Mechanism of the association pathways for a pair of fast and slow binding ligands of HIV-1 protease.

Yu-ming M. Huang[1§], Mark Anthony V. Raymundo[1], Wei Chen[1, 2] and Chia-en A. Chang[1]

[1]Department of Chemistry, University of California, Riverside, Riverside, CA 92521, USA

[2]ChemConsulting LLC, Frederick, MD 21704, USA

[§]Current address: Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093, USA

Supporting Material

Text S1:

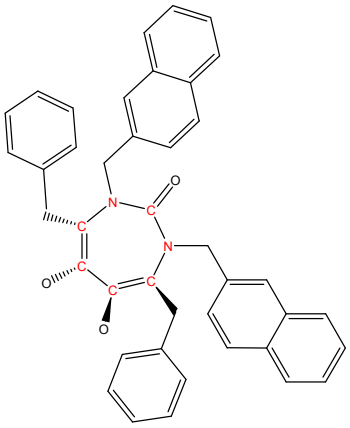
Compare the ligand positions between the MD simulations and crystal coordinates.

1. Align the MD snapshots and crystal coordinates.

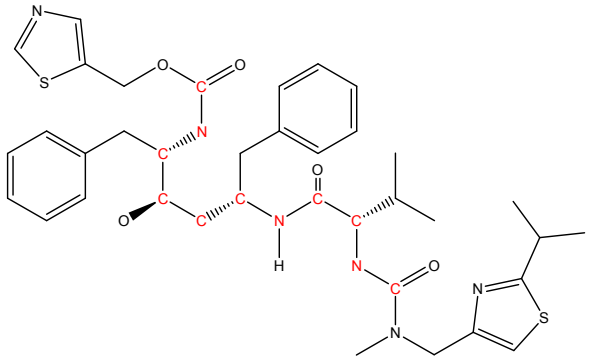
2. Calculate center of mass (COM) of the selected heavy atoms (highlighted in red in the following figure) from the crystal structure as COM_ref.

3. Calculate COM of the selected heavy atoms from MD snapshots as COM_traj.

4. Calculate the distance between COM_ref and COM_traj.

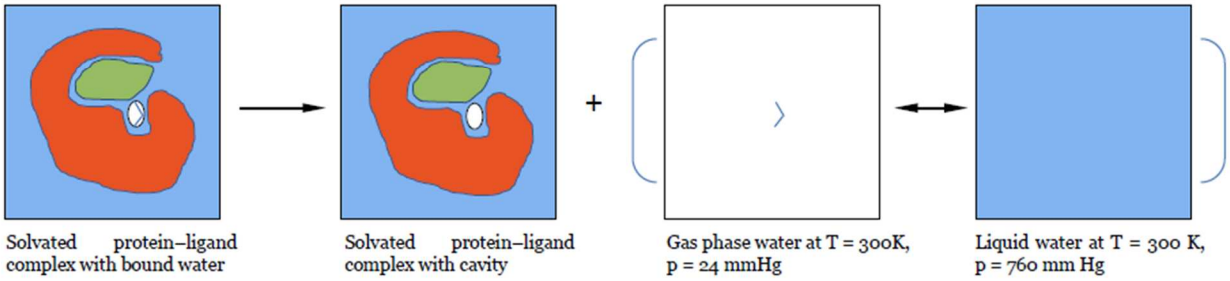


xk263



ritonavir

Text S2: More details about Removal free energy of water molecules



The water-removal free energy is computed as:

$$\Delta G^{o}_{wr-liquid} = \mu^{0}_{PL} + (\mu^{o}_{W-gas} + \Delta G_{gl}) - \mu^{o}_{PLW}$$

ΔG_{gl} is the gas-to-liquid transfer free energy of the water. The vapor pressure of water at 300 K is 24 mm Hg, so the chemical potential of water at 24 mm Hg is equal to that of liquid water. Hence, free energy of gas-liquid transfer ΔG_{gl} is the free energy change from the initial gas-phase concentration C in moles/liter to a new concentration corresponding to an ideal gas at 24 mm Hg. An amount of 1 mol ideal gas at 1 atm occupies 22.4 l, so its concentration is 0.0446 (mol/l-atm). The concentration thus equals [24 mm Hg/760 mm Hg/atm] [0.0446 mol/l-atm] = 0.00141 mol/l. Thus, the free energy change is

$$\Delta G_{gl} = RT \ln\left(\frac{0.00141M}{C}\right)$$

μ^{o}_{PLW}, μ^{o}_{PL} and μ^{o}_{W-gas} are calculated with VM2 software [2]. VM2 belongs to a class of methods that can provide true free energies of binding at a quite moderate computational cost. Because these methods focus on the most stable conformations of the molecules, they are sometimes called "predominant states" methods. They compute the standard chemical potential of each molecular species by finding its M most stable conformations $j = 1 \ldots M$ by low mode

searches, integrating the Boltzmann factor within each energy well j to obtain a local configurational integral z_j, and then combining these local configuration integrals according to the following formula [1, 2], where X stands for the molecular species (PLW, PL and W-gas), as defined above

$$\mu_X^0 = -RT\ln\frac{8\pi^2}{C^0}\sum_{j=1}^{N} z_j \; = -RT\ln\sum_{j=1}^{N} e^{-\frac{\mu_{X,j}^0}{RT}}$$

Here C^0 is the standard concentration, which combined with the factor of $8\pi^2$, accounts for the positional and orientational mobility of the free molecule at standard concentration, and the second form of the summation is given in terms of the chemical potentials of the individual conformations. The calculations are tractable in part because they use implicit solvent models (e.g., Generalized-Born and Poisson-Boltzmann), which are widely accepted as computationally efficient alternatives to more detailed solvent models, and in part because for large systems such as protein-ligand complexes only a subset of atoms (~500-5000 depending on the nature of the active site) are treated as flexible. Predominant states methods can account for changes in entropy associated with the loss in mobility of the ligand and receptor upon binding through the local configuration integrals and the sum over local energy minima.

In this water removal implementation, when a water molecule needs to be removed from the system it is turned to a dummy water by setting the following parameters to zero: the partial charges and van der Waals parameter E_{min} of its atoms, force constants of its bonds, and force constant of its angle. The live set, which defines the mobile atoms in calculations, only includes water molecules. And Flat-bottom energy well constraints, as expressed in the following equation where k is the force constant of the restraint, r is the location of the restrained atom, r_0

is the center of the flat-bottomed well $R_{restraint}$ is the radius of the restraint region and p is a parameter controlling the hardness of the energy wall around the energy well, are applied to all of the water oxygen atoms.

$$E_{constraint} = k\frac{(r - r_0)^p}{R_{constraint}^p}$$

[1] Head, M.S., Given, J.A., Gilson, M.K. *J. Phys. Chem.* **1997**, *101*, 1609.

[2] Chang, C.-E., Potter, M. J., Gilson, M.K. *J. Phys. Chem. B* **2003**, *107*, 1048.

Text S3: LogP caluculations of xk263 and ritonavir

LogP by ACD/LogP [3]: Ritonavir 5.28±0.89, xk263: 10.16±0.57

LogP by ALOGPS2.1 [4]: Ritonavir 4.41, xk263: 6.24

LogP by Virtual LogP [5]: Ritonavir 4.63, xk263: 6.66

[3] Petrauskas, A.A., Kolovanov, E.A. ACD/Log P method description, Perspectives in Drug Discovery and Design, 2000, 19, 99-116.

[4] Tetko, I.V., Tanchuk, V.Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program, J. Chem. Inf. Comput. Sci., 2002, 42, 1136-1145

[5] Gaillard, P., Carrupt, P.A., Testa, B., Boudon, A. Molecular Lipophilicity Potential, a tool in 3D QSAR: Method and applications, Journal of Computer-Aided Molecular Design, 1994, 8(2), 83-96.

System	Initial ligand position	Final ligand position	Final flap orientation	Simulation time (ns)
Xk263 : HIVp (semi-open flaps)	Front	Right	Slightly open	500
	Front	Right-down	Slightly open	200
	Front	Right-down	Slightly open	200
	Front	Front	Widely open	200
	Front	Inside, between flaps; Avg COM dist: 3.970 +/- 0.4 Å	Open	300
	Top	Inside Avg COM Dist: 9.311 +/- 1.3 Å	Close	14000
	Top	Top-right	Open	350
	Top	Near Pro79	Open	400
	Top	Inside, between flaps	Open	400
	Top	Right-down	Open	200
	Top	Left elbow	Open	200
	Left	Left elbow	Slightly switch	200
	Right	Inside with a flap Avg COM Dist: 9.302 +/- 1.3 Å	Widely open	1044
	Right	Left-down	Widely open	200
	Right	Near Pro79	Slightly open	200
	Right	Near Pro79	Open	900
Xk263 : HIVp (slightly-open flaps)	Front	Inside with a flap Avg COM Dist: 4.960 +/- .5 Å	Widely open	200
	Top	Button	Open	200
	Top	Near Pro79	Slightly close	200
	Right	Inside with a flap Avg COM Dist: 8.905 +/- .8 Å	Widely open	200
Xk263 : HIVp (open with a curl flap)	Front	Front, near flaps	Close	300
	Top	Inside Avg COM Dist: 1.489 +/- .5 Å	Close	12400
	Top	Near flaps	Close	300
	Right	Near Pro79	Slightly open	200
Xk263 : HIVp (wide-open flaps)	Front	Left-down	Open	200
	Front	Right-back	Open	200
	Front	Front	Widely open	200
	Top	Left elbow	Slightly open	200
	Top	Between flaps	Slightly open	500

Ritonavir : HIVp (semi-open flaps)	Front	Left-front	Widely open	200
	Front	Left-front	Open	200
	Front	Slightly inside	Open	8000
	Top	Inside with a flap Avg COM Dist: 3.489 +/- .4 Å	Widely open	300
	Top	Near Pro79	Semi-open	200
	Top	Left-down	Semi-open	200
	Top	Near Pro79	Semi-open	200
	Top	Near Pro79	Semi-open	200
	Left	Left elbow	Slightly open	200
	Right	Front-down	Semi-open	200
Ritonavir : HIVp (slightly-open flaps)	Front	Between flaps	Open	200
	Front	Between flaps	Open	200
	Top	Down-right	Close	200
	Top	Inside Avg COM Dist: 5.968 +/- .6 Å	Close	3300
	Top	Inside Avg COM Dist: 3.391 +/- .4 Å	Close	3000
	Top	Near Pro79	Close	200
Ritonavir : HIVp (open with a curl flap)	Front	Back	Close	200
	Front	Inside Avg COM Dist: 3.489 +/- .4 Å	Close	200
	Top	Down, in solvent	Close	200
	Top	Partly inside	Close	200
	Top	Partly inside	Close	300
	Top	Slightly inside	Close	200
Ritonavir : HIVp (wide-open flaps)	Front	Right	Close	200
	Front	Inside Avg COM Dist: 5.112 +/- 1.2 Å	Close	300
	Front	Inside Avg COM Dist: 5.002 +/- .7 Å	Close	300
	Top	Inside Avg COM Dist: 4.281 +/- .9 Å	Close	400
	Top	Inside Avg COM Dist: 3.747 +/- .6 Å	Close	400
	Top	Slightly inside	Slightly open	200
	Top	Slightly inside	Close	200

Table S1: List of all MD simulations for ligand association to HIVp. Red: The ligands successfully entered the binding site. The average COM distance was calculated for 20 ns; during this time, the ligand was closest to the final bound form. Blue: the ligands did not bind successfully inside the enzyme. However, it was located on the periphery, within 10 Å, of the binding site. Black: the ligands did not successfully bind inside the enzyme and was not found along the periphery of the binding site. Regarding the ligand positions, such as top, front and right, please refer Figure 2 for the location description.

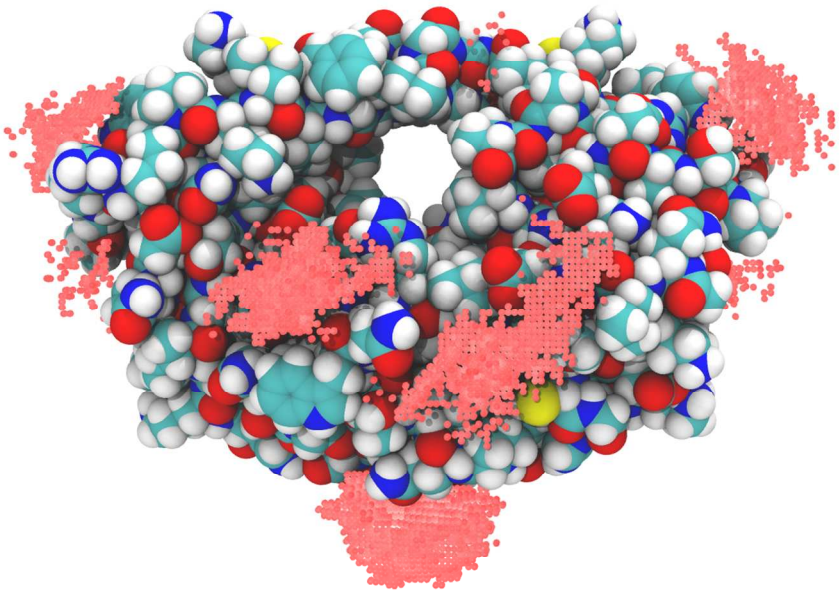


Figure S1: Ligand−protein nonspecific association probability density map of XK263 on the surface of the HIVp modeled by the Brownian dynamics simulations. The simulation was carried out using the program GeomBD2, and the HIVp is held rigid (PDB: 1HHP). Association regions in the elbow, bottom and fulcrum of HIVp were found with a high probability.

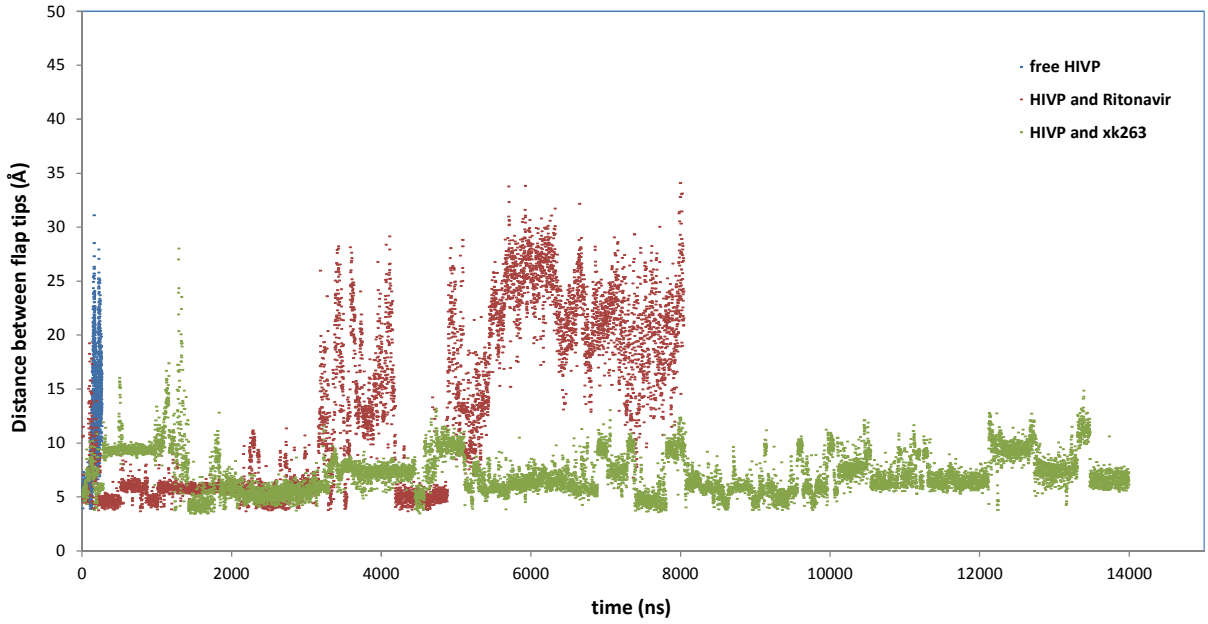


Figure S2: Flap distances along the MD runs. The distance between flap tips is defined as the distance between the alpha carbons of Gly51 and Gly51'. The frames for the complexes were save every 1 ns, while for the free protein they were saved every 100 ps.

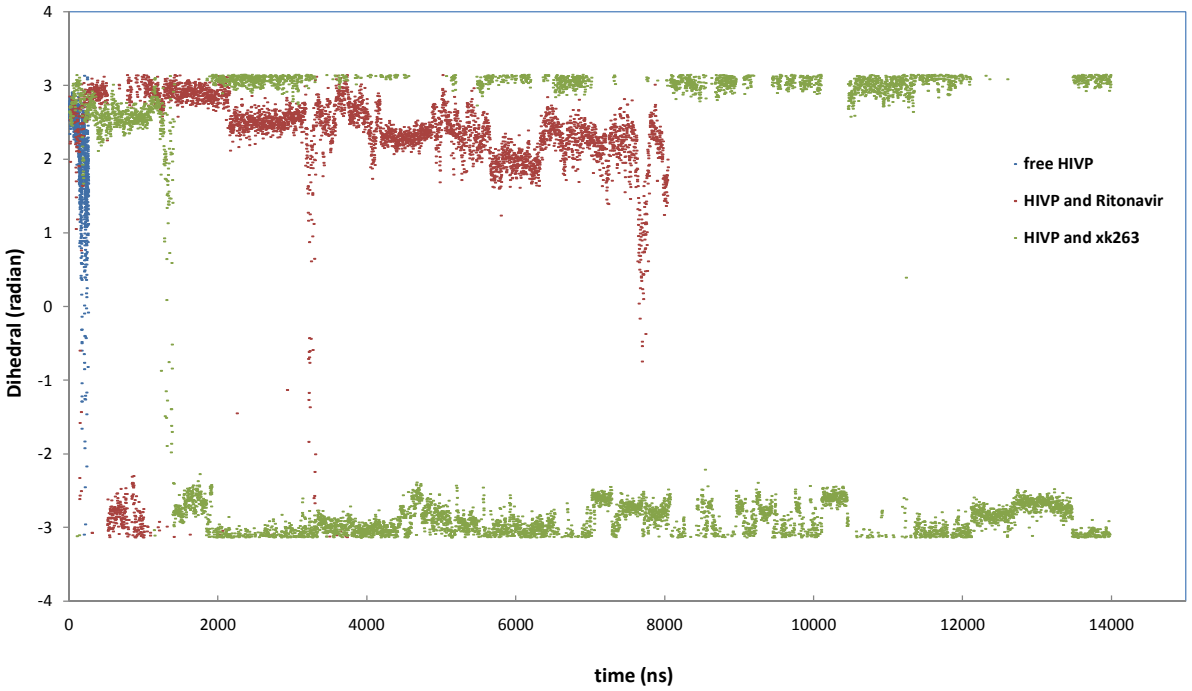


Figure S3: Handedness changes along the MD runs. The handedness is measured by the dihedral defined by the alpha carbons of residues 42, 51, 51' and 42'. The position and negative dihedral number is correlated to semi-open and closed handedness, respectively. The frames for the complexes were save every 1 ns, while for the free protein they were saved every 100 ps.

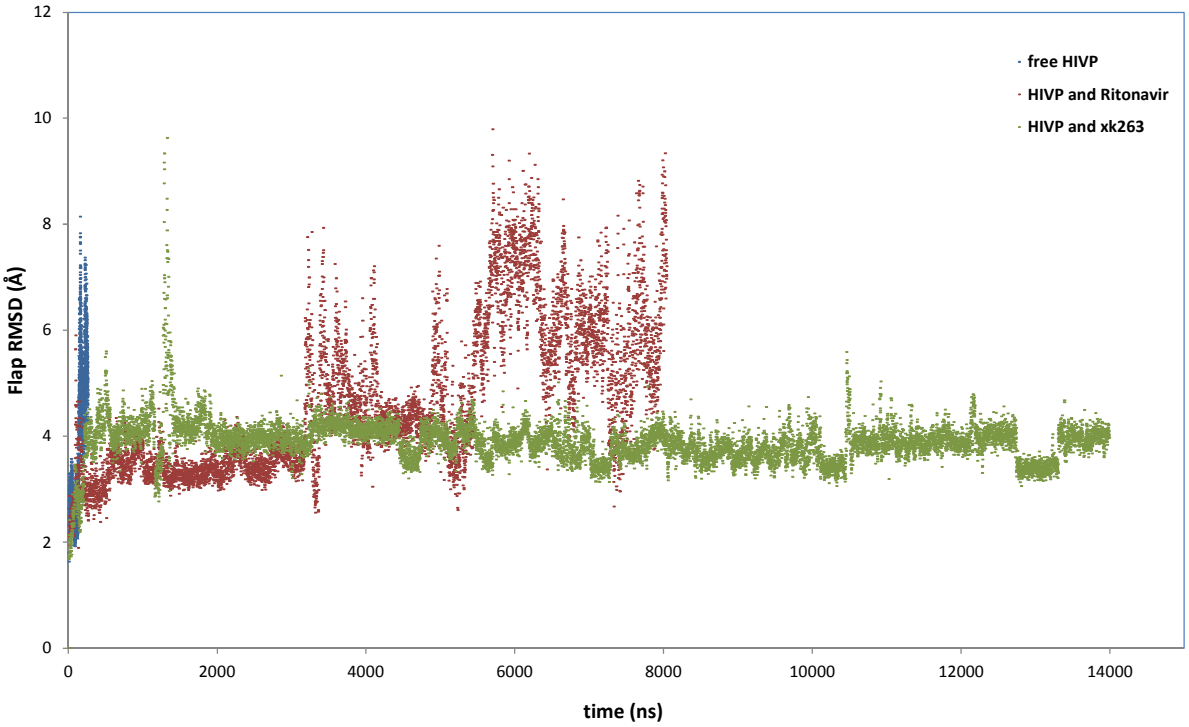


Figure S4: Flap RMSD along the MD runs. Flap RMSD were measured for residues 43-58 and 43'-58'. The frames for the complexes were save every 1 ns, while for the free protein they were saved every 100 ps.

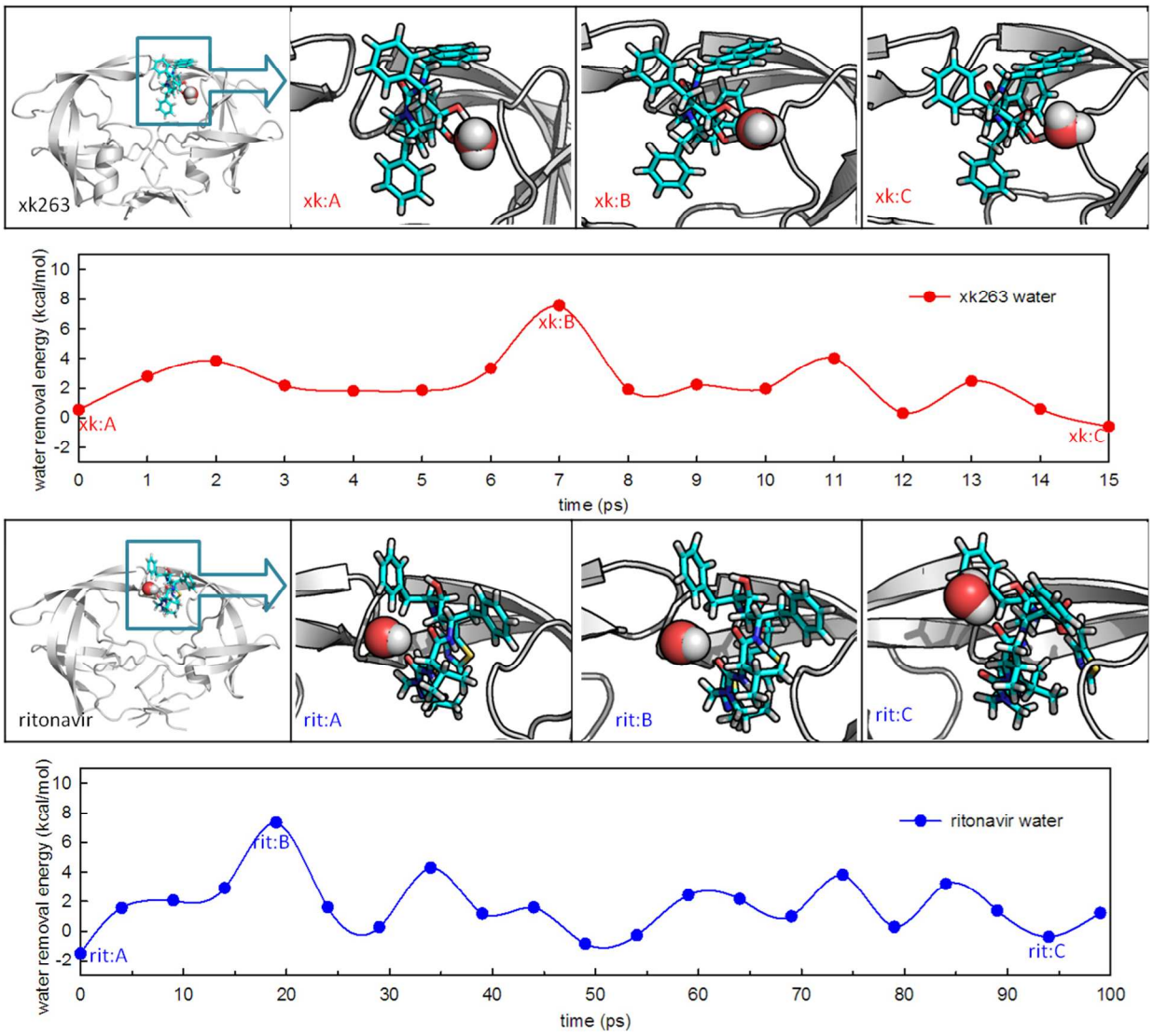


Figure S5: Red and blue lines show the removal free energy of one water molecule. The label for xk:A, xk:B and xk:C indicates the snapshot of the simulation time at 0, 7 and 15 ps. The label for rit:A, rit:B and rit:C indicates the snapshot of the simulation time at 0, 19 and 94 ps.

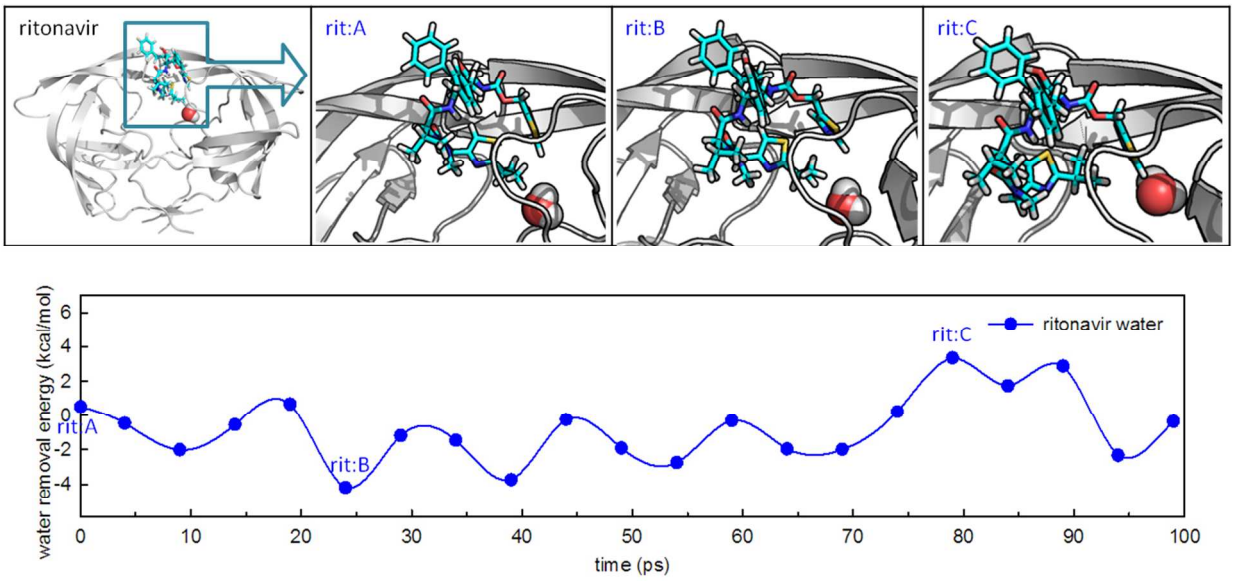


Figure S6: Calculation of the free energy to remove one water molecule when ritonavir associates with the HIVp. The label for rit:A, rit:B and rit:C indicates the snapshot of the simulation time at 0, 24 and 79 ps.

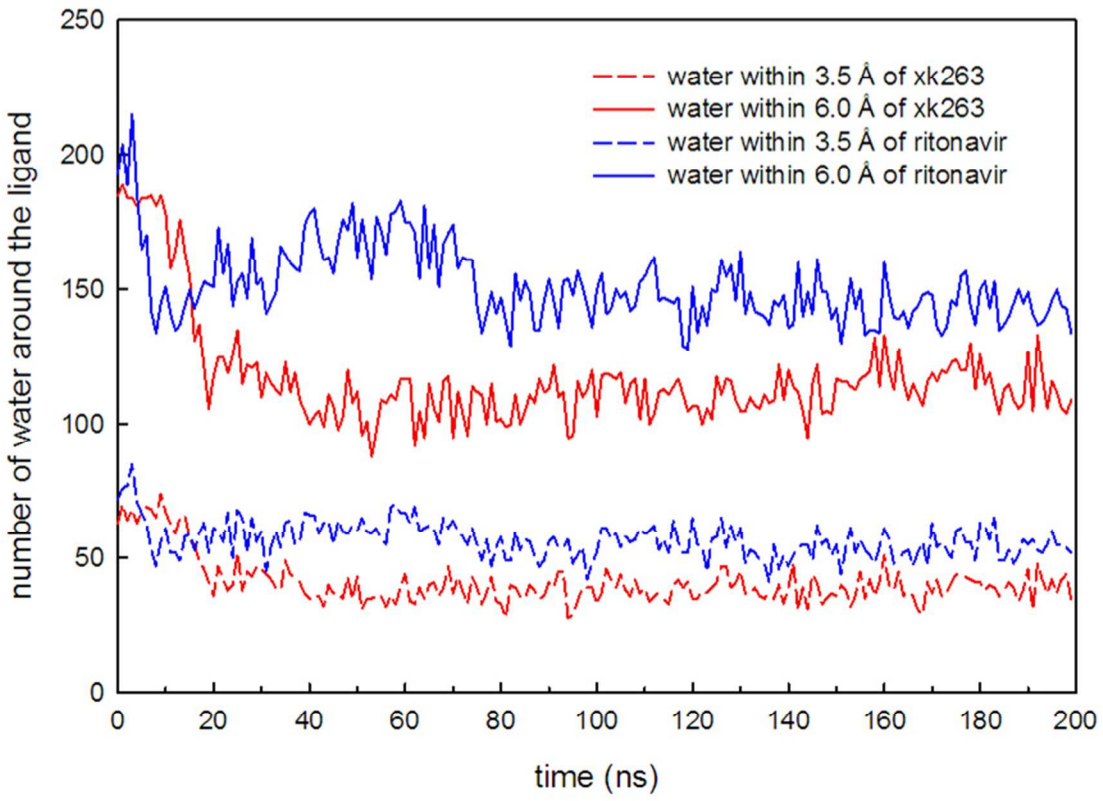


Figure S7: Number of water molecules surrounding xk263 and ritonavir within 3.5 and 6.0 Å water shell.

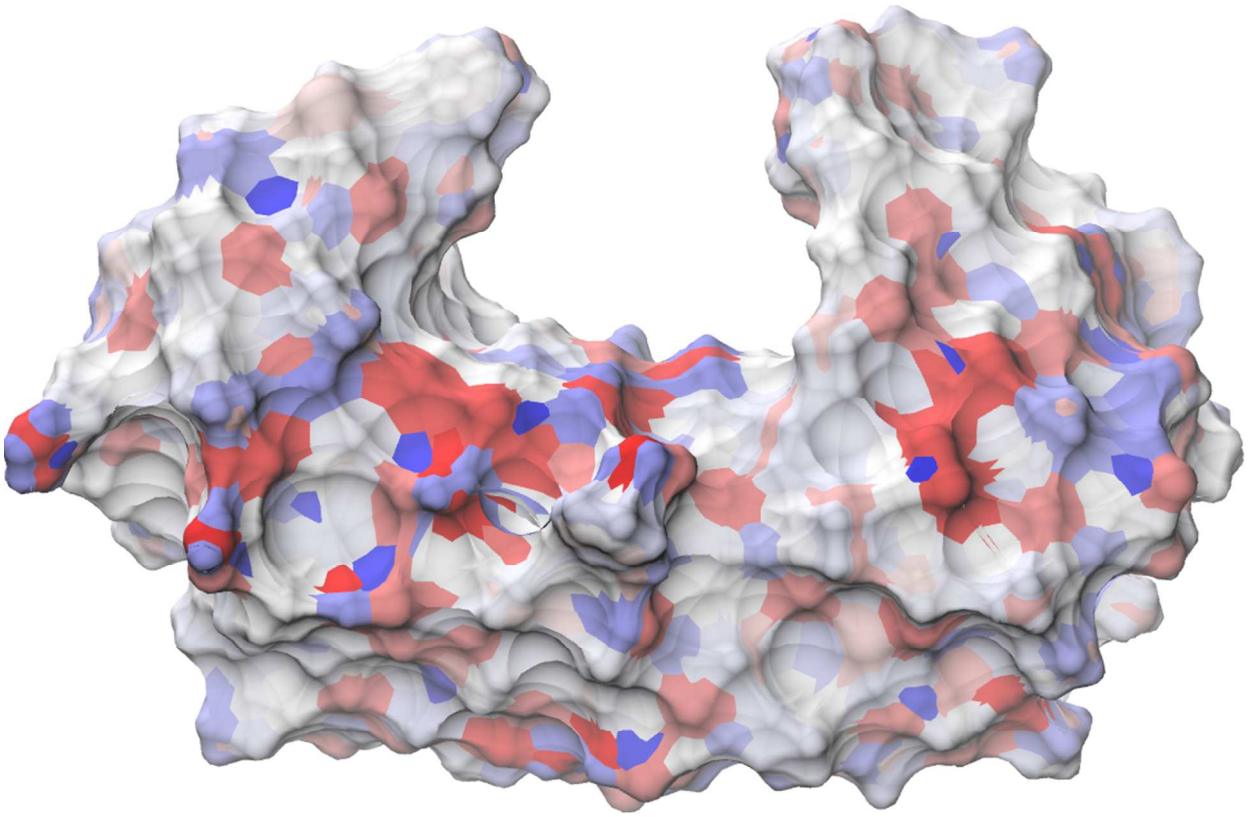


Figure S8: Residue characteristics of HIVp.

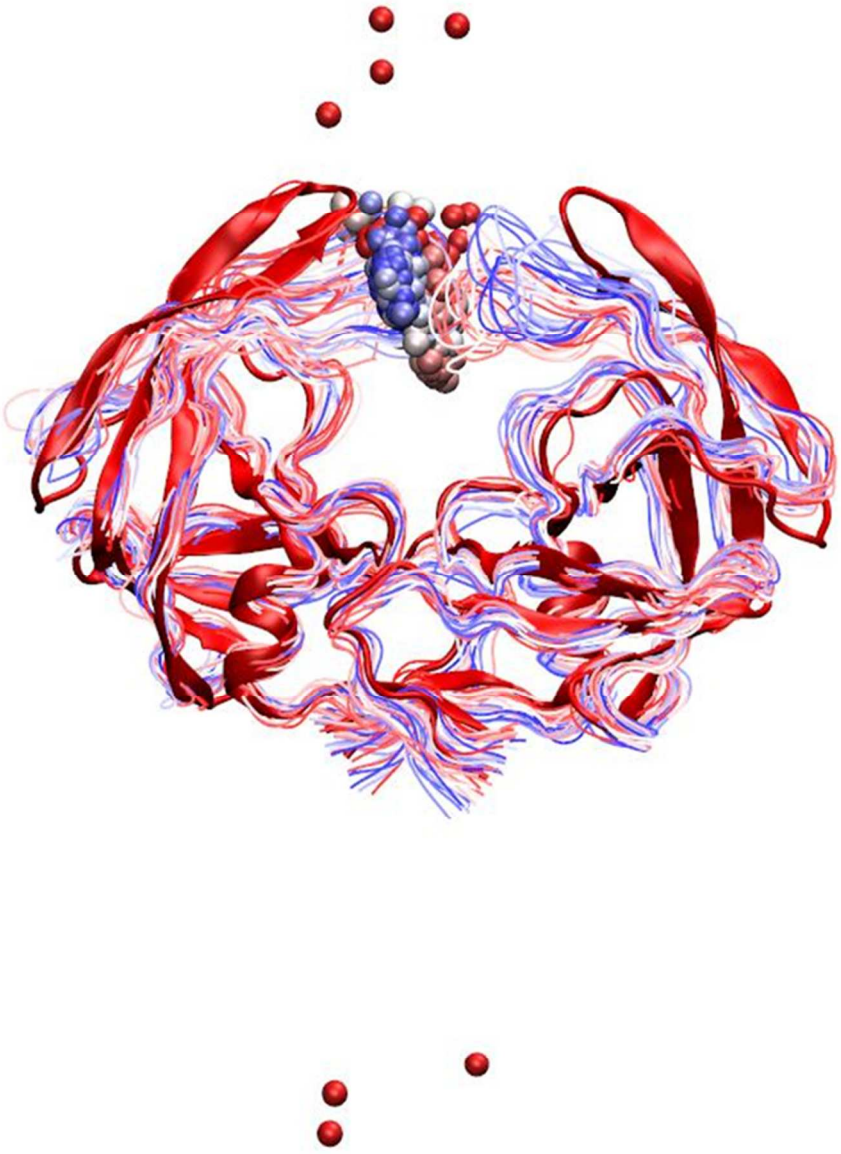


Figure S9: Our MD simulation shows that xk263 is likely to be trapped between the two flaps. The beads represent the center mass of xk263, outlining the binding pathway of xk263 during a 500 ns simulation. The color of beads and tubes change from red, white to blue, which corresponds to the simulation time from 0, 250 and 500 ns, respectively. Similarly, the dominant red color for the cartoon representation of HIVp corresponds to the initial conformation of the enzyme. The white and blue colors of HIVp represents the conformations of the protein near the middle and final time steps of the simulation, respectively.