Supporting Information

for "Ligand Binding Pathways and Conformational Transitions of the HIV Protease"

Yinglong Miao[1,*], Yu-ming M. Huang[2,*], Ross C. Walker[3,4,a], J. Andrew McCammon[2,4], and Chia-en A. Chang[5]

[1]Center for Computational Biology and Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66047; [2]Department of Pharmacology, University of California at San Diego, La Jolla, CA 92093; [3]San Diego Supercomputer Center, University of California at San Diego, La Jolla, CA 92093; [4]Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, CA 92093; [5]Department of Chemistry, University of California, Riverside, Riverside, CA 92521.

[a]Present address: GlaxoSmithKline, Collegeville, PA 19426-0989.

*Correspondence emails: miao@ku.edu and yuh155@ucsd.edu

A

PDB: 1HVR

B

Sim3

Flap
Ligand

RMSD (Å)

Time (ns)

C

PDB: 1HVR

D

Sim4

Flap
Ligand

RMSD (Å)

Time (ns)

E

PDB: 1HVR

F

Sim5

Flap
Ligand

RMSD (Å)

Time (ns)

G

PDB: 1HVR

H

Sim6

Flap
Ligand

RMSD (Å)

Time (ns)
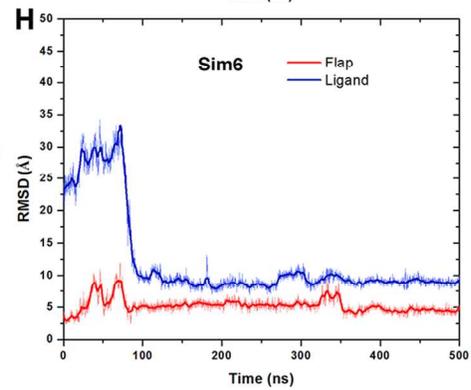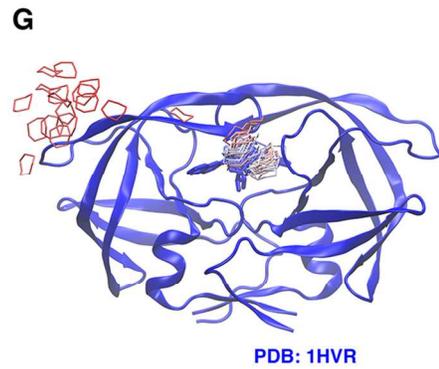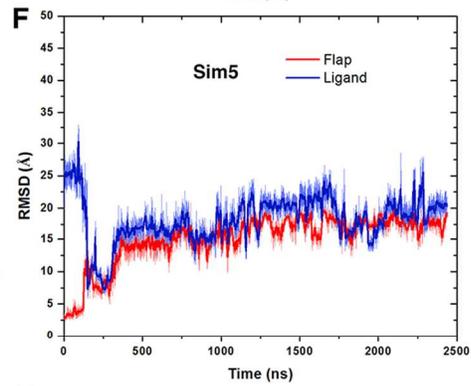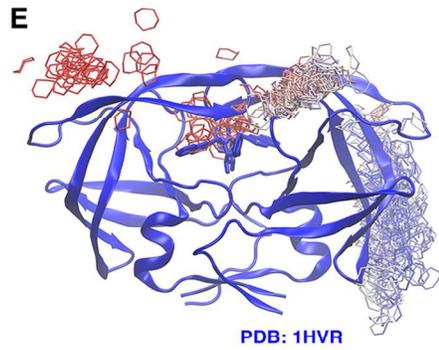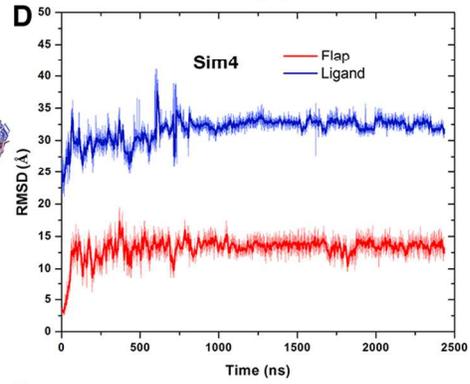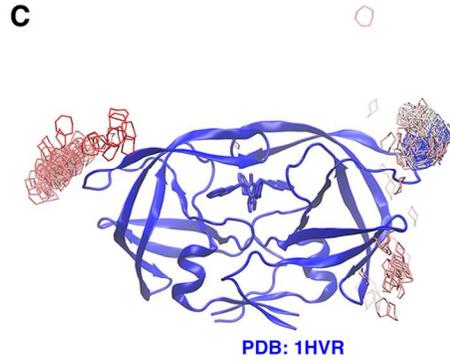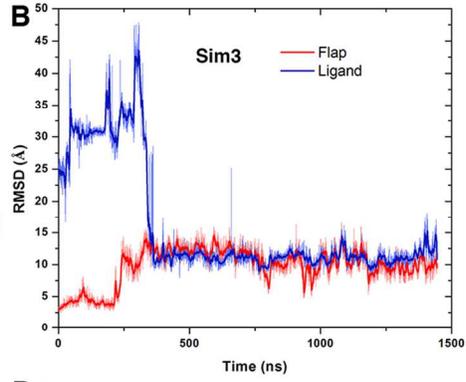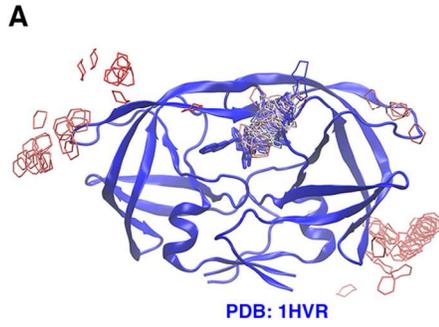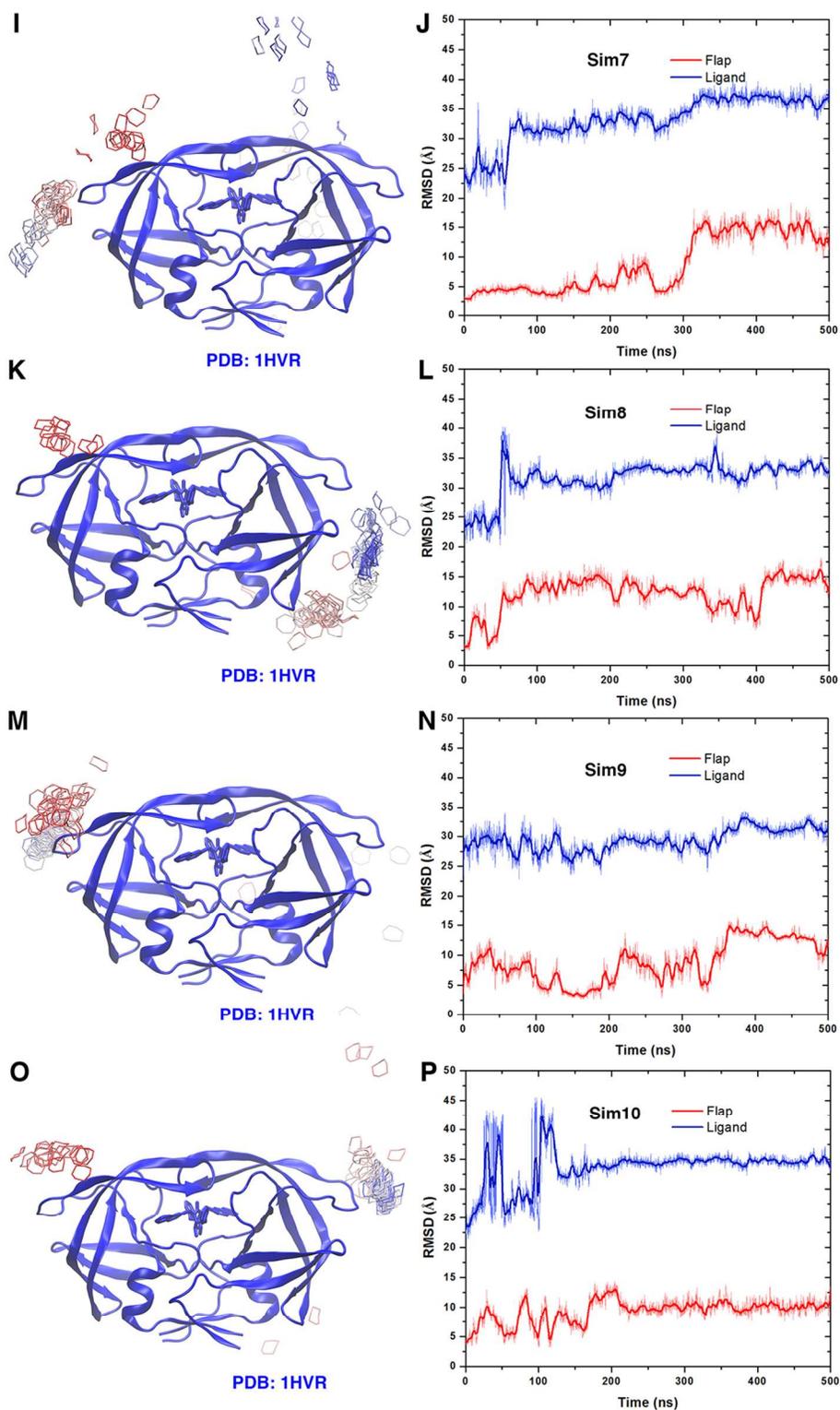
Fig. S1 Pathways of the XK263 ligand molecule in GaMD simulations of the HIV protease: (A) "Sim3", (C) "Sim4", (E) "Sim5", (G) "Sim6", (I) "Sim7", (K) "Sim8", (M) "Sim9" and (O) "Sim10". The core ring of XK263 is represented by lines and colored by simulation time in a BWR color scale. The 1HVR X-ray conformation is shown in blue

for reference. RMSDs of the ligand molecule and protein flaps relative to the 1HVR X-ray conformation are calculated for (B) "Sim3", (D) "Sim4", (F) "Sim5", (G) "Sim6", (J) "Sim7", (L) "Sim8", (N) "Sim9" and (P) "Sim10". Ligand pathways and RMSDs of the ligand molecule and protein flaps are presented in Figure 1 for the other two GaMD simulations "Sim1" and "Sim2", during which the ligand minimum RMSD is smaller than 3 Å (Table 1).
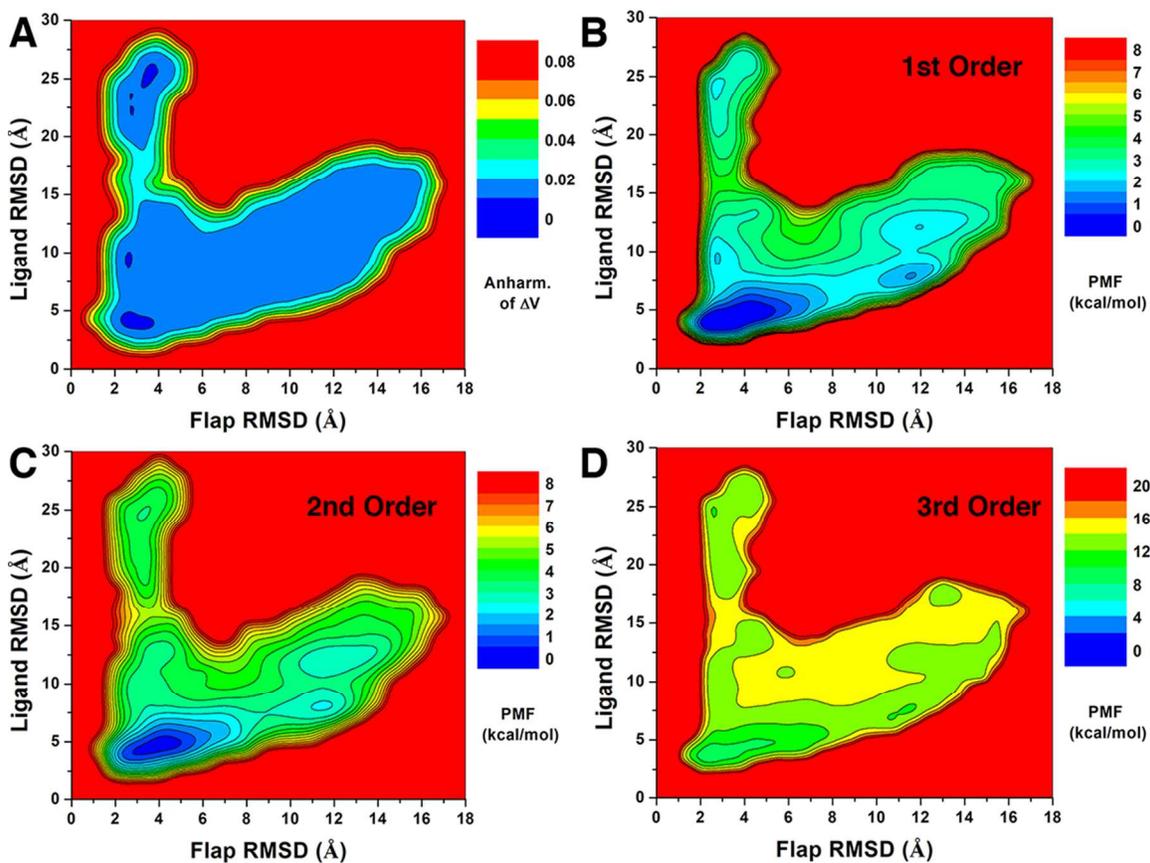
Fig. S2 (A) The distribution anharmonicity of ΔV of frames found in each bin of the 2D PMF profile shown in **Fig. 2A**. (B-D) 2D PMF profiles of the protein flaps and ligand molecule RMSDs combining "Sim1" and "Sim2" GaMD simulations with cumulant expansion to the (B) 1st (C) 2nd (**Fig. 2A**) and (D) 3rd orders.
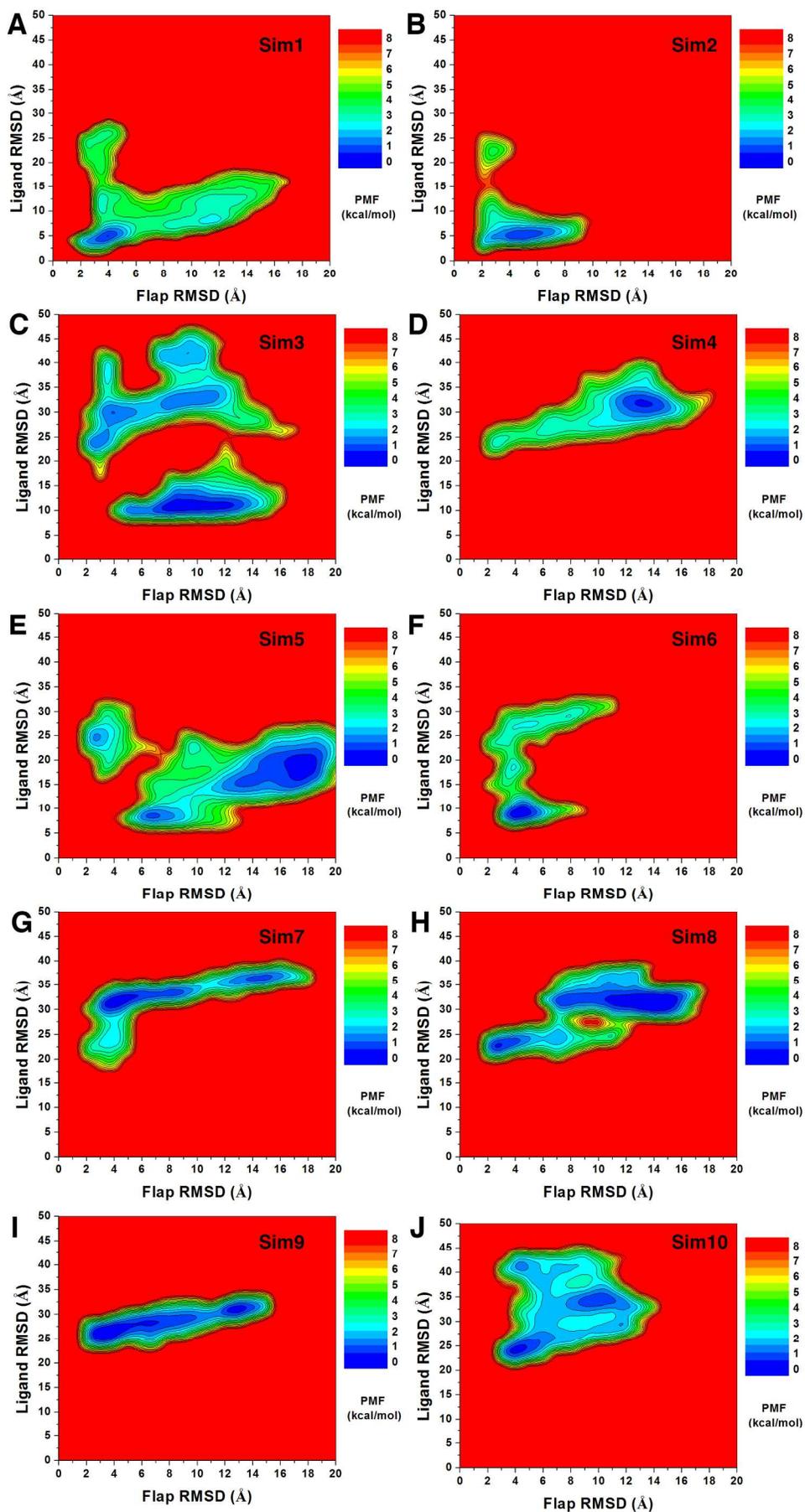
Fig. S3 2D PMF profiles calculated with the protein flaps and ligand molecule RMSDs for ten individual GaMD simulations.
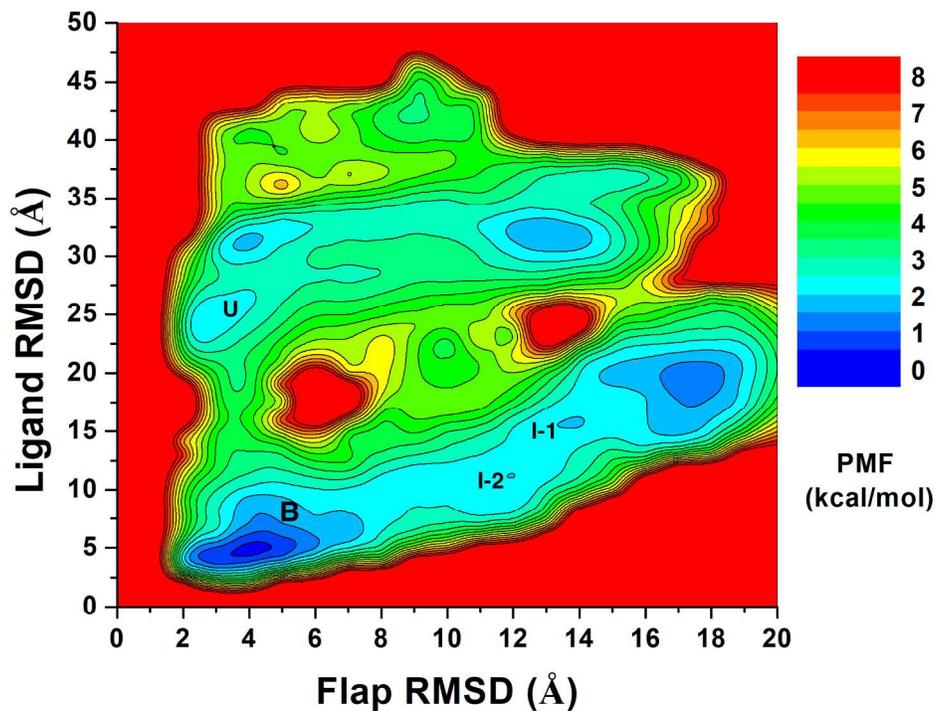
Fig. S4 2D potential of mean force (PMF) calculated with the protein flaps and ligand molecule RMSDs by combining all ten GaMD simulation trajectories.
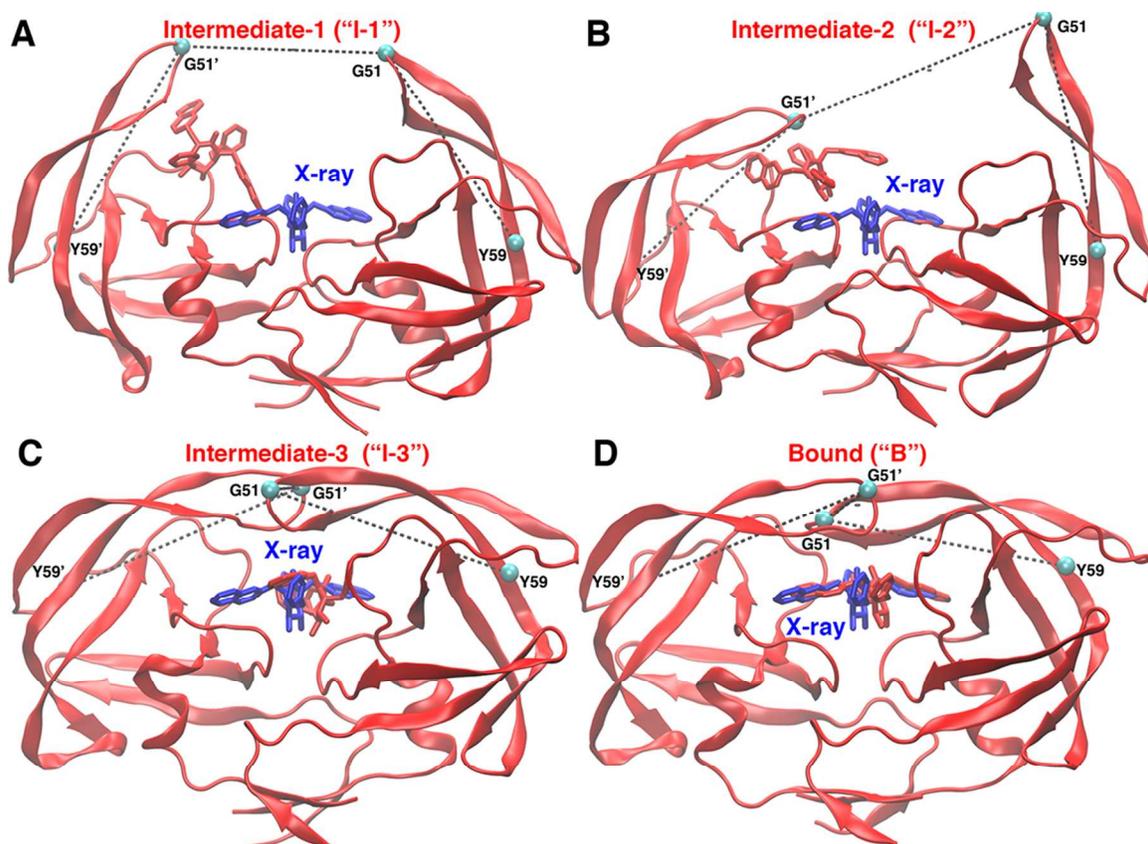
Fig. S5 Front views of the flap handedness, characterized by the dihedral angle of the C_α atoms of residues Tyr59-Gly51-Gly51'-Tyr59', in different conformational states of the HIV protease: (A) Intermediate-1 ("I-1"), (B) Intermediate-2 ("I-2"), (C) Intermediate-3 ("I-3") and (D) Bound ("B") states. The evolving protein (ribbons) and ligand molecule (sticks) are shown in red and the X-ray conformation of bound ligand molecule is shown in blue. The top views of these protein conformations are shown in Fig. 3.
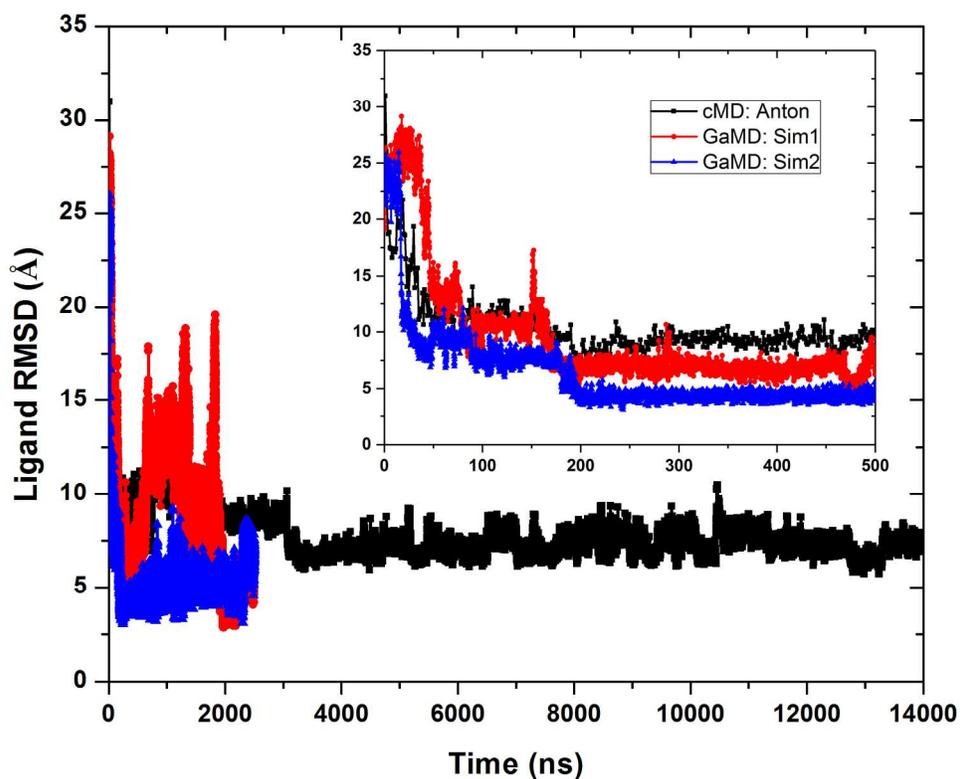
Fig. S6 Comparison of ligand RMSDs relative to the 1HVR X-ray crystal conformation obtained from ~14,000 ns Anton MD simulation and two 2500 ns GaMD simulations ("Sim1" and "Sim2" listed in Table 1).

Movie S1 Binding of the XK263 ligand molecule to the active site of the HIV protease was observed in one of the 2500 *ns* GaMD simulations ("Sim1" in Table 1). Starting from diffusion in the solvent, the ligand molecule attaches to one of the two protein flaps, enters the binding site quickly and induces the two flaps to open. Then the two flaps close back, rearrange their conformations to the closed state and lock the ligand molecule in the active site.

Movie S2 Binding of the XK263 ligand molecule to the active site of the HIV protease was observed in another 2500 *ns* GaMD simulations ("Sim2" in Table 1). A similar pathway was observed compared with "Sim1" (**Movie S1**) except that the protein flaps open for significantly shorter time at a smaller magnitude before closing back to lock the ligand molecule in the active site.