Comparative analysis of the unbinding pathways of antiviral drug Indinavir from HIV and HTLV1 proteases by Supervised Molecular Dynamics simulation

Farzin Sohraby, Hassan Aryapour*

Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran.

*Corresponding author. Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran. Tel No: +98-17-32254161; Fax No. +98-17-32245964; E-mail address: hassan.aryapour@gmail.com.

Abstract

Determining the unbinding pathways of potential small molecule compounds from their target proteins is of great significance for designing efficacious treatment solutions. One of these potential compounds is the approved HIV-1 protease inhibitor, indinavir, which has a weak effect on the HTLV-1 protease. In this work, by employing SuMD method, we reconstructed the unbinding pathways of indinavir from HIV and HTLV-1 proteases in order to compare and to understand the mechanism of the unbinding, and also to discover the reasons for the lack of inhibitory activity of indinavir against the HTLV-1 protease. We achieved multiple unbinding events from both HIV and HTLV-1 proteases in which the RMSD values of indinavir reached over 40 Å. Also, we found that the mobility and the fluctuations of the flap region is higher in the HTLV-1 protease which makes the drug less stable. We realized that critically positioned aromatic residues such as Trp98/Trp98' and Phe67/Phe67' in the HTLV-1 protease can make strong π-Stacking interactions with indinavir in the unbinding pathway which are unfavorable for the stability of indinavir in the active site. The details found in this study can make a good explanation for the lack of inhibitory activity of this drug against HTLV-1 protease. We believe the details discovered in this work can be a great assist for designing more effective and more selective inhibitors for the HTLV-1 protease.

Keywords: HIV protease; HTLV-1 protease; Indinavir; Supervised Molecular dynamics (SuMD).

Introduction

In the last decades, retroviruses have always been a great threat to the lives of human being and since then, countless people lost their lives as a result. Retroviruses are the oldest family of viruses on planet earth. Recent bioinformatics advances have estimated that retroviruses have originated about 500 million years ago [1] and have been inserted into the human genome about a million years ago [2]. Over these years, the genome of retroviruses, like most family of viruses, have been added to the genome of us and as a result about 8 to 10% of the human genome has a viral origin that is the sign of ancient germ line infections. They are known as Endogenous Retroviruses (ERVs) [1, 3] which may have probably evolved from transposable elements. However, RNA enveloped Viruses such as Human Immunodeficiency Virus (HIV) or Human T-cell Leukemia Virus 1 (HTLV-1) have to be transmitted horizontally among hosts and belong to the subfamily of Exogenous Retroviruses (XRVs).

HIV has been the most advertised virus on the planet earth. Since its occurrence in the 1980s, this lethal virus has taken millions of lives. Since the start of the epidemic 32 million people have died from AIDS-related illnesses [4]. This pandemic virus forced the health organizations to speed up the research to find treatment options. Fortunately, the extensive research and development to eradicate HIV led to treatment solutions with good efficacy. In 1987, the FDA approved the first drug for the treatment of HIV called zidovudine (AZT). AZT belongs to a class of drugs known as Nucleoside Reverse Transcriptase Inhibitors (NRTIs) which cause premature determination of the proviral DNA. Many more drugs and solutions were achieved afterwards and now HIV infected individuals are almost able to live a normal life with a usual life expectancy.

The viral proteases are one of the important targets for the treatment of virus infection. Over the last 20 years, HIV protease inhibitors have shown that inhibiting viral protease can be a very good strategy to fight the virus. Nowadays, protease inhibitors are one of the main parts of the combination therapies of HIV treatments. Approved drugs such as indinavir, saquinavir etc. have good efficacy for inhibiting the HIV protease and have been proven to have a great impact on the condition of the patients [5, 6]. The first Anti-Retroviral Therapy (ART) was based on NRTIs

which were not very effective and caused many complications but the development and the emergence of protease inhibitors introduced a new milestone in ART. This milestone gave hope to both the developers and also the patients that this infection may be conquerable. Combination therapies are providing excellent efficacy and show that protease inhibitors are the main parts of the successful treatments [7-12].

The HTLV-1 virus is associated with adult T-cell leukemia (ATL) and an inflammatory disease syndrome called HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) [13-15]. Apparently, 90 percent of infected individuals remain asymptomatic, but due to the permanent entanglement of the virus and the immune system, the infected person will suffer from immune system deficiencies throughout his/her life which can lead to other problematic diseases [16-18]. HTLV-1 has infected about 20 million people worldwide [15], and treatment solutions are urgently needed. Anti-ATL vaccines have been developed recently and are in clinical trials but they are designed to fight the cancerous cells not the virus itself [19]. Unlike HIV which first-line therapies and treatments are available for it, HTLV-1 does not have any treatments to this date. A strong strategy to fight this virus is the direct inhibition of its essential work machines such as the protease enzyme the same as the effective HIV treatments [20]. As it was mentioned before, protease inhibitors have proved their efficacy and in the case of HTLV-1, protease inhibitors may be a potential breakthrough in HTLV-1 treatments [21].

In the last past years, all of the inhibitors of the HIV protease have been tested against the HTLV-1 protease. However, none of the inhibitors exhibited sufficient inhibition. Despite the fact that the two enzymes are very similar in structure, the anti-HIV drug with an inhibitory constant (Ki) of 540 pM [22], indinavir, could only achieve an inhibitory constant of 3.5 μM against the HTLV-1 protease [23, 24]. The only consolation is that the cause of this can only be the difference in their structural behavior and the difference in the sequences of these two enzymes. Specially, the residues of the active site and the flap regions can dictate this behavior [25-28].

Computational methods such as Molecular Dynamics (MD) simulation enables researchers to thoroughly study the structural details of complex mechanisms of biomacromolecules [29]. Studying the unbinding pathway of small molecule compounds in complex with their target proteins is a great feature of MD simulation method that has been made possible in the last decade. Many research groups developed great approaches over these years that empowers us to

understand the fine details of these mechanisms [30-44]. Unraveling the unbinding pathway of a drug from its target protein is a great task that can be very useful to understand the mechanism of the process and also the residues involved [45]. The details achieved can be used to make more effective and more selective drugs. One of the most accurate approaches for this task is the Supervised MD (SuMD) simulations [41]. In this approach, the simulation of the protein-ligand complex is performed in replicas with fixed duration times and at the end of each replica, a specific parameter such the distance of the ligand from the binding site is checked and the frame with the highest distance is extracted and extended as the next replica. This cycle will be continued until the ligand is completely unbinds from the target protein. This atomistic approach is completely unbiased and is very accurate. Herein, by utilizing the SuMD approach, we tried to achieve the unbinding pathway of the anti-HIV drug, indinavir, from both the HIV protease and the HTLV-1 protease as a comparison test in order to understand the details and the behavior of these important protein targets. This can be an effective strategy to find the reasons for the lack of efficacy of the HIV protease inhibitors against the HTLV-1 protease.

Methods

The X-ray crystallography structures of the HTLV-1 (PDB ID: 3WSJ) [24] protease and the HIV protease (PDB ID: 1K6C) [46] in complex with indinavir were obtained from the Protein Data Bank [47]. The structures were prepared by UCSF Chimera software [48]. All of the unnecessary molecules such as water molecules were deleted from the structures and the structures of the protein-ligand complexes were ready for the next step, the MD simulations. There are two catalytic Asp residues in the active site of both proteases and one of them must be protonated during the simulation. The extra hydrogen atom was added to the Asp residue of chain A in both cases. All of the MD simulations were done by GROMACS 2018 package [49] and OPLS-AA force field [50]. The 3D structures of the co-crystallized indinavir was parameterized using ACEPYPE [51] with the default setting for assigning the partial charges and atom types. For the construction of the simulation systems, first, the related protein-ligand complex was placed in the center of a triclinic box with a distance of 1 nm from all edges and then solvated with TIP3P water model[52]. Then, sodium and chloride ions were added to produce a neutral physiological salt concentration of 150 mM. Each system was energy minimized, using steepest descent

algorithm, until the Fmax was found to be smaller than 10 $kJ.mol^{-1}.nm^{-1}$. All of the covalent bonds were constrained using the Linear Constraint Solver (LINCS) algorithm[53] to maintain constant bond lengths. The long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) method[54] and the cut off radii for Coulomb and Van der Waals short-range interactions were set to 0.9 nm for the interaction of the protein-ligand complex. The modified Berendsen (V-rescale) thermostat[55] and Parrinello–Rahman barostat[56] respectively were applied for 100 and 300 ps to keep the system in the stable environmental conditions (310 K, 1 Bar). Finally, SuMD simulations [57] were carried out under the periodic boundary conditions (PBC), set at XYZ coordinates to ensure that the atoms had stayed inside the simulation box, and the subsequent analyses were then performed using GROMACS utilities, VMD[58] and USCF Chimera, and also the plots were created using Daniel's XL Toolbox (v 7.3.2) add-in[59]. The free energy landscapes were rendered using Matplotlib[60]. In addition, to estimate the binding free energy we used the g_mmpbsa package[61].

The difference between the conventional MD (cMD) and the SuMD simulations is the fact that in SuMD, the entire simulation is divided into a series of replicas and a specific parameter is monitored throughout them which is set as the guideline to choose the starting point of the next replica. The original idea was proposed by Giuseppe Deganutti et al. [41] which is a great methodology for achieving unbinding events. For example, in our series of replicas, we considered the distance between the Center Of Mass (COM) of the drug and the COM of the Asp32 and Asp32' in the HTLV-1 and the Asp25 and the Asp25' in the HIV proteases as the guideline for selecting the best frame to be the starting point of the next replica. The duration time of all of the replicas in this study was set to 500 ps. At the end of every 500 ps simulation, the frame with the highest distance was selected as the starting point of the next 500 ps simulation.

Results and Discussion

In this study, by utilizing SuMD simulations, we tried to reveal the unbinding pathway of the anti-HIV drug, indinavir, from the HIV and the HTLV-1 proteases in order to find the details

of the unbinding mechanism, and also to compare the results to understand the lack of inhibitory activity of this drug against the HTLV-1 protease.

The 3D structure of the HTLV-1 and the HIV proteases are very similar. When superimposed, the backbone RMSD value is as low as 0.95 Å (Fig. 1a). Although the positions of the backbone atoms and the protein fold are roughly the same, the sequence identity falls down to 28 percent [62]. The HTLV-1 protease has 125 residues whereas the HIV protease has 99 residues. However, the sequence of the conserved residues in the active sites have an identity value of 45 percent [62]. The various regions of these two enzymes have high structural similarities since their function are the same. But the details of the active sites of the two enzymes differ in some points in order to recognize and accommodate their specific substrates. As a result, none of the HIV protease inhibitors, except for indinavir, can inhibit this enzyme with relatively low concentrations. Kuhnert et al. determined and compared the structures of HIV protease and HTLV-1 protease in complex with indinavir and characterized the role of residues inside the active site and discovered significant deviations with respect to the interaction networks of each moiety (Fig. 1b) with the important sections of the binding pocket [24].

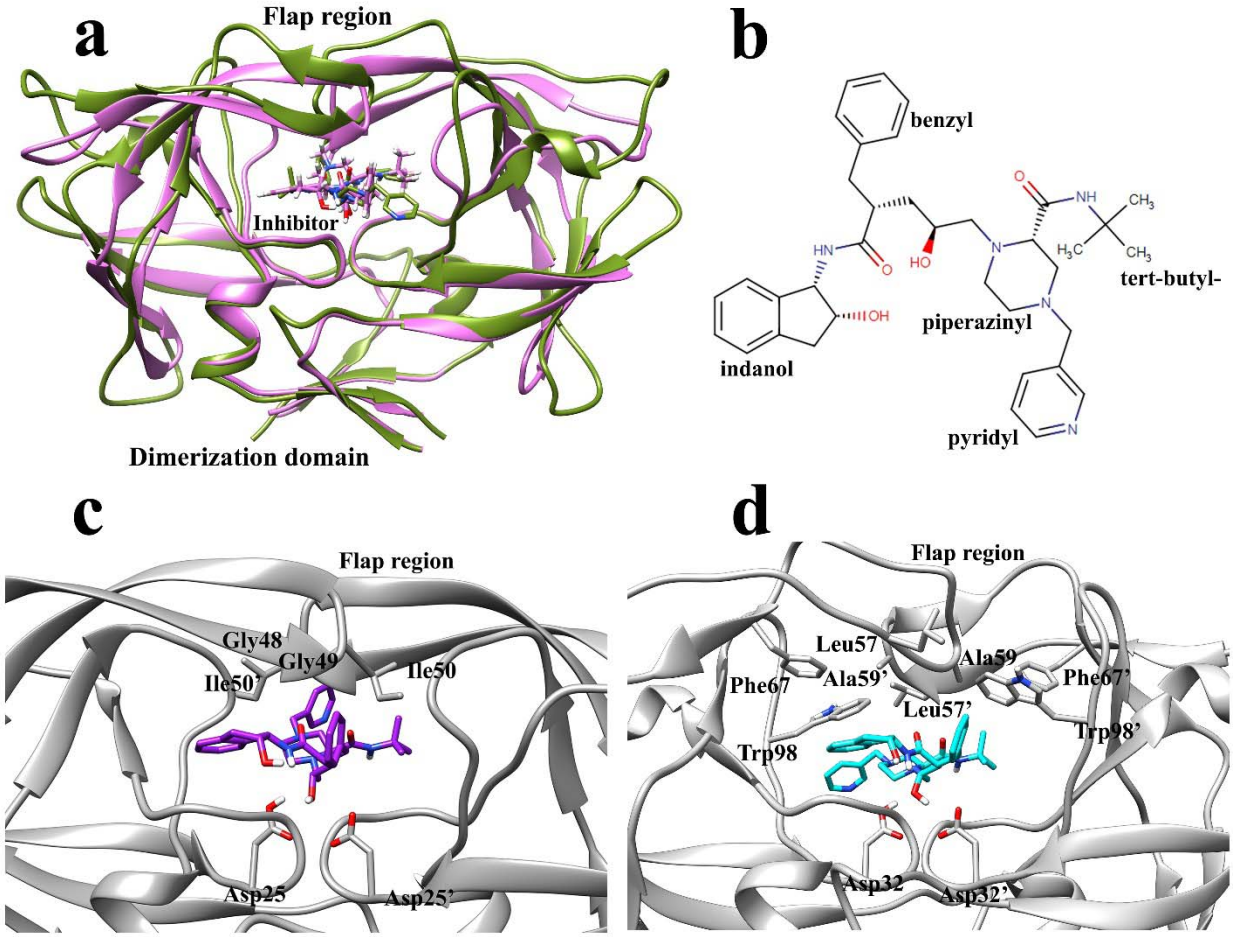

Figure 1. **a,** The superimposed structure of HTLV-1 (green) and HIV (pink) protease in complex with the inhibitor, indinavir, located in the active site of the two enzymes. The backbone RMSD value was calculated by UCSF chimera software. **b,** The 2D structure of indinavir and its corresponding moieties and substitutes. **c,** The position of indinavir and the residues in the active site of HIV protease. **d,** The position of indinavir and the residues in the active site of HTLV-1 protease.

We achieved three unbinding events from each of the complexes in three series of replicas. The replicas were continued until the distance between the two centroids (COMs) reached roughly over 4 nm and the ligand was in the unbound conformation. In this work, the unbinding events happened in 173, 183.5 and 338 ns in the HIV protease-indinavir complex (Fig. 2a), and in the HTLV-1 protease-indinavir complex, the unbinding events happened in 243.5, 552.5 and 671 ns (Fig. 2b). Although indinavir inhibits HIV protease much more selectively than the HTLV-1 protease, the overall time needed for the unbinding events of HTLV-1 protease case

was considerably more than that of the HIV protease case. However, there are a lot of factors governing the activity of an inhibitor and only comparing the overall duration of the unbinding events in three series of replicas is not a correct way of comparing these two cases. This may be due to the inherent limitations of the SuMD method. Because the time window is set to 500 ps, therefore more conformational sampling by ligand is limited.

There are no biasing forces involved in the simulations and they are completely unbiased. The only difference between this method and conventional MD simulations is the automatic supervision at choosing the most appropriate frame in a replica for extending the simulation. As it was explained in the methods section, the most appropriate frame in a 500 ps simulation was the frame with the highest distance between the ligand and the catalytic Asp residues. The SuMD approach is extremely fast and is a remarkable tool for reconstructing the unbinding pathways of small molecule drugs and unraveling their details which are of great significance in the area of drug discovery and design.

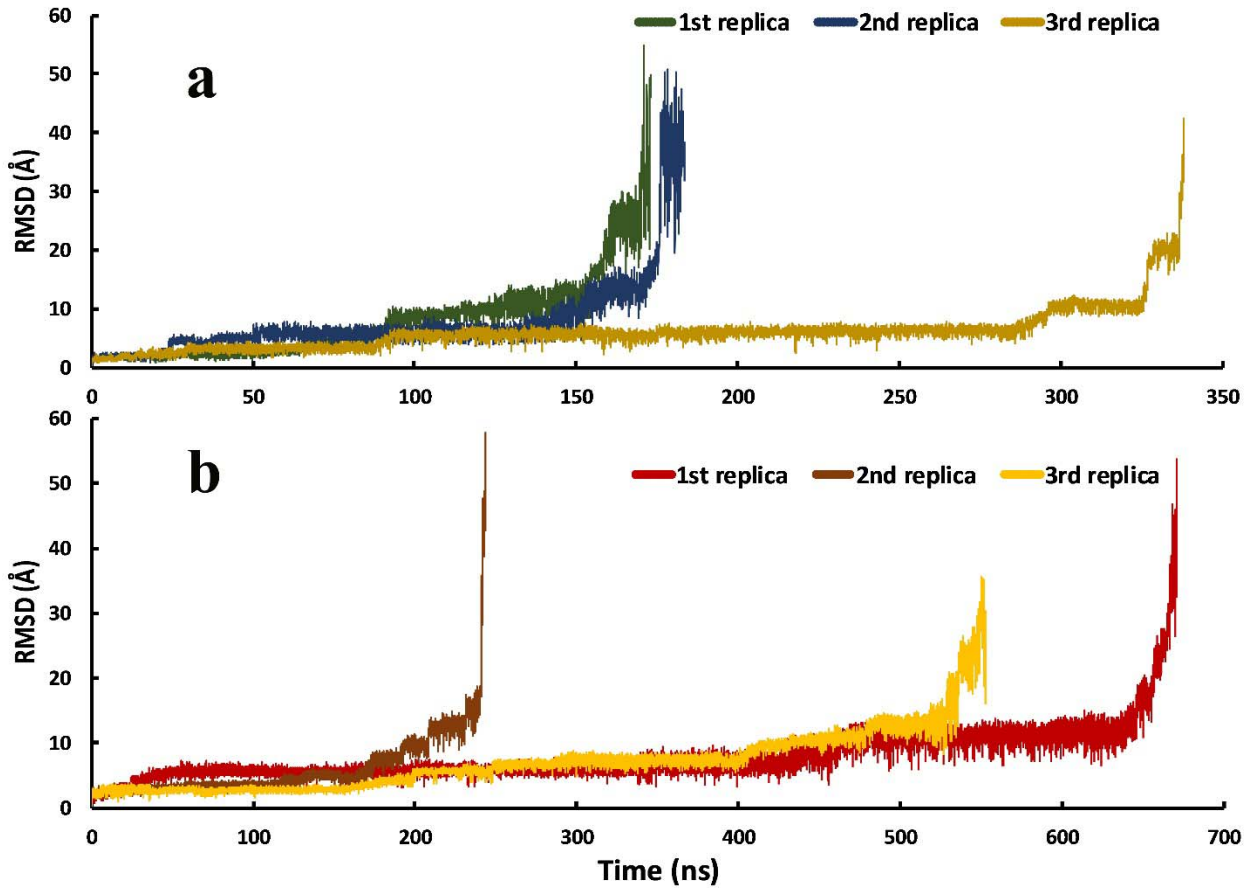

Figure 2. The RMSD values of the indinavir in the unbinding pathways in the three series of replicas for each protein-ligand complex. a, indinavir-HIV protease complex and b, indinavir-HTLV-1 protease complex.

The unbinding pathway of indinavir from the HTLV-1 protease and the HIV protease was very similar. In total, three stable states were observed during the unbinding pathways; (i) the Native state (N), (ii) the Intermediate states (I_1, I_2) and, (iii) the Solvated state (S) (Fig. 3a-f). The native state is where the ligand is in the crystallographic conformation. Indinavir can stay in this state for a long time in the unbinding pathway. In this state, the contact surfaces between the ligand and the protein are at their maximum and the RMSD values are at their minimum. The next state in the unbinding pathway is the intermediate state I_1 and I_2 where the ligand has a tight interaction with the residues of the flap region. The two flaps in both proteases have high mobility and their strong interaction with the ligand can make it get out of the crystallographic conformation in which it has less contact surfaces with the protein and higher RMSD values.

This is the I_1 state. These intermediate states of indinavir in the unbinding pathway are stable and the ligand can stay in them for long times. After the flap region was opened, the ligand can get attached to one of the flaps and get further from the active site, this is the I_2 state.

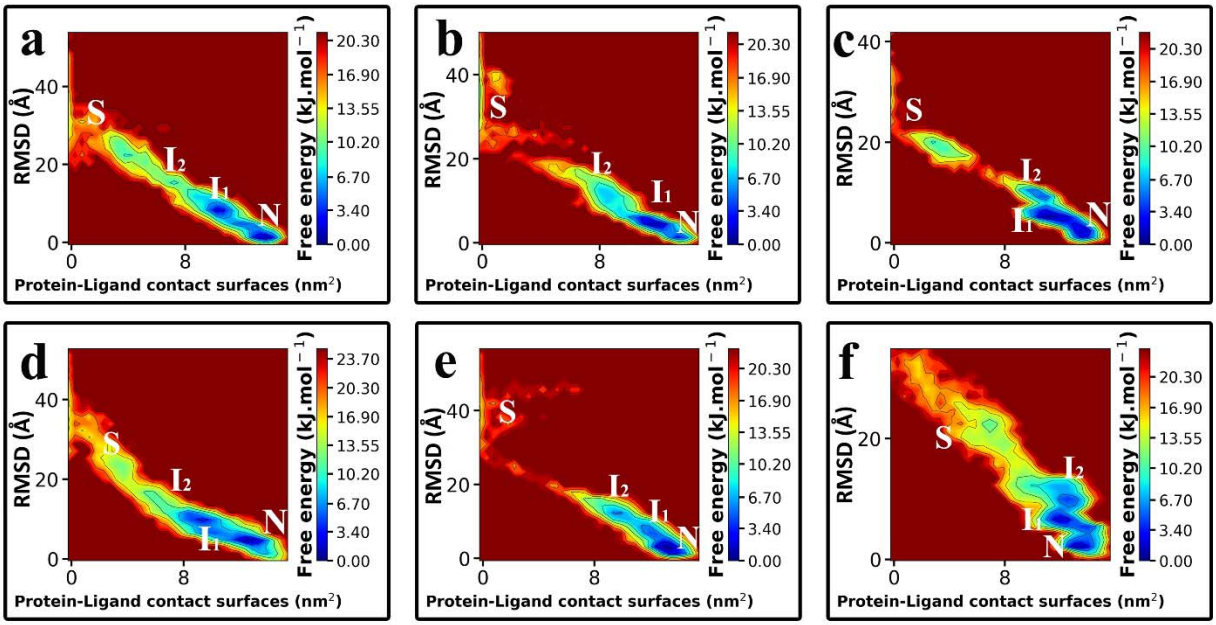

Figure 3. The Free Energy Landscapes (FEL) of the unbinding pathways of indinavir, **a, b, c**, from the HIV protease and, **d, e, f,** from the HTLV-1 protease. The stable states of indinavir during the unbinding pathway is indicated by capital letters; Native (N), Intermediate (I_1, I_2) and Solvated (S).

The total interaction energies of indinavir in the unbinding pathways of both proteases (Fig. 4a, 4b) showed roughly the same amounts. This means that despite the fact that the residues of the active sites of these two enzymes only have 45 percent identity, the interaction profiles are almost the same [23]. The values gradually increased and eventually reached zero when the ligand left the binding pocket and got solvated. Moreover, this analysis also showed that the dominant interactions are the short-range VdW interactions between the aromatic and aliphatic residues and the ligand (Figure S1). This finding indicates that the binding pockets of these two enzymes are rather hydrophobic. Except for the two Asp residues in the deepest part of the active site, almost none of the residues can make direct hydrogen bonds with the ligand using their side

chains. The hydrogen bonds are mainly formed with the backbone of the residues by the mediation of water molecules [24]. There is also a conserved water molecule in the indinavir-HTLV-1 protease complex known as the "flap water" which mediate interactions between the ligand and Ala59 on the tip of the flaps (see Ref 24 for more information) [24].

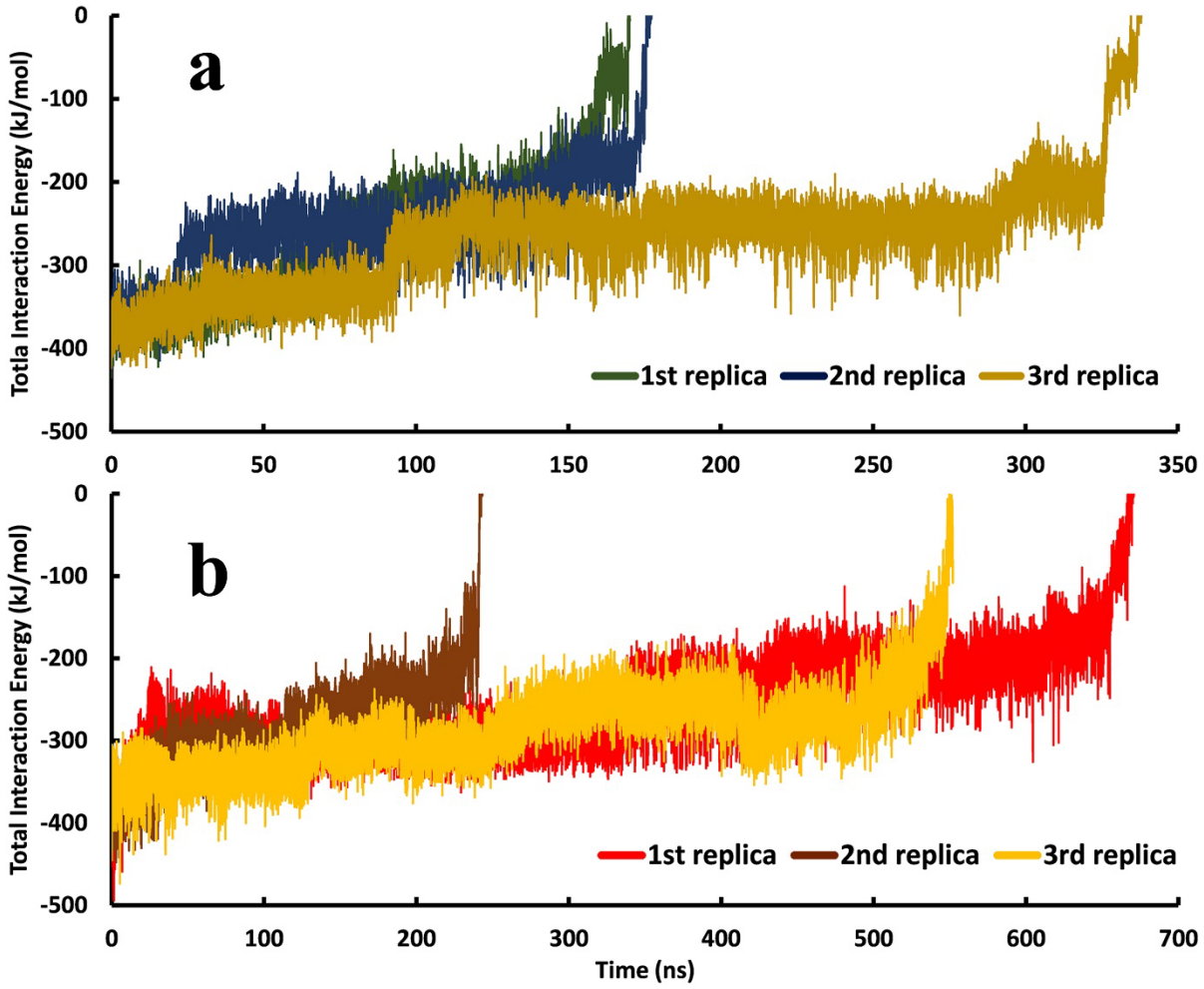

Figure 4. The total interaction energies of indinavir in the unbinding pathways in the three series of replicas for each protein-ligand complex. **a,** indinavir-HIV protease complex. **b,** indinavir-HTLV-1 protease complex.

As it was mentioned above, the interactions between indinavir and the residues of the binding pockets of both enzymes have almost the same amount and nature. However, the

contribution of the active site residues of each enzyme to the total interaction energies differ in some points (Fig. 5a, 5b and Figure S2). In the indinavir-HIV protease complex, Ile50, Gly49, Gly48 from the flap region and Asp25 and Thr80 had the most contribution in the interaction energies (Fig. 1c). In the indinavir-HTLV-1 protease complex, Leu57, Gly58 and Ala59 from the flap region and the Trp98 and Asp32 in the active site had the most contribution in the interaction energies (Fig. 1d). These emphasize the important role of the residues of the flap region. It is also clear that in the HIV protease almost all of the key residues are aliphatic. However, In the HTLV-1, aromatic residues such as Trp98 and Phe67 are concealing the role of aliphatic residues in the unbinding pathway. The existence of these aromatic residues in the active site of the HTLV-1 protease is one of the main reasons for the lack of proper inhibitory activity of indinavir against this enzyme, which will be discussed later [24].

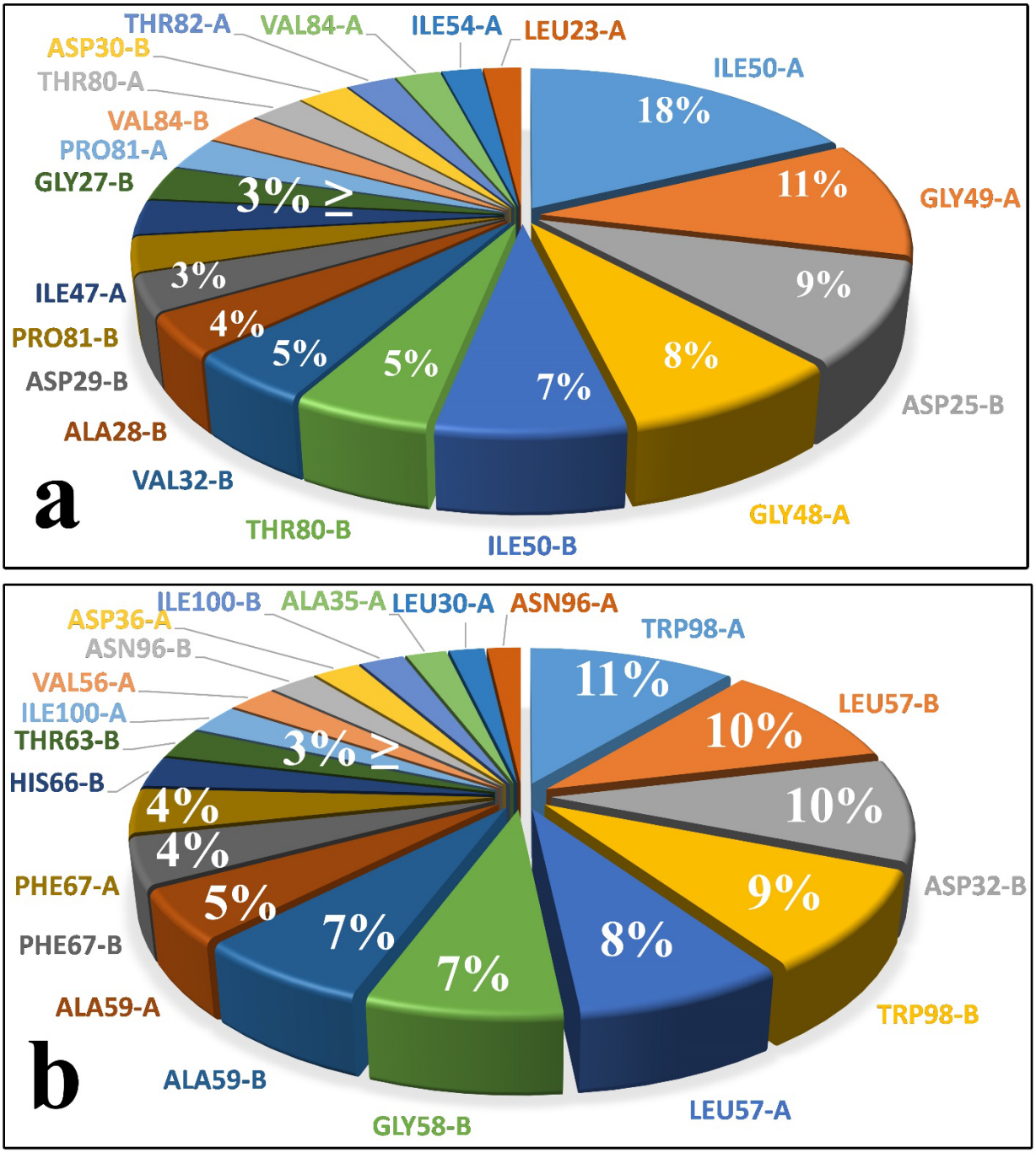

Figure 5. The contribution of each residue from both A and B chains of the proteases in the total interaction energies between indinavir and the enzyme in the unbinding pathways. **a,** indinavir-HIV protease complex, 1st replica. **b,** indinavir-HTLV-1 protease complex, 2nd replica.

In this study, we also found that one of the most important structural differences between the two proteases is the level of mobility of the flap region. In both enzymes, this region has high mobility but in the HTLV-1 protease, the flexibility and the fluctuations of flaps are slightly more than that of the HIV protease, Due to the nature of residues and the larger length of the flaps (Fig. 6a, 6b) [63-66]. The higher mobility of this region is unfavorable for indinavir to stay stable in the active site. In the presence of a ligand inside of aspartic proteases, the fluctuation of the flap region is significantly decreased which makes the ligand very stable in the active site [67, 68]. However, in the case of the indinavir-HTLV-1 protease complex, the higher mobility and fluctuation of this region might play a big part in the lack of sufficient inhibitory activity of indinavir. The handedness feature of the flaps was also visible in the RMSF values in which one of the flaps has more fluctuation than the other [65, 69]. Furthermore, in almost all three replicas of the indinavir-HTLV-1 protease complex, the RMSD values of the flap region is noticeably higher than that of the HIV protease (Fig. 6c, 6d)

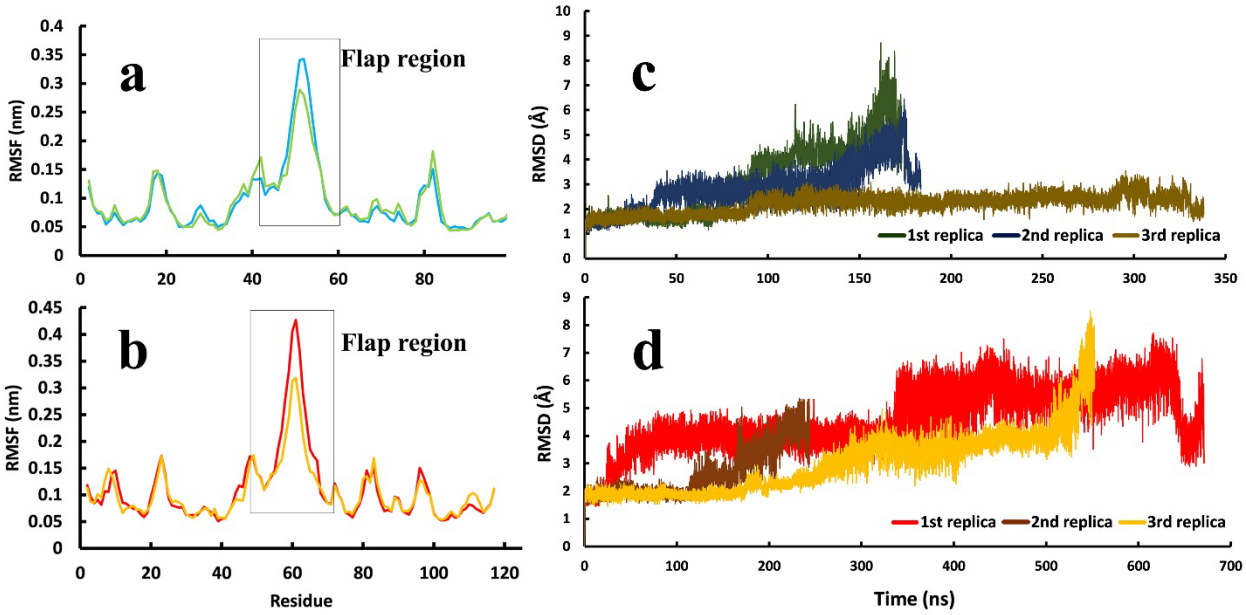

Figure 6. The average RMSF values of residues and the RMSD values of the backbone atoms of the two proteases in three replicas. **a,** RMSF values of HIV protease and, b, RMSF values of HTLV-1 protease. c, RMSD values of the backbone atoms of HIV protease and, d, RMSD values of the backbone atoms of HTLV-1 protease.

The flap region of the two proteases can acquire three different states: closed, semi-open and open [65, 70]. In the 6 SuMD simulations performed in this study, 3 replicas for each complex, all of these states were observed in the unbinding pathways of indinavir (Fig. 7a, 7b). The natural function of proteases is recognizing the polypeptide chain, trap and place it in the active site where the catalytic residue cut the peptide bond at a specific position. The process of recognizing and placing the polypeptide chain is done by the flap region [66]. This function of the flaps is extremely crucial for the whole function of the enzyme [71]. The thigh interaction between the ligand and the residues of the flap region coupled with the high mobility of this region causes the ligand to unbind and leave the enzymes. In almost all of the replicas, indinavir left the active site from the flap region.

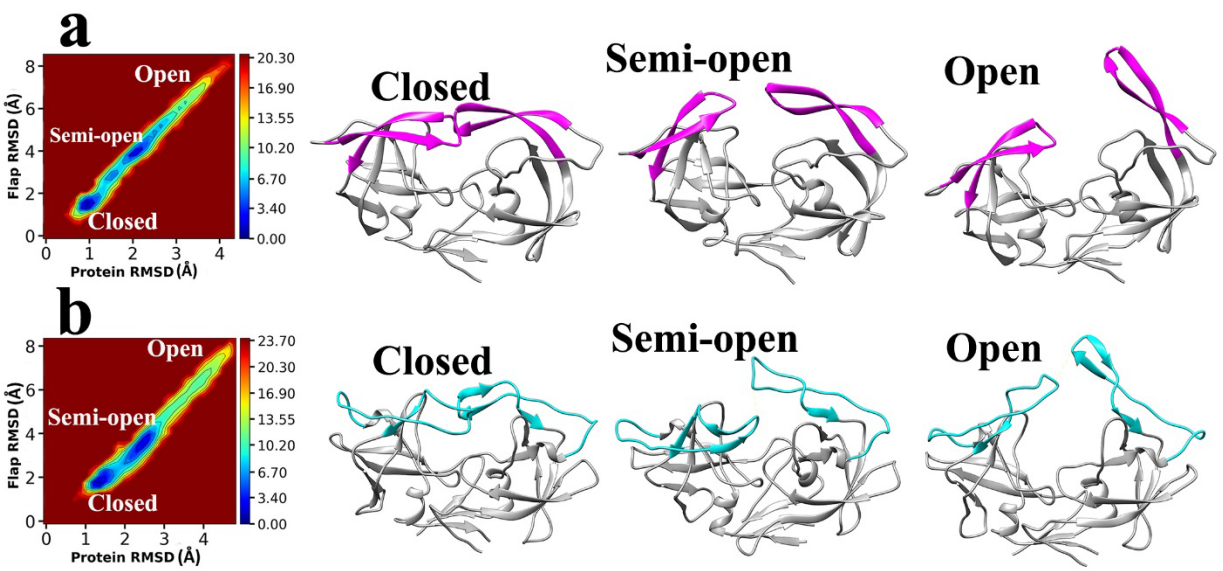

Figure 7. The main conformational states of the flap region which were acquired during the unbinding pathways of indinavir. **a,** The free energy landscape and the 3D structures of HTLV-1 protease with different flap states. **b,** The free energy landscape and the 3D structures of HIV protease with different flap states.

Indinavir interacts with these two enzymes at two separate states, the native and the intermediate states. In the native state, Asp32/Asp32' in the HTLV-1 protease and the Asp25/Asp25' in HIV protease have the role of interacting with the ligand and make it stable.

Although many other residues interact with the ligand in the native state through VdW and electrostatic interactions, the two Asp residues in the deep parts of both enzymes keep indinavir anchored to the native state. One of the two catalytic Asp residues is always protonated which is essential for the function of the enzyme [72-74]. This protonation enables the two Asp residues to make strong hydrogen bonds with the central hydroxyl group of indinavir. The very first step of the unbinding pathway is the breakage of the hydrogen bonds with the catalytic Asp residues. The rotation of the hydroxyl group of indinavir is the main cause of this breakage (Fig. 8a, 8b).

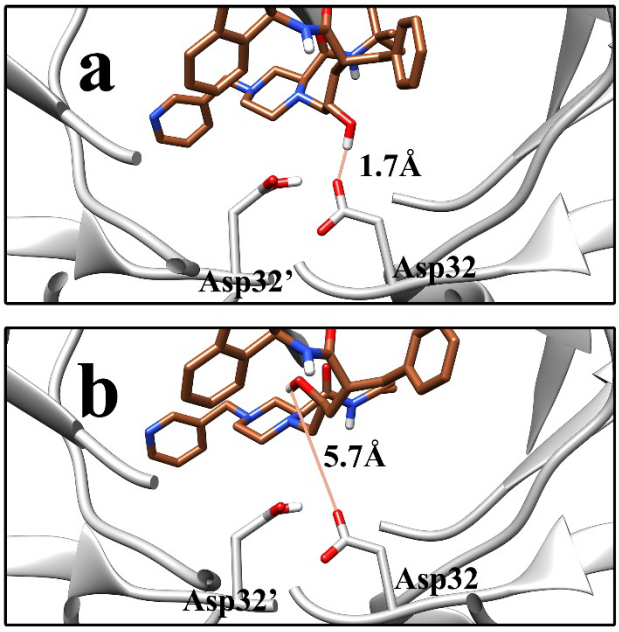

Figure 8. The first step of the unbinding pathway of indinavir. **a,** Indinavir in the native state where a hydrogen bond with the Asp32 in the HTLV-1 protease keep it stable in this state. **b,** the breakage of the hydrogen bond which triggers the unbinding pathway. The two frames are only 50 ps apart which shows that this event needs very little time to happen.

After the first step, the unbinding pathway of indinavir varies in some points between the HIV case and the HTLV-1 case. After the breakage of the above-mentioned hydrogen bond, indinavir started to get further from the deep parts of the active site and interact with higher positioned residues and also the flap region residues. One of the critical points of the active site is the Trp98/Trp98' residue in HTLV-1 protease which is occupied by Val82/Val82' in the HIV

protease at the same position. The Trp residues in the HTLV-1 protease are one of the main contributors to the interaction energies. Phe67/Phe67' was also an aromatic residue in the active site of HTLV-1 protease which has interaction with indinavir in the native and intermediate states. The interactions of these aromatic residues with the ligand make it easier for the ligand to get out of the native state (Fig. 10a, 10b, 10c). They mainly interact with the benzyl and the pyridyl moieties of indinavir but interactions with the indanol and tert-butyl- moieties were also spotted in the unbinding pathway (Fig. 1b, Fig. 9a, 9b). It was observed that they can acquire different orientations and make edge-to-face and face-to-face π-stacking and π-alkyl interactions. After the intermediate state, the ligand in both cases, interacted solely with the residues of the flap region for very short times and then left the enzymes completely and got solvated in the simulation box.

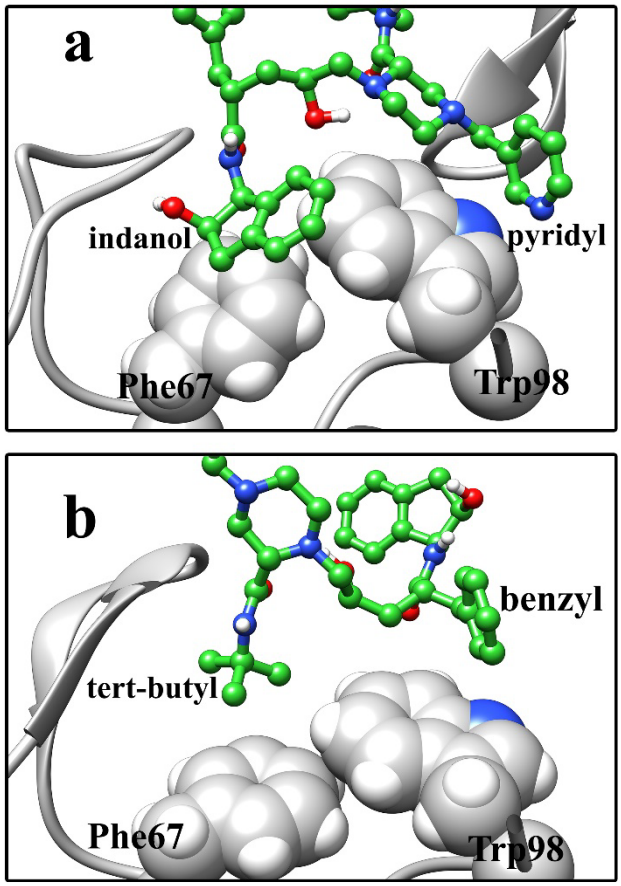

Figure 9. The interactions of indinavir moieties with the Phe 67 and Trp98 from both chains of HTLV-1 protease during the unbinding pathway. **a**, indanol, pyridyl and, **b**, tert-butyl and benzyl moieties of indinavir.

In 2003, Sa-Filho et al. provided a very interesting mutagenesis data about the mutant forms of HIV protease in which V82F was one of the main contributors to drug resistance during treatment by indinavir [75]. This mutation made indinavir and other protease inhibitors act considerably less selectively. An even larger aromatic residue like Trp in this position in HTLV-1 protease can make a huge difference in the activity of the drug. This incredible experimental data guided us to explore the role of the aromatic residues inside the active site of HTLV-1 protease. The aromatic residues in the active site, Trp98 and Phe67 in both chains of this enzyme can make strong interactions with indinavir in the intermediate state. These interactions are entropically unfavorable for the ligand since they make it easier for indinavir to get out of the native binding mode and interact with higher positioned residues. By comparing the RMSD values of the ligand and the RMSD values of these important residues of the HTLV-1 protease binding site during the unbinding pathway (Fig. 10a, 10b, 10c), the role of these aromatic residues, Phe67 and Trp98 was evident. In all three replicas of the Indinavir-HTLV-1 protease complex, the elevation of the RMSD values of these aromatic residues directly correlates to the elevated values of ligand RMSD. The transition of the ligand from the native state to the intermediate states directly corresponds to the conformational changes of these aromatic residues.

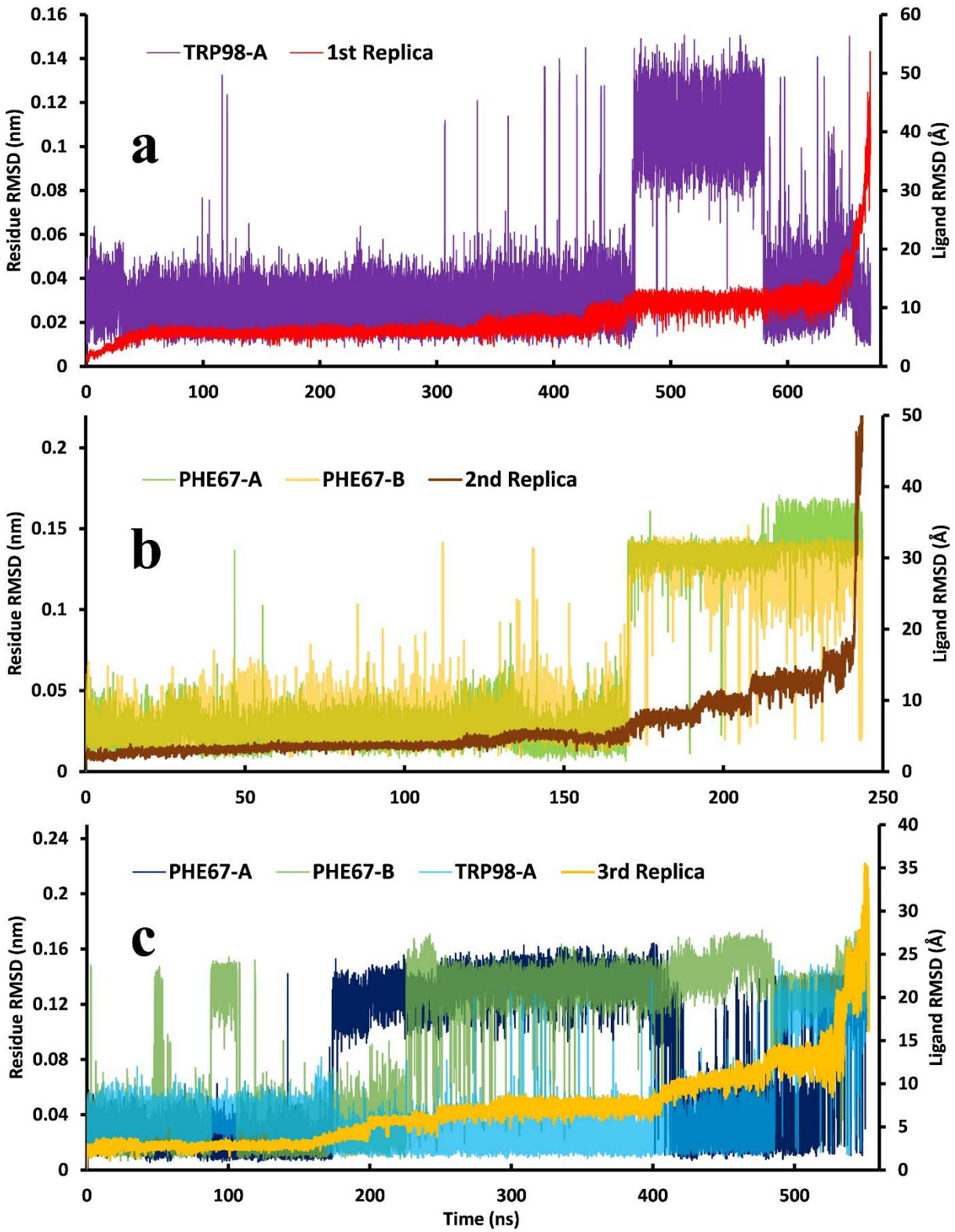
a
TRP98-A
1st Replica
Residue RMSD (nm)
Ligand RMSD (Å)
b
PHE67-A
PHE67-B
2nd Replica
Residue RMSD (nm)
Ligand RMSD (Å)
c
PHE67-A
PHE67-B
TRP98-A
3rd Replica
Residue RMSD (nm)
Ligand RMSD (Å)
Time (ns)

Figure 10, Comparison between the RMSD values of Indinavir (Å), and the RMSD values (nm) of important aromatic residues such as Phe67 and Trp98 in both chains. a, first replica of the unbinding pathway where only Trp98 from the chain A showed elevated values. b, second replica where the two Phe67 from both chains presented their role in the unbinding pathway. c, third replica where Phe67 from both chains and Trp98 from chain A exhibited their role in the unbinding pathway.

As it was described in previous sections, the mobility and the fluctuation of the flap region in the HTLV-1 protease is higher than HIV protease (Fig. 6a, 6b). The residues of the flap region in HIV protease have mostly organized secondary structures, two antiparallel β-strands (Fig. 7b), whereas this region in the HTLV-1 protease is a mixture of short β-strands with random coils (Fig. 7a). We believe that the higher mobility of the flap region coupled with the unfavorable interactions with aromatic residues in the active site of HTLV-1 protease can possibly be the main reasons for the significantly reduced inhibitory activity of indinavir. A more effective inhibitor of HTLV-1 protease must be able to make more interactions with the residues in the deep parts of the active site. It must be able to make more direct or water-mediated hydrogen bonds with the more inner positioned residues to avoid any interaction with the flap region, since high fluctuation of this region makes it easy for the ligand to unbind. The aromatic residues are also a big part of the active site. In order for an inhibitor to stay in the active site, it must be able to reduce the fluctuations of the aromatic residues and also the fluctuations of the flap region residues.

Conclusion

Understanding the details of the unbinding pathway of potential small molecule inhibitors from their target proteins is of significant importance. One of the most promising tools for accomplishing this task is utilizing SuMD simulation method. It is capable of reconstructing the binding and unbinding pathways of small molecule compounds. In this work, by taking advantage of this method, we reconstructed the unbinding pathway of indinavir from the HTLV-1 protease and the HIV protease in order to understand the details of the unbinding pathways from both enzymes and also to compare the two unbinding events to find the reasons for the lack

of inhibitory activity of indinavir against the HTLV-1 protease. We achieved multiple unbinding pathways from both complexes and found that although the structure and the interaction profiles are broadly similar, simple differences in the details cause massive effects in the performance of the inhibitor. It was discovered that in the unbinding pathway of indinavir form HTLV-1 protease, the fluctuation of the flap region is higher than that of HIV protease. We also found that aromatic residues in the active site of HTLV-1 protease, Phe67/Phe67' and Trp98/Trp98', are among the most important residues in the unbinding pathways and had the most contribution to the interaction energies. However, these interactions are unfavorable and cause the ligand to get out of the native binding mode. The unfavorable interactions coupled with the higher fluctuations of the flap region were found to be the reasons for the poor inhibitory activity of indinavir against HTLV-1 protease. We are of the opinion that the details found in this study can be a great assist for designing more effective compounds for inhibiting the HTLV-1 protease as a treatment solution for fighting the HTLV-1 virus.

Acknowledgment

This investigation was supported by Golestan University, Gorgan, Iran.

Conflict of Interest

none declared.

References

[1] Hayward, A. (2017) Origin of the retroviruses: when, where, and how?, *Current Opinion in Virology 25*, 23-27.
[2] Kim, H.-S., Ock, M. S., and Cha, H.-J. (2017) Interactions between human endogenous and exogenous retroviruses, *Genes & Genomics 39*, 923-927.
[3] Meyer, T. J., Rosenkrantz, J. L., Carbone, L., and Chavez, S. L. (2017) Endogenous Retroviruses: With Us and against Us, *Frontiers in chemistry 5*, 23-23.
[4] (2019) UNAIDS.
[5] Lv, Z., Chu, Y., and Wang, Y. (2015) HIV protease inhibitors: a review of molecular selectivity and toxicity, *HIV AIDS (Auckl) 7*, 95-104.

[6] Pokorna, J., Machala, L., Rezacova, P., and Konvalinka, J. (2009) Current and Novel Inhibitors of HIV Protease, *Viruses 1*, 1209-1239.
[7] Hughes, P. J., Cretton-Scott, E., Teague, A., and Wensel, T. M. (2011) Protease Inhibitors for Patients With HIV-1 Infection: A Comparative Overview, *P T 36*, 332-345.
[8] Patick, A. K., and Potts, K. E. (1998) Protease Inhibitors as Antiviral Agents, *Clinical Microbiology Reviews 11*, 614.
[9] Titanji, B. K., Aasa-Chapman, M., Pillay, D., and Jolly, C. (2013) Protease inhibitors effectively block cell-to-cell spread of HIV-1 between T cells, *Retrovirology 10*, 161.
[10] Eron, J. J., Jr. (2000) HIV-1 Protease Inhibitors, *Clinical Infectious Diseases 30*, S160-S170.
[11] Ikezoe, T., Hisatake, Y., Takeuchi, T., Ohtsuki, Y., Yang, Y., Said, J. W., Taguchi, H., and Koeffler, H. P. (2004) HIV-1 Protease Inhibitor, Ritonavir, *Cancer Research 64*, 7426.
[12] Paton, N. I., Stohr, W., Arenas-Pinto, A., Fisher, M., Williams, I., Johnson, M., Orkin, C., Chen, F., Lee, V., Winston, A., Gompels, M., Fox, J., Scott, K., Dunn, D. T., and Protease Inhibitor Monotherapy Versus Ongoing Triple Therapy Trial, T. (2015) Protease inhibitor monotherapy for long-term management of HIV infection: a randomised, controlled, open-label, non-inferiority trial, *Lancet HIV 2*, e417-426.
[13] Fogarty, K. H., Zhang, W., Grigsby, I. F., Johnson, J. L., Chen, Y., Mueller, J. D., and Mansky, L. M. (2011) New insights into HTLV-1 particle structure, assembly, and Gag-Gag interactions in living cells, *Viruses 3*, 770-793.
[14] Kannian, P., and Green, P. L. (2010) Human T Lymphotropic Virus Type 1 (HTLV-1): Molecular Biology and Oncogenesis, *Viruses 2*, 2037-2077.
[15] Gonçalves, D. U., Proietti, F. A., Ribas, J. G. R., Araújo, M. G., Pinheiro, S. R., Guedes, A. C., and Carneiro-Proietti, A. B. F. (2010) Epidemiology, Treatment, and Prevention of Human T-Cell Leukemia Virus Type 1-Associated Diseases, *Clinical Microbiology Reviews 23*, 577.
[16] Satou, Y., and Matsuoka, M. (2010) HTLV-1 and the host immune system: how the virus disrupts immune regulation, leading to HTLV-1 associated diseases, *J Clin Exp Hematop 50*, 1-8.
[17] Goon, P. K. C., and Bangham, C. R. M. (2004) Interference with immune function by HTLV-1, *Clin Exp Immunol 137*, 234-236.
[18] Quaresma, J. A. S., Yoshikawa, G. T., Koyama, R. V. L., Dias, G. A. S., Fujihara, S., and Fuzii, H. T. (2015) HTLV-1, Immune Response and Autoimmunity, *Viruses 8*, 5.
[19] Revaud, D., Bejanariu, A., Loussaief, L., Sarry, E., Zemmar, A., Deplaine, G., Coman, T., Rossignol, J., Hermine, O., and Bauche, C. (2015) Development of an Anti-HTLV-1 Vaccine for the Treatment of Adult T-Cell Leukemia/Lymphoma, *Blood 126*, 4010.
[20] Pasquier, A., Alais, S., Roux, L., Thoulouze, M.-I., Alvarez, K., Journo, C., Dutartre, H., and Mahieux, R. (2018) How to Control HTLV-1-Associated Diseases: Preventing de Novo Cellular Infection Using Antiviral Therapy, *Frontiers in Microbiology 9*.
[21] Tözsér, J., and Weber, I. T. (2007) The protease of human T-cell leukemia virus type-1 is a potential therapeutic target, *Curr Pharm Des 13*, 1285-1294.
[22] Mahalingam, B., Wang, Y. F., Boross, P. I., Tozser, J., Louis, J. M., Harrison, R. W., and Weber, I. T. (2004) Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site, *European journal of biochemistry 271*, 1516-1524.
[23] Kádas, J., Weber, I. T., Bagossi, P., Miklóssy, G., Boross, P., Oroszlan, S., and Tözsér, J. (2004) Narrow substrate specificity and sensitivity toward ligand-binding site mutations of human T-cell Leukemia virus type 1 protease, *J Biol Chem 279*, 27148-27157.
[24] Kuhnert, M., Steuber, H., and Diederich, W. E. (2014) Structural Basis for HTLV-1 Protease Inhibition by the HIV-1 Protease Inhibitor Indinavir, *Journal of Medicinal Chemistry 57*, 6266-6272.

[25] Piana, S., Carloni, P., and Rothlisberger, U. (2002) Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations, *Protein science : a publication of the Protein Society 11*, 2393-2402.
[26] 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, *Structure 7*, 1047-1055.
[27] 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.
[28] 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, *Structure 8*, 1259-1265.
[29] Perilla, J. R., Goh, B. C., Cassidy, C. K., Liu, B., Bernardi, R. C., Rudack, T., Yu, H., Wu, Z., and Schulten, K. (2015) Molecular dynamics simulations of large macromolecular complexes, *Current opinion in structural biology 31*, 64-74.
[30] Gohlke, H., Hergert, U., Meyer, T., Mulnaes, D., Grieshaber, M. K., Smits, S. H. J., and Schmitt, L. (2013) Binding Region of Alanopine Dehydrogenase Predicted by Unbiased Molecular Dynamics Simulations of Ligand Diffusion, *Journal of Chemical Information and Modeling 53*, 2493-2498.
[31] Niu, Y., Li, S., Pan, D., Liu, H., and Yao, X. (2016) Computational study on the unbinding pathways of B-RAF inhibitors and its implication for the difference of residence time: insight from random acceleration and steered molecular dynamics simulations, *Phys Chem Chem Phys 18*, 5622-5629.
[32] Shao, Q., and Zhu, W. (2019) Exploring the Ligand Binding/Unbinding Pathway by Selectively Enhanced Sampling of Ligand in a Protein–Ligand Complex, *The Journal of Physical Chemistry B 123*, 7974-7983.
[33] Rydzewski, J., and Valsson, O. (2019) Finding multiple reaction pathways of ligand unbinding, *The Journal of Chemical Physics 150*, 221101.
[34] Marino, K. A., and Filizola, M. (2018) Investigating Small-Molecule Ligand Binding to G Protein-Coupled Receptors with Biased or Unbiased Molecular Dynamics Simulations, *Methods Mol Biol 1705*, 351-364.
[35] Rydzewski, J., Jakubowski, R., Nowak, W., and Grubmüller, H. (2018) Kinetics of Huperzine A Dissociation from Acetylcholinesterase via Multiple Unbinding Pathways, *Journal of Chemical Theory and Computation 14*, 2843-2851.
[36] Mollica, L., Decherchi, S., Zia, S. R., Gaspari, R., Cavalli, A., and Rocchia, W. (2015) Kinetics of protein-ligand unbinding via smoothed potential molecular dynamics simulations, *Sci Rep 5*, 11539.
[37] Tiwary, P., Limongelli, V., Salvalaglio, M., and Parrinello, M. (2015) Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps, *Proc Natl Acad Sci U S A 112*, E386-391.
[38] D'Annessa, I., Raniolo, S., Limongelli, V., Di Marino, D., and Colombo, G. (2019) Ligand Binding, Unbinding, and Allosteric Effects: Deciphering Small-Molecule Modulation of HSP90, *Journal of Chemical Theory and Computation 15*, 6368-6381.
[39] Schuetz, D. A., Bernetti, M., Bertazzo, M., Musil, D., Eggenweiler, H.-M., Recanatini, M., Masetti, M., Ecker, G. F., and Cavalli, A. (2019) Predicting Residence Time and Drug Unbinding Pathway through Scaled Molecular Dynamics, *Journal of Chemical Information and Modeling 59*, 535-549.
[40] Hu, X., Hu, S., Wang, J., Dong, Y., Zhang, L., and Dong, Y. (2018) Steered molecular dynamics for studying ligand unbinding of ecdysone receptor, *Journal of Biomolecular Structure and Dynamics 36*, 3819-3828.

[41] Deganutti, G., Moro, S., and Reynolds, C. A. (2020) A Supervised Molecular Dynamics Approach to Unbiased Ligand–Protein Unbinding, *Journal of Chemical Information and Modeling 60*, 1804-1817.
[42] Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E. (2011) How does a drug molecule find its target binding site?, *J Am Chem Soc 133*, 9181-9183.
[43] Sohraby, F., Javaheri Moghadam, M., Aliyar, M., and Aryapour, H. (2020) A boosted unbiased molecular dynamics method for predicting ligands binding mechanisms: Probing the binding pathway of dasatinib to Src-kinase, *Bioinformatics*.
[44] Sohraby, F., and Aryapour, H. (2020) Rational drug repurposing for cancer by inclusion of the unbiased molecular dynamics simulation in the structure-based virtual screening approach: Challenges and breakthroughs, *Seminars in Cancer Biology*.
[45] Tiwary, P., Mondal, J., and Berne, B. J. (2017) How and when does an anticancer drug leave its binding site?, *Science Advances 3*.
[46] King, N. M., Melnick, L., Prabu-Jeyabalan, M., Nalivaika, E. A., Yang, S.-S., Gao, Y., Nie, X., Zepp, C., Heefner, D. L., and Schiffer, C. A. (2002) Lack of synergy for inhibitors targeting a multi-drug-resistant HIV-1 protease, *Protein science : a publication of the Protein Society 11*, 418-429.
[47] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res 28*, 235-242.
[48] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J Comput Chem 25*, 1605-1612.
[49] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX 1-2*, 19-25.
[50] Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *Journal of the American Chemical Society 118*, 11225-11236.
[51] Sousa da Silva, A. W., and Vranken, W. F. (2012) ACPYPE - AnteChamber PYthon Parser interfacE, *BMC Res Notes 5*, 367.
[52] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *The Journal of chemical physics 79*, 926-935.
[53] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations, *Journal of Computational Chemistry 18*, 1463-1472.
[54] Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, *The Journal of chemical physics 98*, 10089-10092.
[55] Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling, *The Journal of chemical physics 126*, 014101.
[56] Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method, *Journal of applied physics 52*, 7182-7190.
[57] Sabbadin, D., Salmaso, V., Sturlese, M., and Moro, S. (2018) Supervised Molecular Dynamics (SuMD) Approaches in Drug Design, *Methods Mol Biol 1824*, 287-298.
[58] Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics, *Journal of Molecular Graphics 14*, 33-38.
[59] Kraus, D. (2014) Consolidated data analysis and presentation using an open-source add-in for the Microsoft Excel® spreadsheet software, *Medical Writing 23*, 25-28.
[60] Hunter, J. D. (2007) Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering 9*, 90-95.

[61] Kumari, R., Kumar, R., and Lynn, A. (2014) g_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations, *Journal of Chemical Information and Modeling 54*, 1951-1962.
[62] Kadas, J., Weber, I. T., Bagossi, P., Miklossy, G., Boross, P., Oroszlan, S., and Tozser, J. (2004) Narrow substrate specificity and sensitivity toward ligand-binding site mutations of human T-cell Leukemia virus type 1 protease, *J Biol Chem 279*, 27148-27157.
[63] 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 science : a publication of the Protein Society 11*, 221-232.
[64] Li, M., Laco, G. S., Jaskolski, M., Rozycki, J., Alexandratos, J., Wlodawer, A., and Gustchina, A. (2005) Crystal structure of human T cell leukemia virus protease, a novel target for anticancer drug design, *Proceedings of the National Academy of Sciences of the United States of America 102*, 18332-18337.
[65] Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, *Proceedings of the National Academy of Sciences of the United States of America 103*, 915.
[66] Tóth, G., and Borics, A. (2006) Closing of the flaps of HIV-1 protease induced by substrate binding: a model of a flap closing mechanism in retroviral aspartic proteases, *Biochemistry 45*, 6606-6614.
[67] 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, *Structure 7*, 1047-S1012.
[68] Nicholson, L. K., Yamazaki, T., Torchia, D. A., Grzesiek, S., Bax, A., Stahl, S. J., Kaufman, J. D., Wingfield, P. T., Lam, P. Y. S., Jadhav, P. K., Hodge, C. N., Domaille, P. J., and Chang, C.-H. (1995) Flexibility and function in HIV-1 protease, *Nature Structural Biology 2*, 274-280.
[69] Miao, Y., Huang, Y. M., Walker, R. C., McCammon, J. A., and Chang, C. A. (2018) Ligand Binding Pathways and Conformational Transitions of the HIV Protease, *Biochemistry 57*, 1533-1541.
[70] Hornak, V., and Simmerling, C. (2007) Targeting structural flexibility in HIV-1 protease inhibitor binding, *Drug discovery today 12*, 132-138.
[71] Sadiq, S. K., and De Fabritiis, G. (2010) Explicit solvent dynamics and energetics of HIV-1 protease flap opening and closing, *Proteins 78*, 2873-2885.
[72] Smith, R., Brereton, I. M., Chai, R. Y., and Kent, S. B. (1996) Ionization states of the catalytic residues in HIV-1 protease, *Nat Struct Biol 3*, 946-950.
[73] Torbeev, V. Y., and Kent, S. B. (2012) Ionization state of the catalytic dyad Asp25/25' in the HIV-1 protease: NMR studies of site-specifically 13C labelled HIV-1 protease prepared by total chemical synthesis, *Organic & biomolecular chemistry 10*, 5887-5891.
[74] Wang, Y. X., Freedberg, D. I., Yamazaki, T., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y., and Torchia, D. A. (1996) Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272, *Biochemistry 35*, 9945-9950.
[75] Sa-Filho, D. J., Costa, L. J., de Oliveira, C. F., Guimarães, A. P., Accetturi, C. A., Tanuri, A., and Diaz, R. S. (2003) Analysis of the protease sequences of HIV-1 infected individuals after Indinavir monotherapy, *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology 28*, 186-202.

Supplementary Materials

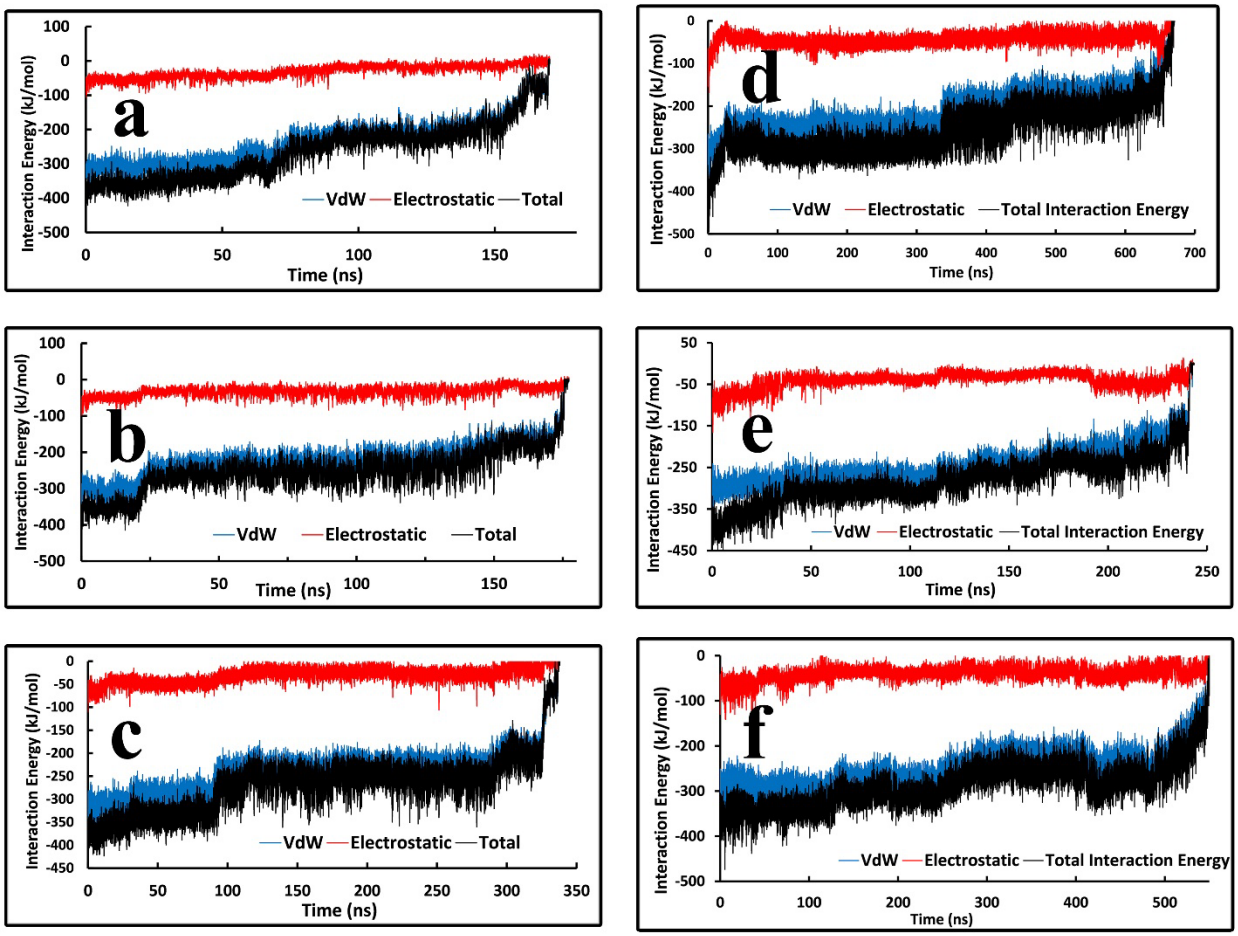

Figure S1. The VdW, Electrostatic and total interaction energies between indinavir and the two proteases during the unbinding pathways. **a, b, c,** The indinavir-HIV complex and, **d, e, f**, The indinavir- HTLV-1 protease complex.

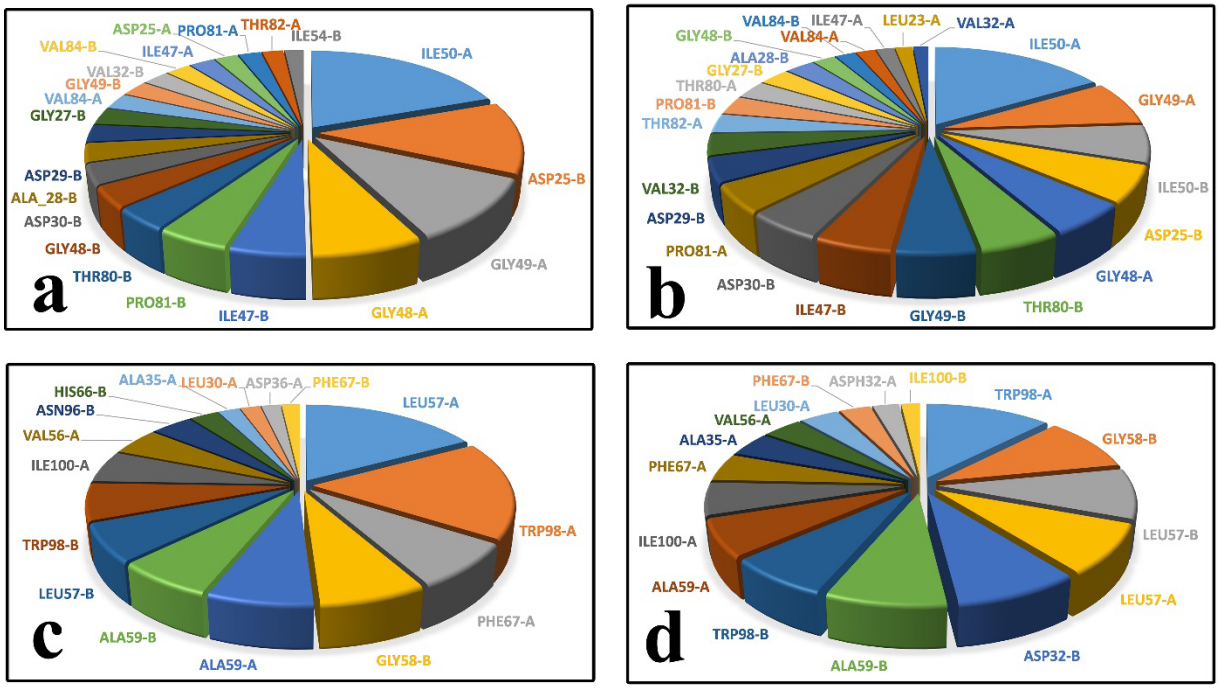

Figure S2. The contribution of the active site residues to the total interaction energies in the unbinding pathways. **a, b,** 1st and 3rd replica of the indinavir-HIV protease case. **c, b,** 1st and 2nd replica of the indinavir-HTLV-1 protease case.