Chapter 4

Binding to the Open Conformation of HIV-1 Protease

4.1 Introduction

Böttcher *et al*. recently reported an interesting crystal structure demonstrating the simultaneous binding of two symmetric pyrrolidine diester inhibitors to the open-flap conformation of HIVp (PDB ID 3BC4).[241] One is bound bridging the traditional active site and the "eye" site[103], while the other is bound between the flaps (Figure 4-1a). The bridged binding pose places two napthyl rings into each eye site (Figure 4-1b), lending support to the possibility of targeting this for inhibitor design. Although the authors did not propose that this was necessarily a 2:1 complex in solution phase, the crystal structure merits investigation.

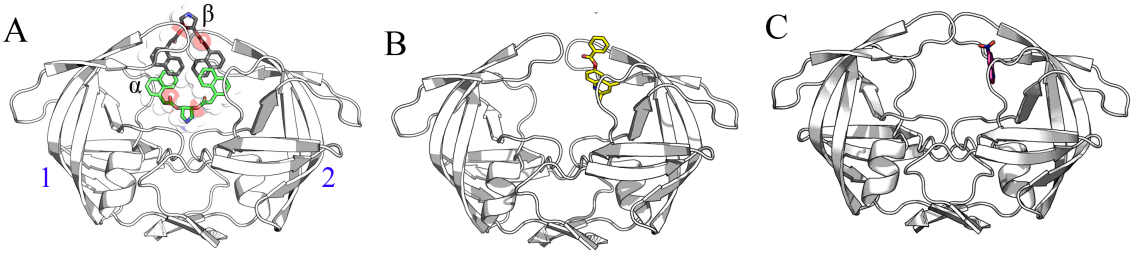

Figure 4-1: A) The crystallized HIV-1 protease uniquely bound by two identical inhibitors, with pose α colored in grey and pose β in black. B) The crystal structure 3BC4 with Damm compound 1[103] (black) bound at the eye site. C) The 5-Nitroindole fragment (black) crystallized in the eye site by Perryman et al.[242] D) A 2-dimensional representation of the pyrrolidine inhibitor that was co-crystallized with 3BC4. For the following figures, we have used a convention of orienting the complex so that a naphthyl occupies the eye position on the right (ie. Monomer 2). We are labeling the monomers as "1" and "2" instead of "A" and "B" to avoid confusion with the α and β notation for the ligands.

We originally proposed the eye site as a possible new mode of HIVp inhibition.[103] Our interest in designing compounds to target the eye site motivated us to study the

conformational states occupied by this receptor-ligand complex in solution. Our goal is to explore the effect of these inhibitors on the conformation of the flaps. It is important to determine whether the conformation seen in the crystal structure is also seen in simulation when symmetry-related contacts of the crystal structure are absent. Retaining the notation used by Böttcher *et al*., we examined the two inhibitors bound to HIVp from the crystal structure (αβ), as well as a single inhibitor bound bridging the active site (α) or in the alternate position against the flaps (β). We hypothesize that the β pose is not stable, nor is the ternary crystal complex, given the poor contacts available to β without the influence of crystal contacts.

4.2 Methods

Our molecular dynamics protocol was based on work by Meagher et al.[243] Our simulations were based on the protein structure crystallized by Böttcher *et al*. (PDB ID 3BC4). PyMOL[237] was used to propagate the asymmetric unit cell. A combination of PyMOL and MolProbity[230] was used to check/flip protonation states while MOE[244] was used to modify the number of bound ligands. The catalytic aspartic acids were both deprotonated, as is appropriate in the presence of the positively charged ligand.

For each inhibitor-binding state (αβ, α-only, or β-only), eight independent, explicit-solvent simulations were performed. Parameters for the inhibitor were generated in antechamber with the Gaff force field[245] and AM1-BCC charges[246]. Hydrogens were built in the *tleap* module of AMBER[245]. TIP3P waters[215] were added as an orthogonal box with a 12 Å buffer to solvate the system. APBS-1.0.0[247] via the plugin for PyMOL[248] was used to calculate an electrostatic surface 10 Å from the vdw surface of the protein, and chloride ions were placed at the most electropositive regions to neutralize the +4*e* charge of the protein and the +1*e* charge of each ligand. MD was performed in the sander module of AMBER using FF99SB[7] and a timestep of 2fs. A non-bonded cutoff was applied at 10 Å. Particle Mesh Ewald[232] was implemented, and bonds to hydrogen were constrained with SHAKE[4]. Water was equilibrated prior to complete system equilibration to prevent protein collapse.[243]

Following minimization of hydrogens, then side chains, then the full system, equilibration was performed with a gradual removal of backbone restraints to achieve a stable trajectory. Over 500 ps, the protein-ligand-solvent system was gradually heated from 10K to 310K, and backbone restraints were gradually softened from 2.0 kcal/mol*Å to 0.1 kcal/mol*Å in a total of five steps. After a two ns equilibration of the protein with no restraints, the production phase lasted 25 ns. This resulted in a total of 8 individual simulations of 25 ns for each of the three sets of bound systems. Therefore, 200 ns of total production time was collected for the HIVp+α complex, HIVp+β complex, and ternary HIVp+αβ complex.

Trajectories were analyzed using the AMBERTOOLS package. *Ptraj* allows for clustering simulations to determine the most prevalent conformations sampled within a specified time period.[249] The trajectories were centered and aligned to the core of the protease (residues 1-45,55-99,1'-45',55'-99'). The last 5ns from each simulation for each complex were then sieved and clustered together with reference to the initial structure. Several clustering protocols were performed to determine the optimal algorithm and family size based on measures of the Davies-Bouldin index (DBI)[250], pseudo F-statistic (pSF)[251], and percentage of variance (SSR/SST). Clustering the simulations into ten families based on the average-linkage algorithm was judged to give the best performance. *Ptraj* was also used to evaluate the degree of flap opening, flap curling, ligand placement, and protein stability.

4.3 Results and Discussion

The existence of the eye site is supported by the recent crystal structure from Klebe and coworkers.[241] The naphthyl rings of the two inhibitors crystallized into each eye site, confining the flaps to the semi-open conformation. Due to the implications of crystal packing effects, we performed MD simulations to determine the conformational behavior of both 1:1 complexes and the 2:1 complex in solution. Of interest in our study was the stability of the different potential complexes, the impact of the inhibitor(s) on flap conformation, and the potential for selective binding at the eye site region.

The impact of bound inhibitors at positions α, β, and αβ was examined over a series of eight unique simulations for each ligand pose. Simulations of α-only and β-only were examined as representatives of possible 1:1 complexes as compared to the 2:1 HIVp+αβ complex. The impact of inhibitor binding on backbone stability over the course of each simulation was measured by determining the Cα RMSD of the core residues from the initial crystal structure (Figure 4-2). For the three different systems, the RMSD of the protease core is 1.65±0.28 Å (HIVp+αβ), 2.25±0.35 Å (HIVp+β), and 1.81±0.25 Å (HIVp+α), signifying core stability. The flaps were mobile as expected (Figure 4-3, 4-4, 4-5).

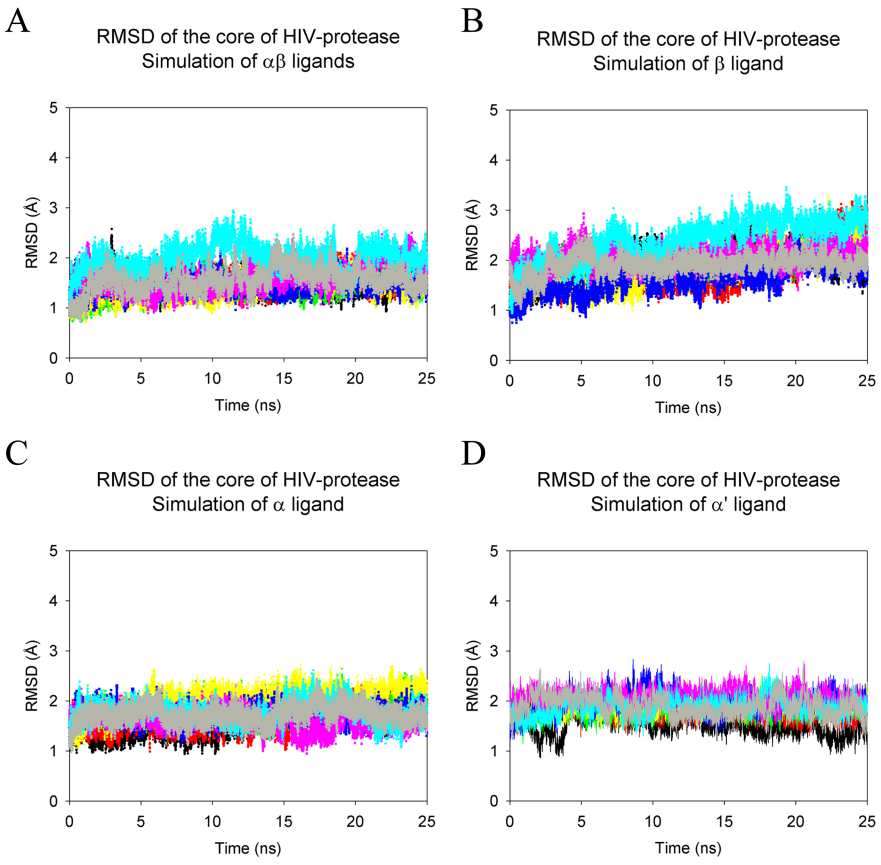

Figure 4-2: The overall RMSD calculated for the core (all residues but the flaps 43-58/43'-58') of HIVp over the length of the production run. The core remains stable for the duration of the trajecories for each protein-ligand complex αβ (A), β (B), α (C), α' (D).

There are several different ways to measure flap opening in HIVp. One of the most common metrics is the distance from the base of the active site (Asp25/25') to the flap tip

(Ile50/50'). The larger the distance, the more open the flaps are considered. A distance of approximately 14 Å is considered closed, while a distance of approximately 18 Å is considered semi-open. An additional metric commonly utilized to measure the extent of flap opening includes the flap tip distance (Ile50-Ile50'). However, due to the nature of a three-dimensional system, this distance measure is heavily influenced by flap curling, which does not necessarily indicate a change in flap conformation. Alternatively, we examined the distance between the Cα atoms' center of mass for the flap residues 48-53/48'-53'. In addition, it was possible to measure the distance from the flap tip to the 80s loop (Ile50-Thr81). This distance can give some insight into the handedness of the flap conformation. A distance of approximately 10 Å is seen in the semi-open structure of 1HHP, while a distance of 15 Å is seen for the bound conformation of 3BC4 as well as the wide-open structure 1TW7. Thus, a distance below 10 Å may indicate a semi-open handedness, while a distance higher than 15 Å may be indicative of a closed handedness.

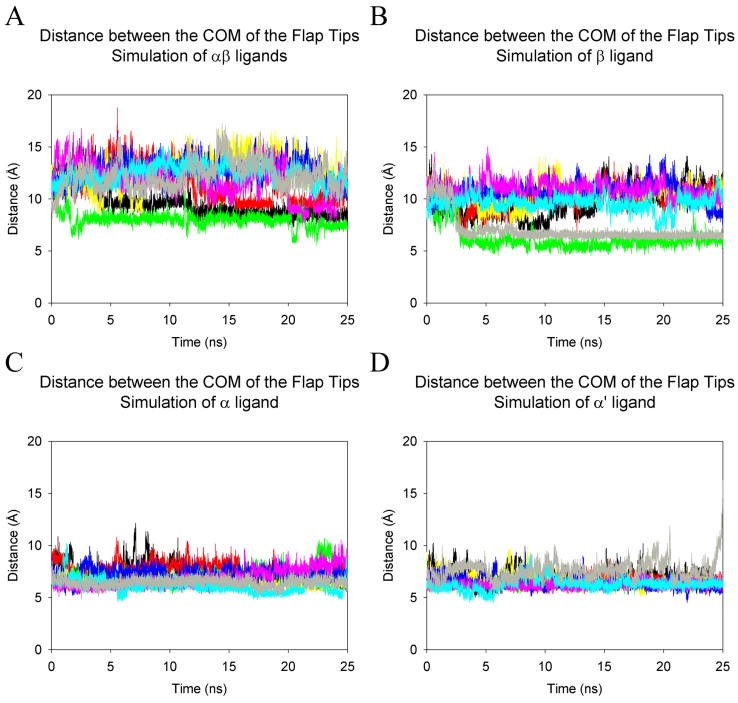

Figure 4-3: The distance calculated between the center of mass (COM) of the two flaps 43-58/43'-58' over the length of the production run. The flaps are less stable for the duration of the trajectories for complexes αβ (A), β (B), as compared to α (C), α' (D).

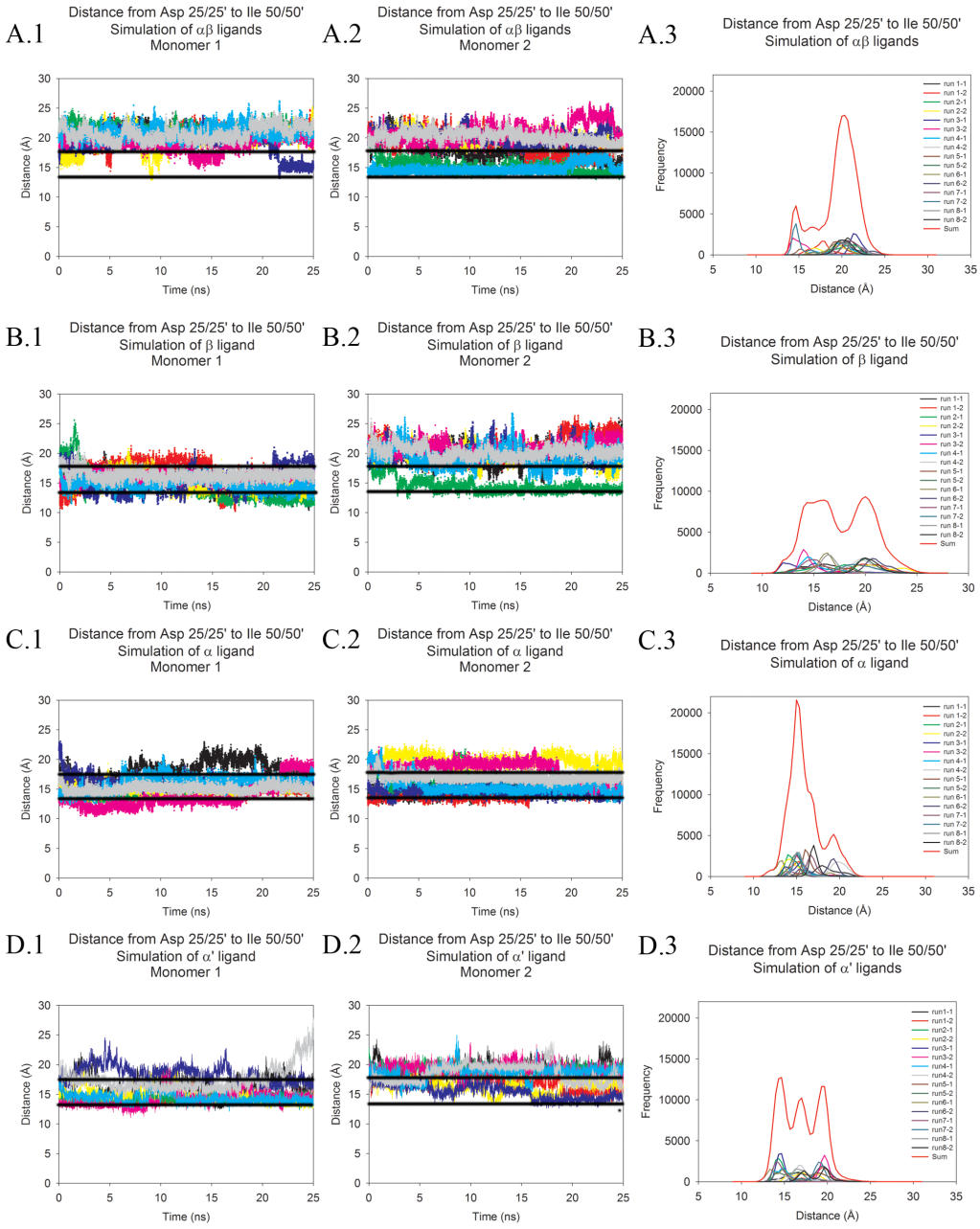

Figure 4-4: The distance between the catalytic Asp25 and the flap tip Ile50 for monomers 1 and 2 of HIVp over the length of the production runs. The flap-to-active-site distances indicate a wide range of motion for both monomers (1-2), with one majority flap conformation seen for simulations of αβ (A) and α (C), while β (B) and α' (D) are characterized by several equally population flap conformations. The typical closed (14 Å) and semi-open (18 Å) distances are indicated with a black line.

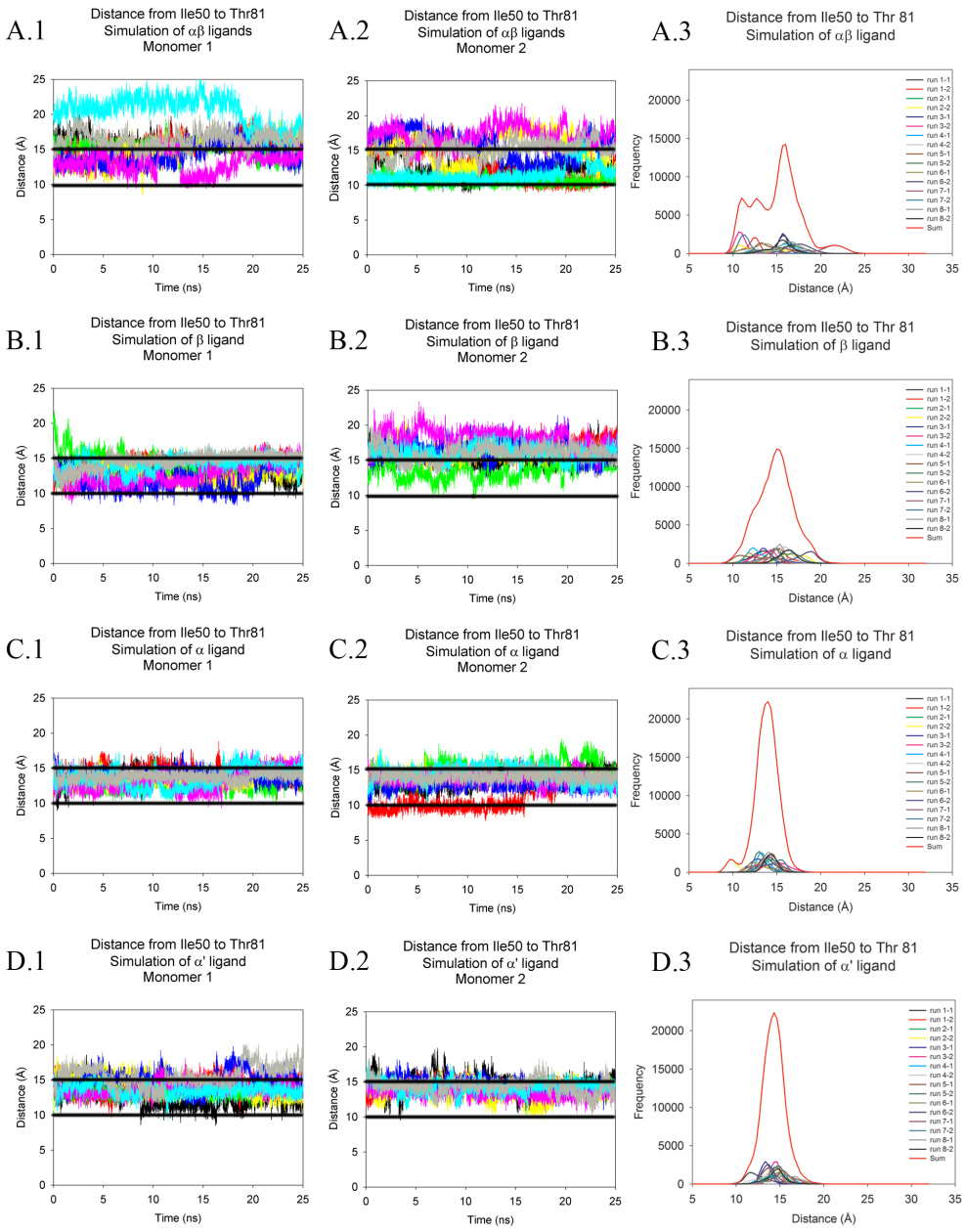

Figure 4-5: The distance between the flap tip Cα to the 80s loop Cα for monomers 1 and 2 over the length of the production run. The distance fluctuates mainly between 10-20 Å throughout our simulations, but on average is much higher for the trajecories of αβ (A) and β (B), while α (C) and α' (D) occupy a narrower range of distances. The typical closed (15 Å) and semi-open (10 Å) distances are indicated with a black line.

The impact of the ligands on the conformational ensemble was examined through clustering of the conformations sampled over the simulation trajectory. The atomic fluctuations of the protein for the complete simulation were calculated with *ptraj* to

define the stable core residues. The trajectory was imaged and rms-fit to the protease backbone, and then *ptraj* was used to cluster the heavy atoms of the stable protease core and the ligand over the last 5ns of each 25ns trajectory. Clustering the final 10ns of the trajectory together required too much system memory; however, conformations observed from the last 10ns of individual runs were in agreement with clustering over all runs. A total of 10 families were generated using the average-linkage algorithm. This was accomplished for all simulations of the HIVp+αβ complex, the HIVp+α complex, and the HIVp+β complex. The representative structures were then examined to determine the similarity between binding modes among the families. In addition, calculations were performed to assess the stability of the bound ligands over time (Figure 4-6,7). The stability of the bound ligands was judged by several metrics. The simplest stability measure is the root-mean-square deviation of the ligand from its crystallographic position over time. This was calculated for each independent simulation. In addition, the placement of the pyrrole moiety of the ligand within the active site was examined based on the distance between the nitrogen of the pyrrole to the center of mass of the carboxylate carbon of the catalytic aspartic acids. The RMSD of each ligand was calculated against the average ligand position as well as the crystallographic pose. All values were calculated for the final 5ns of each trajectory.

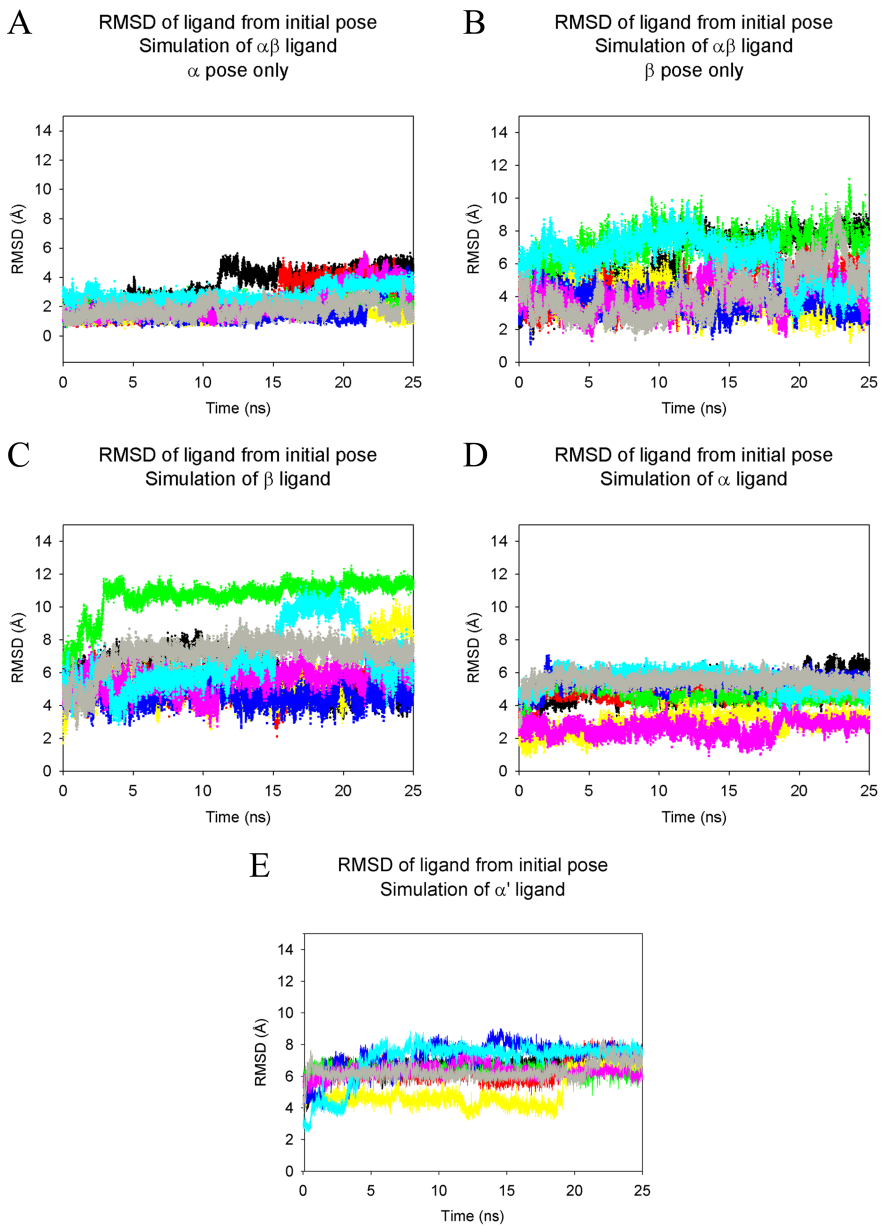

Figure 4-6: The overall RMSD calculated for each ligand within its protein-ligand complex, yielding the RMSD for α-only in HIVp+αβ (A), β-only in HIVp+αβ (B), β in HIVp+β (C), α in HIVp+α (D), and α' in HIVp+α' (E) over the length of the production run. Trajectories were first fit to the Cα core of the 3BC4 crystal structure.

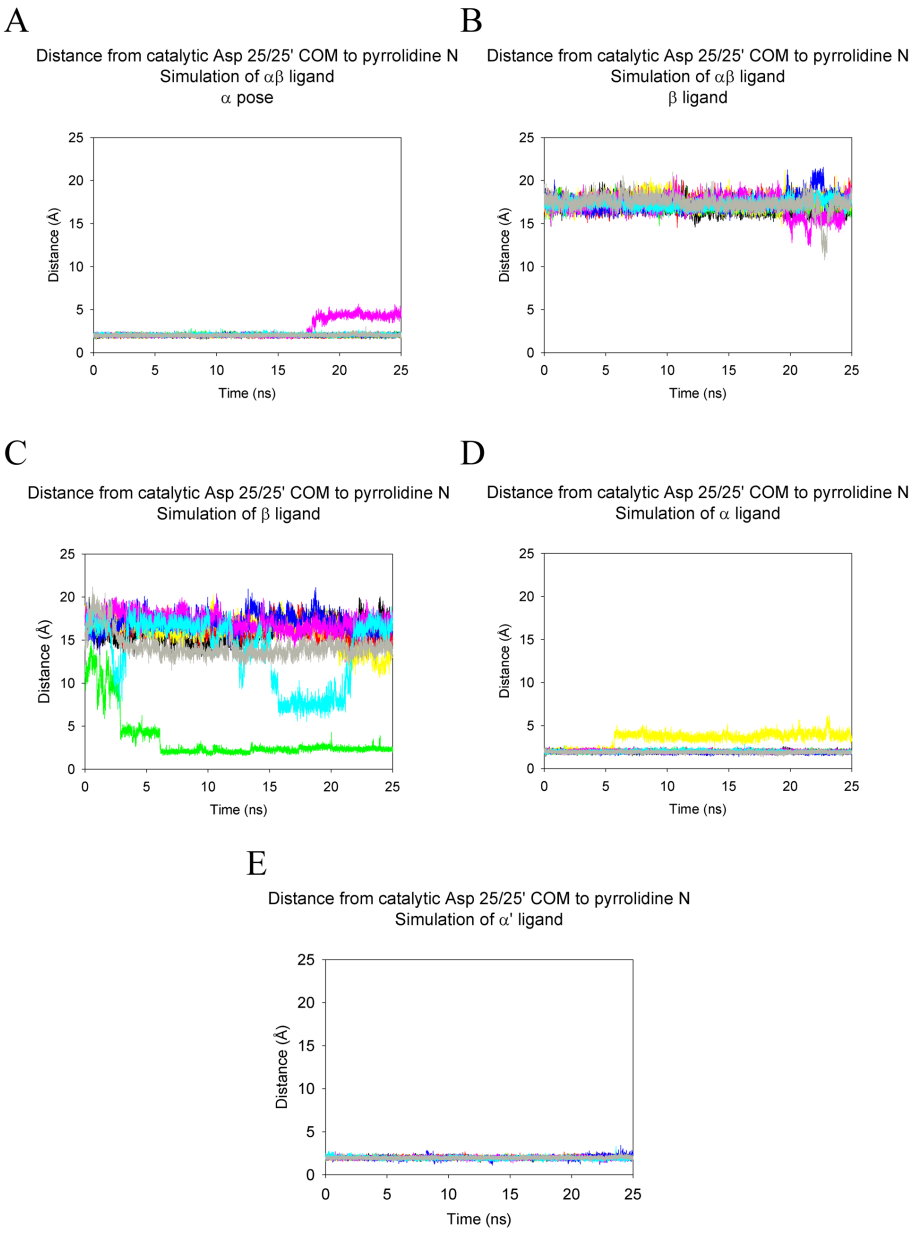

Figure 4-7: The distance from the center of mass (COM) of the catalytic aspartic acids 25/25' to the pyrrole nitrogen on the ligand, for α-only in HIVp+αβ (A), β-only in HIVp+αβ (B), β in HIVp+β (C), α in HIVp+α (D), and α' in HIVp+α' (E). The green line in C clearly depicts the run in which the ligand flips into the active site (the α position), as well as the run with a partial flip.

4.3.1. HIVp+αβ.

The 2:1 complex of HIVp+αβ is unstable. The representative conformations illustrate a wide range of motion sampled by the β ligand, and a moderate amount of sampling by the α ligand. Compared to the average ligand position, the RMSD of the α

ligand over the last 5ns ranged from 1.04 to 15.41 Å and β ligand ranged from 1.07 to 28.28 Å. The large fluctuation in position of the ligands illustrates the instability of the HIVp+αβ complex and hints at the instability of the binary HIVp+β complex as well. Analysis of the conformations present in representative families show that in our HIVp+αβ simulations, the α ligand has one naphthyl that occupies the eye site and the other naphthyl occupies the S1 or S2 pocket in approximately 62% of the trajectory sampled (Figure 4-8). The β ligand in these simulations is quite varied in position; contacts are typically maintained between the pyrrolidine amino group and flap tips or solution.

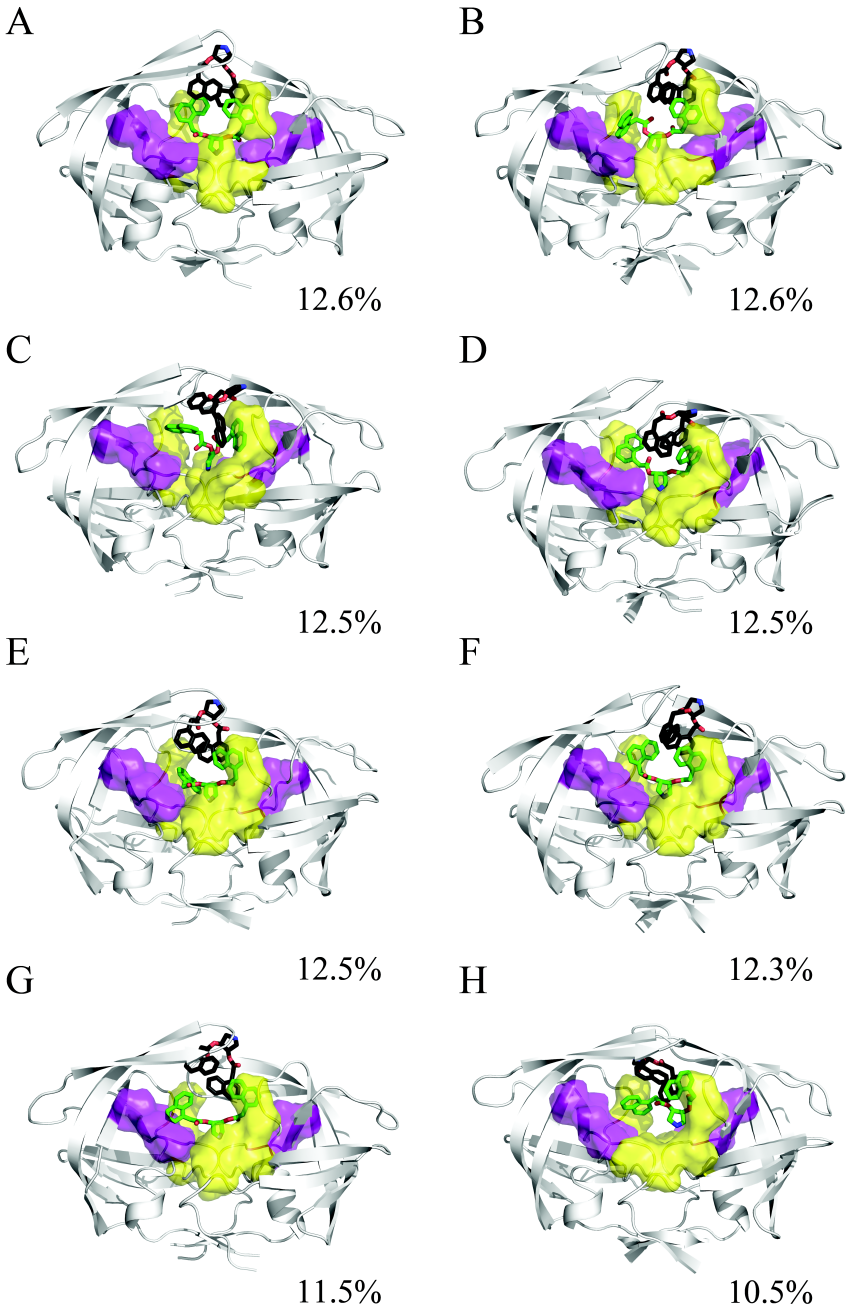

Figure 4-8: Representative structures from the MD simulation of the HIVp+αβ complex, taken from the last 5ns of each 25ns trajectory. The α ligand is shown in green, the β ligand is shown in black, the S1/S1' site is shown in yellow, and the S2/S2' site is shown in purple. The conformational families demonstrate the instability of the 2:1 bound complex. The pyrrolidine ligands find a wide variety of ways to interact with the protease flaps, S1/S1', S2/S2', and/or the eye site.

Since our interest lay in understanding the impact of these inhibitors on flap conformation, we quantified flap motion over time. A common standard for evaluating flap conformation is the distance between the flap tips (Ile50/50') and the catalytic

aspartic acids (Asp25/25'). A typical distance for the closed flap form, based on the crystal structure 1PRO, is 14.1 Å. A typical distance for the semi-open form, based on the crystal structure 1HHP, is 17.8 Å. Over the last 5ns, the HIVp+αβ simulations sampled a median distance of 20.81 Å and 17.18 Å for monomers 1 and 2 respectively, implying an asymmetric, semi-open flap conformation for the duration of production time. The flap RMSD for the 2:1 complex demonstrates a considerable range of motion, with a mean of 3.71±0.98 Å relative to the crystal structure.

As expected, the ternary complex of HIVp+αβ was not stable. Klebe and co-authors did calculate their affinity data using a 1:1 ratio for kinetics in their experimental work. In a 1:1 ratio, the bound state is expected to be more similar to other holo crystal structures than the solved 3BC4 crystal structure. Klebe and co-authors also concluded their paper with a discussion of the possible influence of crystal packing on the crystallographic conformation and its potential instability in solution. In the crystal structure itself, it appears that the majority of symmetry-packing contacts are formed through ligand-ligand stacking. It may be that these contacts stabilize the observed bound handedness while requiring a semi-open flap elevation.

4.3.2. HIVp+β.

Simulations of the β-only complex reveal a wide variation in the population ensemble. None of the representative structures occupy the same conformation as the crystallographic ligand position (Figure 4-9). Over the last 5ns, the range in ligand deviation from the average pose is 1.73 to 7.56 Å, which shows that the β ligand alone is unstable, and this instability is reflected in the diversity of populated states from these simulations. While the β ligand samples widely, it always has a naphthyl ring in the eye site. The other moieties on the ligand are primarily involved in forming hydrophobic interactions with residues in the flap region. The presence of the β ligand skews the flaps asymmetrically, which is consistent with our previous simulation and proposed behavior of the eye site.[11]

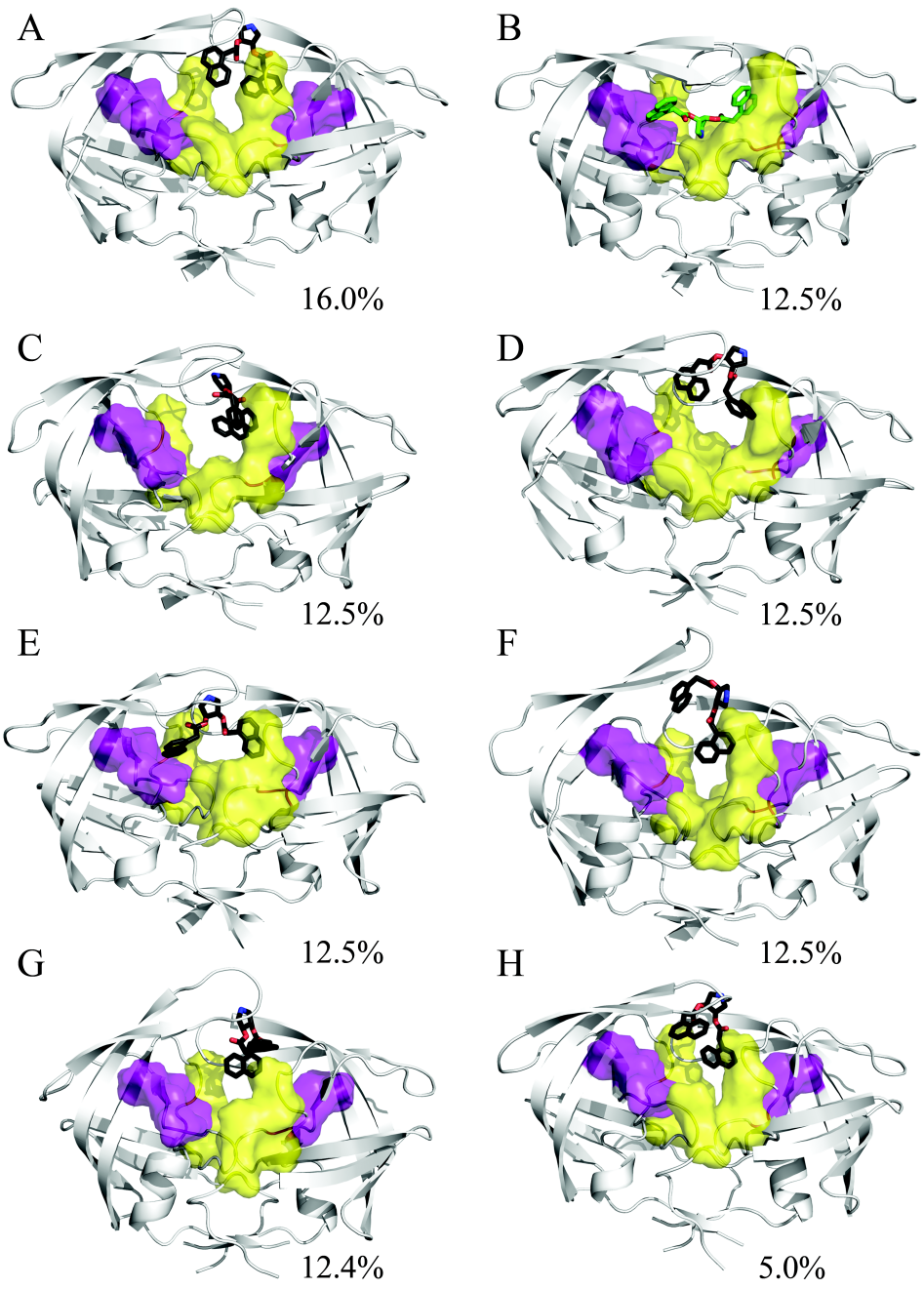

Figure 4-9: Representative structures from the MD simulation of the HIVp+β complex, taken from the last 5ns of each 25ns trajectory. The β ligand is shown in black, the S1/S1' site is shown in yellow, and the S2/S2' site is shown in purple. The conformational variance of the β ligand is exceptionally high. The β ligand does not preferentially interact with the binding site or eye site, except in the case of (B). This family is the result of one run, where the ligand has moved to bind at the active site, displaying a standard α pose, with contact in one eye site and the S2 pocket.

Our simulations show that when bound alone, the pyrrolidine ligand is not likely to bind in the crystallographic β pose. The system is not stably bound, and the ligand moves

to more favorable conformations. This can be seen from both the representative structures of the HIVp+β complex as well as the high RMSD of the ligand from the average conformation.

During the β simulations, the flaps display semi-open to open behavior with averages over the last 5ns for flap tip to catalytic aspartic acid distance of 20.39 Å and 15.43 Å for monomers 1 and 2. The more open flap of monomer 1 is explained by the presence of the β ligand its flap recognition pocket, which prevents traditional flap dynamics. The ligand interactions at the eye site skew the flaps of the protease into an asymmetric conformation. Over the last 5ns, the average RMSD of the flap residues to the crystal position for the β-only simulation is 4.73±1.06 Å, illustrating similar motion to that sampled by the HIVp+αβ complex.

The β-only complex is far less favorable than α-only; in one of the simulations the ligand flips to occupy the α pose (Figure 4-9b) and in other populated states the ligand samples poses approaching α-only.

4.3.3. HIVp+α.

The α simulations most commonly sample a distance between the flap tip to the active site of 15-16 Å. This signifies that over the course of our simulations the flaps sample conformations in between closed and semi-open. They cannot close completely due to the presence of the ligand in an eye site. The flaps themselves have an average RMSD to the crystal structure of 4.44±0.47 Å, over the last 5ns. The low standard deviation signifies the greater stability of flap conformation throughout these simulations of the α-only complex, relative to the αβ and β-only simulations.

It is interesting that the most frequently sampled conformations in the α-only case illustrate a preference for asymmetric binding at the eye region (Fig. 4-4). We find that one of the naphthyl rings of the ligand often reorients in a manner similar to positions of aromatic rings in known protease inhibitors. One naphthyl ring of the Klebe inhibitor remains stably bound at the eye site, while the second naphthyl ring dissociates from an eye site to occupy either the S1 or S2 site in approximately 75% of the simulation time sampled (Fig. 4-4a,b). In addition to demonstrating a possible need to satisfy interactions