

## RESEARCH ARTICLE

# Comparative study of the unbinding process of some HTLV-1 protease inhibitors using unbiased molecular dynamics simulations

Fereshteh Noroozi Tiyoula, Hassan Aryapour<sup>1</sup>\*, Mostafa Javaheri Moghadam

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

\* [hassan.aryapour@gmail.com](mailto:hassan.aryapour@gmail.com)

## Abstract

The HTLV-1 protease is one of the major antiviral targets to overwhelm this virus. Several research groups have developed protease inhibitors, but none has been successful. In this regard, developing new HTLV-1 protease inhibitors to fix the defects in previous inhibitors may overcome the lack of curative treatment for this oncovirus. Thus, we decided to study the unbinding pathways of the most potent (compound 10, PDB ID 4YDF,  $K_i = 15$  nM) and one of the weakest (compound 9, PDB ID 4YDG,  $K_i = 7900$  nM) protease inhibitors, which are very structurally similar. We conducted 12 successful short and long simulations (totaling 14.8  $\mu$ s) to unbind the compounds from two monoprotonated (mp) forms of protease using the Supervised Molecular Dynamics (SuMD) without applying any biasing force. The results revealed that Asp32 or Asp32' in the two forms of mp state similarly exert powerful effects on maintaining both potent and weak inhibitors in the binding pocket of HTLV-1 protease. In the potent inhibitor's unbinding process, His66' was a great supporter that was absent in the weak inhibitor's unbinding pathway. In contrast, in the weak inhibitor's unbinding process, Trp98/Trp98' by pi-pi stacking interactions were unfavorable for the stability of the inhibitor in the binding site. In our opinion, these results will assist in designing more potent and effective inhibitors for the HTLV-1 protease.

## OPEN ACCESS

**Citation:** Tiyoula FN, Aryapour H, Moghadam MJ (2022) Comparative study of the unbinding process of some HTLV-1 protease inhibitors using unbiased molecular dynamics simulations. PLoS ONE 17(7): e0263200. <https://doi.org/10.1371/journal.pone.0263200>

**Editor:** Sandipan Chakraborty, Dr Reddy's Institute of Life Sciences, INDIA

**Received:** January 9, 2022

**Accepted:** June 28, 2022

**Published:** July 14, 2022

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0263200>

**Copyright:** © 2022 Tiyoula et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data and analysis files are uploaded to this address (<https://zenodo.org/record/5633143#.YZXlftBBzIU>) and are available with DOI: [10.5281/zenodo.5633143](https://doi.org/10.5281/zenodo.5633143).

## Introduction

Human T-cell leukemia virus type 1 (HTLV-1) was discovered in 1980 as the first oncogenic retrovirus in the project "War on Cancer" in the United States [1]. According to the latest information, 5–10 million people are infected with this virus worldwide, and only 0.25–5% of them are affected by Adult T-cell Leukemia/Lymphoma (ATLL) and HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) [2], and also HTLV-1-associated ocular diseases. These diseases are known as HTLV-1 uveitis (HU) and ATL-related ocular [3]. Indeed, the reported numbers are not terrible, but there is no standard treatment for all types of diseases [4]. In addition, only a few regions were evaluated, and many unknown infected people could transmit the virus [5]. So, even low-risk areas are in danger because of Global Village. After HTLV-1 discovery, all its components were identified gradually, and its protease was

**Funding:** Yes Golestan university. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

discovered in 1989 [6]. HTLV-1 protease is a homodimer protein containing 125 residues in each subunit, which is one of the A2 family of aspartic proteases, with two critical aspartates in the catalytic dyad. This enzyme is essential for viral growth because it cleaves the Gag-Pro-Pol-Env polyprotein, a necessary viral replication component [7]. Since this part is vital for the viral life cycle, it is an interesting target for HTLV-1 demise.

Toward this end, many research groups in different countries succeeded in designing and synthesizing various compounds with inhibitory effects in the micromolar to nanomolar ranges [8, 9]. Finally, some German scientists considered the structural similarities between HTLV-1 and HTLV-3 (HIV) and determined the X-ray structure of Indinavir complexed with HTLV-1 protease, which is the only AIDS protease drug that has an inhibitory effect on HTLV-1 protease in low micromolar concentration. Unfortunately, this drug failed to be used to eradicate HTLV-1 [10]. After being frustrated with AIDS drugs, this team, in 2015, succeeded in synthesizing ten inhibitors that contain the most potent nonpeptidic inhibitor of HTLV-1 protease [11].

All reported HTLV-1 protease compounds only remain as inhibitors, and we do not have any specialized FDA-approved drug for inhibiting this virus. It is evident that experimental research alone is not sufficient. *In silico* methods, like unbiased molecular dynamics (UMD), are needed to provide valuable information for rational drug design, which is the primary goal of all researchers in this field. So, various research groups have used *in silico* methods for different studies, involving: carrying out the docking simulation to investigate the substrate-binding cavity of the IDH1 enzyme [12], using dynamic simulations and binding free energy studies to design new potent inhibitors against CDK2 [13], prediction of the binding pocket in the interface of the aurora-A-TPX2 complex using molecular docking [14], finding specific naturally originated antagonist of the benzodiazepine binding site by performing docking experiments and molecular mechanics Poisson-Boltzmann surface area analysis [15]. MD simulation offers information about the reaction pathways of the ligand-protein complexes, and it has been considered by many research groups over these years and led to effective drug design [16, 17]. Therefore, besides the importance of one particular drug's binding affinity to a target protein in the traditional drug design, the binding and unbinding processes and the residence time of the compound that interacts with the protein in each intermediate state are just as important. So by a complete understanding of the unbinding mechanism, we can uncover the key elements in the protein-ligand complex interactions [18], ligand flexibility, and solvation effects that are more critical in the rational drug design. The vital information will ultimately appear in a scenario with fully atomistic details [19]. For investigating unbinding pathways of inhibitors, some advanced MD simulation approaches like metadynamics and supervised metadynamics (suMetaD) simulation have been used before [20, 21], and one of the newest MD approaches is the supervised molecular dynamics (SuMD) method. This method performs simulation in replicas with fixed parameters. It gets information regarding metastable intermediate ligand-bound states using a tabu-like supervision algorithm. It is possible to fully unbind small molecules from their molecular targets without any biasing force or potential. In this regard, the SuMD has been utilized to discover the reaction pathways of various ligands in molecular targets [22].

The previous work of our team attempted to achieve Indinavir's interaction with HIV protease and HTLV-1 protease as a comparison test to better understand the complexity of these key proteins [23]. In the following, we decided to examine the unbinding pathways of the most potent and one of the weakest HTLV-1 protease inhibitors retrieved from the last designed compounds using SuMD. In this study, the simulations will reveal the dynamic behavior of HTLV-1 protease as well as the strengths and weaknesses of the selected inhibitor. So, the results will guide the design of new potent HTLV-1 protease inhibitors. Also, this unbiased

comparative study caused monitoring of all important factors in the two different inhibitor-protein interactions. So the outcomes were more reliable.

## Methods

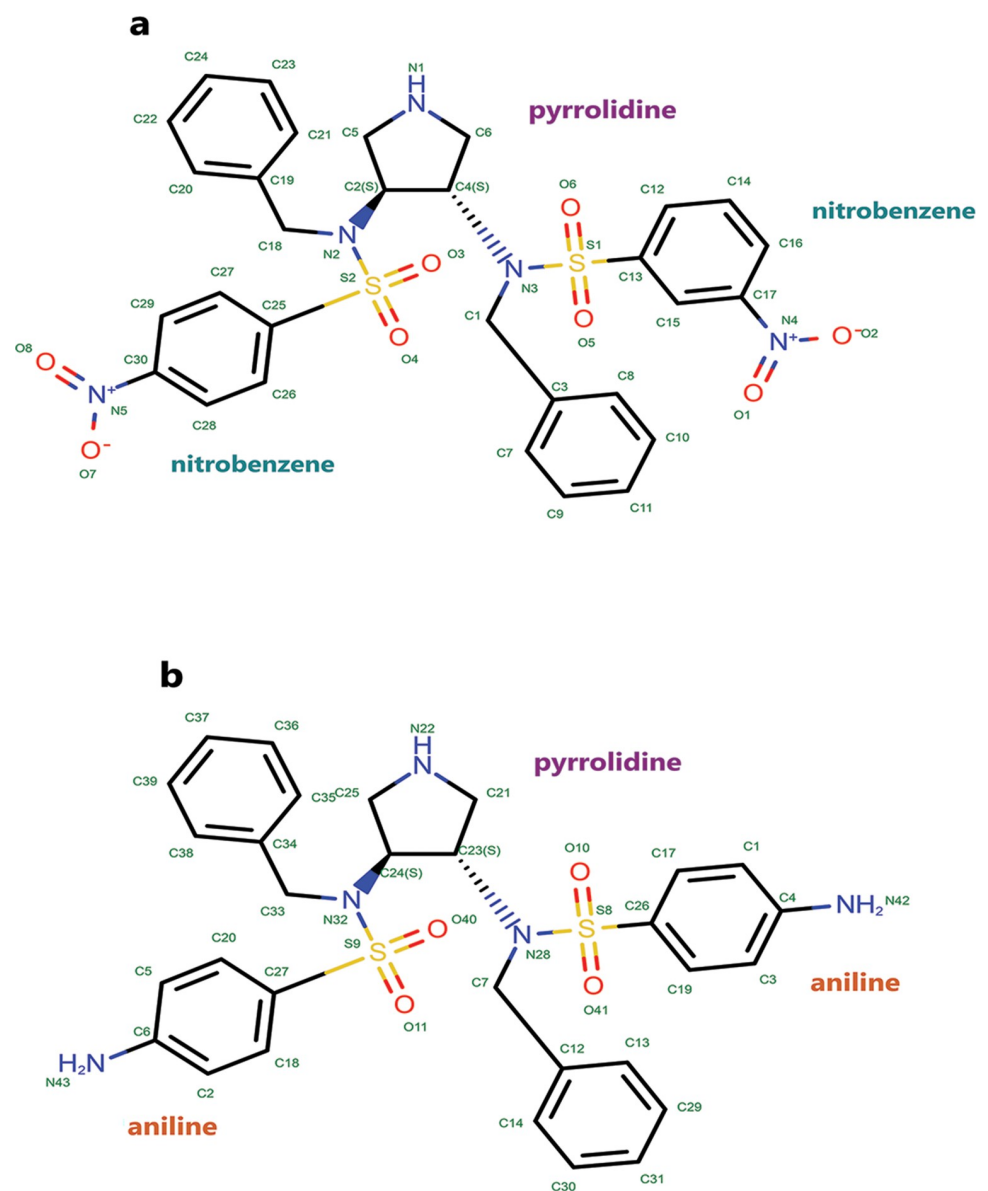
The X-ray crystallography structures of the HTLV-1 protease-ligand complex (PDB IDs: 4YDG, 4YDF [11]) were obtained from the Protein Data Bank. At first, for protein preparation, all protein missing residues and atoms in 4YDG were remodeled and fixed using UCSF Chimera software [24]. Then, for the ligand preparation, according to the practical information, the nitrogen of the pyrrolidine ring was protonated in both compounds [11] and parameterized by ACEPYPE using default settings (the GAFF atom type and BCC partial charges) [25]. HIV protease is an aspartic protease, and there were no studies on HTLV-1 protease, so we used a pattern from HIV research. In some research, protonation is not applied or not mentioned, or in some research, protonation state is applied in chain A. As we wanted a more complete outcome, we used protonation separately in both chains [26, 27]. So after preparing complexes, each catalytic Asp was considered separately as an ionization state based on the monoprotonated (mp) form of the catalytic dyad Asp32-Asp32' in the active site [28]. Finally, we constructed our systems in GROMACS 2018 [29] using the OPLS all-atom force field [30] and with the TIP3P water model [31]. The considered holoproteins were centered in a triclinic box with a distance of 1 nm from each edge. The next step was to provide a 150 mM neutral physiological salt concentration, sodium, and chloride ions. Then all systems were relaxed in energy minimization using the steepest descent algorithm and reached Fmax of less than 1000  $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-1}$ . All covalent bonds by Linear Constraint Solver (LINCS) algorithm were constrained to maintain constant bond lengths [32]. The long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) method [33]. The cut-off radii for Coulomb and Van der Waals (VdW) short-range interactions were set to 0.9 nm for all systems.

At last, the modified Berendsen (V-rescale) thermostat [34] and Parrinello-Rahman barostat [35] were applied for 100 and 300 ps, respectively, for keeping the system in stable environmental conditions (310 K, 1 Bar). To reach complete unbinds, we performed 12 separate replicas (three replicas for each type of mp form) with fixed duration times using the SuMD method with some modifications [36]. During the simulation, the distance between the center of masses of the ligands and selected residues was monitored until complete unbind occurred. This method is based on a tabu-like supervision algorithm without applying any human or non-human biasing force or potential. Herein, we set the center of mass (COM) of ligands as the first spot and the COM of the catalytic aspartic acids (Asp32, Asp32') as the second spot and ran all simulations with a time window of 500 ps and a time step of 2 fs. After finishing each run, the frame with the longest distance between COMs was selected automatically to extend the next 500 ps simulation. These processes were continued until complete unbind was obtained, which is equal to a distance of 50 Å between the mentioned spots. Finally, all events in every concatenated trajectory file were investigated carefully with GROMACS utilities for data analysis. To picture the important interactions, we used UCSF chimera and used Daniel's XL Toolbox (v7.3.4) to create plots [37], and using Matplotlib to show free energy landscapes (S1 Fig) [38]. The free energy landscapes plots were made base on three variables time, ligand RMSD, and protein RMSD. The ligand and protein RMSD values were selected because they were meaning full and had sharp changes as a function of time during unbindings. Analyzing these plots can reveal the stable states of inhibitors, as well as the residence time of inhibitors in each state over unbinding. Areas that tend to turn blue indicate that the inhibitor has been present in this area for a longer time.

## Results and discussion

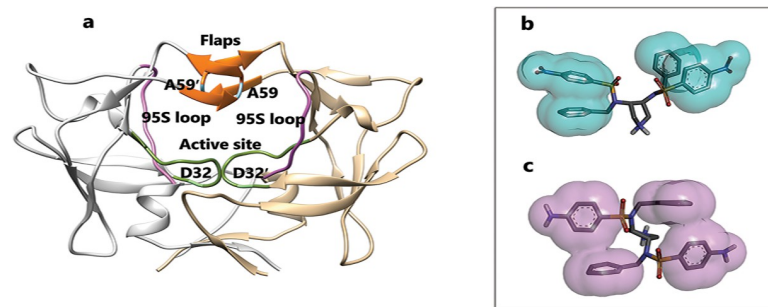
Since the only structural difference between compounds 9 and 10 is in the amino and nitro groups on the benzene ring (Fig 1A and 1B), compound 10 ( $K_i = 15$  nM) is approximately 526 times more potent in complex with HTLV-1 protease [11]. Therefore, a proper understanding of the unbinding pathways of these compounds is vital to unveiling secrets that a minor structural difference can have a dramatic effect on inhibitory effects. Therefore, understanding the unbinding process of these inhibitors may provide an insight into the design of next-generation inhibitors.

For a complete understanding of unbinding mechanisms of these two compounds, it is better to get more familiar with this less-known virus's protease structure and inhibitors features



**Fig 1.** The 2D structures of selected compounds were obtained from PDB. A, Compound 9, the inhibitor in (PDB ID 4YDF). B, Compound 10, the inhibitor in (PDB ID 4YDG).

<https://doi.org/10.1371/journal.pone.0263200.g001>



**Fig 2. The 3D structure of HTLV-1 protease (PDB ID 4YDF) and intramolecular interactions of inhibitors.** A, All important domains of HTLV-1 protease: the green area is the active site region, the purple area is the lateral loops or 95S loops part of protease, the orange area is the flaps region, and the blue area is the flap tips part. B, Compound 9. C, Compound 10.

<https://doi.org/10.1371/journal.pone.0263200.g002>

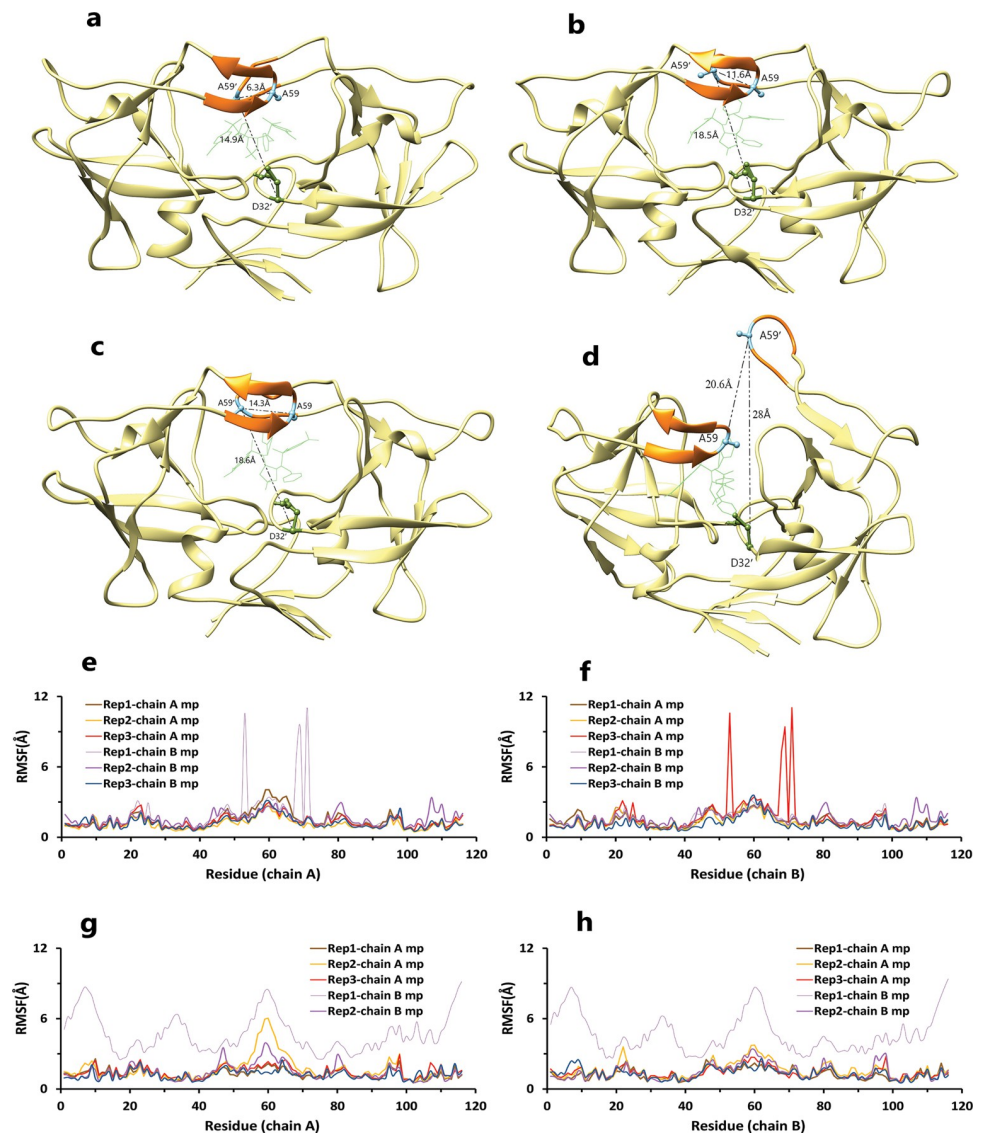
at first. This homodimer protein has some particular regions with strategic effects in keeping ligands in the protein's binding pocket that are obtained from the analysis of trajectories. The active site region (Leu31-Val39 and Leu31'-Val39') contains catalytic dyad aspartate residues (Asp32, Asp32') that are so important in protein-inhibitor interactions. The second essential region is the flaps (Val56-Thr63 and Val56'-Thr63'). The specific residues of Ala59-59' consider as flap tips in the region of the flaps. Finally, Lateral Loops or 95S loops part of protease (Lys95-Gly102 and Lys95'-Gly102') are other key regions in this aspartic protease (Fig 2A). For the inhibitors, both compounds have pi-pi intramolecular interactions. With more details, in compound 9, the nitrobenzene ring can form face-to-face pi-pi interaction with the benzene ring (Fig 2B), and in compound 10, the aniline ring can form T-shaped edge-to-face pi-pi interaction with the benzene ring (Fig 2C).

As mentioned, one of the essential parts of this protein is the region of the flaps, which showed high flexibility during simulations (Fig 3E–3H). So, during our simulations, four modes were observed for the flaps. Herein, we considered two factors to show these modes: the distance between COMs of Ala59 and Ala59' (d1) and the second one is the distance between COMs of Ala59' and Asp32' (d2). The second factor can be even between COMs of Ala59 and Asp 32 due to flaps' handedness opening. In the close form, the maximum amount of d1 and d2 are 10 and 15 Å, respectively (Fig 3A). In the semi-open form, the maximum amount of d1 and d2 is 14 and 20 Å, respectively (Fig 3B). In the open state, the minimum amount of d1 is 14 Å, and the maximum amounts of d1 and d2 are 20 Å (Fig 3C). In the wide-open form, d1 and d2 must be more than 20 Å (Fig 3D).

For our purpose, we had various simulations, that in total, we could have 12 successful unbindings (a total of 14.800 μs) for the two compounds from a minimum of 94 ns to a maximum of 4.4 μs in both mp forms. At last, for providing comprehensive information, all the events in each frame of trajectories were investigated carefully, and different analyses were performed on them.

Like the weak compound, we had two mp forms of the potent compound (AspH32 and AspH32'). In this regard, in the duration times of 4.4 μs (Fig 4A) and 260 and 305 ns (Fig 4B), which were in the chain A, Asp32 protonated state, we saw a uniform mechanism to unbind with some essential differences that caused a significant difference in one of simulation time. So, in the first state of rep1, 2, and 3 (Fig 4C–4E), Asp32', which had salt bridge interaction with the positive charge of the pyrrolidine ring, plays a crucial role in preserving ligand in the binding pocket of protease. This acidic residue is essential because it is located almost at the bottom and center of the binding pocket. This residue considers a strategic residue due to the

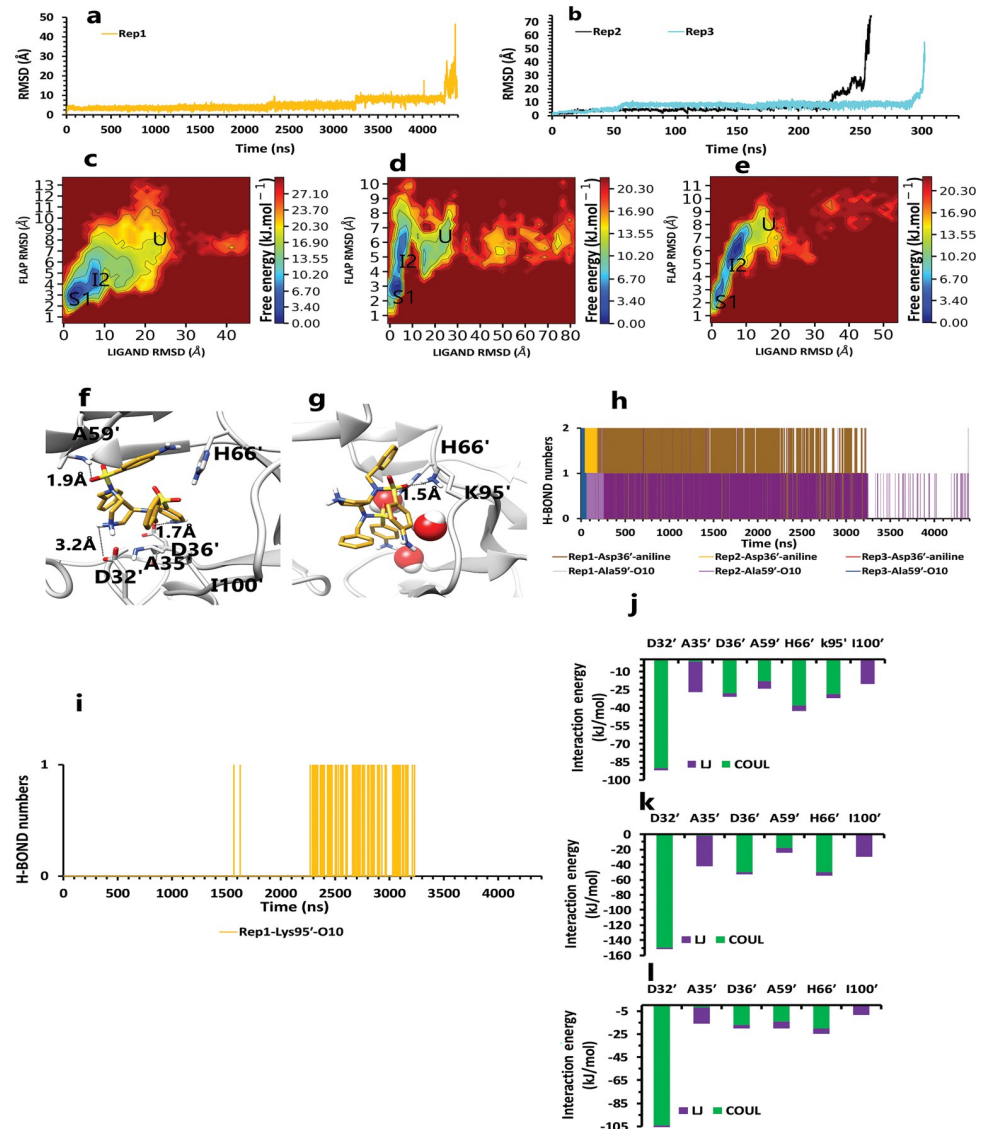




**Fig 3. Different modes of the flap.** A, Close form of the flap. B, Semi-open form of the flap. C, Open form of the flap. D, Wide-open form of the flap. E, F, RMSF values of HTLV-1 protease in the 4YDG PDB code, during simulations. G, H, RMSF values of HTLV-1 protease in the 4YDF PDB code during simulations.

<https://doi.org/10.1371/journal.pone.0263200.g003>

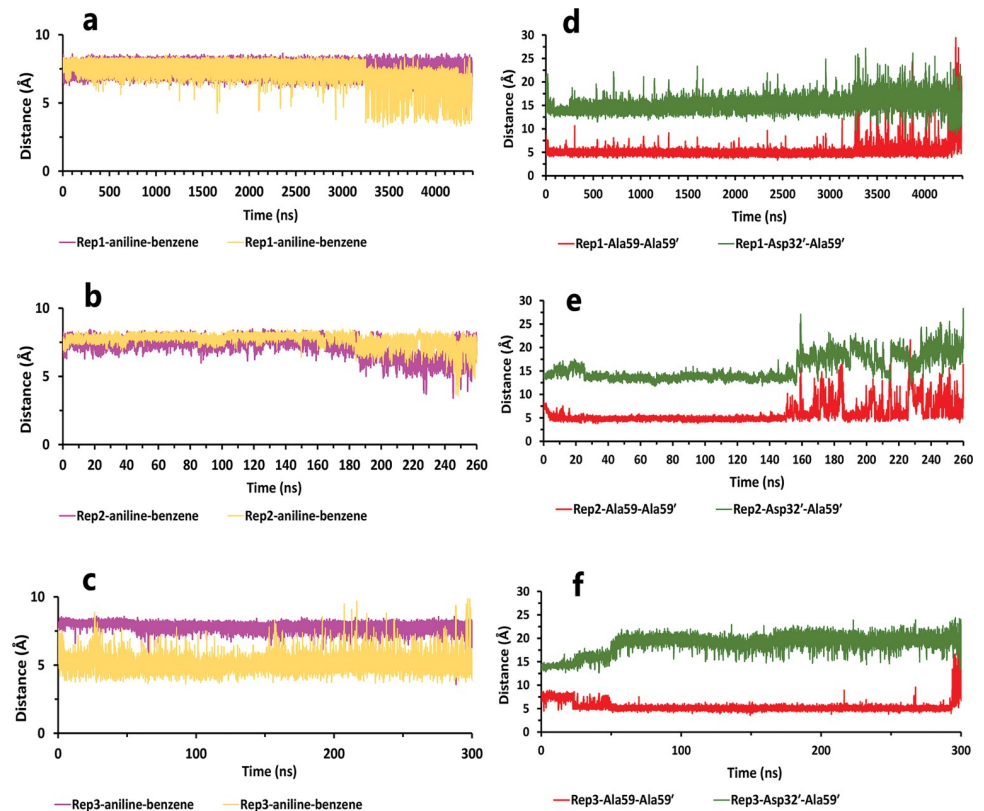
positive charge of pyrrolidine. Parallel to that, His66' by cation- $\pi$  interaction [39, 40] with an aniline ring was the second important preserving factor. In addition, Ala59' in the flap tip by forming H-bond with the atom of O<sub>10</sub> (Fig 4H) and also Ala35' by VdW interactions with a benzene ring, Asp36' by forming H-bond with an aniline fragment in the active site, and finally Ile100' in 95S loop (Fig 4J–4L), with VdW interaction, blocked all the exit routes, like the fence (Fig 4F). As mentioned before, His66' was the second essential residue in this state, which was a supporter of Asp32' to fix the inhibitor in the binding pocket. As time passing, Asp32' loosed its superpower of preserving, and the inhibitor entered the second intermediate state. In this state, Lys95', forming a hydrogen bond with the atom of O<sub>10</sub>, along with His66' cation- $\pi$  interaction with the benzene ring, was a third essential residue. This residue increased protein-ligand interactions time in the rep1 and was absent in rep2 and rep3 (Fig 4G). According to



**Fig 4. The details of compound 10 unbinding pathways in complex with HTLV-1 protease when Asp32 of chain A was protonated in three replicas.** A, RMSD value of the ligand from binding pose to complete unbinding in the rep1. B, RMSD values of the ligand from binding pose to complete unbinding in the rep2 and rep3. C, D, and E, The free energy landscape of rep1, 2, and 3 during the unbinding process (state (S), intermediate state (I), unbound (U)), respectively, which was calculated by using "gmx sham". F, The interactions between the ligand and essential residues in the binding pose of rep1, 2, and 3. G, The new interactions between the inhibitor and particular residues in the second intermediate state of rep1. H, Hydrogen bond numbers of Asp36' and Ala59' with the inhibitor in rep1, 2, and 3. I, Hydrogen bond numbers of Lys95' with the inhibitor in rep1. J, K, and L, The average of most important interaction energies of the protein-ligand complex in rep1, 2, and 3, respectively.

<https://doi.org/10.1371/journal.pone.0263200.g004>

significant differences in replicas simulation times, the effects of the Lys95' hydrogen bond (Fig 4I) appear more pronounced. Finally, ligand pi-pi intramolecular interactions, observed during the whole time of simulations (Fig 5A–5C), slowly weakened all-important protein-inhibitor interactions. The critical point was that, over the entire simulation time, flaps positioning impacted the ligand's behaviors, so the exit process started when the flaps began to open, and Ala59' loosed its effect (Fig 5D–5F) in the second intermediate state gradually with the help of water mediation (Fig 6A–6C).



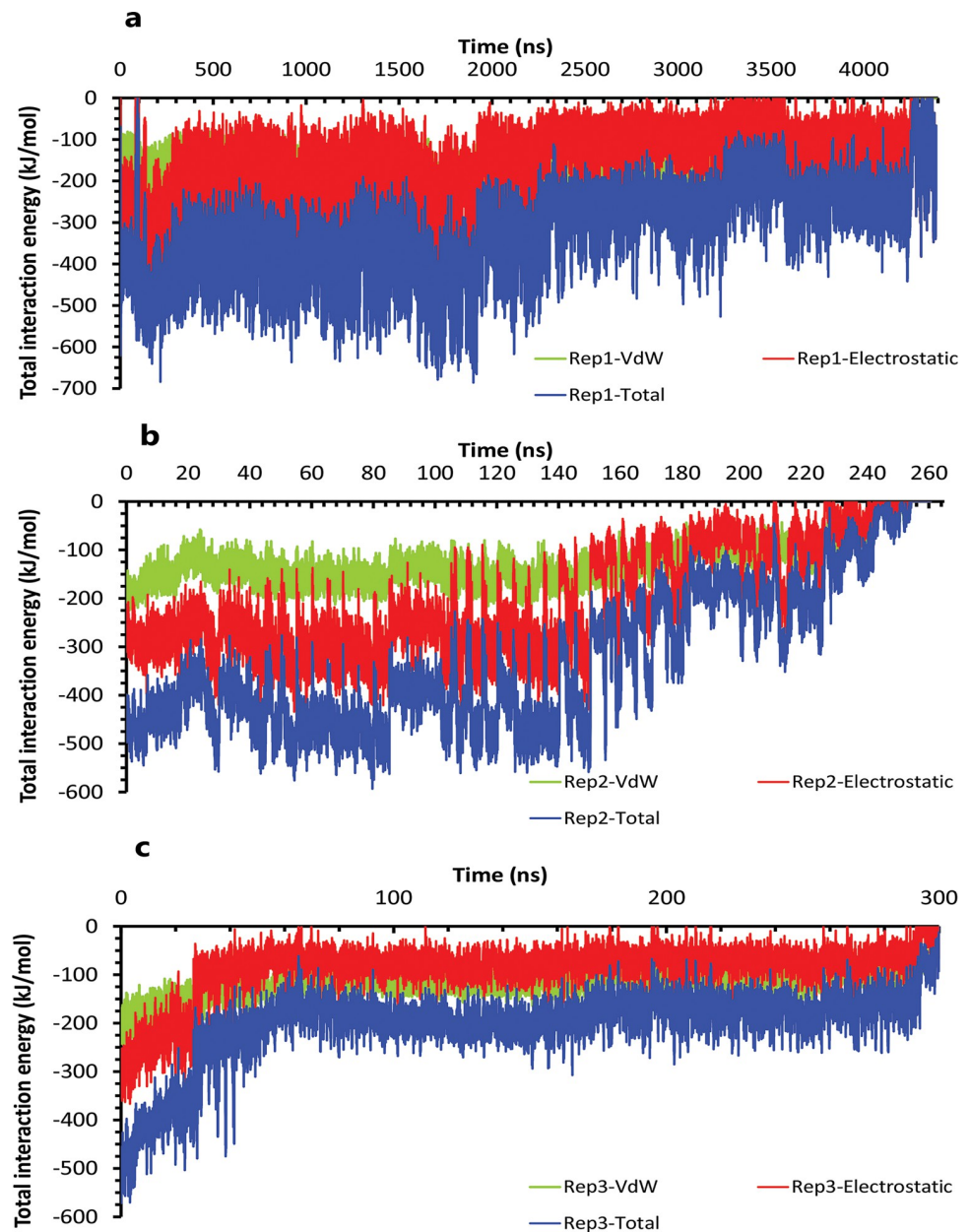
**Fig 5. The details of distances between particular parts of compound 10 complexed with HTLV-1 protease when Asp32 of chain A was protonated in three replicas.** A, B, and C, The distance between COMs of both aniline rings and benzene rings, which were in a position that could form pi-pi intramolecular interactions in all replicas. These plots prove that these fragments were so close together during the simulation. D, E, and F, The distance between COMs of Ala59 and Ala59', and also Asp32' and Ala59' in all replicas (these plots should be checked along with Fig 3).

<https://doi.org/10.1371/journal.pone.0263200.g005>

Conversely, in the other state of protonation (AspH32'), we saw a uniform pathway that was dissimilar to the previous unbinding process with the different lengths of times involving: 94, 320, and 790 ns (Fig 7A and 7B). In the first state (S1 Fig) of these pathways (Fig 7C–7E), Asp32 was as important as expected. Asp36 and Asp36', Leu57, and Ala59' are the residues that acted as auxiliary agents (Fig 7J–7L) to the pivotal amino acid (Asp32). At the first simulation times, along with the salt bridge of Asp32 and pyrrolidine fragment (Fig 7F), both aniline rings had H-bonds with Asp36 and Asp36' in the active site (Fig 7H). Along with these residues, Leu57 and Ala59' formed a hydrogen bond with an aniline fragment and O<sub>41</sub> atom of inhibitor, respectively (Fig 7G and 7I). In the following, in the lack of His66 and Lys95 effects, after time passing with the help of pi-pi ligand intramolecular interactions and water molecules effect (Fig 8A–8C), active site and flaps' important residues lost their effects, and full unbind was observed between the flaps (Figs 8D–8F and 9A–9C).

Depending on close and open flaps and mp states, we had different unbinding mechanisms for the weaker inhibitor. Accordingly, compound 9 was unbound in 148 ns (Fig 10A), 3.5  $\mu$ s, and 3  $\mu$ s (Fig 10B) when Asp 32 of chain A was protonated. In the first state (S1 Fig) (Fig 10C) of the rapid unbinding pathway (rep1), a repulsive force occurred between the pyrrolidine ring of ligand and AspH32 of the binding pocket, and because of no attractive interactions in this area, AspH32 forced the ligand to push out. In addition to this interaction, VdW interaction between both nitrobenzene rings and one of the benzene rings of inhibitor and Leu57,



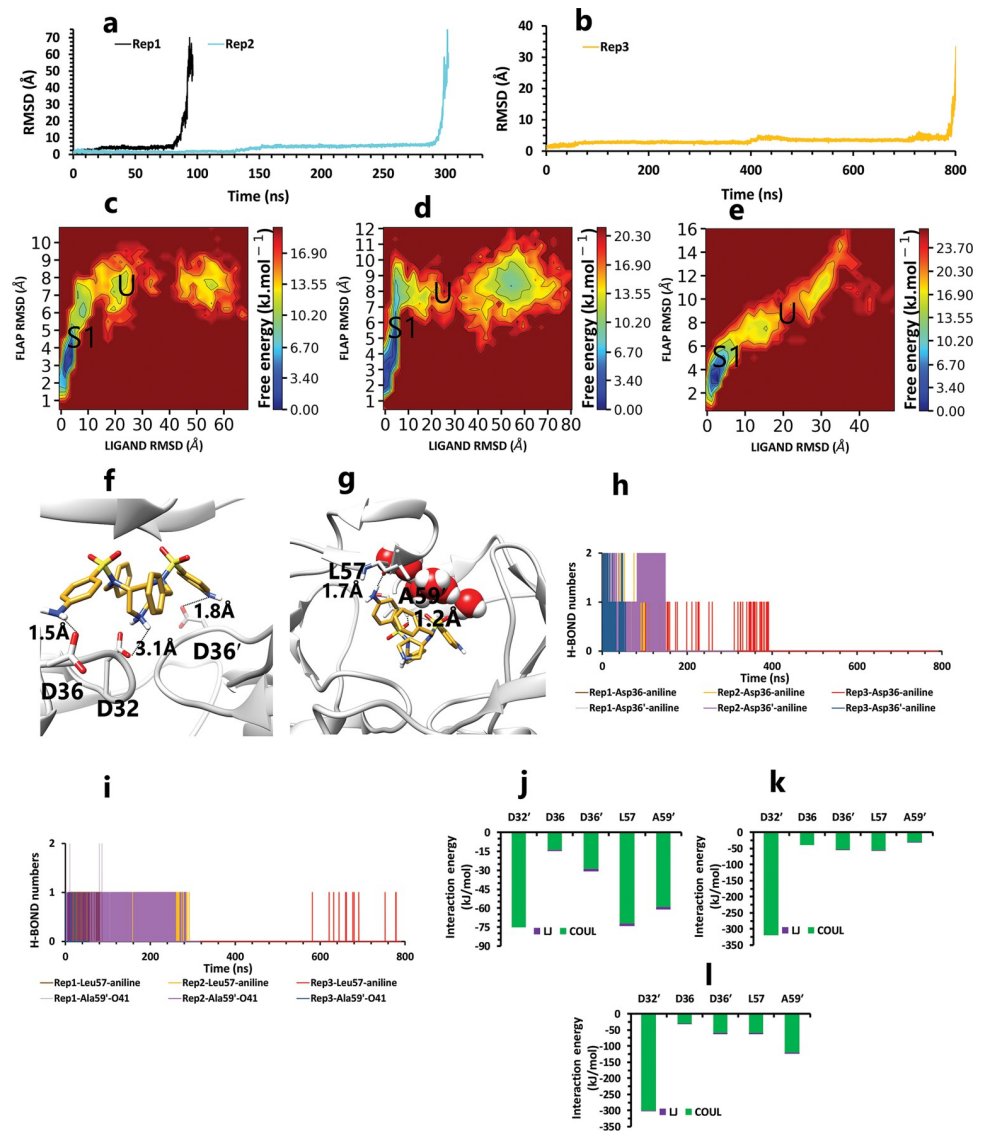


**Fig 6. The interaction energies plots of compound 10 in complex with HTLV-1 protease when Asp32 of chain A was protonated in three replicas. A, B, and C, The total VdW and electrostatic interactions energies of protein-inhibitor complexes in rep1, 2, and 3.**

<https://doi.org/10.1371/journal.pone.0263200.g006>

Gly58, and Ala59 in the close flap region and pi-pi stacking interaction of Trp98' and nitrobenzene fragment and also pi-alkyl interaction of Ile100' with the benzene ring, were other protein-inhibitor significant interactions, which were not potent enough to prevent from repulsive interaction effect (Fig 10F and 10K). In the two other longer simulations, compound 9 was unbound in 3.5  $\mu$ s in wide-open flaps (rep2) and 3  $\mu$ s in close and semi-open flaps (rep3).

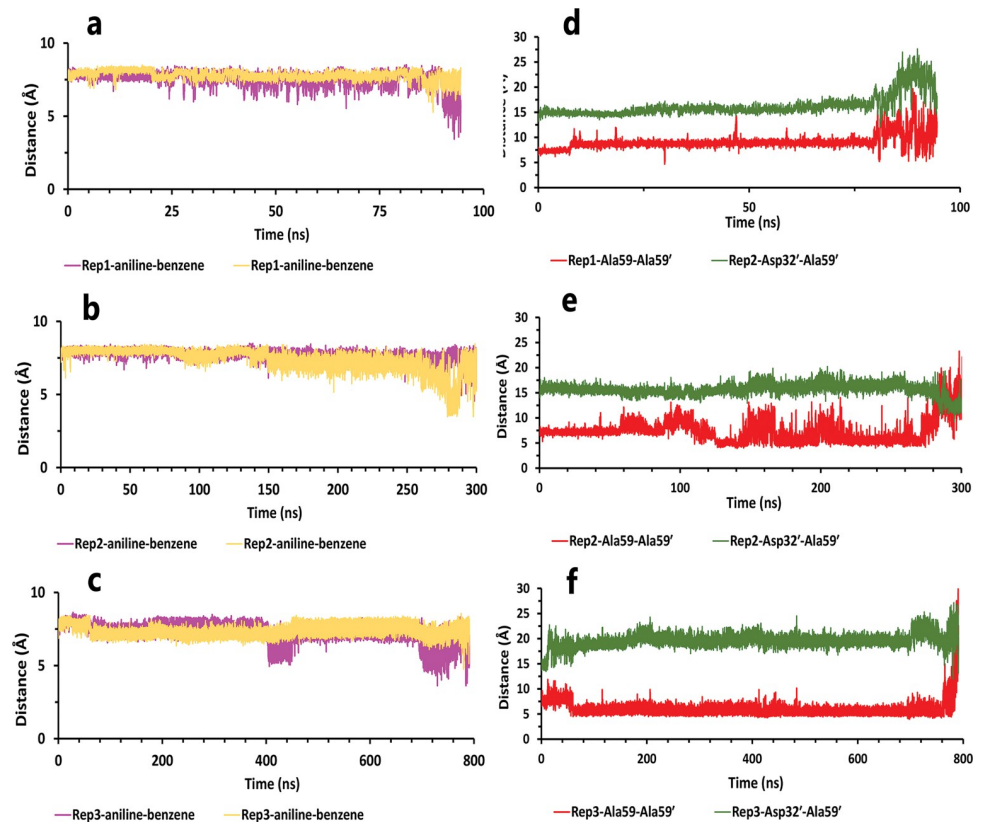
Similarly, Asp32' was the most important amino acid with its salt bridge and the only common point in both pathways. In the first state of rep2 (S1 Fig) (Fig 10D), due to handedness



**Fig 7. The details of compound 10 unbinding pathways in complex with HTLV-1 protease when Asp32' of chain B was protonated in three replicas.** A, RMSD values of the ligand from binding pose to complete unbinding in the rep1 and rep2. B, RMSD value of the ligand from binding pose to complete unbinding in the rep3. C, D, and E, The free energy landscape of rep1, 2, and 3 during the unbinding process (state (S), intermediate state (I), unbound (U)), respectively, which was calculated by using "gmx sham". F, The interactions between the ligand and essential residues of the active site in the binding pose of rep1, 2, and 3. G, The interactions between the ligand and essential residues of the region of the flaps in the binding pose of rep1, 2, and 3. H, Hydrogen bond numbers of Asp36 and Asp36' with aniline fragment in rep1, 2, and 3. I, Hydrogen bond numbers of Leu57 and Ala59' with the inhibitor in rep1, 2, and 3. J, K and L, The average of most important interaction energies of the protein-ligand complex in rep1, 2, and 3, respectively.

<https://doi.org/10.1371/journal.pone.0263200.g007>

opening, only one of the flaps had forward and backward motions, so Leu57', Gly58', Ala59' by VdW interactions kept the ligand in exposing to Asp32'. Also, in this state, Trp98 in the lateral loop built up pi-pi stacking interaction with the nitrobenzene ring of the ligand, and Trp98' built up pi-pi stacking interaction with the benzene ring of another side of the inhibitor (Fig 10G and 10L). So even with enough space for the exit, the inhibitor was still in blockage. These important protein-inhibitor interactions were maintained until the effect of the Asp32'



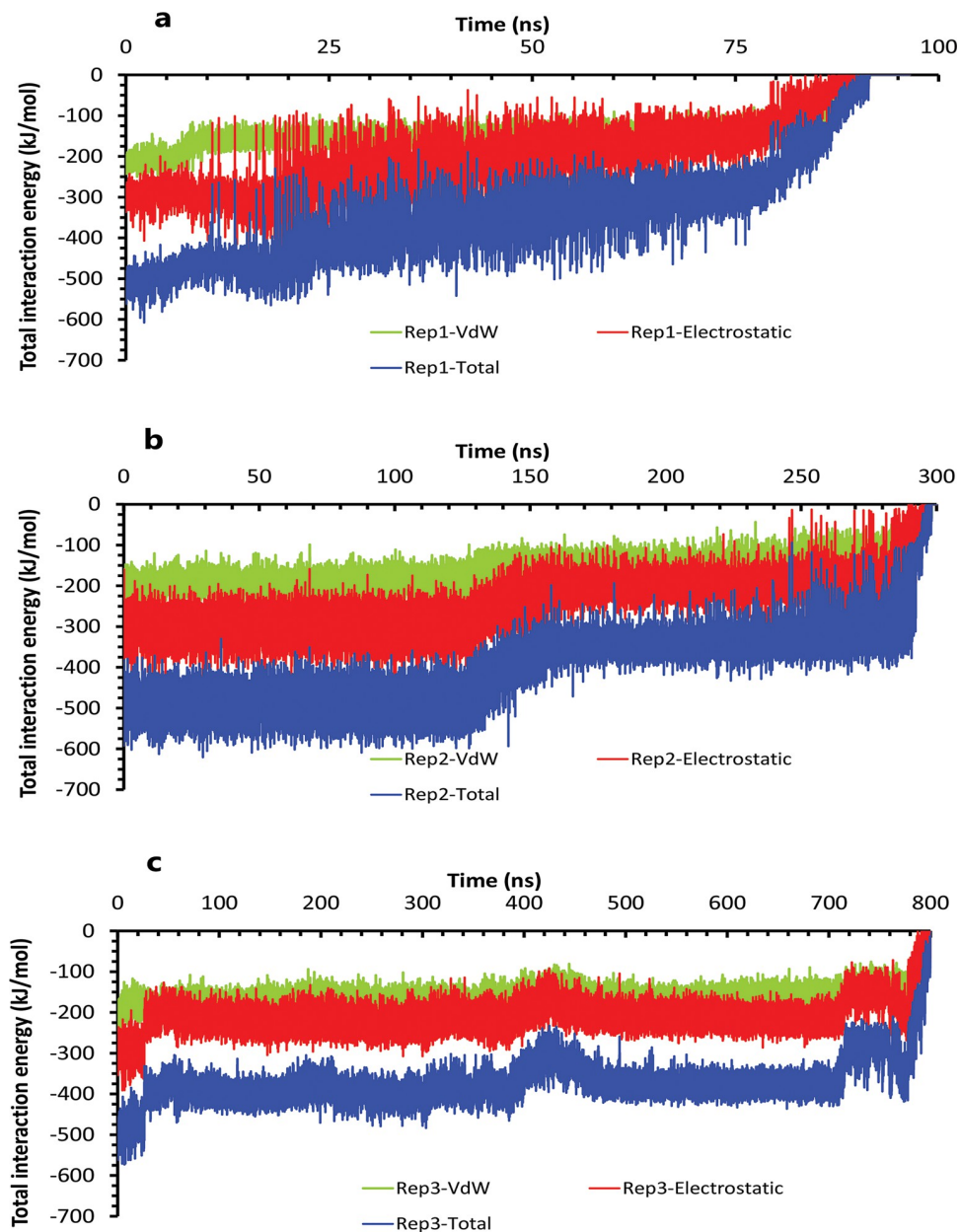
**Fig 8. The details of distances between particular parts in compound 10 in complex with HTLV-1 protease, when Asp32 of chain B was protonated, in three replicas. A, B, and C, The distance between COMs of both aniline rings and benzene rings, which were in a position that could form pi-pi intramolecular interactions in all replicas. These plots prove that these fragments were so close together during the simulation. D, E, and F, The distance between COMs of Ala59 and Ala59' and also Asp32' Ala59' in all replicas (these plots should be checked with Fig 3).**

<https://doi.org/10.1371/journal.pone.0263200.g008>

became faded, and other agents lost their effect. Unexpectedly, the interesting point was that the complete unbinding process does not occur from the region of the flaps. In the rep3 pathway, that the flaps were close or semi-open the whole time, from the first state (Fig 10E), not only Asp32' was necessary, and Asp36 in a close position to Asp32' was powerful too (Fig 10M).

On the other hand, during the first two states, Asp36, by forming pi-anion interaction [41] with nitrobenzene fragment, was momentous as a second ligand preserving residue (Fig 10H), which was promoted to the first important factor in the following intermediate state by replacing pi-anion interaction with the salt bridge with pyrrolidine fragment (Fig 10I). From a holistic view, even though Asp32' was more critical for protein, it was effective until the second intermediate state or until 2  $\mu$ s, but Asp36 (Fig 10J) was effective until complete unbind. Actually, the ligand in all replicas showed face-to-face pi-pi intramolecular interactions between mentioned fragments that caused weakened important protein-ligand interactions gradually with the help of the water mediation effect (Fig 11A–11C). Finally, for the flaps behaviors in all replicas, we saw a new opening form for the rep2 as it was opened from chain A (Fig 11E), and for rep1 and rep3 wide opening (Fig 11D and 11F) was not seen until complete unbound (Fig 12A–12C).

On the contrary, when the Asp32 of chain B was protonated, we saw the same mechanism during the 1.2  $\mu$ s, 410, and 450 ns of simulations (Fig 13A and 13B). In the first state of these

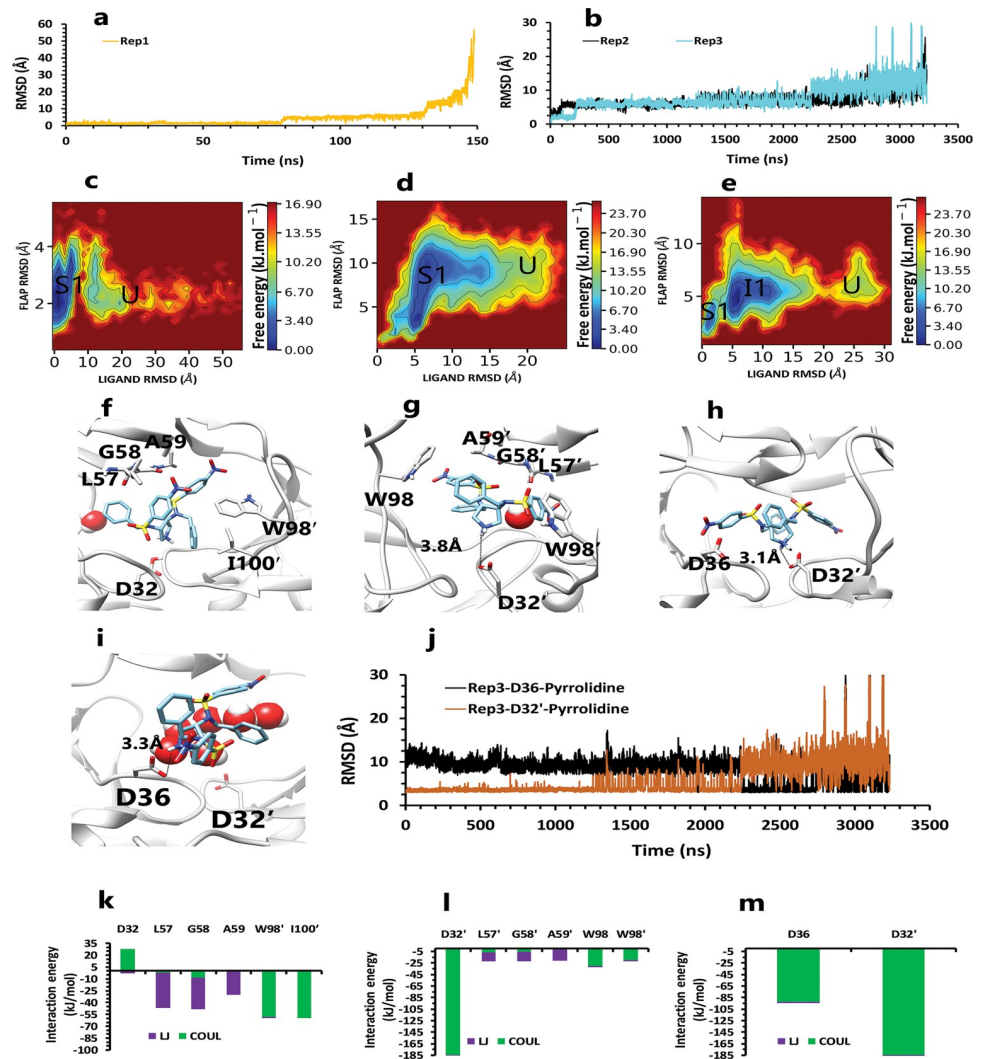


**Fig 9.** The interaction energies plots of compound 10 in complex with HTLV-1 protease when Asp32 of chain B was protonated in three replicas. A, B, and C, The total VdW and electrostatic interactions energies of protein-inhibitor complexes in rep1, 2, and 3.

<https://doi.org/10.1371/journal.pone.0263200.g009>

replicas (Fig 13C–13E), the ligand was surrounded by interactions of some residues in both chains (Fig 13G–13I). Asp32 had salt bridge interaction with pyrrolidine fragment as a most important interaction. In more detail, this fragment also had VdW interaction with Gly34, one of the nitrobenzene rings was in VdW interactions with Leu57' and Gly58' in the flaps regions for the other fragments. The benzene rings were in important interactions involving: pi-pi interaction with Trp98 and pi-alkyl interaction with Ile100 on one side, and pi-pi interactions with Trp98' and pi-alkyl interaction with Ile100' on the other side (Fig 13F). It may be due to



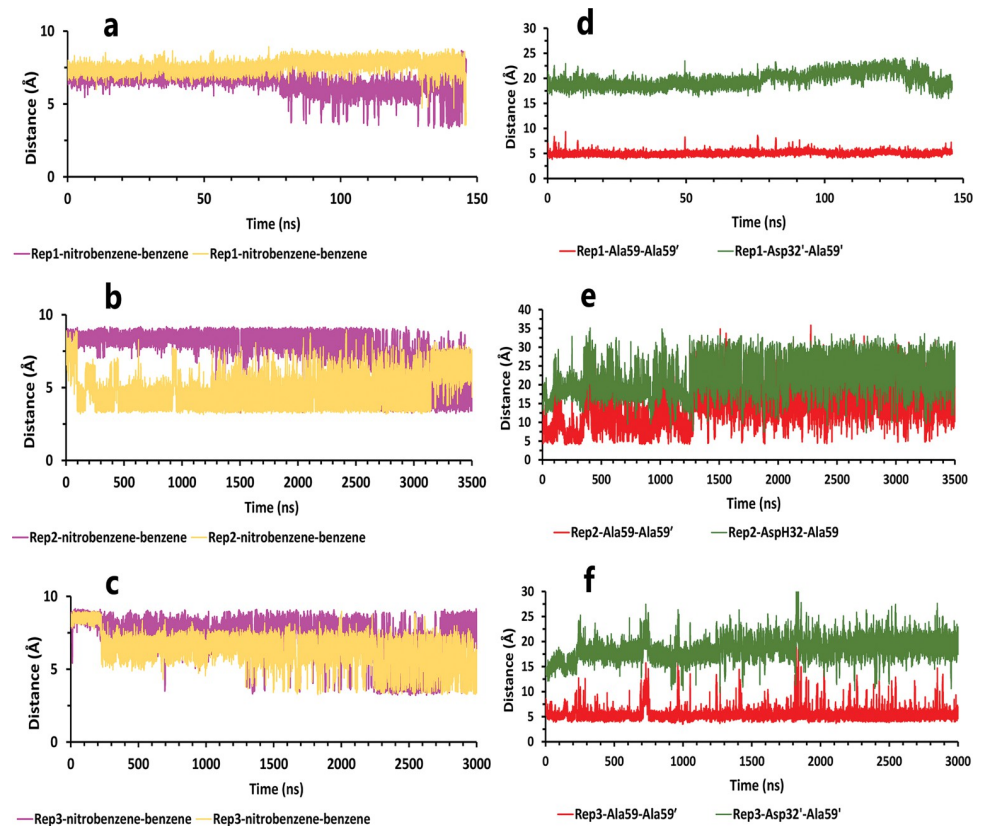


**Fig 10.** The details of compound 9 unbinding pathways in complex with HTLV-1 protease when Asp32 of chain A was protonated in three replicas. A, RMSD value of the ligand from binding pose to complete unbinding in the rep1. B, RMSD values of the ligand from binding pose to complete unbinding in the rep2 and rep3. C, D, and E, The free energy landscape of rep1, 2, and 3 during the unbinding process (state (S), intermediate state (I), unbound (U)), respectively, which was calculated by using "gmx sham". F, The interactions between the ligand and essential active site residues in the rep1. G, The interactions between the ligand and essential active site residues in the rep2. H, The interactions between the ligand and essential active site residues in the binding pose of rep3. I, The new interactions between the inhibitor and particular residues in the second intermediate state of rep3. J, The distance between COMs of pyrrolidine ring and Asp36 and Asp32' in rep3, to show after 2us of simulation this fragment get closer to Asp36 and get farther from Asp32'. K, L, and M, The average of most important interaction energies of the protein-ligand complex in rep1, 2, and 3, respectively.

<https://doi.org/10.1371/journal.pone.0263200.g010>

the high number of important factors; it seems that compound 9 is potent, but except for Asp32 other agents did not have any significant effect. So, they could not keep the ligand after disappearing the Asp32 effect. Thus, as time passed, intramolecular interactions of ligand and water mediation contributed to full unbinding in these three replicas (Fig 14A–14C). Ultimately for the flaps effects, in rep1, the flaps showed high motions, and even though the flaps were wide open (Fig 14D), the full unbind did not occur from this region. In rep2 and 3, the inhibitor unbound between semi-open flaps forms (Figs 14E, 14F and 15A–15C).





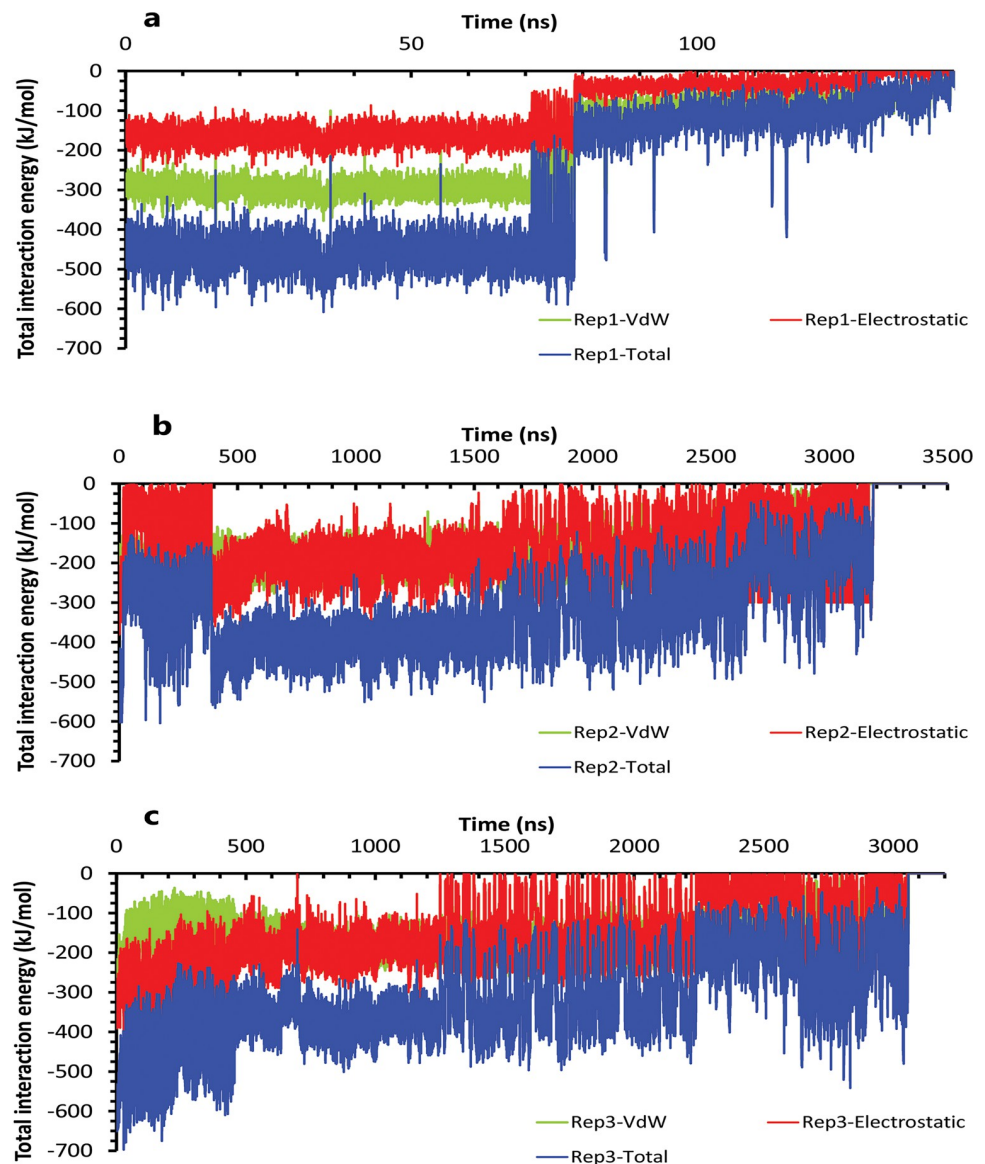
**Fig 11. The details of distances between particular parts in compound 9 in complex with HTLV-1 protease, when Asp32 of chain B was protonated, in three replicas.** A, B, and C, The distance between COMs of both nitrobenzene rings and benzene rings, which were in a position that could form pi-pi intramolecular interactions in all replicas. These plots prove that these fragments were so close together during the simulation. D, The distance between COMs of Ala59 and Ala59' and also Asp32' Ala59' in the rep1 (these plots should be checked with Fig 3). E, The distance between COMs of Ala59 and Ala59' and also AspH32 Ala59' in the rep2. F, The distance between COMs of Ala59 and Ala59' and also Asp32' Ala59' in the rep3. G and H, The total interactions energies of protein-inhibitor complexes in rep1, 2, and 3.

<https://doi.org/10.1371/journal.pone.0263200.g011>

## Conclusion

The atomistic details of unbinding pathways of selected inhibitors in all replicas with various times, the importance of Asp32' in chain A protonation state and Asp32 in chain B protonation state are pretty straightforward. Due to its strategic position, this effective residue could play a critical role in keeping the ligand in the binding pocket for a long time, so the more exposed to Asp32 or Asp32', the more inhibitory effects. The pyrrolidine fragment was held well by Asp32 or Asp32' from the native binding pose of the two compounds, which correlates with experimental research [11]. Thus the interactions of other fragments with other residues in different protein regions caused significant differences.

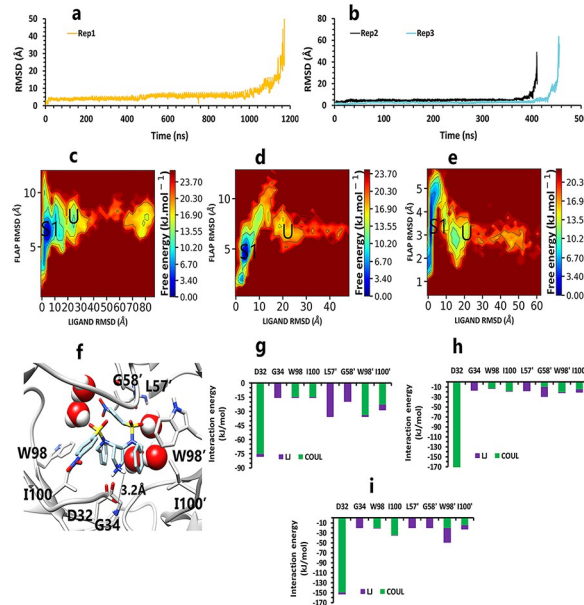
Herein, we cannot conclude which state of protonation occurs, so with our obtained information for the potent compound in chain A protonation state, His66' with its cation- $\pi$  interaction with an aniline ring of inhibitor was a perfect supporter to Asp32'. This residue's effect was absent in the other form of protonation state and caused a significant difference in simulation time. In the weak inhibitor unbinding pathways, Trp98 and Trp98' with pi-pi interactions, due to their close position to one of the exit areas, were not good supporters for Asp32 or Asp32', like His66'. His66', due to its far position from the bottom and center of the binding



**Fig 12.** The interaction energies plots of compound 9 in complex with HTLV-1 protease when Asp32 of chain A was protonated in three replicas. A, B, and C, The total VdW and electrostatic interactions energies of protein-inhibitor complexes in rep1, 2, and 3.

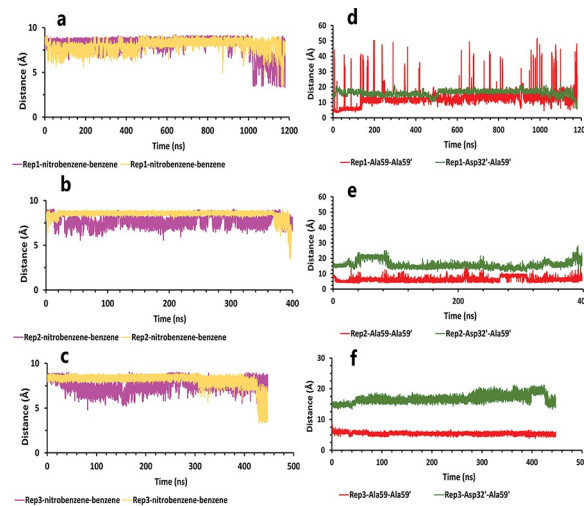
<https://doi.org/10.1371/journal.pone.0263200.g012>

pocket, could fix aniline fragments and decrease ligand fluctuations. The two mentioned tryptophan were closer to the essential aspartic acids, and there was enough space for ligand fluctuations. For this reason, Asp36 in the active site that was close to the exit area could be a competitor with Asp32' and was not a good interaction for keeping the ligand in the binding pocket. Similarly, attenuating effect of Trp98/Trp98' residues in unbinding pathways of the weak inhibitor correlates with our other research result [23]. These residues' interactions are unfavorable to Indinavar's stability in complex with HTLV-1 protease and result in it being a weak inhibitor. As we said before, both compounds had intramolecular interactions that caused weakening critical protein-ligand interactions as time passing. These two compounds did not have the same intramolecular interactions type, so in the weak inhibitor, face-to-face



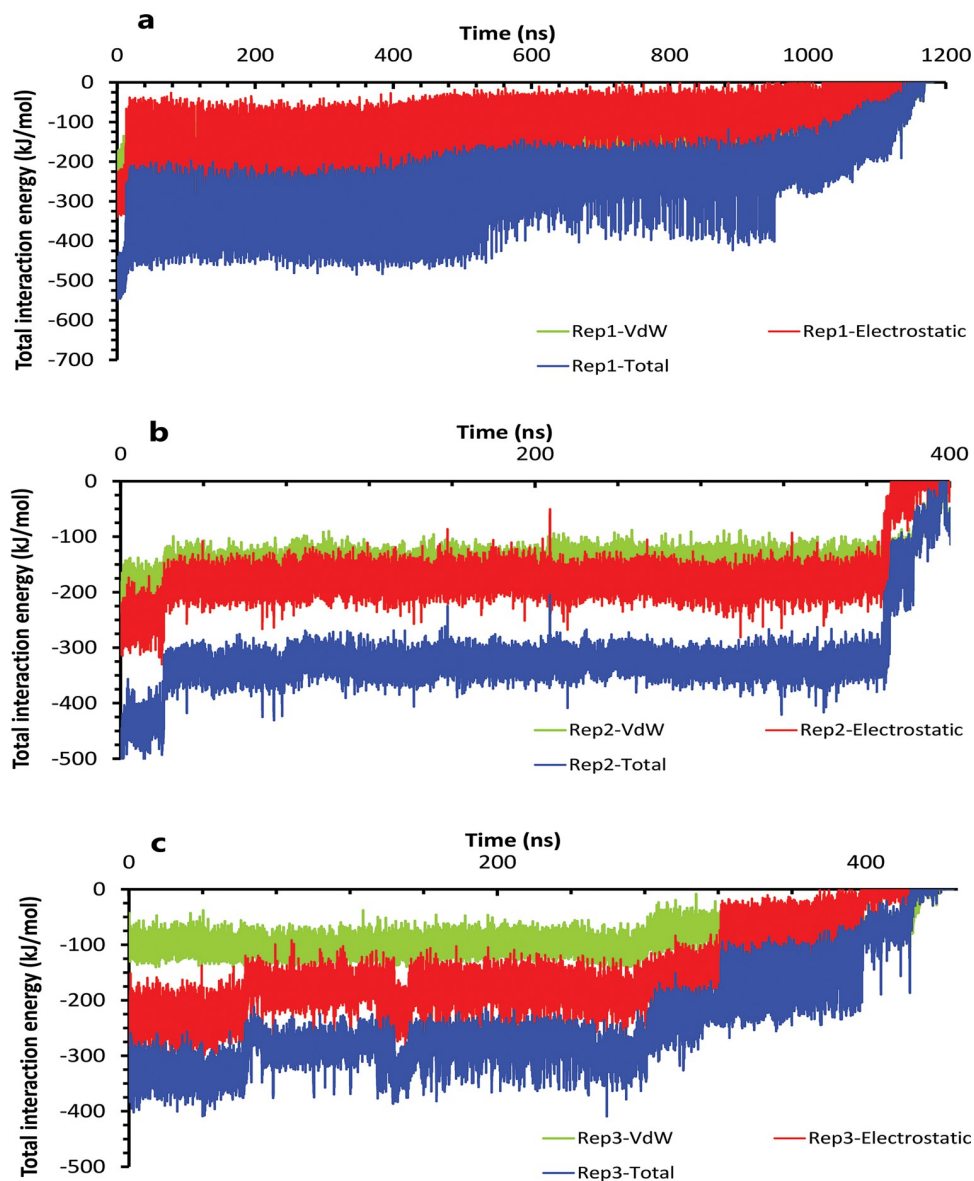
**Fig 13. The details of compound 9 unbinding pathways in complex with HTLV-1 protease when Asp32 of chain B was protonated in three replicas.** A, RMSD value of the ligand from binding pose to complete unbinding in the rep1. B, RMSD values of the ligand from binding pose to complete unbinding in the rep2 and rep3. C, D, and E, The free energy landscape of rep1, 2, and 3 during the unbinding process (state (S), intermediate state (I), unbound (U)), respectively, was calculated using "gmx sham". F, The interactions between the ligand and essential active site residues in the rep1, 2, and 3. G, H, and I, The average of most important interaction energies of the protein-ligand complex in rep1, 2, and 3, respectively.

<https://doi.org/10.1371/journal.pone.0263200.g013>



**Fig 14. The details of distances between particular parts of compound 9 complexed with HTLV-1 protease when Asp32 of chain B was protonated in three replicas.** A, B, and C, The distance between COMs of both nitrobenzene rings and benzene rings, which were in a position that could form pi-pi intramolecular interactions in all replicas. These plots prove that these fragments were so close together during the simulation. D, E, and F, The distance between COMs of Ala59 and Ala59' and also Asp32' Ala59' in all replicas (these plots should be checked with Fig 3).

<https://doi.org/10.1371/journal.pone.0263200.g014>



**Fig 15.** The interaction energies plots of compound 9 in complex with HTLV-1 protease when Asp32 of chain B was protonated in three replicas. A, B, and C, The total VdW and electrostatic interactions energies of protein-inhibitor complexes in rep1, 2, and 3.

<https://doi.org/10.1371/journal.pone.0263200.g015>

pi-pi interactions resulted in the loss of significant pi interactions with the protein. However, the potent inhibitor could have formed more important pi interactions with the protein along with intramolecular interactions. Overall, this obtained information is valuable for designing a new generation of inhibitors against this molecular target.

Because no similar simulation has been performed on the interaction of these inhibitors with HTLV-1 protease, we were only able to compare the results of native states with the crystallographic binding poses. So, except Asp32, Asp32', in the potent inhibitor and Trp98, and in the weaker inhibitor, Asp36, Trp98 and Trp98' were the only common critical residues [11].

## Supporting information

**S1 Fig. The free energy landscape plots of all replicas.**

(TIF)

**S2 Fig.**

(TIF)

## Author Contributions

**Conceptualization:** Hassan Aryapour.

**Data curation:** Fereshteh Noroozi Tiyoula, Hassan Aryapour.

**Formal analysis:** Fereshteh Noroozi Tiyoula, Mostafa Javaheri Moghadam.

**Methodology:** Hassan Aryapour, Mostafa Javaheri Moghadam.

**Project administration:** Hassan Aryapour.

**Software:** Mostafa Javaheri Moghadam.

**Supervision:** Hassan Aryapour.

**Validation:** Hassan Aryapour.

**Writing – original draft:** Fereshteh Noroozi Tiyoula.

**Writing – review & editing:** Hassan Aryapour.

## References

1. Vahlne A., "A historical reflection on the discovery of human retroviruses," *Retrovirology*, vol. 6, pp. 1–9, 2009.
2. Haziot M. E. et al., "Detection of clinical and neurological signs in apparently asymptomatic HTLV-1 infected carriers: Association with high proviral load," *PLoS Negl. Trop. Dis.*, vol. 13, no. 5, pp. 1–9, 2019. <https://doi.org/10.1371/journal.pntd.0006967> PMID: 31042700
3. Kamoi K., "HTLV-1 in Ophthalmology," *Front. Microbiol.*, vol. 11, no. March, pp. 1–12, 2020.
4. Tagaya Y., Matsuoka M., and Gallo R., "40 years of the human T-cell leukemia virus: past, present, and future," *F1000Research*, vol. 8, p. 228, 2019.
5. Edlich R. F., Arnette J. A., and Williams F. M., "Global epidemic of human T-cell lymphotropic virus type-I (HTLV-I)," *J. Emerg. Med.*, vol. 18, no. 1, pp. 109–119, 2000. [https://doi.org/10.1016/s0736-4679\(99\)00173-0](https://doi.org/10.1016/s0736-4679(99)00173-0) PMID: 10645850
6. Hatanaka M. and Nam S. H., "Identification of HTLV-I gag protease and its sequential processing of the gag gene product," *J. Cell. Biochem.*, vol. 40, no. 1, pp. 15–30, 1989. <https://doi.org/10.1002/jcb.240400103> PMID: 2663887
7. S. B. Shuker, V. L. Mariani, B. E. Herger, and K. J. Dennison, "Pii: S1074-5521(03)00002-4," vol. 10, pp. 1–9, 2003.
8. Nguyen J. T. et al., "Truncation and non-natural amino acid substitution studies on HTLV-I protease hexapeptidic inhibitors," *Bioorganic Med. Chem. Lett.*, vol. 18, no. 1, pp. 366–370, 2008. <https://doi.org/10.1016/j.bmcl.2007.10.066> PMID: 18006315
9. Satoh T., Li M., Nguyen J. T., Kiso Y., Gustchina A., and Wlodawer A., "Crystal structures of inhibitor complexes of human T-Cell leukemia virus (HTLV-1) protease," *J. Mol. Biol.*, vol. 401, no. 4, pp. 626–641, 2010. <https://doi.org/10.1016/j.jmb.2010.06.052> PMID: 20600105
10. Kuhnert M., Steuber H., and Diederich W. E., "Structural basis for HTLV-1 protease inhibition by the HIV-1 protease inhibitor indinavir," *J. Med. Chem.*, vol. 57, no. 14, pp. 6266–6272, 2014. <https://doi.org/10.1021/jm500402c> PMID: 25006983
11. Kuhnert M., Blum A., Steuber H., and Diederich W. E., "Privileged Structures Meet Human T-Cell Leukemia Virus-1 (HTLV-1): C2-Symmetric 3,4-Disubstituted Pyrrolidines as Nonpeptidic HTLV-1 Protease Inhibitors," *J. Med. Chem.*, vol. 58, no. 11, pp. 4845–4850, 2015. <https://doi.org/10.1021/acs.jmedchem.5b00346> PMID: 26000468



12. Rajendran V., "Structural analysis of oncogenic mutation of isocitrate dehydrogenase 1," *Mol. Biosyst.*, vol. 12, no. 7, pp. 2276–2287, 2016. <https://doi.org/10.1039/c6mb00182c> PMID: 27194485
13. Singh R., Bhardwaj V. K., Sharma J., Das P., and Purohit R., "Identification of selective cyclin-dependent kinase 2 inhibitor from the library of pyrrolone-fused benzosuberene compounds: an in silico exploration," *J. Biomol. Struct. Dyn.*, vol. 0, no. 0, pp. 1–9, 2021.
14. Bhardwaj V. K. and Purohit R., "Targeting the protein-protein interface pocket of Aurora-A-TPX2 complex: rational drug design and validation," *J. Biomol. Struct. Dyn.*, vol. 39, no. 11, pp. 3882–3891, 2021. <https://doi.org/10.1080/07391102.2020.1772109> PMID: 32448055
15. Sharma J., Bhardwaj V. K., Das P., and Purohit R., "Identification of naturally originated molecules as  $\gamma$ -aminobutyric acid receptor antagonist," *J. Biomol. Struct. Dyn.*, vol. 39, no. 3, pp. 911–922, 2021. <https://doi.org/10.1080/07391102.2020.1720818> PMID: 31980008
16. Niu Y., Li S., Pan D., Liu H., and Yao X., "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.*, vol. 18, no. 7, pp. 5622–5629, 2016. <https://doi.org/10.1039/c5cp06257h> PMID: 26862741
17. Rydzewski J. and Valsson O., "Finding multiple reaction pathways of ligand unbinding," *J. Chem. Phys.*, vol. 150, no. 22, 2019. <https://doi.org/10.1063/1.5108638> PMID: 31202231
18. Halder R. and Jana B., "Exploring and Engineering the Conformational Landscape of Calmodulin through Specific Interactions," *J. Phys. Chem. B*, vol. 123, no. 44, pp. 9321–9327, 2019. <https://doi.org/10.1021/acs.jpcc.9b06343> PMID: 31613097
19. Tiwary P., Limongelli V., Salvalaglio M., and Parrinello M., "Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 5, pp. E386–E391, 2015. <https://doi.org/10.1073/pnas.1424461112> PMID: 25605901
20. Barducci A., Bonomi M., and Parrinello M., "Metadynamics," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 1, no. 5, pp. 826–843, 2011.
21. Tiwary P., Mondal J., and Berne B. J., "How and when does an anticancer drug leave its binding site?," *Sci. Adv.*, vol. 3, no. 5, 2017.
22. Sabbadin D., Salmaso V., Sturlese M., and Moro S., "Supervised molecular dynamics (SuMD) approaches in drug design," *Methods Mol. Biol.*, vol. 1824, no. Md, pp. 287–298, 2018. [https://doi.org/10.1007/978-1-4939-8630-9\\_17](https://doi.org/10.1007/978-1-4939-8630-9_17) PMID: 30039414
23. Sohraby F. and Aryapour H., "Comparative analysis of the unbinding pathways of antiviral drug Indinavir from HIV and HTLV1 proteases by supervised molecular dynamics simulation," *PLoS One*, vol. 16, no. 9, p. e0257916, 2021. <https://doi.org/10.1371/journal.pone.0257916> PMID: 34570822
24. F. T. J. C. C. 2004 UCSF Chimera—a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, "UCSF Chimera."
25. Sousa Da Silva A. W. and Vranken W. F., "ACPYPE—AnteChamber PYthon Parser interfacE," *BMC Res. Notes*, vol. 5, pp. 1–8, 2012.
26. Sun J., Raymundo M. A. V., and Chang C. E. A., "Ritonavir and xk263 Binding-Unbinding with HIV-1 Protease: Pathways, Energy and Comparison," *Life*, vol. 12, no. 1, 2022. <https://doi.org/10.3390/life12010116> PMID: 35054509
27. Cong Y. et al., "Exploring the reasons for decrease in binding affinity of HIV-2 against HIV-1 protease complex using interaction entropy under polarized force field," *Front. Chem.*, vol. 6, no. AUG, pp. 1–18, 2018. <https://doi.org/10.3389/fchem.2018.00380> PMID: 30197882
28. Hofer F., Kraml J., Kahler U., Kamenik A. S., and Liedl K. R., "Catalytic Site pKa Values of Aspartic, Cysteine, and Serine Proteases: Constant pH MD Simulations," *J. Chem. Inf. Model.*, vol. 60, no. 6, pp. 3030–3042, 2020. <https://doi.org/10.1021/acs.jcim.0c00190> PMID: 32348143
29. Abraham M. J. et al., "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1–2, pp. 19–25, 2015.
30. Monvall E., "Statsbudgeten: de stora reformerna gäller arbetslivet," *Tidskr. Sver. Sjukskot.*, vol. 43, no. 2, pp. 54–58, 1976.
31. Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., and Klein M. L., "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, 1983.
32. Hess B., Bekker H., Berendsen H. J. C., and Fraaije J. G. E. M., "LINCS: A Linear Constraint Solver for molecular simulations," *J. Comput. Chem.*, vol. 18, no. 12, pp. 1463–1472, 1997.
33. Darden T., York D., and Pedersen L., "Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems," *J. Chem. Phys.*, vol. 98, no. 12, pp. 10089–10092, 1993.
34. Bussi G., Donadio D., and Parrinello M., "Canonical sampling through velocity rescaling," *J. Chem. Phys.*, vol. 126, no. 1, 2007. <https://doi.org/10.1063/1.2408420> PMID: 17212484

35. Parrinello M. and Rahman A., "Polymorphic transitions in single crystals: A new molecular dynamics method," *J. Appl. Phys.*, vol. 52, no. 12, pp. 7182–7190, 1981.
36. Sabbadin D. and Moro S., "Supervised molecular dynamics (SuMD) as a helpful tool to depict GPCR-ligand recognition pathway in a nanosecond time scale," *J. Chem. Inf. Model.*, vol. 54, no. 2, pp. 372–376, 2014. <https://doi.org/10.1021/ci400766b> PMID: 24456045
37. 25–28. Kraus D. (2014) Consolidated data analysis and presentation using an open-source add-in for the Microsoft Excel® spreadsheet software. Medical Writing, 23, "Daniel's XL Toolbox."
38. 2007. Hunter J. D., "Matplotlib: A 2D Graphics Environment", *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, "Matplotlib."
39. Orabi E. A. and Lamoureux G., "Cation- $\pi$  and  $\pi$ - $\pi$  Interactions in aqueous solution studied using polarizable potential models," *J. Chem. Theory Comput.*, vol. 8, no. 1, pp. 182–193, 2012. <https://doi.org/10.1021/ct200569x> PMID: 26592880
40. Halder R., Manna R. N., Chakraborty S., and Jana B., "Modulation of the Conformational Dynamics of Apo Adenylate Kinase through a  $\pi$ -cation Interaction Modulation of the Conformational Dynamics of Apo Adenylate Kinase through a  $\pi$ -cation Interaction Ritaban Halder, Rabindra Nath Manna, Sandipan Chakraborty," 2017. <https://doi.org/10.1021/acs.jpcc.7b01736> PMID: 28534408
41. Wang D. X. and Wang M. X., "Anion- $\pi$  interactions: Generality, binding strength, and structure," *J. Am. Chem. Soc.*, vol. 135, no. 2, pp. 892–897, 2013. <https://doi.org/10.1021/ja310834w> PMID: 23244296