorphisms by also analyzing the connection and RNase H domains.

1. Introduction

The number of people living with HIV/AIDS in Thailand at the end of 2008 was 532,500 (Ministry of Public Health, 2008). In resource-poor countries such as Thailand, the recommended first-line regimen for treating HIV/AIDS is a combination of two nucleoside reverse transcriptase inhibitors (NRTIs) and one non-nucleoside reverse transcriptase inhibitor (NNRTIs) (WHO, 2003). HIV/AIDS patients in Thailand have been treated since 2002 through the national antiretroviral treatment (ART) program with a generic, fixed-dosed single tablet (GPOvir) with 3 antiretroviral agents: stavudine (d4T), lamivudine (3TC) and nevirapine (NVP). The major reason for plasma viral load rebound and treatment failure remains the emergence of drug resistance. Therefore, HIV drug-resistance genotypic testing (HIV genotyping) has become an important tool in deciding about appropriate treatment regimens. HIV genotyping, i.e., the determination of mutations that confer drug resistance, is now widely established as the standard of care to guide treatment in the context of both primary infection and virological failure (Hirsch et al., 2003). To date, the design and development of antiretroviral drugs, research on drug resistance, and interpretation systems have been largely based on the HIV-1 subtype B virus, the major subtype in developed countries. However, the findings on subtype B may not always be

* Corresponding author. Present address: Clinical Research Center, Nagoya Medical Center, 4-1-1 Sannomaru, Nakaku, Nagoya 4600001, Japan.
Tel.: +81 52 951 1111; fax: +81 52 963 3970.
E-mail address: wsugiura@nih.go.jp (W. Sugiura).

0166-3542/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2010.04.001

Table 1
Demographic and clinical characteristics of GPOvir virologic-failure cases at 6 or 24 months ($n = 59$).

Variable	ART-naïve cases ($n = 29$) n (%)	ART-experienced cases ($n = 30$) n (%)	Odds ratio
Age (years)			
<35	19 (65.5)	22 (73.3)	0.70
≥35	10 (34.5)	8 (26.7)	
Gender			
Male	20 (69.0)	18 (60.0)	1.47
Female	9 (31.0)	12 (40.0)	
CD4 at baseline (cells/μl)			
<50	16 (55.2)	15 (50.0)	1.27
≥50	10 (34.5)	12 (40.0)	
Unknown	3 (10.3)	3 (10.0)	
AIDS symptoms			
Asymptomatic	3 (10.3)	2 (6.7)	1.67
AIDS/symptomatic	25 (86.2)	28 (93.3)	
Unknown	1 (3.5)	0 (0.0)	
Route of transmission			
Heterosexual	27 (93.1)	29 (96.7)	0.47
Homosexual	2 (6.9)	1 (3.3)	

applicable to other subtypes, and some minor mutations, which are recognized as drug-resistant mutations in subtype B, exist as natural variants in non-B subtypes (Kantor and Katzenstein, 2003).

Furthermore, under antiretroviral treatment certain subtypes select specific mutations that are different from those of subtype B (Brenner et al., 2003; Grossman et al., 2004; Loemba et al., 2002). For example, in data on GPOvir-failure cases collected from 7 hospitals in Thailand, where the most prevalent subtype is CRF01_AE, the most commonly reported drug-resistant mutations were G190S/A and Y181C/I; and K103N, Y181C/I, M184V/I were significantly associated with efavirenz, NVP, 3TC, respectively (Chetchotisakd et al., 2006). The reported pattern was identical to that of subtype B. Interestingly 26% of cases in that study had received dual-NRTI treatment before the GPOvir regimen, but ART-naïve and ART-experienced patients were not analyzed in detail for differences in resistance-acquisition patterns. However, another study of drug-resistance mutation patterns among GPOvir-failure cases in Thailand found that the most frequent resistance mutation was M184V, with higher frequencies of K65R (6%) and Q151M (8%) than for subtype B (Sungkanuparph et al., 2007). Thus, there are differences in the reported drug-resistance patterns after GPOvir treatment, indicating the need for further data on drug resistance of GPOvir-resistant cases to better understand drug-resistance acquisition patterns in CRF01_AE.

Therefore, the aim of this study was to clarify drug-resistance mutation pattern in GPOvir treatment-failure cases from the Lampang cohort (Tsuchiya et al., 2009). To understand the effect of previous antiretroviral exposure in GPOvir-resistance acquisition, we analyzed data not only from ART-naïve cases but also from those previously treated with mono- or dual therapies. Recently, several studies demonstrated that resistance to NRTI and/or NNRTI therapies is enhanced and the balance between nucleotide excision and template RNA degradation is affected by mutations in the connection domain and RNase H region (Brehm et al., 2007; Delviks-Frankenberry et al., 2007; Ehteshami et al., 2008; Julias et al., 2003; Nikolenko et al., 2007; Ntemgwa et al., 2007; Santos et al., 2008; Waters et al., 2009; Yap et al., 2007; Zelina et al., 2008). However, these studies were mostly on subtype B, with less information on non-B subtypes. Therefore, to clarify the effect of CRF01_AE mutations in the connection domain and RNase H region, we analyzed the sequences of these domains in GPOvir treatment-failure cases.

2. Materials and methods

2.1. Samples

Plasma samples were collected from patients in Lampang Hospital, a government referral hospital in Lampang province of northern Thailand. In total, 345 HIV-1-infected Thai patients agreed and started GPOvir therapy at the hospital's Day Care Center clinic between 1 April 2002 and 31 January 2004. Of these 345 cases, 244 cases were ART-naïve, and 101 cases had been exposed to ART before initiating GPOvir treatment (baseline). Their plasma samples were collected and analyzed for HIV-1 sequences at baseline and at different time points until the end-point of 24-month follow-up or a switch in therapy. Treatment-failure cases were defined as cases with a detectable viral load (>50 copies/ml) despite having received GPOvir therapy for at least 3 months; this criterion was met by 78 cases. However, 19 cases were excluded from the study for the following reasons: 9 changed to other treatment, 1 had poor adherence, 6 changed to undetectable viral load, and 3 had unknown treatment histories. Samples from the remaining 59 cases were sequenced. Their demographics and clinical variables are summarized in Table 1. These cases were separated into two groups: ART-naïve and ART-exposed. The ART-naïve cases ($n = 29$, 49.1%) had never been exposed to antiretrovirals prior to GPOvir treatment and the ART-exposed cases ($n = 30$, 50.9%) had been exposed to antiretrovirals. These groups did not differ significantly in terms of clinical variables. Most patients were infected with HIV-1 through heterosexual contact ($n = 56$, 94.9%), which did not differ from the HIV-1 transmission pattern in the whole GPOvir study population (Tsuchiya et al., 2009). In the ART-exposed group, the most common treatment regimen was dual therapy with AZT and ddC or ddI (Table 2); none were previously exposed to NNRTIs.

This study was conducted according to principles of the Declaration of Helsinki, the Lampang HIV study was approved by the Thai government ethics committee, and written informed consent was obtained from patients who agreed to join this study.

2.2. Sequencing of the RT and RNase H genes

All samples were determined for viral load, and when it exceeded 1000 copies/ml, the RT region (residues 1–240) was tested for drug resistance using an in-house genotyping protocol reported elsewhere (Myint et al., 2002; Saeng-Aroon et al., 2007).

Table 2
Previous antiretroviral treatment histories of drug-experienced cases ($n = 30$).

Treatment regimen	n
AZT/ddC	14
AZT/ddI	4
AZT	1
AZT/ddC/RTV	1
AZT/ddI/RTV	1
AZT/ddI/IDV	1
SQV/RTV	1
AZT/ddI–AZT/ddC	1
GPOvir	1
Unknown	5

Note: AZT: azidothymidine; ddC: dideoxycytidine; ddI: dideoxyinosine; RTV: ritonavir; IDV: indinavir; SQV: saquinavir; GPOvir: stavudine, lamivudine and nevirapine.

In brief, HIV RNA was extracted from 140 μl plasma using the NucleoSpin viral RNA extraction kit (NucleoSpin, Duren, Germany) following the Manufacturer's instructions. cDNA and PCR product were then obtained using the SuperScript III One Step RT-PCR kit (Invitrogen, Carlsbad, CA) and primers listed in Table 3.

For amplifying the extended RT regions, connection and RNase H domains, new primers were designed. Outer PCR was performed with RT1 and GPR2M, while nested PCR was performed with primers RT7L and GPR3L. The amplification profile for outer PCR was 40 min at 55 °C, 2 min at 95 °C followed by 40 cycles of 15 s at 95 °C, 15 s at 55 °C and 1.30 min at 72 °C, and 7 min at 72 °C. The reaction mixture for the nested PCR contained 3.5 μl of the product from the first PCR. The amplification profile in the second PCR was 2 min at 92 °C followed by 30 cycles of 10 s at 94 °C, 4 s at 60 °C and 15 s at 74 °C, and 7 min at 72 °C. The amplicon (1700 bps) represented the HIV-1 pol region, spanning the RT region, connection domain, and RNase H domain. Both strands of the PCR product were sequenced using six different primers and BigDye® Terminator v3.1 chemistry on an ABI 3100 Genetic Analyzer. SeqScape software version 2.5 was used for editing and assembling sequence fragments, and the assembled sequences were compared with the reference strain HBX2 from the Los Alamos HIV sequence database. To obtain maximum prevalence of drug-resistance mutations, codons with wild type and resistant mixtures were counted as resistance positive.

In addition to sequencing and analyzing the 5′ 240-amino acid RT region, we clarified the substitution patterns of the connection and RNase H domains under GPOvir treatment by sequencing and analyzing 49 plasma samples of GPOvir-failure cases collected at their last visit. As naïve control sequences, the connection and RNase H domains were sequenced from 76 randomly selected HIV-1 CRF01_AE ART-naïve cases from the same hospital.

2.3. Data analysis, determination of subtypes, drug-resistant mutations and polymorphisms

To confirm that patients were infected with HIV-1 CRF01_AE, all nucleotide sequences were aligned using Clustal W, version 2.0.10 and BioEdit, version 7.0.9.0. Phylogenetic tree and bioinformatics analyses were conducted using MEGA, version 4 (Tamura et al., 2007). The genetic distances were calculated using Kimura's 2-parameter analysis (Kimura, 1981), and phylogenetic trees were constructed by the neighbor-joining method. The overall polymorphism of RT genes was analyzed by comparing 76 CRF01_AE sequences from our cohort and reference sequences, all treatment-naïve cases, from Los Alamos HIV sequence database (http://www.hiv.lanl.gov). The following sequences were obtained from the Los Alamos HIV sequence database: 42 subtype B (accession no. EF637056, DQ837381, DQ676874, EF637057, DQ676870, DQ676877, DQ676880, EF363122, DQ127537, BD455696, K03455, EF363124, U69584, EF637053, DQ487190, AY314044, EF363122, DQ007902, DQ007903, DQ990880, AY945710, EF363127, DQ396398, EF637054, DQ207940, EF637048, EF175209, EF637051, DQ853436, EF637050, EF637049, DQ676886, DQ127548, DQ207942, AB428551, AB428557, AB428556, AB428553, AB428554, AB428552, AB428555, AB428561), 26 CRF01_AE (DQ859178, DQ859179, DQ859180, EF036527, EF036528, EF036529, EF036530, EF036531, EF036532, EF036533, EF036534, AY945712, AY945713, DQ789392, AY945716, AY945717, AY945719, AY945720, AY945721, AY945722, AY945724, AY945727, AY945728, AY945730, AY945731, AY945732).

Resistance-related mutations were based on guidelines published by the International AIDS Society United States (IAS-USA) HIV Resistance Testing Guideline Panel 2008 (Johnson et al., 2008) according to the subtype B consensus strain. To compare the distribution of qualitative variables according to groups, χ^2-test was used or the Fisher exact test when the sample was too small. All statistical tests were interpreted at the 5% significance level.

Table 3
Primers used for amplification and sequencing.

Name	Region	Usage	Primer sequence 5′–3′	Position
RT1L	RT	Outer forward	ATGATAGGGGGAATTGGAGGTTT	2388–2410
RT4L	RT	Outer reverse	TACTTCTGTTAGTGCTTTGGTTCC	3402–3425
RT7L	RT	Inner forward	GACCTACACCTGTCAACATAATTGG	2485–2509
RT6L	RT	Inner reverse	TAATCCCTGCATAAATCTGACTTGC	3348–3372
RT7L	RT	Sequencing	GACCTACACCTGTCAACATAATTGG	2485–2509
RT26	RT	Sequencing	CAAAAATTGGGCCTGAAAATCC	2692–2713
RT28	RT	Sequencing	TGGAATATTGCTGGTGATCC	3012–3031
RT6L	RT	Sequencing	TAATCCCTGCATAAATCTGACTTGC	3348–3372
RT1L	RT	Outer forward	ATGATAGGGGGAATTGGAGGTTT	2388–2410
GPR2M	Connection and Rnase H	Outer reverse	GGACTACAGTCYACTTGTCCATG	4380–4402
RT7L	RT	Inner forward	GACCTACACCTGTCAACATAATTGG	2485–2509
GPR3L	Connection and Rnase H	Inner reverse	TTAAAATCACTARCCATTGYTCTCC	4285–4309
RT7L	RT	Sequencing	GACCTACACCTGTCAACATAATTGG	2485–2509
RT26	RT	Sequencing	CAAAAATTGGGCCTGAAAATCC	2692–2713
RT28	RT	Sequencing	TGGAATATTGCTGGTGATCC	3012–3031
RT31	RT	Sequencing	GAGCTCATCTATTGAGCTGG	3166–3185
RT32	RT	Sequencing	GAACCTCCATTCCTTTGGATGGG	3219–3241
RT6L	RT	Sequencing	TAATCCCTGCATAAATCTGACTTGC	3348–3372
RT35	Connection and Rnase H	Sequencing	GCAGAAGTACAGAAACAAGG	3528–3547
GPR3L	Connection and Rnase H	Sequencing	TTAAAATCACTARCCATTGYTCTCC	4285–4309

Table 4
Known mutations associated with drug resistance (residues 1–240).

HXB2	Residue	Mutation frequency		p
		Naïve group ($n=76$)	Failure group ($n=49$)	
M	41	M(76)	M(34),L(15)	<0.001
A	62	A(76)	A(47),V(2)	–
K	65	K(76)	K(47),R(2)	–
D	67	D(76)	N(26),D(22),G(1)	<0.001
K	70	K(76)	K(38),R(10),G(1)	<0.001
L	74	L(76)	L(47),I(2)	–
V	75	V(76)	V(46),I(3)	–
F	77	F(76)	F(48),L(1)	–
A	98	A(76)	A(45),G(3),S(1)	–
L	100	L(76)	L(49)	–
K	101	K(76)	K(36),E(12),H(1)	<0.001
K	103	K(76)	K(36),N(11),S(2)	<0.001
V	106	V(74),I(2)	V(45),I(4)	–
V	**108**	**V(76)**	**V(44),I(5)**	**<0.05**
Y	115	Y(76)	Y(49)	–
F	116	F(76)	F(44),Y(5)	<0.05
V	118	V(76)	V(43), I(6)	<0.05
Q	151	Q(76)	Q(43),M(6)	<0.05
Y	181	Y(76)	Y(22),C(23),V(4)	<0.001
V	179	V(68),I(6),IV(2)	V(35),I(10),D(1),IV(3)	–
M	184	M(76)	V(37),I(8),M(4)	<0.001
Y	188	Y(76)	Y(48),L(1)	–
G	190	G(76)	G(29),A(18),S(2)	<0.001
G	**196**	**G(76)**	**G(44),E(5)**	**<0.05**
L	210	L(76)	L(37),W(12)	<0.001
T	215	T(76)	T(27),F(12),Y(10)	<0.001
K	219	K(76)	K(42),Q(6),E(1)	<0.05
H	**221**	**H(76)**	**H(42),Y(7)**	**<0.05**
P	225	P(76)	P(48),H(1)	–

Subtype B consensus residues are displayed on the left side of each position.
Bold represents new substitutions not previously reported for CRF01_AE.

3. Results

3.1. New patterns of drug-resistance mutations emerge in CRF01_AE GPOvir-failure cases

Drug-resistance mutations related to GPOvir failures are summarized in Table 4. Sequences were compared between 49 samples of treatment-failure cases collected at their last visit and 76 randomly selected treatment-naïve samples at baseline. Almost all of the known mutations associated with d4T/3TC/nevirapine treatment were significantly higher in the GPOvir-failure cases, except the 4 following mutations: K65R, L100I, V106M/A, and Y188C/L/H. Other than known mutations, V118I and H221Y were observed in significantly higher prevalence in the GPOvir-failure group ($p<0.05$). Interestingly, cases with the H221Y mutation all had Y181C, and the linkage of the two mutations was statistically significant ($p<0.05$), as previously reported (Liu et al., 2007). A likely role for these mutations in resistance to NRTIs has been suggested by a report of pre- and post-treatment frequencies of H221Y (0–13.7%) in subtype C isolates from India (Deshpande et al., 2007). In subtype B isolates, H221Y and D223E/Q were associated with therapy only if individuals receiving both NRTI and NNRTI were included. G196E was also reported to be significantly different in subtype B ($p<0.05$) (Gonzales et al., 2003).

3.2. Drug-resistance mutations, especially d4T resistance-related mutations, are more prevalent in the ART-exposed group than the naïve group

The frequencies of drug-resistance mutations in the ART-naïve and -exposed groups at baseline, 6, 12, 18, and 24 months are shown in Fig. 1. The two groups differed significantly in their resistance mutation-acquisition patterns. The most apparent difference was the frequency of d4T resistance-related mutations. The exposed group had significantly higher frequencies of mutations M41L, D67N, K70R, L210W, T215Y/F and K219Q/E, most of which already existed at baseline. As for K70R, T215Y/F and K219Q/E, their frequencies at difference time points did not change during the observation period, whereas the frequencies of M41L and D67N increased from 10% (4 of 40) to 20% (9 of 46) and from 10% (4 of 40) to 26% (12 of 46), respectively.

In contrast to the exposed group, a few cases in the naïve group acquired d4T resistance-related mutations, similar to a pattern previously reported (Arts et al., 1998; Lacey and Larder, 1994). Frequencies of d4T resistance-related mutations did not increase during the observation period in the naïve group, except for the D67N mutation. The prevalence of D67N in the naïve group increased from 0% (baseline) to 15% (7 of 46) at 24 months. Thus, GPOvir appears to be selecting the D67N mutation. As described above, the exposed group showed similar findings; D67N prevalence increased from 10% (4 of 40) at baseline to 26% (12 of 46) at 24 months. The 10% of cases at baseline in the exposed group can be explained by previous AZT exposure, and the additional 16% might result from induction and selection by d4T administration.

Another notable finding of our study is the detection of the Q151M multi-drug-resistant mutation. In the ART-naïve group, 5 cases acquired Q151M during the observation period. As Q151M prevalence increased over the treatment period, from 0% at baseline to 11% (5 of 46) at 24 months, it is clear that GPOvir treatment selected the Q151M mutation. Interestingly, the prevalence of Q151M in our study is similar to that of two previous reports on CRF01_AE, i.e., 8% (Sungkanuparph et al., 2007) and 11% (Chetchotisakd et al., 2006). As the CRF01_AE prevalence of Q151M is higher than that of subtype B, CRF01_AE appears to be more prone to acquire this mutation.

In both the ART-naïve and -exposed groups, the most frequently observed mutation was lamivudine-resistant M184V, suggesting that this mutation has a low genetic barrier. Comparing the two

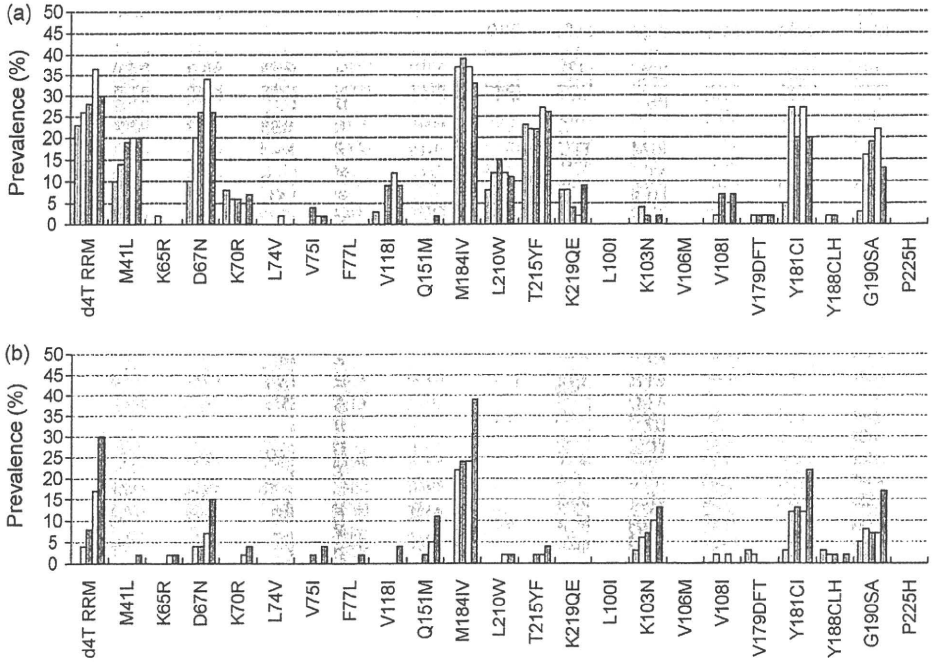

Fig. 1. Summary of drug resistant mutations detected in (a) antiretroviral exposed group, and (b) naive group. Light blue, gray, green, yellow and red bars indicate prevalence of mutation detected at baseline, 6, 12, 18 and 24 months after initiation of GPOvir treatment, respectively. d4T related resistance mutations (d4T RRM) include M41L, K65R, D67N, K70R, Q151M, L210W, T215YF and K219QE. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

groups, M184V was detected earlier and at higher prevalence in the exposed group (0% at baseline and 37% [18 of 49] at 6 months) than in the ART-naïve group (0% at baseline and 22% [11 of 49] at 6 months).

The ART-naïve and -exposed groups also had an interesting difference in their patterns of NNRTI-resistance acquisition. The ART-naïve group had a higher frequency of K103N mutation (13% [6 of 46]) than the exposed group (2% [1 of 46]). Instead of acquiring K103N, the exposed group tended to develop Y181C/I and G190A mutations. Since neither group had a history of NNRTI treatment prior to GPOvir, this difference in drug-resistance acquisition patterns cannot be explained by previous NNRTI treatment or differences in NNRTI-mutation patterns at baseline.

3.3. *Polymorphisms and drug-resistant mutations in the connection domain and RNase H mutations in CRF01_AE*

To determine subtype-specific polymorphisms in CRF01_AE, we compared sequences from 76 randomly selected CRF01_AE ART-naïve cases at baseline from our cohort with 42 subtype B reference sequences from the Los Alamos database. CRF01_AE-specific polymorphisms were determined at 9 residues in the connection domain and at 10 residues in the RNase H domain (Table 5). Interestingly, G335D and A371V, which have been recognized as NRTI-related resistance mutations in subtype B (Brehm et al., 2007; Nikolenko et al., 2007), were observed as natural polymorphisms among CRF01_AE. However, the contribution of these mutations to GPOvir resistance is not yet clear. Further studies are needed to clarify their role in NRTI resistance.

To determine treatment-specific mutations related to GPOvir administration in the connection and RNase H domains, sequences were compared between 49 samples from the last visit of GPOvir-failure cases and 76 baseline samples. The results (Table 6) show 13 mutations in the connection and RNase H domains: 9 mutations in the connection domain (Y318F, G335C/D, N348I, A360I/V, V365I, T369I, A371V, T376S, and E399D) and 4 in the RNase H domain (N447S, Q509L, P537S, and I542M). Among these mutations, N348I ($p < 0.001$) and E399D ($p < 0.001$) in the connection domain, and P537S ($p < 0.05$) and I542M ($p < 0.001$) in the RNase H domain were observed at significantly higher prevalence in GPOvir-failure cases. E312Q, G335C/D, N348I, A360I/V, V365I, and A376S were previously reported to confer AZT resistance in subtype B (Nikolenko et al., 2007). In addition, A371V and Q509L were reported to be selected *in vitro* by AZT and to confer greater AZT resistance and cross-resistance to other nucleoside RT inhibitors in combination with thymidine analogue-related mutations (TAM) (Brehm et al., 2007). However, as the number of cases in our study is small, the significance of E312Q, Y318F, G333D, A360I/V, V365I, T376S, and Q509L in CRF01_AE drug resistance is not well understood.

4. Discussion

Here we compared two patient groups, ART-naïve or -exposed at baseline, and analyzed their differences in their responses to nevirapine + 3TC + d4T (GPOvir) and drug-resistance acquisition patterns. This drug combination is widely used in the developing world today, and drug-resistant mutation patterns induced by this combination have been described in different countries and subtypes (Chetchotisakd et al., 2006; Kumarasamy et al., 2003; Pujari et al., 2004; Sungkanuparph et al., 2007; Zhou et al., 2007), but most reports describe drug-resistance mutation patterns in naïve cases, and few in exposed cases.

By comparing ART-naïve and -exposed groups, we observed the following notable findings. First, our study demonstrated that GPOvir is effective in exposed cases as well as naïve cases. Of the 101 exposed cases in our study, many had previously been treated with several nucleoside analogue inhibitors. Nonetheless, 64.4% of cases were successfully treated with GPOvir at the 24-month readout time point, despite our earlier finding that previous exposure to antiretrovirals was associated with virological failure (Tsuchiya et al., 2009). This finding suggests that though the priority of antiretroviral usage in resource-limited settings should be considered, pre-exposure history may not be an excuse to limit use of nevirapine + 3TC + d4T.