Mechanism of Drug Resistance Due to N88S in CRF01_AE HIV-1 Protease, Analyzed by Molecular Dynamics Simulations

Hirotaka Ode, *† Shou Matsuyama, † Masayuki Hata, † Tyuji Hoshino, †‡ Junko Kakizawa, § Wataru Sugiura§

† Graduate School of Pharmaceutical Sciences, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

CORRESPONDING AUTHOR FOOTNOTE

* Corresponding author

Tel: +81-43-290-2926

Fax: +81-43-290-2925

E-mail: odehir@graduate.chiba-u.jp

INSTITUTIONAL ADDRESS FOOTNOTE

‡ also PRESTO, JST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

§AIDS Research Center, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-1011, Japan
Abstract

Nelfinavir (NFV) is a currently available HIV-1 protease (PR) inhibitor. Patients in whom NFV treatment has failed predominantly carry D30N mutants of HIV-1 PRs if they have been infected with the subtype B virus. In contrast, N88S mutants of HIV-1 PRs predominantly emerge in patients in whom NFV treatment has failed and who carry the CRF01_AE virus. Both D30N and N88S confer resistance against NFV. However, it remains unclear why the non-active site mutation N88S confers resistance against NFV.

In this study, we examined the resistance mechanism through computational simulations. The simulations suggested that despite the non-active site mutation, N88S causes NFV resistance by reducing interactions between PR and NFV. We also investigated why the emergence rates of D30N and N88S differ between subtype B and CRF01_AE HIV-1. The simulations suggested that polymorphisms of CRF01_AE PR are involved in the emergence rate of the drug-resistant mutants.
**Introduction**

Human immunodeficiency virus type 1 (HIV-1) is one of the most hazardous viruses for humans, and there is still a risk of a worldwide HIV-1 pandemic. HIV-1 has high genetic variability and has been classified into three groups, labeled M, N, and O. Viruses in group M are further divided into subtypes, sub-subtypes, and circulating recombinant forms (CRFs). The subtype B virus is commonly found in HIV-1-infected patients in the Americas, Europe, and Japan. In contrast, developing countries suffer from a growing epidemic of non-subtype B viruses.

HIV-1 proliferates with the assistance of its own aspartic protease, so-called HIV-1 protease (HIV-1 PR), in its life cycle. \(^1\) HIV-1 PR is an enzyme composed of two identical polypeptides, each consisting of 99 amino acid residues, and its function is to process the viral Gag and Gag-Pol polyprotein precursors (Fig. 1A). Because this processing is essential for viral maturation, inhibition of PR function leads to incomplete viral replication and prevents the transfer to other cells. \(^2\) Therefore, HIV-1 PR is an attractive target for anti-HIV-1 drugs. Nine PR inhibitors (PIs) \(^3\)\(^-\)\(^11\) have been approved by the FDA and have successfully lowered the death rate due to acquired immune deficiency syndrome (AIDS) in advanced countries during the past decade. However, the currently available PIs were developed and tested only against subtype B PRs. Few studies have examined the susceptibility of non-subtype B viruses to those PIs, and no standard protocol of chemotherapy for non-subtype B viruses has been established. \(^12\)\(^-\)\(^17\)
Recently, Ariyoshi et al. reported that the pattern of drug-resistant mutations differed between subtype B and CRF01_AE (subtype AE) HIV-1. Mutations of L10F, K20I, L33I, and N88S in PR were more frequently seen in patients infected with subtype AE HIV-1 than in patients infected with subtype B HIV-1. Mutations of D30N, A71V, and N88D in PR were found in patients with subtype B HIV-1. Most of the characteristic mutation patterns in that study were associated with a history of treatment with nelfinavir (NFV, Fig. 1B), an FDA-approved PIs. D30N and N88S are particularly related to resistance against NFV. Interestingly, N88S is also known to cause hypersensitivity to another PI, amprenavir. We have previously suggested by computational simulations that D30N in subtype B PR confers resistance against NFV by canceling hydrogen bonds between NFV and N30. In addition, we and other groups have proposed an explanation of why some mutations confer resistance against PIs by not only X-ray crystallography but also computational studies. However, it has not been clarified why N88S in subtype AE PR confers resistance against NFV. Since N88S occurs at a non-active site of PR, it is difficult to speculate on the mechanism of resistance. Furthermore, it has not been understood why N88S emerges more predominantly than D30N in patients with subtype AE HIV-1, or why D30N emerges more predominantly than N88S in patients with subtype B HIV-1. Subtype AE HIV-1 has natural polymorphisms, K20R, E35D, M36I, R41K, H69K, L89M and I93L, in PR unlike subtype B HIV-1 (Fig. 1C). These polymorphisms are also located at the non-active site of PR. It is still uncertain whether or not the polymorphisms affect the emergence rates of those mutations.
In this study, we investigated the mechanism of resistance against NFV due to N88S in subtype AE HIV-1 PR through computational simulations. Our simulations indicated that the N88S mutation creates hydrogen bonds between the D30 and S88 side chains. Therefore, N88S mutation reduces interactions between D30 and NFV. We also investigated the reason for the difference between the two subtypes in the emergence rate of D30N as well as that of N88S. The results indicated that, in subtype B HIV-1, D30N PR has a lower affinity for NFV than does N88S PR. In subtype AE HIV-1, on the other hand, D30N PR has a higher affinity than does N88S. Our findings suggest that despite the non-active site mutations, the polymorphisms regulate the emergence rates of these drug-resistant mutants.
Results

Reconsideration of torsional force field parameters for benzamide. Before carrying out molecular dynamics (MD) simulations, we reconsidered torsional force field parameters for benzamide: CA-CA-C –N and CA-CA-C –O (Supporting Information Fig. S1). The benzamide group comprises a part of NFV. The benzamide moiety in NFV has an important interaction with D30 of PR.\textsuperscript{22,42} Nevertheless, the AMBER ff03\textsuperscript{43} and general AMBER force fields\textsuperscript{44} each cause a much higher energy barrier around the rotatable bond between the benzene and amide groups in benzamide than that based on quantum chemical calculations. This is a serious problem for our simulations. The force field parameters for benzamide need to be carefully examined and preferably changed from the original AMBER force fields, as described in the AMBER Archive in 2003.\textsuperscript{45} Since these force field parameters have not been changed in the AMBER force fields yet, we improved the torsional force field parameters for the benzamide moiety in NFV. The torsional parameters were generated in the same manner as that for the development of the AMBER ff03 force field. The obtained parameters are listed in Table 1. We executed MD simulations using these newly developed force field parameters.

Hydrogen bonds between NFV and PRs. Hydrogen bonds play an important role in protein-ligand bindings. First, we examined the hydrogen bonds between NFV and PR in each complex: wild-type (WT) PR; D30N PR; N88S PR of subtype B HIV-1 (labeled B(WT), B(D30N), and B(N88S), respectively); and the reference (Ref) PR, D30N PR, N88S PR of subtype AE HIV-1 in complex with NFV (AE(Ref), AE(D30N), and
AE(N88S)). We examined 1000 snapshot structures for the last 1.0 ns and identified direct or one-water-molecule-mediated hydrogen bonds (Table 2 and Supporting Information Table S1 and Fig. S2). All six PRs create similar hydrogen bond networks. The side chains of both D25 and D25’ interact with the central hydroxyl group of NFV (the atom corresponding to O3 in Fig. 1B). One water molecule mediates the interaction between the main chains of I50/I50’ and NFV. Furthermore, another water molecule mediates the interaction between D29’ and NFV. However, the interaction between NFV and the 30th residue has variations among the six PRs. In B(WT) and AE(Ref), either the main chain or the side chain of D30 makes a direct hydrogen bond with NFV. D30N and N88S models show different interactions between subtype B and AE PRs. B(D30N) has no hydrogen bond between N30 and NFV, whereas AE(D30N) has direct or one-water-molecule-mediated hydrogen bonds. B(N88S) frequently creates a direct hydrogen bond between the main chain of D30 and NFV. On the other hand, AE(N88S) mainly creates one-water-molecule-mediated hydrogen bonds between the main chain of D30 and NFV. Interestingly, the side chain of N30 in AE(D30N) is clearly closer to the phenol group of NFV than that of B(D30N) (Fig. 2). In contrast, the side chain of D30 in AE(N88S) is more distant from the phenol group of NFV than that of B(N88S).

Hydrogen bonds of the side chain of the 30th residue with PR residues. In B(D30N) as well as in AE(N88S), the side chain of the 30th residue does not create any hydrogen bonds with NFV. To clarify the effects of the D30N and N88S mutations in detail, the interactions of the side chain of the 30th residue with other residues of PR were investigated as shown in Table 3. B(WT) and AE(Ref) have an interaction between the
side chains of D30 and K45. B(D30N) has direct hydrogen bonds from the side chain of N30 to T31 and T74. On the other hand, AE(D30N) has one-water-molecule-mediated hydrogen bonds from N30 to T31, T74, and N88. T31, T74, and N88 also create hydrogen bond networks at the non-active sites in B(WT) and AE(Ref), although D30 is not involved in the networks (Supporting Information Table S2). The side chains of D30 in both B(N88S) and AE(N88S) have either a direct hydrogen bond with the side chain of S88 or one-water-molecule-mediated hydrogen bonds with T31, T74, and S88. The mutations D30N and N88S affect those hydrogen bond networks.

**Comparison of the structures with B(WT).** To clarify the effects of mutations at the 30th and the 88th residues on the active site conformations, the average structure of each model for the last 1.0 ns was compared with that of B(WT). Each model was fitted to B(WT) using the coordinates of main chain atoms N, Cα, and C, and the root mean squared deviation (RMSD) value was calculated (Fig. 3). When the active site residues of each PR are compared with those of B(WT), conformational changes are observed only on the active site residues around the 30th residue. AE(N88S) shows a large conformational change at D30 (RMSD = 1.5±0.4Å, Fig. 4). The other four models -- B(D30N), B(N88S), AE(Ref), and AE(D30N) -- each show a slight conformational change at D30. When each subtype B PR is compared with the corresponding subtype AE PR, B(D30N) is found to have larger conformational changes on N30 than AE(D30N). AE(N88S) shows larger conformational changes on D30 than B(N88S). Next, we compared the location of NFV in each model with that of B(WT). The benzamide group of NFV, which interacts with the 30th residue, shows a
larger positional deviation than do other parts of NFV in every model (Supporting Information Fig. S3).

**Binding free-energy calculations.** The influence of mutation or polymorphism on the binding free energy $\Delta G_b$ was examined for each model. Table 4 shows the results of MM/PBSA calculations for all of the PRs in complex with NFV. In subtype B HIV-1, B(D30N) reduces the binding energy with NFV from B(WT) more than B(N88S) does. On the other hand, in subtype AE HIV-1, AE(D30N) shows affinity with NFV, similar to that of AE(Ref), and has a higher affinity with NFV than AE(N88S). The results correspond to the emergence rates of subtypes B and AE variants in patients in whom NFV treatment has failed. D30N predominantly emerges in patients with subtype B HIV-1, whereas N88S predominantly emerges in patients with subtype AE HIV-1. We also investigated the contributions of the respective residues to binding free energy (Fig. 5). In all six models, the active site residues stabilize the complex of each PR and NFV. When we focus on the binding energy due to the 30th residue, D30 or N30, B(D30N) reduces the contribution to the binding free energy in comparison with that of B(WT) (Fig. 6). AE(N88S) also reduces the contribution to the binding energy compared with that of AE(Ref). B(N88S) shows a contribution similar to that of B(WT), and AE(D30N) shows a contribution similar to that of AE(Ref).
Discussion

In this study, we performed MD simulations of HIV-1 PRs in complex with NFV for the purpose of clarifying 1) the mechanism of resistance against NFV due to N88S in subtype AE PR and 2) the reason why the emergence rates of D30N and N88S differ between subtypes B and AE HIV-1.

The 88th residue is located at a non-active site of HIV-1PR. Thus, it is difficult to speculate on the mechanism of resistance due to N88S. Our simulations indicate that N88S mutant PR has a lower affinity with NFV than does Ref PR in subtype AE HIV-1, owing to the following mechanism. First, a hydrogen bond between the side chain of D30 and the side chain of S88 is created (Fig. 7). Second, the location of D30 is changed. Finally, the interaction between D30 and NFV is reduced. N88S indirectly affects the binding between NFV and D30. Accordingly, both N88S and D30N are thought to confer specific resistance against NFV because the interaction with D30 is an essential factor for NFV binding. Indeed, it has been reported that the emergence of the N88S mutation is highly related to resistance against NFV. 18-21 N88D is another frequently observed mutation at the 88th residue of PR. It has also been reported that N88D changes the interactions of the 88th residue with D30, T31, and T74.22, 26 However, N88D hardly affects the ligand binding at the active site and does not cause resistance against NFV. N88S changes the interactions in a different manner from that of N88D.

We then pose another question: Why does N88S emerge more frequently in patients with subtype AE HIV-1 in whom NFV treatment has failed than in patients
with subtype B HIV-1? Ariyoshi et al. reported that D30N emerged predominantly in patients with subtype B HIV-1, whereas N88S appeared predominantly in patients with subtype AE HIV-1. Subtype AE HIV-1 PR has some natural polymorphisms — K20R, E35D, M36I, R41K, H69K, L89M and I93L — unlike subtype B PR. These amino acids are located at non-active sites of PR. To reveal whether the polymorphisms affect NFV binding or what causes the difference in the emergence rates of D30N and N88S, we carried out simulations of NFV complexes of WT PR, D30N PR, and N88S PR of subtype B (B(WT), B(D30N), B(N88S)) and Ref PR, D30N PR, and N88S PR of subtype AE (AE(Ref), AE(D30N), AE(N88S)). AE(Ref) has an interaction with NFV similar to that of B(WT). On the other hand, D30N and N88S mutations show different effects between subtypes B and AE PRs. D30N in subtype B PR greatly reduces the binding affinity with NFV, because the hydrogen bonds between N30 and NFV are canceled as we previously reported. In contrast, D30N in subtype AE PR hardly affects the affinity with NFV. AE(D30N) has direct or one-water-molecule-mediated hydrogen bonds between N30 and NFV. On the other hand, N88S in subtype AE PR significantly reduces the binding affinity with NFV, whereas N88S in subtype B PR hardly affects the affinity with NFV. In both B(N88S) and AE(N88S), a hydrogen bond is created between the side chain of D30 and the side chain of S88. However, the interactions of NFV with D30 differ between subtype B and AE PRs. B(N88S) has a direct hydrogen bond between the main chain of D30 and NFV, whereas AE(N88S) mainly has one-water-molecule-mediated hydrogen bonds between the main chain of D30 and NFV. D30N PR has lower affinity with NFV than does N88S PR in subtype B
HIV-1. In contrast, D30N PR has higher affinity than N88S PR in subtype AE HIV-1. These results are compatible with the results of a study by Ariyoshi et al. Both D30N and N88S mutations in HIV-1 PRs exhibit significant losses of viral fitness. Therefore, D30N and N88S mutants of HIV-1 have low growth kinetics relative to WT or Ref variants under the condition without any PIs. Nevertheless, it is frequently observed that the D30N mutant emerges in patients with subtype B HIV-1 in whom NFV treatment has failed, and that the N88S mutant emerges in patients with subtype AE HIV-1. These results indicate that the effectiveness of NFV is significantly reduced for these mutants. In contrast, N88S mutants of subtype B PR and D30N mutants of subtype AE PR have rarely been seen clinically. This is thought to be due not only to their low degree of fitness but also to their affinities with NFV comparable to those of subtype B WT or subtype AE Ref variants. Our simulations suggest that the natural polymorphisms of subtype AE PR, in spite of the non-active site mutations, reduce the emergence rate of D30N and increase that of N88S.

The polymorphisms in subtype AE PR increase the emergence rate of N88S. However, there remains the question of which is the key mutation that affects the emergence rate of N88S among the polymorphisms: K20R, E35D, M36I, R41K, H69K, L89M, and I93L. In this study, we focused on M36I for three reasons. First, M36I is related to the resistance against NFV. Second, N88S has been observed in combination with mutations at various positions, including 20, 36, 46, 63, and 77. Third, M36I is frequently observed as a polymorphism in other subtypes, namely A and C. We executed additional simulations of M36I PR, M36I/N88S PR, and
L10F/M36I/N88S PR of subtype B HIV-1 in complex with NFV (labeled B(M36I), B(M36I/N88S), and B(L10F/M36I/N88S), respectively). L10F is a mutation that is frequently seen in CRF01_AE HIV-1 accompanied by N88S. Our simulations suggest that the single M36I mutation in subtype B PR does not affect NFV binding. B(M36I) has stable hydrogen bonds between NFV and D30 (Supporting Information Table S3 and Fig. S4). In contrast, the combination of M36I and N88S mutations in subtype B PR reduces the binding affinity with NFV. B(M36I/N88S) has fewer hydrogen bonds with NFV than does B(M36I) or B(N88S). Furthermore, the conformational change at D30 is larger in B(M36I/N88S) than in B(M36I) or B(N88S) (Supporting Information Fig. S5 and Fig. S6). B(L10F/M36I/N88S) also creates fewer hydrogen bonds between NFV and D30 and causes conformational alteration at D30. The polymorphism M36I reduces the contribution of D30 to the binding with NFV (Supporting Information Fig. S7). Our simulations suggest that N88S in subtype B PR reduces the binding affinity with NFV when it appears together with M36I.

It is interesting that both D30N and N88S confer resistance against NFV by decreasing the interaction between the 30th residue and NFV. Both D30N and N88S affect the active site residues around the 30th residue. Other active site residues hardly change their interaction with NFV or their conformations. As can be seen in Fig. 2 and Fig. S4, the NFV-resistant PRs -- B(D30N), AE(N88S), B(M36I/N88S), and B(L10F/M36I/N88S) --each show an increase in distance between the 30th residue and NFV. The NFV-resistant N88S mutants -- AE(N88S), B(M36I/N88S) and B(L10F/M36I/N88S) --each have a stable direct hydrogen bond between the side chain
of S88 and the side chain of D30. Therefore, N88S does not appear simultaneously with D30N clinically.

Prior to the MD simulations, we reconsidered torsional force field parameters for the benzamide moiety in NFV. This moiety has essential hydrogen bonds with D30 of HIV-1 PR. Thus, those torsional parameters are expected to greatly affect the results of the simulations. Nevertheless, the AMBER ff03 and general AMBER force fields each cause a much higher energy barrier around the rotatable bond between the benzene and amide groups in benzamide than that based on quantum chemical calculations (Supporting Information Fig. S1). This was a serious problem for our simulations. Therefore, we improved the torsional force field parameters for the benzamide moiety in NFV by fitting them to the energy curve obtained from quantum chemical calculations. Our newly developed parameters enabled us to carry out precise simulations of HIV-1 PR in complex with NFV.

In this study, we not only proposed the mechanism of resistance against NFV of N88S in subtype AE PR but also examined the influence of the polymorphisms in subtype AE PR on the emergence rates of D30N and N88S mutations. N88S and the polymorphisms in subtype AE PR are all classified as non-active site mutations. Nevertheless, these mutations affect the binding of NFV. We and other groups have reported that the non-active site mutations affect the binding affinity of some inhibitors, the emergence rate of mutants, and the catalytic activity of the protease. For example, the polymorphisms in subtype C HIV-1 enhance the catalytic efficacy. However, there have been few studies on the influence of non-active site mutations
from structural viewpoints. There have also been few studies on the differences between HIV-1 subtypes. Clarification of the roles of non-active site mutations and polymorphisms will enable us to design potent drugs, since the currently available PIs were developed and tested only against subtype B PRs. Accumulation of data on the susceptibilities of non-subtype B viruses to the currently available PIs is also needed in order to establish an effective HIV-1 therapy strategy. Clarification of these susceptibilities will also be useful for selecting more appropriate drugs for patients.
Conclusions

We proposed a mechanism of resistance against NFV due to the non-active site mutation N88S in CRF01_AE PR through computational simulations. CRF01_AE PR has polymorphisms at non-active sites, unlike subtype B PR. Nevertheless, the polymorphisms affect the binding affinities between NFV and PR variants that have the D30N or N88S mutation. The simulations suggest that N88S in CRF01_AE PR confers NFV resistance by reducing interaction energy between D30 and NFV. N88S creates hydrogen bonds between the D30 and S88 side chains, and causes conformational changes at D30. These changes reduce the interactions between D30 and NFV. Furthermore, we proposed an explanation of why the emergence rates of D30N and N88S differ between subtypes B and AE HIV-1. The M36I mutation seen in the natural polymorphisms of CRF01_AE PR is particularly involved in the difference in the emergence rates of D30N and N88S.
Acknowledgments

This work was supported by a Health and Labor Science Research Grant for Research on HIV/AIDS from the Ministry of Health and Labor of Japan, by JSPS Research Fellowships for Young Scientists, and by a grant-in-aid for JSPS Fellows. A part of this work was also supported by a grant from the Japan Science and Technology Agency.

Supporting Information Available: A list of hydrogen bond networks in each model, determination of protonation states of catalytic aspartates, parameter development for the torsional parameters CA-CA-C-N and CA-CA-C-O, 3D plot of RMSD of the average structure from that of B(WT), contributions of each individual residue to the binding energy, RMSD plots during MD simulations, results of principal component analyses. This material is available free of charge via the Internet at http://pubs.acs.org.
Experimental section

**Force field parameters.** Before carrying out molecular dynamics (MD) simulations, we improved the torsional force field parameters for benzamide and determined the restrained electrostatic potential (RESP) charges for NFV. The improved torsional parameters were generated in the same manner as that for the development of the AMBER ff03 force field. First, RESP charges of benzamide were determined based on data from quantum chemical calculations. Geometric optimization was performed at the HF/6-31G(d,p) level, and the electrostatic potential was calculated at the B3LYP/cc-pVTZ level under the solvation condition with ether ($\varepsilon$=4) by the IEFPCM method using the Gaussian03 program. The partial atom charges were determined using the RESP method so that the atom charges could reproduce the values of the calculated electrostatic potential at the surrounding points of the benzamide. Charges were equalized between two atoms if they were the same element and had the same bond coordination. Second, a potential energy curve was obtained by repeating the energy calculations with 5-degree stepwise changes in the torsional angle around the torsional axis. Energy calculations were executed at the MP2/cc-pVTZ level in the ether phase after geometric optimizations at the HF/6-31G(d,p) level. Third, the torsional parameters were obtained by fitting them to the potential energies from quantum chemical calculations. The torsional parameter for CA-CA-C –N was assumed to be equal to that for CA-CA-C –O. RESP charges for NFV were also determined in the same manner as that described above.
**Molecular dynamics (MD) simulations.** Minimizations and MD simulations were carried out using the Sander module of AMBER 8. The AMBER ff03 force field was used as the parameters for proteins, ions, and water molecules. The general AMBER force field and our developed force field were used as the parameters for NFV.

We examined the structure of each of the six PRs in complex with NFV: wild-type (WT) PR; D30N PR; N88S PR of subtype B HIV-1 (labeled B(WT), B(D30N), and B(N88S), respectively); and the reference (Ref) PR, D30N PR, N88S PR of subtype AE HIV-1 in complex with NFV (AE(Ref), AE(D30N), and AE(N88S)). Additionally, we investigated the structures of M36I PR, M36I/N88S PR, and L10F/M36I/N88S PR of subtype B HIV-1 in complex with NFV (labeled B(M36I), B(M36I/N88S), and B(L10F/M36I/N88S), respectively). We used HXB2 as the WT sequence of subtype B HIV-1 and NH1 as the reference sequence of subtype AE HIV-1. Each initial structure for the PR in complex with NFV was modeled from the atom coordinates of an X-ray crystal structure (PDB code: 1OHR) using the LEaP module. Each model was placed in a rectangular box filled with about 8000 TIP3P water molecules, with all of the crystal water molecules remaining. The cutoff distance for the long-range electrostatic and the van der Waals energy terms was set to 12.0 Å. The expansion and shrinkage of all covalent bonds connecting to hydrogen atoms were constrained using the SHAKE algorithm. Periodic boundary conditions were applied to avoid the edge effect in all calculations. Energy minimization was achieved in three steps. First, movement was allowed only for water molecules and ions. Next, the ligand and the mutated residues were allowed to move in addition to the water
molecules and ions. Finally, all atoms were permitted to move freely. In each step, energy minimization was executed by the steepest descent method for the first 10000 steps and the conjugated gradient method for the subsequent 10000 steps. After a 0.1 ns heating calculation until 310 K using the NVT ensemble, a 3.0 ns equilibrating calculation was executed at 1 atm and at 310 K using the NPT ensemble, with an integration time step of 2.0 fs. In the present calculations, the MD simulations showed no large fluctuations after about 2.0 ns equilibrating calculations (Supporting Information Fig. S8 and Fig. S9). Hence, atom coordinates were collected at the interval of 1.0 ps for the last 1.0 ns to analyze the structure in detail.

The protonation states of catalytic aspartates D25 and D25’ vary depending on the binding ligands or PRs.\textsuperscript{57} Hence, the appropriate protonation states of catalytic aspartates should be determined for each model. Because NFV mimics a transition state of catalytic reaction by HIV-1 PR, we considered two kinds of protonation states.\textsuperscript{58-60} One represented a combination of protonated D25 and unprotonated D25’ states, and the other represented the opposite combination. In order to determine the protonation states when NFV binds to each PR, the free energies of these two kinds of protonation states were compared using the calculation data obtained during 2.0-3.0 ns MD simulations. The free energies were calculated on the basis of the MM/PBSA method.\textsuperscript{61,62} We used the same parameter set for electrostatic and van der Waals energy terms as that was used in the MD simulations, and no cutoff was applied for the calculation. Since the dielectric constants for the interior of proteins are considered to be in the range of 2 to 4, the interior dielectric constant was set to 2.0.\textsuperscript{63} The outer dielectric
constant was set to 80.0. The pbsa program was used to solve the Poisson-Boltzmann (PB) equation. B(D30N), B(M36I), B(L10F/M36I/N88S), and AE(D30N) were found to prefer the combination of protonated D25 and unprotonated D25’. The other five PRs -- B(WT), B(N88S), B(M36I/N88S), AE(Ref) and AE(N88S) -- preferred the combination of unprotonated D25 and protonated D25’ (Supporting Information Table S4).

**Hydrogen bond criteria.** The formation of a hydrogen bond was defined in terms of distance and orientation. The combination of donor D, hydrogen H, and acceptor A atoms with a D - H … A configuration was regarded as a hydrogen bond when the distance between donor D and acceptor A was shorter than 3.5 Å and the angle H-D-A was smaller than 60.0 degrees.

**Binding free-energy calculation.** The binding free energy was calculated by the following equation:

\[
\Delta G_b = \Delta G_{\text{int}}^{\text{ele}} + \Delta G_{\text{int}}^{\text{vdw}} + \Delta G_{\text{sol}} - T\Delta S,
\]

where \( \Delta G_b \) is the binding free energy in solution, \( \Delta G_{\text{int}}^{\text{ele}} \) and \( \Delta G_{\text{int}}^{\text{vdw}} \) are electrostatic and van der Waals interaction energies between a ligand and a protein, \( \Delta G_{\text{sol}} \) is the solvation energy, and \(-T\Delta S\) is the contribution of conformational entropy to the binding. In this study, assuming that the contribution of conformational entropy to the change in \( \Delta G_b \) is negligible among mutants, we neglected the entropy term in the energy estimation. \( \Delta G_{\text{int}}^{\text{ele}} \) and \( \Delta G_{\text{int}}^{\text{vdw}} \) were computed using the same parameter set as that used in the MD simulation, and no cutoff was applied to the calculation. Solvation energy \( \Delta G_{\text{sol}} \) was calculated using the pbsa program. The interior dielectric constant was set to 2.0, and the outer dielectric constant was set to 80.0. Furthermore, the
contribution of each residue to the binding free energy was calculated. The total binding free energy was decomposed into the contribution from each individual residue by the MM/GBSA method. The modified GB model developed by Onufriev, Bashford and Case\textsuperscript{66} was used to calculate the solvation energy term. To ascertain whether or not the MM/GBSA results were consistent with the MM/PBSA results, we compared the total binding free energy obtained by the MM/PBSA method with that obtained by the MM/GBSA method for all coordinates acquired through the MD simulation. The MM/GBSA results were confirmed to be highly correlated with the MM/PBSA results (correlation coefficient: $r \geq 0.998$) (Supporting Information Fig. S10).
References


(6) Livingston, D. J.; Pazhanisamy, S.; Porter, D. J.; Partaledis, J. A.; Tung, R. D.;


(32) Prabu-Jeyabalan, M.; Nalivaika, E. A.; King, N. M.; Schiffer, C. A. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1


(38) Wittayanarakul, K.; Aruksakunwong, O.; Saen-oon, S.; Chantratita, W.; Parasuk, V.; Sompornpisut, P.; Hannongbua, S. Insights into saquinavir resistance in the
G48V HIV-1 protease: quantum calculations and molecular dynamic simulations.


(44) Wang, J.; Wolf, R. M.; Cladwell, J. W.; Kollman, P. A.; Case, D. A. Development


(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.;

*Gaussian 03, 2004*, Gaussian, Inc., Wallingford CT.


(61) Srinivasan, J.; Cheatham, T. E. III; Kollman, P.; Case, D. A. Continuum solvent


Table 1. Force field parameters for the torsional parameters of the benzamide part of NFV.

<table>
<thead>
<tr>
<th>CA-CA-C –N</th>
<th>CA-CA-C- O</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_n/2$</td>
<td>$n$</td>
</tr>
<tr>
<td>Developed parameters</td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>2</td>
</tr>
<tr>
<td>0.05</td>
<td>4</td>
</tr>
<tr>
<td>Cf.) AMBER ff03 force field</td>
<td></td>
</tr>
<tr>
<td>3.63</td>
<td>2</td>
</tr>
</tbody>
</table>

Torsional energy is given by $E = \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$. 
### Table 2. Hydrogen bond networks of NFV with D30 or N30 in PR.

<table>
<thead>
<tr>
<th></th>
<th>Donor</th>
<th>Acceptor</th>
<th>Subtype B (%)</th>
<th>Acceptor</th>
<th>Subtype AE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>D30</td>
<td>O1</td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>B(WT)</td>
<td>O1</td>
<td>NFV</td>
<td>30.7</td>
<td>O1</td>
<td>NFV</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>OD1/OD2</td>
<td>31.6</td>
<td>O1</td>
<td>OD1/OD2 D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>O</td>
<td>42.9</td>
<td>O1</td>
<td>O D30</td>
</tr>
<tr>
<td>B(D30N)</td>
<td>O1</td>
<td>NFV</td>
<td>31.6</td>
<td>O1</td>
<td>WAT766</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>O</td>
<td>69.2</td>
<td>O1</td>
<td>WAT1770 OD1</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>O</td>
<td>9.0</td>
<td>O1</td>
<td>WAT8063 OD1</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>25.2</td>
<td>O1</td>
<td>OD1 N30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT1770</td>
<td>6.8</td>
<td>O1</td>
<td>N30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT8063</td>
<td>13.7</td>
<td>O1</td>
<td>OD1 N30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>12.4</td>
<td>O1</td>
<td>N30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>12.4</td>
<td>O1</td>
<td>D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT7886</td>
<td>5.2</td>
<td>O1</td>
<td>OD1 D30</td>
</tr>
<tr>
<td>B(N88S)</td>
<td>O1</td>
<td>NFV</td>
<td>26.7</td>
<td>O1</td>
<td>NFV</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>NFV</td>
<td>7.4</td>
<td>O1</td>
<td>NFV</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>OD2</td>
<td>12.4</td>
<td>O1</td>
<td>OD2 D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>O</td>
<td>56.7</td>
<td>O1</td>
<td>O D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>28.8</td>
<td>O1</td>
<td>WAT6715 N D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>28.3</td>
<td>O1</td>
<td>WAT6715 O D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>NFV</td>
<td>31.8</td>
<td>O1</td>
<td>NFV O D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT7886</td>
<td>13.6</td>
<td>O1</td>
<td>OD1 D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>5.2</td>
<td>O1</td>
<td>OD1 D30</td>
</tr>
<tr>
<td></td>
<td>O1</td>
<td>WAT6715</td>
<td>25.5</td>
<td>O1</td>
<td>OD1 D30</td>
</tr>
</tbody>
</table>

*a Occupancy of hydrogen bonds during 2.0-3.0 ns MD simulation.

*b The atom names of NFV are shown in Fig. 1.
Table 3. Hydrogen bond networks of the side chain of D30 or N30 with PR residues.

<table>
<thead>
<tr>
<th>Donor</th>
<th>Acceptor</th>
<th>%</th>
<th>Donor</th>
<th>Acceptor</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>NZ</td>
<td>K45</td>
<td>OD1/OD2</td>
<td>D30</td>
<td>NZ</td>
<td>K45</td>
</tr>
<tr>
<td>B(WT)</td>
<td></td>
<td></td>
<td>67.0</td>
<td>AE(Ref)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AE(D30N)</td>
<td></td>
</tr>
<tr>
<td>ND2</td>
<td>N30</td>
<td>O</td>
<td>T74</td>
<td>ND2</td>
<td>N30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>T31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>WAT24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>WAT24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ND2</td>
<td>N88</td>
</tr>
<tr>
<td>B(N88S)</td>
<td></td>
<td></td>
<td></td>
<td>OG</td>
<td>S88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>WAT1142</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>T31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>WAT1142</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OG</td>
<td>S88</td>
</tr>
</tbody>
</table>

*a* Occupancy of hydrogen bonds during 2.0-3.0 ns MD simulation.
<table>
<thead>
<tr>
<th></th>
<th>$\Delta G_{\text{int}}^{\text{ele}}$</th>
<th>$\Delta G_{\text{int}}^{\text{vdw}}$</th>
<th>$\Delta G_{\text{mol}}$</th>
<th>$\Delta G_b$</th>
<th>$\Delta \Delta G_b$</th>
<th>$\Delta \Delta G_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT (B)</td>
<td>-12.5±1.4</td>
<td>-71.8±3.8</td>
<td>15.1±1.4</td>
<td>-69.2±3.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ref (AE)</td>
<td>-12.8±1.7</td>
<td>-70.8±4.0</td>
<td>15.1±1.5</td>
<td>-68.6±3.7</td>
<td>0.6</td>
<td>-</td>
</tr>
<tr>
<td>D30N (B)</td>
<td>-6.9±1.2</td>
<td>-70.5±4.1</td>
<td>10.9±0.9</td>
<td>-66.5±3.9</td>
<td>2.7</td>
<td>-</td>
</tr>
<tr>
<td>(AE)</td>
<td>-7.5±1.3</td>
<td>-70.2±3.9</td>
<td>9.5±1.0</td>
<td>-68.2±3.7</td>
<td>1.0</td>
<td>0.4</td>
</tr>
<tr>
<td>N88S  (B)</td>
<td>-12.0±1.3</td>
<td>-71.7±3.8</td>
<td>15.0±1.2</td>
<td>-68.7±3.7</td>
<td>0.5</td>
<td>-</td>
</tr>
<tr>
<td>(AE)</td>
<td>-10.6±1.4</td>
<td>-67.8±4.0</td>
<td>12.8±1.9</td>
<td>-65.6±3.9</td>
<td>3.6</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Energy is presented in units of kcal/mol.

$a$ TΔS is not included.

$b$ Difference from B(WT).

$c$ Difference from AE(Ref).
Figure legends

Fig. 1. A: Structure of HIV-1 PR. Locations of two catalytic aspartates, the 30th and the 88th residues, are shown in ball and stick representation. Locations of polymorphisms in subtype AE PR (K20R, E35D, M36I, R41K, H69K, L89M, and I93L) are shown in stick representation. B: Chemical structure of NFV. C: Amino acid sequences of a wild-type (WT) subtype B HIV-1 PR (HXB2) and a reference sequence of subtype AE HIV-1 PR (NH1). The polymorphisms in subtype AE PR are highlighted in red letters.

Fig. 2. Distance between NFV and the 30th residue. Each red and green solid line corresponds to the distance between N of the 30th residue and the O1 atom of NFV and the distance between O of the 30th residue and the O1 atom of NFV. Blue solid lines of B(WT), AE(Ref), B(N88S), and AE(N88S) show the distances between O1 of NFV and OD1/OD2 of D30, while those of B(D30N) and AE(D30N) show the distances between O1 of NFV and OD1/ND2 of N30.

Fig. 3. 3D plot of RMSD of the average structure of each model from that of B(WT). PR in each model is shown in colored tube representation. The color refers to the magnitude of RMSD shown in the bottom bar. Each model was fitted to B(WT) using the coordinates of main chain atoms N, Cα, and C of PR. The superimposed gray sticks and tubes represent the structure of B(WT).

Fig. 4. The RMSD value on the 30th residue of the average structure of each PR from that of B(WT). Error bars show root mean squared fluctuations (RMSF).

Fig. 5. A: Contribution of each individual residue to binding free energy. B: Difference in contribution of each residue to the binding energy between the respective
mutant and B(WT). The energies of contributions of the residues correspond to red solid lines, and those of B(WT) to green lines. The bottom black lines indicate the locations of the active site residues (R8, L23-V32, I47-I50, P81-I84, R8’, L23’-V32’, I47’-I50’, P81’-I84’).

**Fig. 6.** Contribution of the 30th residue to binding free energy in each model. Error bars stand for standard deviation.

**Fig. 7.** Hydrogen bond networks around the 88th residue of PR. (A) The hydrogen bond networks in B(WT) and (B) those in AE(N88S).
Fig. 1.
Fig. 2.
Fig. 3.
Fig. 4.
Fig. 5.
Fig. 6.
Fig. 7.