Closing of the Flaps of HIV-1 Protease Induced by Substrate Binding: A Model of a Flap Closing Mechanism in Retroviral Aspartic Proteases

Gergely Tóth*,‡,§ and Attila Borics‖,⊥

Locus Pharmaceuticals, Four Valley Square, 512 Township Line Road, Blue Bell, Pennsylvania 19422, and Department of Biomedical Sciences, School of Medicine, Creighton University, 2500 California Plaza, Omaha, Nebraska 68178

Received January 27, 2006; Revised Manuscript Received March 24, 2006

ABSTRACT: The active site of aspartic proteases is covered by one or more flaps, which control access to the active site and play a significant role in the binding of the substrate. An extensive conformational change of the flaps takes place upon binding of substrate to the active site. A long molecular dynamics simulation was performed on the complex consisting of a peptide (CA-p2) from a natural substrate cleavage site of the gag/pol polyprotein placed in the active site of HIV-1 protease (PR) with an open flap conformation. During the simulation, the substrate induced the closing of the flaps into the closed conformation in an asymmetrical way through a hydrophobic intermediate state cluster. The nature of the residues of HIV-1 PR identified to be important in the flap closing mechanism is conserved across known structures of retroviral aspartic proteases family. The flap closing mechanism described in HIV-1 PR is proposed to be a general model for flap closing in retroviral aspartic proteases.

The active site of aspartic proteases is covered by flaps, flexible β-hairpin structures, which are believed to control substrate access to the active site and take part in substrate recognition. In retroviral proteases, the active site is covered by two flaps, while in eukaryotic proteases, it is covered by only one. The catalytic mechanism of aspartic proteases is considered to consist of the following steps (*1*): (1) binding of the substrate to form a loose complex, (2) closing of the flaps down upon the substrate to set all components of the substrate into the correct geometry for the catalytic process, (3) a bond cleavage event, (4) opening of the flaps, (5) release of the products, and (6) re-formation of catalytic activity in the active site. The flaps have a dual role in this mechanism. They have a structural role, because they form interactions with the substrate to stabilize the substrate−protease complex. They also have a kinetic role, because the flaps need to close for substrate binding and open for product release (*2*, *3*).

Human immunodeficiency virus 1 (HIV-1)[1] protease (PR) is a member of the retroviral aspartic protease family of enzymes. HIV-1 PR cleaves the *gag* and *pol* viral polyproteins in the maturation step during the replication cycle of the virus. Inhibition of HIV-1 PR prevents the virus from taking its mature and infectious form (*4*). HIV-1 PR is an important drug target. More than half a dozen commercially available drugs are HIV-1 PR inhibitors, which are used for the treatment of acquired immune deficiency syndrome (*5*). The current generation of inhibitors targets the active site of HIV-1 PR.

HIV-1 PR is a homodimer. The active site of HIV-1 PR is located at the homodimer interface and is covered by two flexible antiparallel β-hairpins, termed flaps, consisting of residues 45−55. X-ray crystallography studies (*6*−*8*) observed that in the free enzyme the flaps are loosely packed to each other in a semi-open conformation in a way in which they sterically restrict substrate access to the active site. NMR investigation of the substrate-free HIV-1 PR observed that the flap region moves on a microsecond time scale, while the Gly-rich flap tips are more flexible, moving on a subnanosecond time scale (*9*−*12*). Recent molecular dynamics (MD) simulations of HIV-1 PR proposed that flap separation from the semi-open flap conformation to open flap conformations can occur in the absence of substrate at 300 K (*13*) or due to a small external impulse on the flaps (*14*). Flap opening was observed to be a complex dynamic process as two distinct mechanisms of flap opening, symmetric and asymmetric, were detected during previous MD simulations (*15*). The flaps were also observed to open and return to the semi-open position during the same long molecular dynamics trajectory (*16*). Figure 1 illustrates the different flap conformations of HIV-1 PR. The flaps reversed into a closed conformation in the crystal structures of the substrate-bound HIV-1 PR (*8*, *17*) compared to the semi-open flap conformation of the crystal structures of the substrate-free protease (Figure 1). The overall structure of the substrate-bound HIV-1 PR is more compact compared to that of the substrate-free HIV PR (*18*, *19*). The evidence presented above supports the general model of binding of the substrate to HIV-1 PR,

* To whom correspondence should be addressed. E-mail: gergely.toth@elan.com.

‡ Locus Pharmaceuticals.

§ Present address: Elan Pharmaceuticals, 800 Gateway Blvd., South San Francisco, CA 94080.

‖ Creighton University.

⊥ Present address: Institute of Biochemistry, Biological Research Center, Hungarian Academy of Sciences, Temesvári körút 62, H-7620 Szeged, Hungary.

[1] Abbreviations: HIV, human immunodeficiency virus; PR, protease; MD, molecular dynamics; R_g, radius of gyration; rmsd, root-mean-square deviation; SIV, simian immunodeficiency virus; RSV, Rous sarcoma virus; FIV, feline immunodeficiency virus; EIAV, equine infectious anemia virus.

10.1021/bi060188k CCC: $33.50
Published on Web 05/03/2006

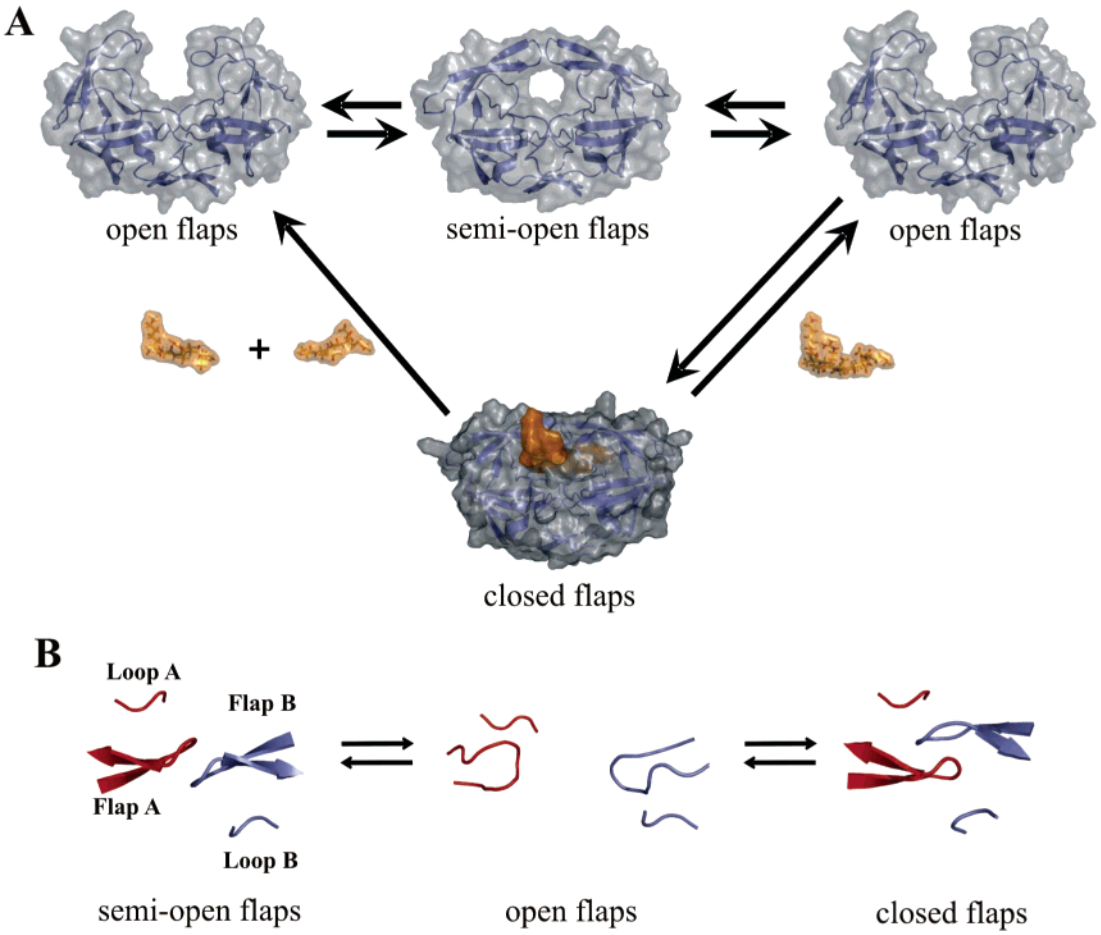

FIGURE 1: Different flap conformations of HIV-1 PR observed during its catalytic cycle. (A) HIV-1 PR is colored gray, and the substrate peptide (Ca-p2) from gag and pol polyprotein is colored with an orange surface representation. (B) Flap A (residues 45−55 of monomer A), flap B (residues 45−55 of monomer B in cyan), loop A (residues 79−82 of monomer A), and loop B (residues 79−82 of monomer B) in ribbon representation shown from the top looking down into the active site (active site not shown).

wherein the substrate first enters the active site through open flaps and induces the closing of the flaps. An extensive conformational change must take place upon substrate binding for the flaps to transition from the open to the closed conformation (Figure 1).

Although more than 180 crystal structures of ligand-bound HIV-1 PR are available in the Protein Data Bank, the mechanism of flap closing is still unknown. The structural basis for flap closing has been difficult to investigate using experimental techniques at an atomic resolution due to the high flexibility of the flaps and the short time scales associated with their closing. Computational techniques such as MD simulation, however, can provide key structural and dynamic details for the understanding of their function in the substrate binding process. Since the closing of the flaps of aspartic proteases is a critical step in the catalytic activity of the protease, the elucidation of the mechanism of closing of the flaps provides an improved understanding of the catalytic mechanism of aspartic proteases and could support more efficient inhibitor design. This study sets out to investigate the flap closing mechanism due to substrate binding in HIV-1 PR. A long MD simulation was performed on a molecular complex consisting of a peptide (CA-p2) from the natural substrate cleavage site of the gag/pol polyprotein positioned in the active site of HIV-1 PR with the open flap conformation. During the simulation, the substrate induced the transitioning of the flaps into the closed conformation. Our results shed light on the complex mechanism of flap closing induced by substrate binding and lead us to propose a general model for the flap closing mechanism in retroviral aspartic proteases.

MATERIALS AND METHODS

In this study, two molecular dynamics simulations were performed. The starting structure for the first simulation was the crystal structure of the CA-p2−HIV-1 PR complex (PDB entry 1f7a) (*20*). CA-p2 is a 10-residue peptide (Lys-Ala-Arg-Val-Leu-Ala-Glu-Ala-Met-Ser) derived from the HIV-1 gag/pol polyprotein. CA-p2 contains one of the sites within the HIV-1 gag/pol polyprotein which is cleaved by HIV-1 PR. Calculations performed on this complex are called the substrate−HIV-1 PR complex simulation. The second simulation was performed on the CA-p2−open HIV-1 PR complex and is called the substrate−open HIV-1 PR simulation. The starting structure for the CA-p2−open HIV-1 PR complex simulation was prepared in the following way. The open HIV-1 PR structure was taken from a previous study (*15*), in which MD simulation of the substrate-free HIV-1 PR (PDB entry 1hhp) was performed at 306 K. During this MD simulation, the flaps of the substrate-free HIV-1 PR transitioned from a semi-open to an open conformation. Several conformational families were identified on the basis of the different open flap conformations of HIV-1 PR. Among the open flap conformational families, a representa-

tive structure from one of the conformational families was used as a starting structure for the substrate−open HIV-1 PR simulation in this study. This conformational family was chosen for this study because the active site was most accessible for the substrate in this family among the identified conformational families. To mimic the first step of the catalytic cycle of aspartic proteases (the binding of substrate by the protease to form a loose complex), the substrate was placed into the active site of the open HIV-1 PR in the following way. First, the structure of the open HIV-1 PR was superposed onto the structure of the CA-p2−HIV-1 PR complex (PDB entry 1f7a). The CA-p2 substrate was then placed into the active site of the superposed open HIV-1 PR structure without the coordinates of CA-p2 being changed. Calculations were performed in torsion space using the Imagiro molecular mechanics simulation platform (*15*, *21*) (developed in-house at Protein Mechanics and later at Locus Pharmaceuticals) and the OPLS-AA (*22*) molecular force field. The OPLS-AA force field was projected into internal coordinates using a scheme similar to one proposed previously (*23*). The GB/SA continuum solvent model (*24*) was used to mimic the aqueous environment. Nonbonded interactions were calculated between all atoms.

Initially, each complex was subjected to a short energy minimization utilizing dynamic relaxation (*25*). The energy-minimized structure served as input for Langevin dynamics (LD) simulations. Simulations were performed with the temperature of the system set to 306 K for 5 and 10 ns for the substrate−HIV-1 PR and substrate−open HIV-1 PR complexes, respectively. An error-controlled, adaptive time step integrator, Runge−Kutta−Merson (*26*), was used to integrate Newton's equations of motion, with an average time step of approximately 9 fs. The frictional coefficient used in LD simulation was set to 2 ps^{-1}, to facilitate the observation of otherwise slow, large-scale motions in a reasonable time frame (*27*). Some trajectory analysis was performed using GROMACS version 3.0 (*28*). The figures showing molecular structures were generated using PyMOL (*29*).

RESULTS

Mechanism of Conformational Change in Open HIV-1 PR Induced by Substrate Binding

The first step of the catalytic cycle of aspartic proteases, the binding of substrate by the protease to form a loose complex, was mimicked via placement of the substrate into the active site of the open HIV-1 PR (see more details in Materials and Methods). The introduction of the substrate into the active site of open HIV-1 PR induced a series of conformational changes in the flaps and the loops of the protease during the substrate−open HIV-1 PR simulation. Ultimately, this cascade of conformational changes led to the transition of the flaps from an open to a closed state. Below, this transition is described in chronological order as observed during the simulation.

From 0 to 6 ps. In the first 6 ps of the simulation, the N-terminus of the substrate moved toward flap A and lined up parallel with residues Gly48A, Gly49B, and Ile50A in flap A. The side chain of Phe53A turned above the opposite strand of the β-sheet, forming aromatic−amide (*30*) and CH−π (*31*) interactions with the backbone of Gly48A and Gly49A.

From 6 to 200 ps. During the next 30 ps, flap A separated from loop A and shifted toward the substrate, forming hydrogen bonds between the carbonyl of Gly48A and the amide of Ala2C and between the amide of Ile50A and the carbonyl of Ala2C (Figure 2B).

From 200 to 2400 ps. By ∼200 ps, Ile50A of flap A reached loop B and the formation of a hydrophobic cluster, consisting of Ile50A, Ile50B, and Pro81B of loop B, began (Figure 2C). This cluster, termed the intermediate state cluster, existed for 2500 ps.

From 2400 to 5400 ps. At ∼2400 ps, hydrogen bonds started to form between flap B and the C-terminus of the substrate, starting with the carbonyl of Gly48B and the amide of Ala8C, and between the amide of Gly48B and the carbonyl of Ala8C. The side chain of Phe53B flipped [from a gauche (+) side chain conformation] from pointing away from the flap to being above the opposite strand of the flap [in a gauche (−) side chain conformation], forming aromatic−amide and CH−π interactions with the backbone of Gly48B and Gly49B. From the looplike conformation of flap B, the formation of a β-hairpin structure was observed. At this point, the tip of flap B was still curved and the intermediate state cluster was still intact (Figure 2D). In the next 2000 ps, the intermediate cluster was in a metastable state as the Ile50B−Pro81B side chain contacts were broken and formed several times. At the end of this period, ∼5400 ps, the intermediate cluster formed for the last time for 200 ps.

From 5400 to 10000 ps. Ile50B slowly separated from Pro81B, and the tip of flap B slowly straightened. Now, hydrophobic contacts existed between the side chains of Ile50A and Ile50B (Figure 2E). At this point, flap B reached a conformation that matched its conformation in the crystal structure of the protease−substrate complex. The C_α atoms of the simulated open HIV-1 PR structure superimposed with a rmsd of 2.8 Å from the crystal structure of HIV-1 PR in complex with the substrate, while the C_α atoms of flap B superimposed with a rmsd of 1.4 Å. Flap A, however, remained in an intermediate state, for more than 5000 ps, until the end of the simulation, with the side chain of Ile50A contacting Pro81B. Flap A did not reach its crystal structure conformation during the 10 ns simulation. It appears that this last step of the binding process, the straightening of flap A, is slow, and the simulation was not long enough to sample this last step of flap closing.

Overall Dynamics

Substrate binding induced the formation of a more compact complex structure of open HIV-1 PR during the substrate−open HIV-1 simulation. In the course of the simulation, the radius of gyration (R_g) of the substrate−open HIV-1 PR complex gradually decreased (Figure 3). At ∼3000 ps, the substrate−open HIV-1 PR complex had R_g values comparable to the R_g values from the substrate−HIV-1 PR complex simulation. This is consistent with the observation that the crystal structure of the substrate-bound form of HIV-1 PR is more compact compared to that of the substrate-free form (*18*, *19*).

During the substrate−open HIV-1 simulation, the rmsd

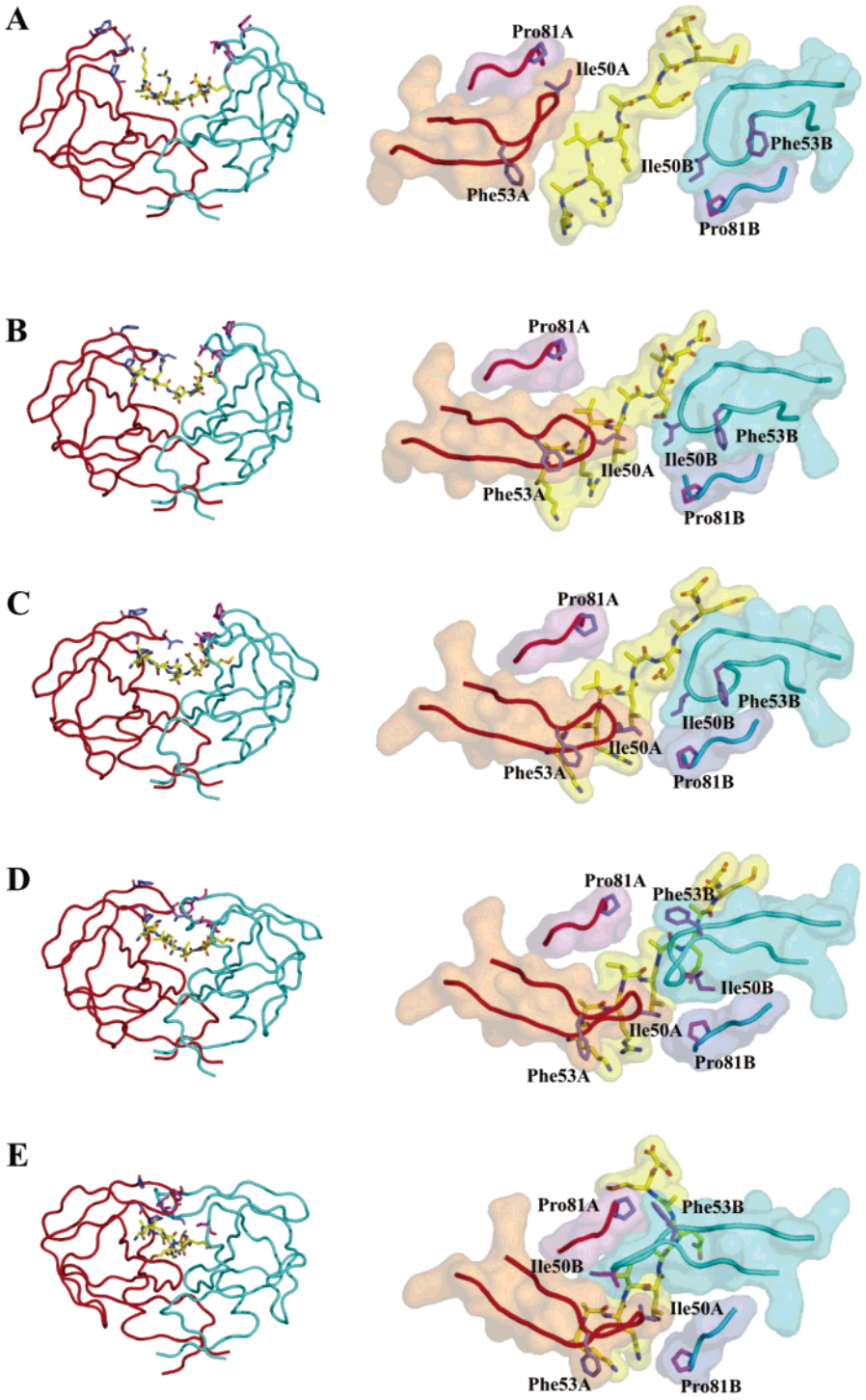

FIGURE 2: Ribbon drawing of the transition of the flaps from an open to a closed state during the simulation of the substrate−open HIV-1 PR complex. The column on the left shows the entire HIV-1 PR, while the column on the right shows only flap A, flap B, loop A, loop B, and the substrate peptide from the top looking down into the active site (active site not shown). The surfaces are shown in transparent representation: flap A in red, flap B in cyan, loop A and loop B in blue, and substrate peptide in yellow. The backbone of HIV-1 PR is shown in tube representation: monomer A in red, monomer B in cyan, and substrate peptide in yellow. Residues Ile50A, Ile50B, Pro81A, Pro81B, Phe53A, and Phe53B are shown in stick representation: (A) substrate−open HIV-1 PR structure representing the starting structure, (B) substrate−open HIV-1 PR structure representing the 0−200 ps period, (C) substrate−open HIV-1 PR structure representing the 200−2400 ps period, (D) substrate−open HIV-1 PR structure representing the 2400−5400 ps period, and (E) substrate−open HIV-1 PR structure representing the 5400−10000 ps period. Further explanation of the figure is given in the text.

of the C_α atoms between the crystal structure of the substrate−HIV-1 PR complex and the simulated substrate−open HIV-1 PR complex decreased from the starting value of 4.9 to 2 Å by 3400 ps, and then gradually increased to ~3.6 Å by 8000 ps (Figure 4A). The increase in the rmsd is due to the high mobility of loop C, residues 66−74, during the second half of the simulation. The average rmsd of C_α atoms between the crystal structure of the substrate−HIV-1 PR complex and the simulated substrate−HIV-1 PR complex was 2.4 Å during the 5 ns of simulation.

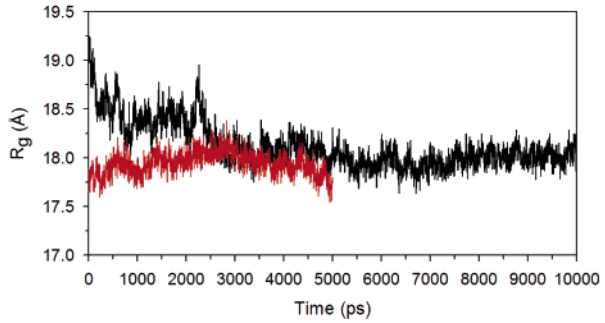

FIGURE 3: Evolution of R_g of HIV-1 PR during the simulations of the substrate−open HIV-1 PR complex (black) and the substrate−HIV-1 PR complex (red).

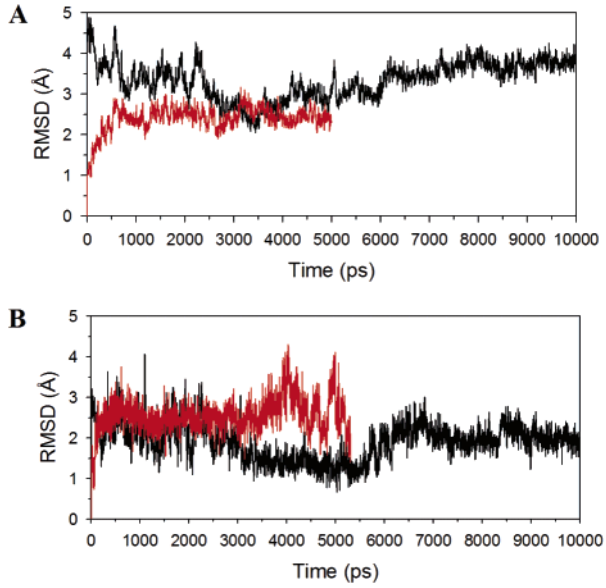

FIGURE 4: (A) Evolution of the rmsd of C_α atoms of HIV PR between the simulated substrate−HIV-1 PR complex (red) and the simulated substrate−open HIV-1 PR complex (black) and the crystal structure of the substrate−HIV PR complex. (B) Evolution of the rmsd of C_α atoms of the substrate (for residues 3−8 in positions P3−P3′, when P1−P1′ is the cleavage site) between the simulated substrate−HIV-1 PR complex (red) and the simulated substrate−open HIV-1 PR complex (black) and the crystal structure of the substrate−HIV PR complex. The rmsd of the substrate was calculated by superposing the simulated substrate with the substrate in the crystal structure by fitting the C_α atoms of residues 1−45 and 55−99 of both monomers.

The fluctuation and position of the substrate peptide in both simulations existed to similar extents. The rmsd values of C_α atoms between the crystal structure of the substrate peptide and the simulated substrate, for residues 3−8 in positions P3−P3′ (where P1−P1′ is the cleavage site), in the HIV-1 PR complex and the open HIV-1 PR simulations were comparable (Figure 4B). This indicates that the substrate remained close to its substrate−HIV-1 PR complex position in the bottom of the active site throughout the substrate−open HIV-1 PR simulation. Interestingly, somewhat lower rmsd values were observed for the substrate−open HIV-1 PR simulation. The increase in the rmsd of the substrate in the substrate−open HIV-1 PR simulation after 6000 ps was attributed to the high mobility of loop C (see Figure 4 for fitting details). The conserved hydrogen bonds described by Prabu-Jeyabalan and co-workers (*20*) between the substrate and the protease were observed to be mostly stable during the substrate−HIV-1 PR simulation.

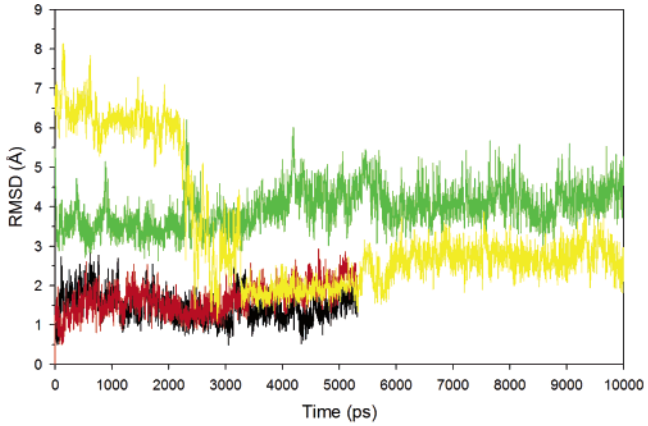

FIGURE 5: Evolution of the rmsd of C_α atoms of the flaps (residues 45−55) between the simulated substrate−HIV-1 PR complex (red and black for flaps A and B, respectively) and the simulated substrate−open HIV-1 PR complex (green and yellow for flaps A and B, respectively) and the crystal structure of the substrate−HIV PR complex. The rmsd of the flaps was calculated by superposing the simulated HIV-1 PR with HIV-1 PR in the crystal structure by fitting the C_α atoms of residues 1−45 and 55−99 of both monomers.

Only a single hydrogen bond between the flaps (Gly51A NH···Ile50B CO) was observed during the substrate−HIV-1 PR simulation (*17*). It was found to be partially stable, since it broke and re-formed several times. The formation of this hydrogen bond was also observed in the substrate−open HIV-1 PR simulation after 5800 ps. Because of the low number of interactions between the flaps, the stability of the closed flap conformation appears to originate from the interactions between the flaps and the substrate.

Conformational Changes in the Flaps Induced by Substrate Binding

The extent of conformational change of the flaps during the simulation is illustrated in Figure 5 by plotting the rmsd of C_α atoms of the flaps between the crystal structure of the substrate−HIV-1 PR complex and the simulated substrate−open HIV-1 PR complex. Initially, the rmsd values of flaps A and B between the crystal structure of the substrate−HIV-1 PR complex and the open HIV-1 PR were 5.5 and 6.8 Å, respectively. While the rmsd of flap A dropped 2 Å to 3.5 Å in the first few picoseconds and stayed around this value for the rest of the simulation, the rmsd of flap B gradually decreased and converged to the rmsd values of flap B of the substrate−HIV-1 complex.

The structure of flap B in the open HIV-1 PR was slightly different compared to the β-hairpin structure of the semi-open and closed HIV-1 PR crystal structures. The β-hairpin structure was broken by a bend between residues 47 and 52, which caused the tip of the flap to twist ~75° compared to a straight β-hairpin structure. Although flap A of the open HIV-1 PR had a β-hairpin structure, this flap was slightly curved, resulting in the similar contacts between the tip of the flap and loop B as observed between the tip of flap B and loop A. During the simulation of the binding of the substrate to the open HIV-1 PR, the looplike conformation of flap B transitioned into a β-hairpin structure gradually in 2500 ps at the same time that the tip of flap B formed hydrogen bonds with the substrate and the side chain of Phe53 flipped over and stacked above Gly48 and Gly49 (Figure

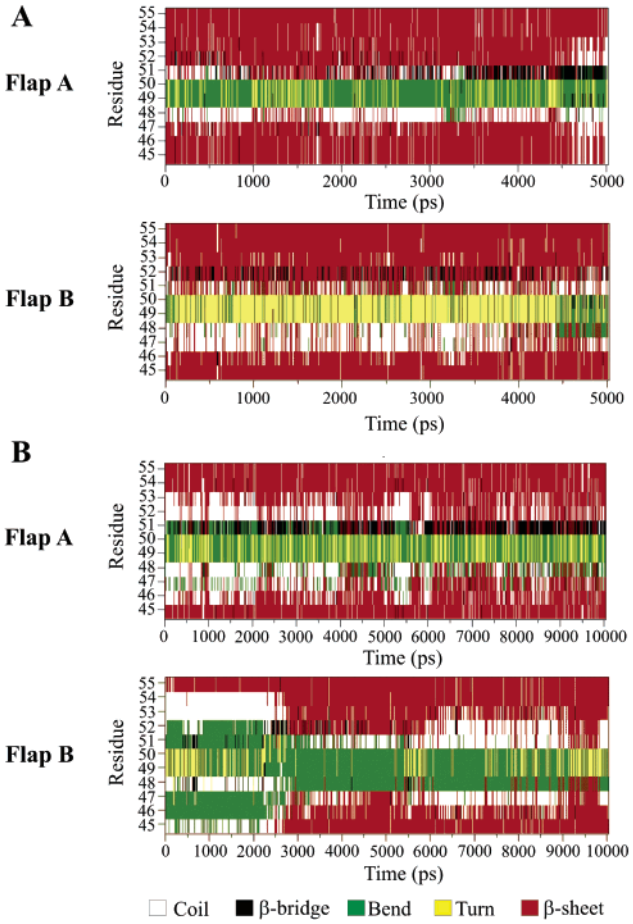

FIGURE 6: Evolution of the secondary structure of the flaps (residues 45−55) in the (A) substrate−HIV-1 PR and (B) substrate−open HIV-1 PR simulations. The following color code was used for the illustration of secondary structure: black for random coil, green for β-bridge, yellow for turn, and red for β-sheet.

6). After a few picoseconds of simulation, flap A lost some of its bridging hydrogen bonds between the two strands of the β-sheets and only regained them after 4000 ps.

The flexibility of the flaps and the stability of the β-hairpin structure may be also partially related to the side chain conformational transition of Phe53 during binding of the substrate to open HIV-1 PR. In the substrate-free, semi-open flap conformation, the aromatic rings of Phe53 point outward on both flaps, while in the substrate−HIV-1 PR complex structure, these aromatic side chains are positioned above the flaps forming aromatic−amide and CH−π interactions with the backbone amide of Gly49 and the HC_α group of Gly49. Such aromatic−glycyl interactions in β-sheets have been shown to have a stabilizing affect on the secondary structure (*32*, *33*). During the simulation, the side chains of Phe53 on both flaps moved above the opposite strand of the β-sheet, forming aromatic−amide and CH−π interactions with the backbone of Gly48 and Gly49. The overall effect of these interactions on the flaps may be to reduce the flexibility of the flaps by stabilizing the β-sheet structure.

DISCUSSION

Closing Mechanism of the Flaps. The following mechanism is proposed for flap closing induced by substrate binding on the basis of our simulations. The closing of the flaps was observed to proceed in an asymmetrical way. First, the closing of flap A was initiated by the formation of flap A−substrate interactions, which led to the assembly of a hydrophobic intermediate state cluster consisting of Ile50A, Ile50B, and Pro81B. The presence of the intermediate state cluster slowly triggered the formation of flap B−substrate interactions and the closing of flap B. The last step of the mechanism appears to be the final closing of flap A.

Initial contacts between the flaps and the substrate were established through backbone hydrogen bonds between the Gly48 residues of the flaps and the substrate. This is consistent with the observation that HIV-1 PR recognizes a particular shape rather than a substrate sequence and that the substrate must have an extended conformation to take part in the numerous backbone−backbone hydrogen bonds necessary for the binding of the different substrates (*20*). The stabilization of the β-sheet structure of the flaps and the subsequent reduction of their flexibility during the flap closing process can be attributed to Phe53, since its side chain aromatic ring flipped over the opposite strand and stacked over the backbone of Gly48 and Gly49. This appears to be a necessary step for the formation of flap−substrate interactions and for achieving the closed conformation of the flaps.

Interestingly, the same intermediate state cluster was observed in the opening of flaps from a semi-open flap conformation of HIV-1 PR (*15*). These results suggest that this arrangement of the flaps is an intermediate state, which can be accessed by the flaps in their closing and opening transitions. The role of the intermediate state cluster in flap closing was dual as revealed by the simulations. It stabilized a flap conformational state which ensured proper geometry for the closing of the second flap, while it also initiated the closing of the second flap.

A water molecule (Wat3) was observed to serve as a bridge between the backbone amide groups of Ile50A and -B of the flaps and the backbone carbonyl groups of ValP1′ and AlaP1′ of the substrate in the crystal structure of the HIV-1 PR−CA-p2 substrate complex (*19*). Although no explicit water molecules were included in our simulations, the aqueous environment was mimicked by a continuum solvent model. The observation from our simulations that the flaps close in a continuum water environment suggests that the presence of bulk water is important while single water molecules, like Wat3, have no role or a limited role in the early stages of the flap closing mechanism. However, water molecules inherently become part of the complex and affect both the final conformation of the HIV-1 PR−substrate complex and the strength of their interaction.

Mutagenesis Studies Support the Proposed Mechanism of Flap Closing. Direct verification at atomic resolution of the mechanism of flap closing proposed herein is not possible with currently available experimental methods to the best of our knowledge. However, much biochemical data is available on mutants of HIV-1 PR which may be related to the proposed flap closing mechanism. Flaps A and B and loops A and B are two of the three regions most sensitive to mutations in HIV-1 PR (*34*). Nonconservative mutagenesis studies showed that any replacement of residues Ile47, Gly49, Ile50, Gly51, and Gly52 results in a loss of protease activity (*34*). Although the backbone carbonyl and amide groups of Gly48 were shown to be important in the initiation of flap closing, mutagenesis studies found that HIV-1 PR is able to accommodate nonconservative mutations in this position. Our

evaluation of the effect of nonconservative mutations of Gly48 found that a mutation of Gly48 results in the positioning of the side chain of the newly mutated residue pointing into the solvent and away from the active site and the substrate, given that the β-sheet structure of the backbone is held. Therefore, the introduction of a side chain in this position should not affect the formation of backbone−substrate interactions. The mutagenesis of Phe53 to other aromatic and hydrophobic residues conserved protease activity (*34*). Both of these types of mutations conserve β-sheet stability (*32*, *33*) and thus conserve the nature and role of Phe53. In loops A and B, only the nonconservative mutation of Val82 is accommodated.

The flap closing mechanism proposed in this study for wild-type HIV-1 PR also prevails among drug resistant mutants of HIV-1 PR. Scott and Schiffer (*13*) provided a detailed summary of observed resistant mutations in the flaps and loops A and B, which appear in patients treated with HIV-1 PR inhibitors. According to their summary, only the following residues mutate and confer drug resistance to the protease in the tips of the flaps: Gly48 (to Val), Ile50 (to Val or Leu), and Phe53 (to Leu), and Val82 (to Ala, Thr, Phe, or Ser) in loops A and B. As discussed above, these mutations of Gly48 and Phe53 are consistent with the structural basis of the flap closing observed during the simulation. The Ile50 to Val or Leu mutations conserve the nature of the side chain, and therefore, these residues are able to substitute for Ile50 in the flap closing mechanism.

Flap Closing in Retroviral Aspartic Proteases. According to the MEROPS database, 131 sequences currently belong to retroviral protease family A2 of aspartic peptidase clan AA (*35*). The amino acid sequences of retroviral proteases are significantly similar, particularly in the locations of residues that are important in preserving both structure and function (*36*). The crystal structures of six different inhibitor-bound retroviral proteases have been determined to date. Wlodawer and Gustchina reported (*37*) a detailed description of the superposition of these structures. They found that the structures superimpose well onto HIV-1 PR with a C_α rmsd between 1.08 and 1.78 Å.

To examine weather the flap closing mechanism in HIV-1 PR may be a general model for flap closing in retroviral aspartic proteases, the following detailed analysis across the retroviral aspartic protease family was performed. The following residues were observed to be important during the flap closing in HIV-1 PR: tip of the flaps (residues 48−52), Phe53, and Pro81. The conservation of the nature of these residues in retroviral aspartic proteases with known crystal structure indicates that the flap closing may happen with a mechanism similar to that in HIV-1 PR (Figure 7). In a review, Wlodawer and Gustchina present an alignment of amino acid sequences on the basis of three-dimensional structure of six retroviral proteases compared to HIV-1 PR (*37*). This alignment reveals that the tips of the flaps contain mostly Gly in all six proteases, suggesting that the tips of the flaps remain flexible in the family. The hydrophobic nature of Ile50 is conserved in all six proteases. Ile50 is conserved in human immunodeficiency virus type 2 (HIV-2) PR, in simian immunodeficiency virus (SIV) PR, and in Rous sarcoma virus (RSV), but it is replaced with Val in feline immunodeficiency virus (FIV) PR and equine infectious anemia virus (EIAV) PR. The hydrophobic nature of

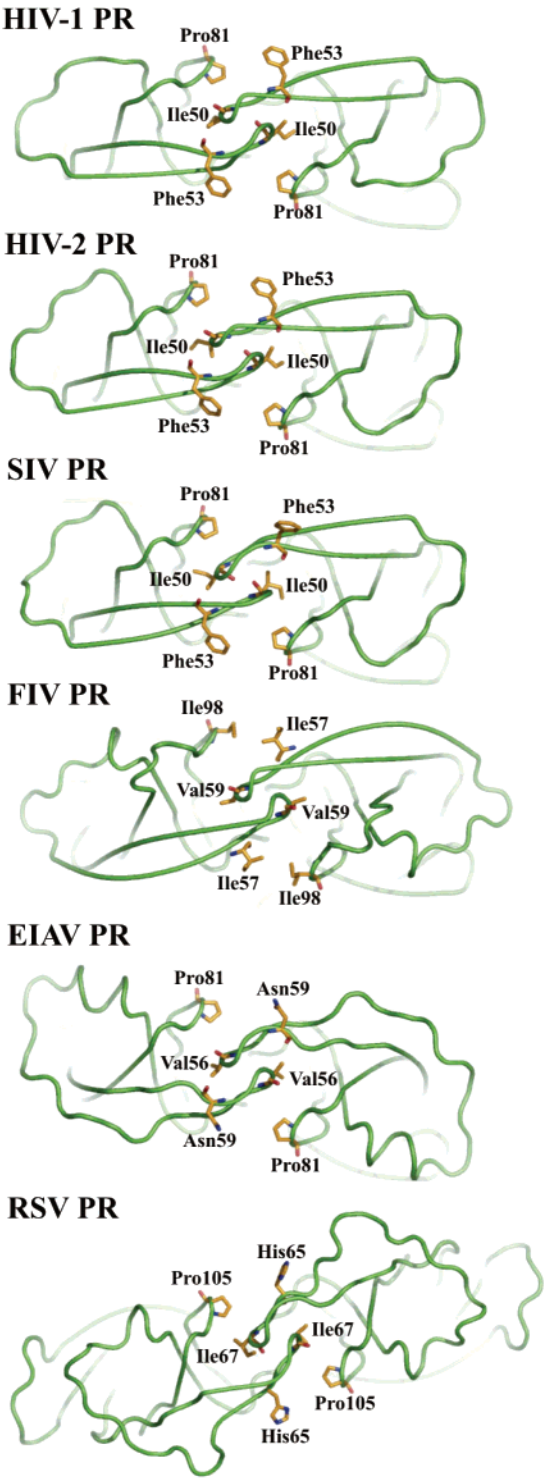

FIGURE 7: Illustration of the conservation of the nature of the residues (residues 48−52, Phe53, and Pro81, which were observed to be important in flap closing in HIV-1 PR due to substrate binding, in the structure of retroviral aspartic proteases with known crystal structure. The following abbreviations are used: HIV-2 for human immunodeficiency virus type 2 (PDB entry 1ida), SIV for simian immunodeficiency virus (PDB entry 1siv), RSV for Rous sarcoma virus (PDB entry 1bai), FIV for feline immunodeficiency virus (PDB entry 4fiv), and EIAV for equine infectious anemia virus (PDB entry 2fmb).

Pro81 is also conserved in all six proteases. Pro81 is conserved in HIV-2 PR, SIV PR, and EIAV PR, while it is replaced with Ile98 in FIV PR and Val104 in RSV PR. It appears that the role of Phe53 is also conserved among the

six proteases, although sometimes not by the Phe residue and not in the same position of the flap. Phe53 is conserved in HIV-2 PR and SIV PR, but it is replaced by Ile57, Asn57, and His65 in FIV PR, EIAV PR, and RSV PR, respectively. All this evidence indicates that the flap closing mechanism described in HIV-1 PR may be a general model for flap closing in retroviral aspartic proteases.

CONCLUSION

Flap closing is considered to be a general phenomenon in substrate binding in the catalytic mechanism of aspartic proteases. Nevertheless, little has been known about the structural nature and dynamics of flap closing. The results of this study shed light on the complex mechanism of flap closing due to substrate binding in wild-type HIV-1 PR and drug resistant mutants. Detailed analysis across the structure of the retroviral aspartic proteases revealed that the nature of the residues of HIV-1 PR identified to be important in the flap closing mechanism is conserved across known structures of the retroviral aspartic protease family. On this basis, the flap closing mechanism described in HIV-1 PR is proposed to be a general model for flap closing in retroviral aspartic proteases.

The results of this study offer a unique perspective into the structural insights of substrate binding by retroviral aspartic proteases. Generally, it has been difficult to interpret such binding mechanisms because crystal structures provide information about only substrate-free and substrate-bound forms of proteins. The mechanism of conformational change between these two forms of the protein has been, therefore, difficult to establish. Long MD simulations, such as the one in this study, should become an effective way of exploring such conformational changes in the future.

The ability to control flap opening or closing could offer a novel and alternative strategy for the inhibition of aspartic proteases. The structural insights provided by the results of this study may bring us closer to the development of such a strategy and aid efforts toward the design of a new generation of HIV-1 PR inhibitors.

NOTE ADDED IN REVIEW

During the review of this paper, a communication titled HIV-1 Protease Flaps Spontaneously Close to the Correct Structure in Simulations Following Manual Placement of an Inhibitor into the Open State was published by Hornak et al. (*38*), which reports on the closing of the flaps of HIV-1 PR induced by a cyclic urea inhibitor manually docked into the active site of the protease observed by molecular dynamics simulation. The report by Hornak and co-workers is compelling evidence that our observation on the closing of the flaps of HIV-1 PR is not dependent on the starting conformation or model.

ACKNOWLEDGMENT

We gratefully acknowledge the work of co-workers at Protein Mechanics and Locus Pharmaceuticals who have contributed to the development of Imagiro.

REFERENCES

1. Dunn, B. M. (2002) Structure and Mechanism of the Pepsin-Like Family of Aspartic Peptidases, *Chem. Rev. 102*, 4431−4458.
2. Hong, L., and Tang, J. (2004) Flap position of free memapsin 2 (*β*-secretase), a model for flap opening in aspartic protease catalysis, *Biochemistry 43*, 4689−4695.
3. Patel, S., Vuillard, L., Cleasby, A., Murray, C. W., and Yon, J. (2004) Apo and inhibitor complex structures of BACE (*β*-secretase), *J. Mol. Biol. 343*, 407−416.
4. Katz, R. A., and Skalka, A. M. (1994) The retroviral enzymes, *Annu. Rev. Biochem. 63*, 133−73.
5. Wlodawer, A., and Vondrasek, J. (1998) Inhibitors of HIV-1 protease: A major success of structure-assisted drug design, *Annu. Rev. Biophys. Biomol. Struct. 27*, 249−284.
6. Spinelli, S., Liu, Q. Z., Alzari, P. M., Hirel, P. H., and Poljak, R. J. (1991) The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU, *Biochimie 73*, 1391−1396.
7. Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, P. E., Geoghegan, K. F., Hawrylik, S. J., Lee, S. E., Scheld, K. G., and Hobart, P. M. (1989) X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes, *Nature 342*, 299−302.
8. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. (1989) Conserved folding in retroviral proteases: Crystal structure of a synthetic HIV-1 protease, *Science 245*, 616−621.
9. Ishima, R., Freedberg, D. I., Wang, Y. X., Louis, J. M., and Torchia, D. A. (1999) Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function, *Struct. Folding Des. 7*, 1047−1055.
10. Freedberg, D. I., Ishima, R., Jacob, J., Wang, Y. X., Kustanovich, I., Louis, J. M., and Torchia, D. A. (2002) Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations, *Protein Sci. 11*, 221−232.
11. Katoh, E., Louis, J. M., Yamazaki, T., Gronenborn, A. M., Torchia, D. A., and Ishima, R. (2003) A solution NMR study of the binding kinetics and the internal dynamics of an HIV-1 protease-substrate complex, *Protein Sci. 12*, 1376−1385.
12. Nicholson, L. K., Yamazaki, T., Torchia, D. A., Grzesiek, S., Bax, A., Stahl, S. J., Kaufman, J. D., Wingfield, P. T., Lam, P. Y., Jadhav, P. K., Hodge, C. N., Domaille, P. J., and Chang, C. (1995) Flexibility and function in HIV-1 protease, *Nat. Struct. Biol. 2*, 274−280.
13. Scott, W. R., and Schiffer, C. A. (2000) Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance, *Struct. Folding Des. 8*, 1259−1265.
14. Collins, J. R., Burt, S. K., and Erickson, J. W. (1995) Flap opening in HIV-1 protease simulated by 'activated' molecular dynamics, *Nat. Struct. Biol. 2*, 334−338.
15. Toth, G., and Borics, A. (2005) Flap Opening Mechanism of HIV-1 Protease, *J. Mol. Graphics Modell.* (in press).
16. Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, *Proc. Natl. Acad. Sci. U.S.A. 103*, 915−920.
17. Rick, S. W., Erickson, J. W., and Burt, S. K. (1998) Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease, *Proteins 32*, 7−16.
18. Louis, J. M., Dyda, F., Nashed, N. T., Kimmel, A. R., and Davies, D. R. (1998) Hydrophilic peptides derived from the transframe region of Gag-Pol inhibit the HIV-1 protease, *Biochemistry 37*, 2105−2110.
19. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2000) How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease, *J. Mol. Biol. 301*, 1207−1220.
20. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) Substrate shape determines specificity of recognition for HIV-1 protease: Analysis of crystal structures of six substrate complexes, *Structure 10*, 369−381.
21. Carnevali, P., Toth, G., Toubassi, G., and Meshkat, S. N. (2003) Fast protein structure prediction using Monte Carlo simulations with modal moves, *J. Am. Chem. Soc. 125*, 14244−14245.
22. Kaminski, G. R. A. F., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides, *J. Phys. Chem. B 105*, 6474−6487.

23. Katritch, V., Totrov, M., and Abagyan, R. (2003) ICFF: A new method to incorporate implicit flexibility into an internal coordinate force field, *J. Comput. Chem. 24*, 254−265.
24. Qui, D., Shenkin, P. S., Hollinger, F. P., and Still, W. C. (1997) The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii, *J. Phys. Chem. A 101*, 3005−3014.
25. Underwood, P. (1983) Dynamic Relaxation, in *Computational Methods for Transient Analyses* (Belytschko, T. A., Ed.) pp 245−265, Elsevier Science Publishers, Amsterdam.
26. Butcher, J. C. (1987) *The Numerical Analysis of Ordinary Differential Equations: Runge−Kutta and General Linear Methods*, Wiley, Chichester, U.K.
27. Zagrovic, B., and Pande, V. (2003) Solvent viscosity dependence of the folding rate of a small protein: Distributed computing study, *J. Comput. Chem. 24*, 1432−1436.
28. Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis, *J. Mol. Model. 7*, 306−317.
29. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlo, CA.
30. Toth, G., Watts, C. R., Murphy, R. F., and Lovas, S. (2001) Significance of aromatic-backbone amide interactions in protein structure, *Proteins 43*, 373−381.
31. Nishio, M., Umezawa, Y., Hirota, M., and Takeuchi, Y. (1995) The CH/π interaction: Significance in molecular recognition, *Tetrahedron 51*, 8665−8701.
32. Merkel, J., and Regan, L. (1998) Aromatic rescue of glycine in β sheets, *Folding Des. 3*, 449−455.
33. Hutchinson, E. G., Sessions, R. B., Thornton, J. M., and Woolfson, D. N. (1998) Determinants of strand register in antiparallel β-sheets of proteins, *Protein Sci. 7*, 2287−2300.
34. Loeb, D. D., Swanstrom, R., Everitt, L., Manchester, M., Stamper, S. E., and Hutchison, C. A. (1989) Complete mutagenesis of the HIV-1 protease, *Nature 340*, 397−400.
35. Rawlings, N. D., Tolle, D. P., and Barrett, A. J. (2004) MEROPS: The peptidase database, *Nucleic Acids Res. 32*, D160−D164.
36. Dunn, B. M., Goodenow, M. M., Gustchina, A., and Wlodawer, A. (2002) Retroviral proteases, *Genome Biol. 3*, 3006.1−3006.7.
37. Wlodawer, A., and Gustchina, A. (2000) Structural and biochemical studies of retroviral proteases, *Biochim. Biophys. Acta 1477*, 16−34.
38. Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 Protease Flaps Spontaneously Close to the Correct Structure in Simulations Following Manual Placement of an Inhibitor into the Open State, *J. Am. Chem. Soc. 128*, 2812−2813.

BI060188K