at either the S1 or S2 binding pocket, this also agrees with two recent studies that illustrate a requirement for asymmetric binding to the eye site.[103,242]

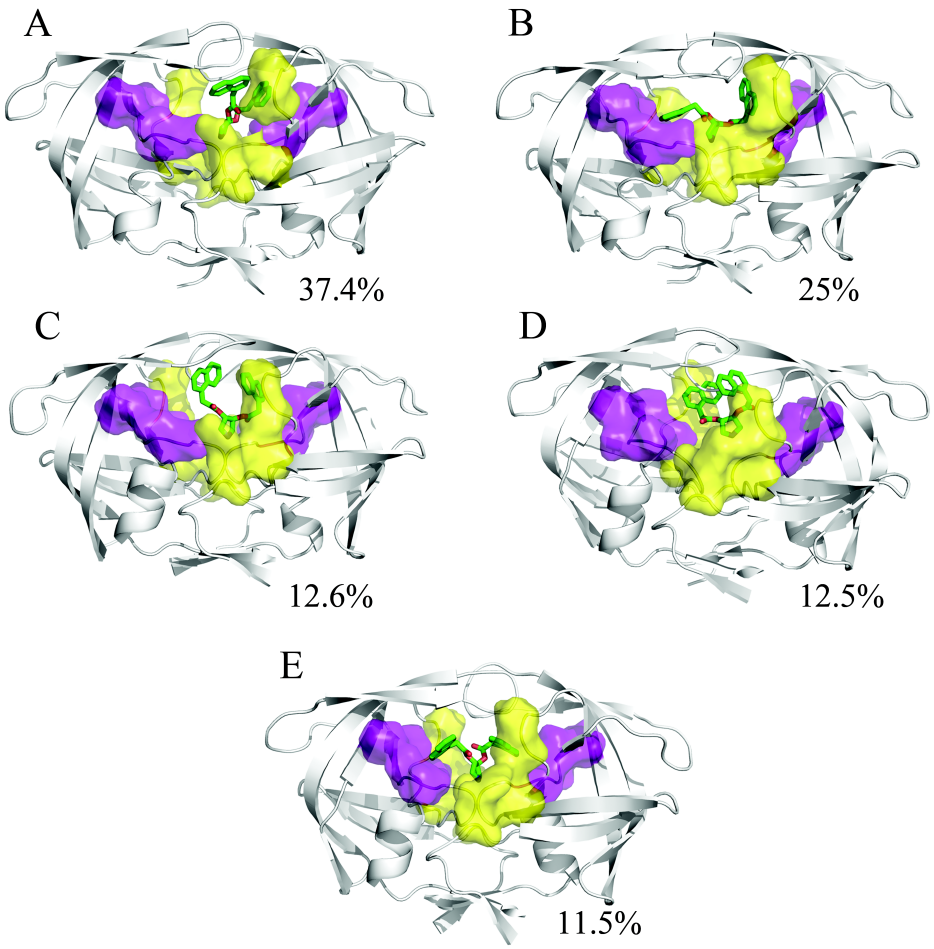

Figure 4-10: Representative structures from the MD simulation of the HIVp+α complex, taken from the last 5ns of each 25ns trajectory. The α ligand is shown in green, the S1/S1' site is shown in yellow, and the S2/S2' site is shown in purple. The conformational families for the α ligand illustrate its strong preference for forming one interaction between the naphthyl ring and the eye site, while the other naphthyl ring flips to interact at the S1/S1' or S2/S2' site, and the pyrrole maintains a hydrogen-bonding interaction with the catalytic aspartic acids.

There is moderate deviation from the average ligand position in simulations of HIVp+α over the last 5ns, with a range of 1.78 to 5.86 Å. The loss of one naphthyl ring from an eye site results in the difference from the crystallographic pose. A low standard deviation of 0.73 Å validates the greater stability of this binding pose. RMSD values for the two naphthyl rings were calculated to better demonstrate alterations in ligand binding over time for the α-only case. RMSD traces were created to detail the ligand's naphthyl

ring position over time. The α ligand was fairly stable over time (Figure 4-10), and the RMSD of the naphthyl rings clearly shows the continued binding of one ring in the eye site (Figure 4-11a) and the absence of binding at the other eye site (Figure 4-11b). In the least populated conformational family, both sides of the ligand do flip down into the traditional binding pocket (Figure 4-10e). This indicates that sufficient sampling has occurred and that this pose is less preferred than asymmetric binding to the eye and S1 or S2 site.

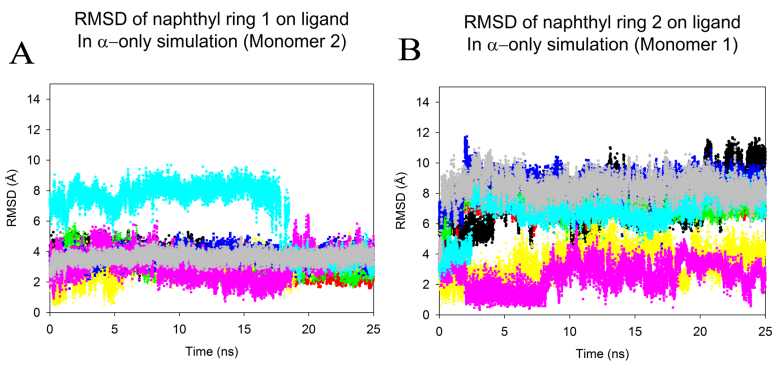

Figure 4-11: The overall RMSD from the crystal pose calculated for each naphthyl ring of the ligand in HIVp+α over the length of the production run. Trajectories were first fit to the Cα core of the 3BC4 crystal structure. Each color represents a single production run; and denotes the same run for each plot. A) highlights the RMSD of the first naphtyhl ring over time, (B) highlights the RMSD of the second naphthyl ring over time. As noted in figure 1, we have used the convention of labeling monomers 1 and 2 based on the behavoir of the ligand, where better agreement with the initial position in the eye is oriented to the right in the figures and labeled as monomer 2 in the graphs. An RMSD of 6.2-7.9 Å indicates occupation of the S2/S2' site, while an RMSD of 7.8-10.1 Å indicates occupation of the S1/S1' site.

4.3.4. HIVp+α'.

Of course, it is possible that we observed incomplete sampling, and the inhibitor could actually prefer to be "extended" with both naphthyl rings occupying S1 and/or S2 sites. To determine this, we conducted a fourth series of MD. This system, HIVp+α', was obtained by modifying the HIVp+α crystal pose to flip both naphthyls into the S2/S2' pockets. HIVp+α' was subjected to hydrogen minimization in the gas phase with AMBER to ensure that the MD simulations commenced from an unstrained system. After this, the set-up and simulation of the HIVp+α' complex followed the previously described protocol for complexes HIVp+α, HIVp+β, and HIVp+αβ. Again, eight independent simulations were conducted, and the last 5 ns of each simulation were examined.

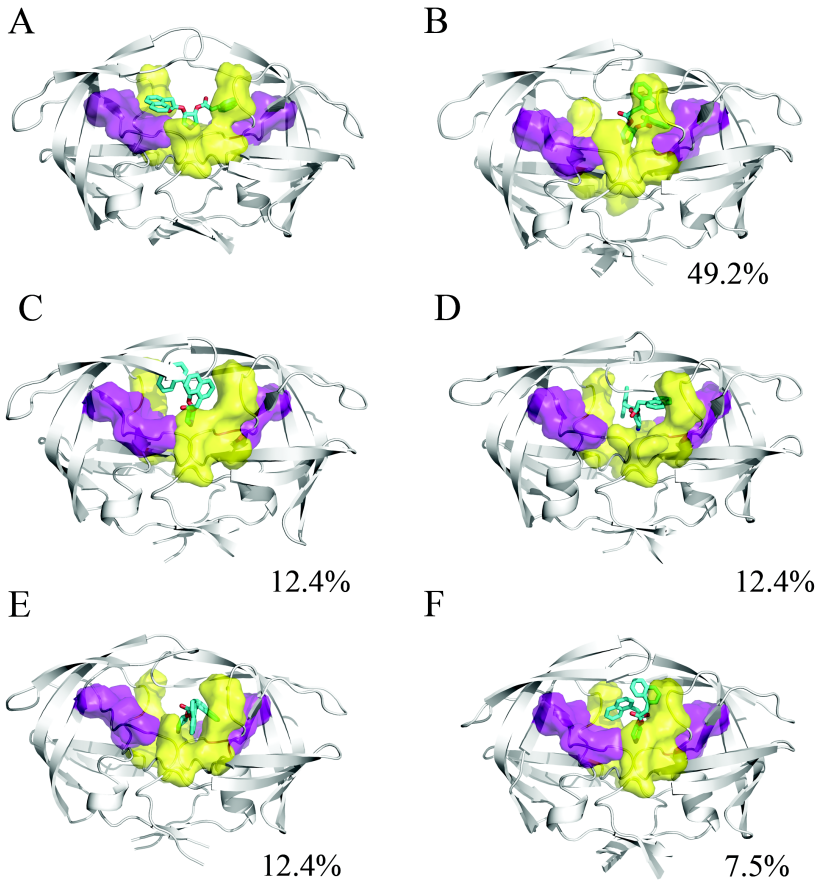

Figure 4-12: A) The initial minimized conformation of the α' ligand. B-F) Representative structures from the 200ns MD simulation of the HIVp-α' complex. The α' ligand is shown in cyan, the S1/S1' site is shown in yellow, and the S2/S2' site is shown in purple. Although the simulations were initiated with the naphthyl rings occupying traditional subsites of the active site, one naphthyl ring moves to form interactions at the eye site over the course of all eight independent simulations. The second ring remains in contact with the S1 or S2 site.

These additional simulations, beginning with both naphthyl rings interacting at the S2/S2' pockets, resulted in at least one ring altering its position during simulation to interact with the eye region (Figure 4-12). The most populated family type is extremely similar to the most populated families from HIVp+α complex simulations; wherein one ring interacts at the eye region while the other interacts at the S2 pocket. It is possible that the naphthyl rings are positioned one up, one down in solution. NMR data might show whether or not there is symmetry of the two rings' environment in solution. We found that in the ternary complex, even in the presence of the β ligand, one side of α flips down to occupy a similar position to known inhibitors.

It is interesting to note that all of the protease conformations in the clustered families display flaps with the same handedness as the closed state. This signifies that the protease is occupying similar conformational space of the bound form, even though the flaps are typically in the semi-open position (also seen in the close-handed, wide-open structure 1TW7)[252]. However, this does not mean the flaps cannot flip handedness during the simulations, only that we do not observe it in the majority conformations. The flaps do display the type of curling commonly observed during flap transitions. It may require more simulation time for the flaps to flip handedness simply because of the presence of an inhibitor molecule bound in the flap region.

Despite the relative instability of the crystal conformation during MD simulations, the placement of moieties in the eye site intrigued us due to our previous work. We find that αβ and β probably do not exist because of the instability of β and poor contacts available to β. Considerable flexibility in the flap region is observed in simulations with the alternate ligand (αβ and β) as compared to the simulations with α-only. Although these molecules have a unique crystallographic conformation, the structure in solution most likely resembles a conformer similar to the HIVp+α complex. The α pose is far more stable, and it most likely contacts one eye site as well as the traditional active site. Our results provide strong support for further exploration of the eye site as a new mode of inhibition for HIVp.

4.4 Conclusion

Although the original crystal structure of the pyrrolidine inhibitors is unlikely to exist in solution, we were interested in exploring the potential shown by this mode of binding because of its relationship to the eye site. Naphthyl groups are not ideal because of solubility and metabolic issues, but these inhibitors show that we can take advantage of the eye site in inhibitor design. The binding assays performed by Klebe and co-authors[1] show the potential of these compounds for targeting HIVp. Investigating all of the potential bound states of this complex – HIVp+αβ, HIVp+β, HIVp+α, and HIVp+α' – allows for an accurate study of the impact these ligands may have on flap conformation, and therefore, protease activity.

Our study utilized 200ns of simulation time per system to examine the conformational stability of several HIVp-ligand complexes based on a symmetric inhibitor from Klebe and co-authors[10]. Our present results support previous findings that indicate the existence of an alternate binding site for HIVp: the eye site[11]. Furthermore, our data supports a preference of asymmetric binding at the eye site, as previously suggested.[11,26] The representative structures of the HIVp+α and HIVp+α' complexes illustrate that only one eye site tends to be occupied, while the other naphthyl ring prefers binding at the S1 or S2 site. This implies that traditional inhibitors could be modified to take advantage of this interaction and/or targeting the eye site may be improved by including some traditional S1 or S2 contacts. Inhibitors with improved contacts would be an important step towards demonstrating the viability of the eye site as a target for protease inhibition.

This work was published as: Lexa KW and Carlson HA. Binding to the Open Conformation of HIV-1 Protease. *Proteins*. **2011**, *79*, 2282-90.

Chapter 5

Clarifying Allosteric Control of Flap Conformations in the 1TW7 Crystal Structure of HIV-1 Protease

5.1 Introduction

A recent crystal structure of a multidrug-resistant HIVp (1TW7) showed the flaps were wider and more open than in other apo, semi-open structures.[252] Crystal packing created contacts between the flap tips in the neighboring unit cell and the elbow region of HIVp (Figure 5-1). These contacts were proposed to be experimental corroboration of allosteric control. LD simulations by Layten *et al*. showed that the conformation of HIVp seen in 1TW7 relaxed into the typical semi-open conformation in the absence of the crystal contacts.[253] When all packing neighbors within 15 Å of the central dimer were replicated and restrained to their location, the unrestrained central HIVp sampled only the wide-open conformation seen in 1TW7. Layten *et al*. definitively showed that the crystal contacts cause the deformation, but they were astute not to claim that those contacts proved or disproved allosteric modulation of the flaps.

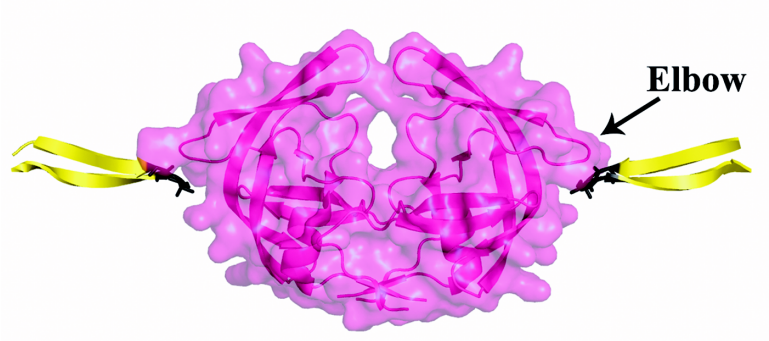

Figure 5-1: The 1TW7 structure of HIVp, showing crystal contacts between neighboring flap tips and the elbow. The flap-tip residues in direct contact with the elbow cleft (residues 49-52) are shown in black.

The tips of the flaps are the regions that make contact with elbow residues in neighboring cells. The structure may be deformed through "pulling" the flap tips into the

next cell, or they may be "pushed" through allosteric contact with the elbow. In order to provide evidence of allosteric control, altered dynamics of the flaps must be demonstrated when the contacts are made solely with the elbows. Furthermore, these contacts should be allowed full conformational freedom.

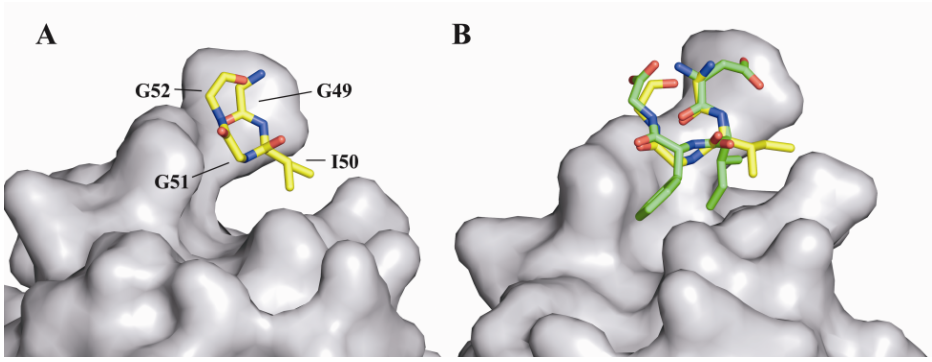

Figure 5-2: (A) The GIGG sequence makes no contacts into the base of the elbow cleft, but (B) using D-Ile50 and D-Phe51 (green) increases the peptide's contact with the sides and bottom of the cleft by ~100 Å².

In this study, we truncated the points of contact to create small peptides associated to the elbow region of dimeric HIVp. The peptides failed to restrict the conformation of the flaps. When the peptides were restrained from dissociating from the elbow, the flaps still sampled the semi-open and open conformations. Even modifying the peptides to create more contact within the cleft failed to improve their control of the flaps (Figure 5-2). When unrestrained, all peptides quickly dissociated from the elbow in multiple simulations, showing that the contact seen in the 1TW7 crystal structure is simply opportunistic crystal packing, not allosteric control. Lastly, experimental testing of short model peptides failed to inhibit HIVp.

5.2 Methods

Two small peptides were created based on the flap residues of HIVp that were in contact with the elbows in the 1TW7 structure. The first peptide consisted only of the residues in direct contact with the protease: Ac-Gly49-Ile50-Gly51-Gly52-NMe (GIGG). The complementarity is relatively poor and the contact skims the surface with no functional groups placed in the elbow cleft itself (Figure 5-2a). Therefore, another tetrapeptide, Asp-D-Ile-D-Phe-Gly (DifG), was designed from the backbone of residues

49-52, using D-amino acids to better orient side chains directly into the cleft region and increase complementary contact (Figure 5-2b). The use of D-Ile and D-Phe was suggested by solvent-mapping aromatic and hydrophobic functional groups into the binding cleft of the protease elbow.[103,211-213] The initial L-Asp was used to improve solubility and facilitate subsequent experimental testing of the model peptide.

The two GIGG tetrapeptides occupied 1087 $Å^2$ of the total solvent accessible surface area (SASA) of HIVp, and the DifG peptides occupied 1283 $Å^2$. Changing the chirality of the Ile increased the contact of each peptide by ~22 $Å^2$, and adding the Phe side chain in the D- orientation added ~77 $Å^2$ of contact, resulting in each DifG peptide having almost 100 $Å^2$ of increased contact with the elbow cleft of HIVp. SASA was measured in NACCESS2.1.[254]

Both implicit and explicit solvent simulations were performed. A total of 16 simulations were carried out. Twelve independent implicit-solvent LD simulations were performed, six for each tetrapeptide and initiated from random number seeds. Two simulations of GIGG restrained in the elbow cleft were run, one implicit LD and one explicit-solvent MD. As a control, one implicit LD and one explicit MD simulation of apo HIVp were also performed; these simulations were unrestrained and started from the 1TW7 crystal pose. Each simulation with GIGG or DifG was constructed as a 2:1 complex such that each HIVp dimer contained two peptide ligands, one in each elbow. This generated trajectory data for 12 DifG-HIVp associations (the six unrestrained LD) and 16 GIGG-HIVp associations (six unrestrained LD, one restrained LD, and one restrained MD) for analysis. Figure 5-3 shows the individual restraints applied for the simulations of GIGG restrained in the elbow. In the restrained simulations, an upper bound of 32 kcal/mol-$Å^2$ and a lower bound of 0 kcal/mol-$Å^2$ were used (weight was increased from 0.1 to 1.0 during the first phase of equilibration and then held constant at 1.0 for the extent of the simulation). Restraints kept the peptides associated to the elbow region but still allowed some flexibility for the peptides. It was desirable for the peptides, not the artificial restraints, to control the conformational sampling of HIVp.

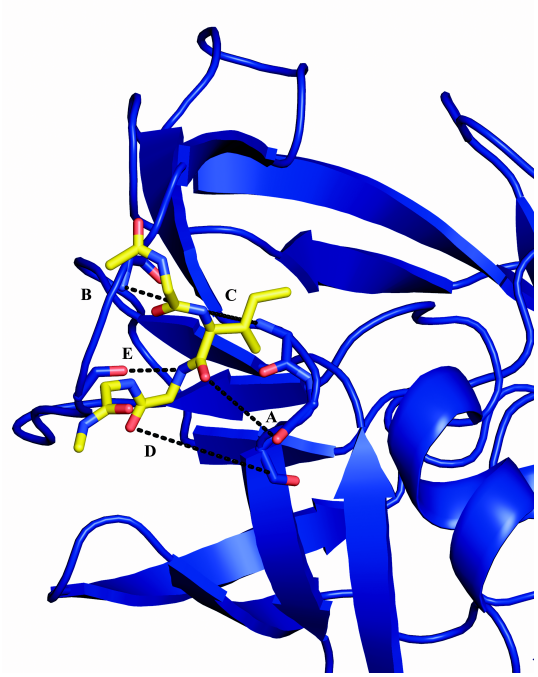

Figure 5-3: Distance restraints employed in the LD and MD simulations of the GIGG-HIVp complex. The Ac-G1-I2-G3-G4-NMe ligand was held in the elbow cleft with restraints woven between the elbow and the cantilever regions: A) I2(O) – Q61(O) = 6.17Å, B) I2(N) – R41(N) = 4.83 Å, C) I2(N) – D60(N) = 7.55 Å, D) G3(O) – V62(O) = 10.42 Å, E) G3(N) – P39(O) = 4.08 Å. These values are upper limits of the allowed distances; there is no penalty for forming closer contacts. This prevents dissociation but allows for some freedom in sampling and adaptation outside the crystalline environment of 1TW7.

The implicit-solvent simulations used the LD protocol of Simmerling and co-workers,[253,255,256] while explicit-solvent protocol was based upon work by Meagher *et al*.[243] All simulations were initiated from 1TW7 crystal structure of apo HIVp. PyMol[257] was used to propagate the unit cells and obtain the two protein chains in contact with the elbow region. Truncation of those chains into peptides was performed in MOE2006.08;[258] the side chains modifications to create DifG were also done in MOE. Hydrogens were added in the tLEaP module in AMBER8,[259] and the FF99SB force field[260] was used. The time step was 1 fs and bonds to hydrogens were constrained with SHAKE. Temperature was controlled for LD simuations with a collision frequency of 1 ps^{-1}. A modified GB solvation model[9] was implemented to represent aqueous solvation in the LD simulations.

The explicit-solvent MD protocol was similar to the implicit setup. TIP3P waters were used to solvate the system as an octagonal box (14,190 waters in the GIGG-HIVp complex and 10,492 waters around apo HIVp from 1TW7). Chloride ions were added to neutralize, and the Particle Mesh Ewald (PME) method was used to calculate long-range

electrostatic interactions.[232] A cutoff of 10 Å for non-bonded vdW interactions was employed.

Equilibration was accomplished over a series of six phases. The system was gradually heated from 100-300 K over the first two steps and remained at 300 K thereafter. Restraints were placed on all heavy atoms and gradually removed over the first four phases using force constants from 2.0 to 0.1 kcal/mol-Å^2. In the fifth phase, only the backbone atoms were restrained with a force constant of 0.1 kcal/mol-Å^2. Phases one through three were each 10 ps; phases four and five were 50 ps. In the last phase of equilibration, all atomic force restraints were removed, and the system sampled 200 ps at 300 K. For the unrestrained simulations, the subsequent production phase was performed under the same conditions, sampling 1.5 ns for unrestrained DifG and 3 ns for unrestrained GIGG. For the restrained LD and MD simulations of GIGG and the LD and MD simulations of apo HIVp from 1TW7, setup and equilibration occurred as before, except that the final temperature was 310 K. These systems were equilibrated during phase six for 2 ns and the production run lasted 16 ns.

Snapshots were taken every 1 ps for analysis in the ptraj module of AmberTools 1.0.[10,249] For each snapshot, the RMSD to the 1TW7 conformation was calculated to the Cα core of the protease (all residues except 43-58 and 43'-58'). The RMSD of flap Cα atoms was measured from the core-overlaid frame of reference. Snapshots from the simulation were also manually viewed to confirm that the peptides were dissociating and not simply finding an alternative-binding mode. The percentage of native contacts between HIVp and the tetrapeptides were calculated over the course of the unrestrained LD, using the MMTSB code.[261] The root-mean-squared fluctuation (RMSF) of the backbone heavy atoms was calculated in *ptraj* for each residue (Figure 5-4).

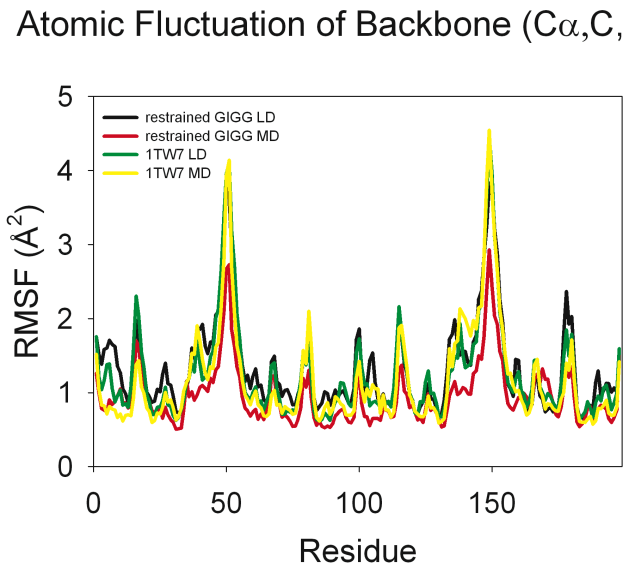

Figure 5-4: The rmsf of the backbone heavy atoms is presented for the simulations of restrained GIGG-HIVp and apo HIVp (noted 1TW7). The restrained MD is often the most restricted, but the restrained LD is often the least - flaps are highly mobile in each simulation.

A FRET-based assay was used to determine the inhibition constants of GIGG and DifG against HIVp.[103,262,263] The substrate in the assay was a labeled oligopeptide, RE(EDANS)SQNYPIVQK(DABCYL)R, purchased from Molecular Probes (Cat. No. H-2930); recombinant HIVp was purchased from BaChem Biosciences (Cat. No. H-9040.0100), and the tetrapeptides GIGG and DifG were synthesized by the Peptide Core at the University of Michigan Medical School. Pepstatin A (PepA) was used as a positive control for inhibition of HIVp (USB, lot #110018). The fluorimetric assays were performed in triplicate in 384 well plates (Corning No. 3676) and read using a SpectraMax M5 (Molecular Devices). Protease cleavage of the substrate released EDANS from DABCYL, and EDANS fluorescence was monitored with excitation/emission wavelengths of 340/490nm with a cutoff filter at 475nm.

To help prevent peptide and protease precipitation, PEG-400 was diluted in Buffer A (20mM phosphate, 1 mM DTT, 1 mM EDTA, 20% glycerol, and 0.1% CHAPS at pH 5.1); 1 µL was added to each well (PEG-400 final concentration, 0.1%). 2 µL of compound was diluted in water and then added to the wells to provide final concentrations ranging 50-250 µM, followed by dilution of 5 µL HIVp in Buffer A (final concentration of 30 nM). The peptide and protease were incubated for 45 min at room

temperature; then, 12 μL of substrate (diluted in Buffer A to a final concentration of 2 μM) was added to initiate the assay.

5.3 Results and Discussion

5.3.1. Unrestrained LD simulations of peptide-HIVp complexes

Twelve independent LD simulations of HIVp complexed with peptide ligands were initiated from the 1TW7 crystal structure (6 for GIGG and 6 for DifG). The tetrapeptides were unrestrained and allowed to freely associate with the protease. HIVp itself remained stable, with flap tips sampling freely. Throughout the first five steps of equilibration, the peptides remained in contact with the structure, with a maximum RMSD of 2.5 Å to their initial location based on 1TW7. However, all of the peptides dissociated from the protease during the production phase, as demonstrated by the percentage of native contacts lost over the simulation (Figure 5-5). Although run 2 and 4 of the GIGG-HIVp complex still retain some of the native contacts at 3 ns, the peptides are no longer positioned in the protease elbow (Figure 5-6). None of the unrestrained tetrapeptides remained in the binding cleft throughout the simulation. The additional contacts provided by the DifG modifications provided no improvement.

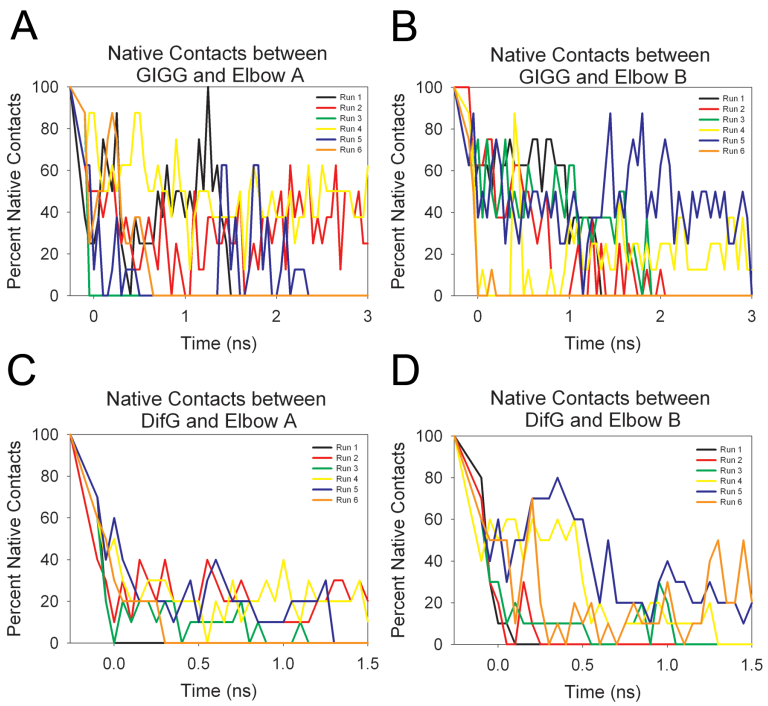

Figure 5-5: Percent of native contacts over the course of the unrestrained LD for (A,B) GIGG or (C,D) DifG bound to HIVp.

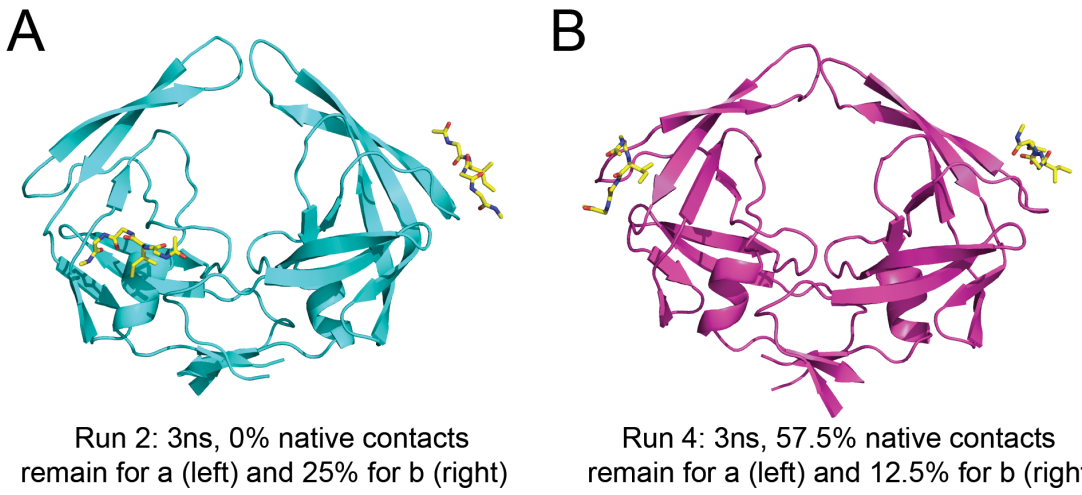

Figure 5-6: The 3-ns snapshot of runs 2 and 4 from the unrestrained LD simulations of the GIGG-HIVp complex. These images demonstrate that while some of the native contacts are retained, the tetrapeptide is no longer associating with the elbow pocket (compare to contacts in Figure 5-1).

5.3.2. LD and MD of restrained GIGG-HIVp and unrestrained apo HIVp

There are many reasons a ligand can be unstable in a simulation, so it was important to establish if conformational control of the flaps was possible in the event that the

peptides remained in the elbow. As outlined in methods, we restrained GIGG to remain at the elbow for an explicit MD and implicit LD simulation of 16 ns each. As a control system, two 16 ns simulations of apo HIVp (initiated from 1TW7 with ligands removed) were performed and analyzed in comparison to the complex. The LD simulations were performed to access greater sampling of conformational states, while the MD were generated to more accurately sample the motion of the system in the native state.

Over the course of sampling, all trajectories demonstrated stability of the core HIVp residues. However, both the free and bound systems displayed considerable motion in the flap region. In these simulations, the flaps moved away from their wide-open position in the crystal structure of 1TW7 and sampled the semi-open conformations. In fact, the apo handedness of the semi-open state was obtained for both the restrained and free LD simulations (the wide-open flaps have the handedness of the bound state). Sampling was reduced in explicit-solvent, and neither the restrained or unrestrained simulation changed flap handedness.

To quantify the sampling, snapshots from every 1 ps of the 16-ns simulation were clustered into 20 distinct conformational families, using the means algorithm within ptraj.[249] Our analysis of the conformational behavior of the families focused on the flexible flaps. The core of each conformation was overlaid to the 1TW7 conformation based on Cα RMSD. The flexibility across families was then measured as the Cα RMSD (without fitting) for the flap residues 43-58/43'-58'. The core of the protein was very stable and similar across all of the simulations, but the flaps were quite mobile. For both the LD and MD simulations, the conformations from the restrained GIGG-HIVp complex showed nearly identical sampling to those from the unrestrained apo HIVp. The RMSD of the Cα of the flaps were 5.26 ± 0.56 Å and 5.32 ± 0.66 Å for the LD of the restrained GIGG-HIVp and unrestrained apo HIVp, respectively (Table 5-1). For the MD, flap RMSDs were 3.17 ± 0.60 Å and 3.61 ± 0.74 Å for the restrained and unrestrained simulations, respectively (

Table 5-2). Although the restrained simulations showed slightly less sampling, this was insignificant when compared to the range of RMSD across the 20 conformational families. We also considered the impact of only the most densely populated states. The conformational families with the most occupants were chosen such that the largest families that represented ~85% of each trajectory were used (11 conformational families from the MD and LD of the restrained complex; 12 from the unrestrained LD and MD). The conformational sampling across 85% of the LD and MD trajectories was quite marked (Figure 5-7; note that the RMSD values are slightly different than those given above, as these reflect the variation across a subset of the conformational sampling).

Table 5-1: Snapshots from the implicit-solvent LD simulations were clustered into 20 conformational families. Below, Cα-rmsd is used to compare each representative to the dimer of HIVp in the 1TW7 structure. The "RMSD All" is the standard comparison by global overlay of all residues in the dimer; the other two metrics are obtained by overlaying only the core of HIVp to the 1TW7 structure and then measuring RMSD for the flaps and core separately (measurement, not additional overlays). The flap residues are 43-58 and 43'-58', and all other residues are included in the core.

Restrained, GIGG-HIVp Complex LD				Unrestrained, Apo HIVp LD			
Occupancy	**RMSD All**	**RMSD Flap**	**RMSD Core**	**Occupancy**	**RMSD All**	**RMSD Flap**	**RMSD Core**
14.9%	2.82	5.01	2.25	13.4%	2.78	5.39	2.00
10.6%	3.18	5.72	2.53	12.0%	2.65	5.03	1.94
9.8%	2.88	4.85	2.40	11.0%	2.93	5.20	2.31
9.0%	2.54	4.46	2.04	9.5%	2.83	5.56	2.01
8.5%	2.88	5.11	2.25	9.2%	3.33	6.29	2.43
7.6%	2.55	5.06	1.78	5.9%	2.90	4.77	2.40
6.6%	2.49	4.57	1.88	4.6%	2.85	6.14	1.75
6.5%	2.97	4.98	2.47	4.3%	2.87	4.84	2.36
5.4%	3.02	5.63	2.37	4.0%	2.41	5.25	1.87
3.6%	2.64	5.03	1.93	4.0%	2.64	4.08	1.98
2.9%	3.48	6.25	2.70	3.8%	2.80	4.28	2.45
2.4%	3.15	5.71	2.50	3.0%	2.64	4.96	1.97
2.4%	2.79	4.69	2.34	2.6%	3.05	5.73	2.28
2.2%	2.83	4.74	2.40	2.5%	3.02	5.06	2.48
1.8%	3.11	4.87	2.73	2.3%	2.96	5.96	2.03
1.6%	3.31	6.09	2.54	2.1%	3.13	5.82	2.34
1.2%	2.81	5.18	2.23	1.9%	3.30	5.37	2.76
1.1%	3.52	6.21	2.85	1.5%	3.08	5.71	2.31
1.1%	2.89	5.04	2.40	1.4%	3.14	6.49	2.08
0.9%	3.30	5.92	2.69	1.2%	2.42	4.38	1.90
Average	2.96	5.26	2.36	**Average**	2.89	5.32	2.18
St. Dev	0.30	0.56	0.29	**St. Dev**	0.25	0.66	0.26

Table 5-2: Snapshots from the explicit-solvent MD simulations were clustered into 20 conformational families. Below, Cα-rmsd is used to compare each representative to the dimer of HIVp in the 1TW7 structure. The "RMSD All" is the standard comparison by global overlay of all residues in the dimer; the other two metrics are obtained by overlaying only the core of HIVp to the 1TW7 structure and then measuring RMSD for the flaps and core separately (measurement, not additional overlays). The flap residues are 43-58 and 43'-58', and all other residues are included in the core.

Restrained, GIGG-HIVp Complex MD				Unrestrained, Apo HIVp MD			
Occupancy	**RMSD All**	**RMSD Flap**	**RMSD Core**	**Occupancy**	**RMSD All**	**RMSD Flap**	**RMSD Core**
12.3%	2.07	3.60	1.65	22.5%	2.49	4.72	1.78
10.7%	1.94	3.61	1.44	11.7%	2.31	4.19	1.75
10.3%	1.75	3.11	1.37	10.9%	2.16	3.80	1.68
9.9%	1.90	3.34	1.50	10.4%	2.51	4.51	1.90
7.6%	1.71	3.00	1.34	6.5%	2.08	3.73	1.59
7.5%	1.72	3.04	1.35	5.2%	2.46	4.61	1.79
6.6%	1.85	3.60	1.32	3.9%	1.34	2.18	1.13
5.5%	2.17	4.30	1.52	3.5%	2.16	4.52	1.35
5.2%	1.52	2.78	1.13	3.3%	1.88	4.11	1.12
4.8%	1.55	2.98	1.12	2.9%	1.49	2.60	1.17
3.3%	1.88	3.95	1.24	2.7%	1.57	2.97	1.12
3.1%	1.89	3.58	1.41	2.5%	1.53	2.54	1.26
2.6%	1.67	3.22	1.20	2.5%	1.75	3.27	1.34
2.0%	1.65	3.50	1.04	2.5%	1.84	3.35	1.40
1.8%	1.27	2.33	0.99	2.1%	1.98	4.01	1.40
1.6%	1.47	2.81	1.05	1.7%	1.70	3.03	1.36
1.6%	1.53	2.89	1.10	1.6%	2.06	3.90	1.50
1.5%	0.94	1.45	0.82	1.5%	1.72	3.29	1.24
1.1%	1.70	3.18	1.23	1.3%	1.90	3.82	1.26
0.9%	1.60	3.06	1.15	0.9%	1.58	3.06	1.10
Average	1.69	3.17	1.25	**Average**	1.93	3.61	1.41
St. Dev	0.28	0.60	0.21	**St. Dev**	0.35	0.74	0.26

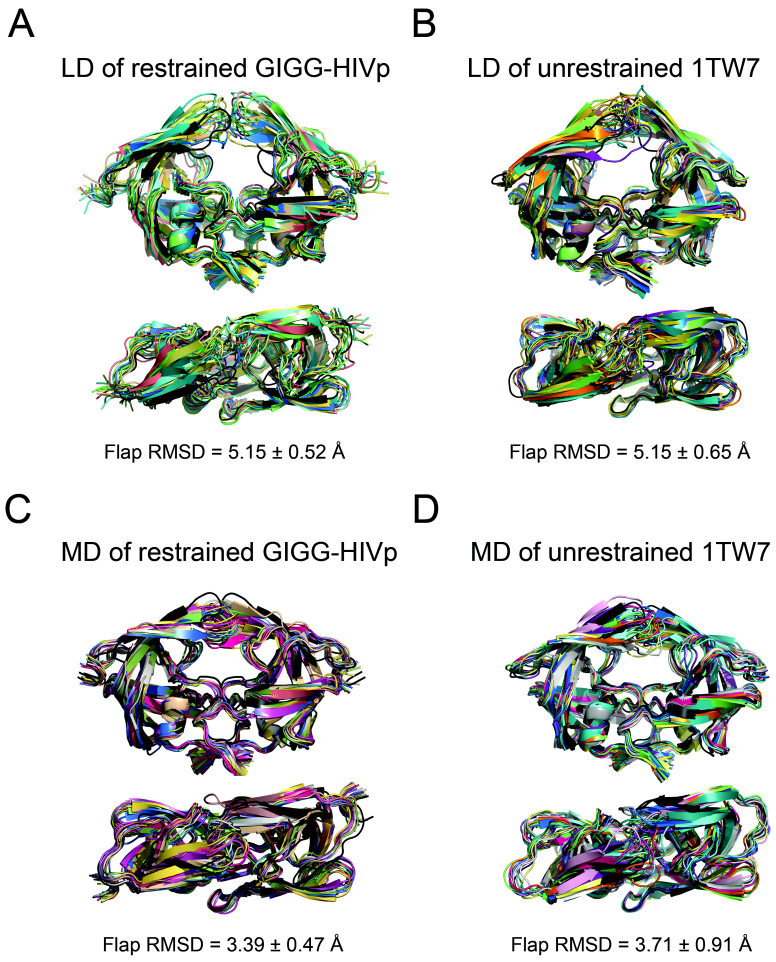

Figure 5-7: Representative cluster families display the conformations sampled in 85% of each simulation. Each set is overlaid to the crystal structure conformation of 1TW7 (in black): (A) LD of the restrained GIGG-HIVp, (B) LD of unrestrained apo HIVp, (C) MD of restrained GIGG-HIVp, and (D) MD of unrestrained apo HIVp. The structures are overlaid by the Cα of the core residues (all residues except the flaps 43-58 and 43'-58'), and the RMSD of the flap region is noted.

There was little difference in flap mobility between the unrestrained apo HIVp and HIVp-GIGG complex; both freely sampling flap conformations between 3-11 Å RMSD of the placement in the 1TW7 structure. One of the standard metrics for assessing the conformational state of the flaps has been the distance between Asp25 and Ile50. In 1TW7, the skewed-open structure has a distance of 18.8 Å. The semi-open structure 1HHP has a distance of 17.2 Å and the closed structure 1PRO has a distance of 14.1 Å. The flaps in both the restrained and unrestrained LD simulations clearly sample semi-open and open conformations (Figure 5-9). The presence of explicit solvent reduced the degree of sampling and biased flap conformations towards the closed position (Figure

5-10). However, the restrained and unrestrained simulations were not significantly different in their conformational sampling.

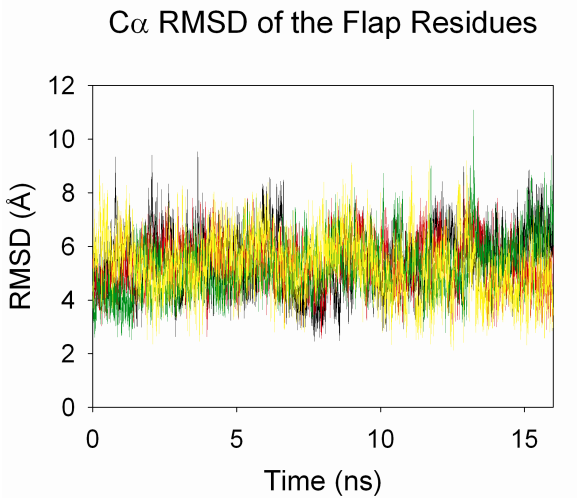

Figure 5-8: A wide degree of conformational sampling is seen in the LD simulations, whether GIGG is present or not. Both flaps of the restrained GIGG-HIVp (yellow and green lines) and unrestrained apo HIVp (red and black lines) simulations are shown. The snapshots were overlaid with respect to the Cα atoms of the core of the dimer in 1TW7, and the RMSD of only the flap Ca (residues 43-58 and 43'-58') are shown above.

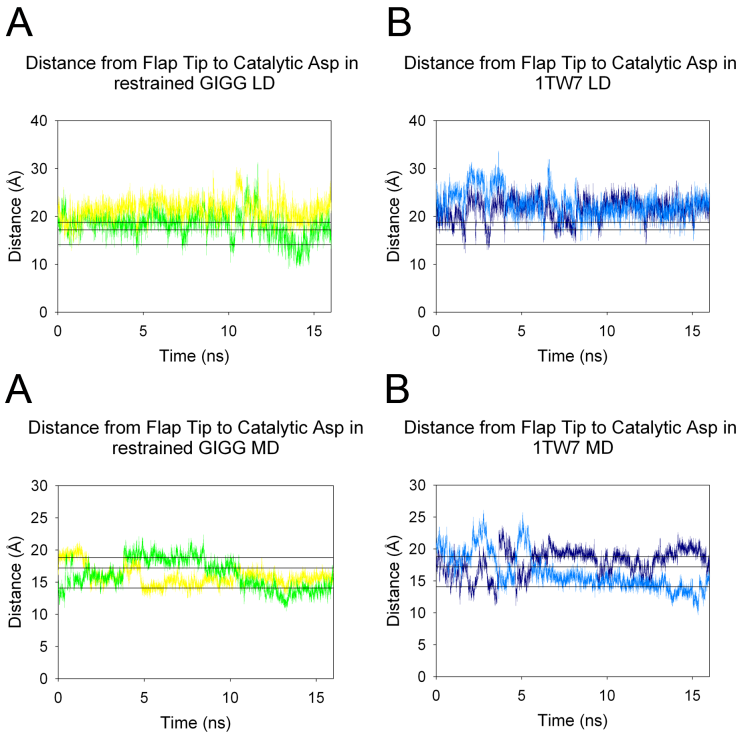

Figure 5-9: The distance from the flap tips Ile50/50' to the catalytic Asp25/25' during the implicit-solvent LD of (A) the restrained, GIGG-HIVp complex and (B) the unrestrained, apo HIVp. The individual flaps have different colors in each plot. Horizontal, black lines note the distances seen in different

conformational states: the skewed-open 1TW7 is 18.8 Å, the semi-open 1HHP is 17.2 Å, and the closed 1PRO is 14.1 Å.

5.3.3. Experimental testing of peptides

To further support our conclusions, we conducted experimental inhibition assays. A fluorimetric assay was used to discern the inhibitory potency of the GIGG and DifG peptides. Consistent with our simulations, we found no inhibition of HIVp by either tetrapeptide at concentrations of 250 μM (Fig. 2-7).

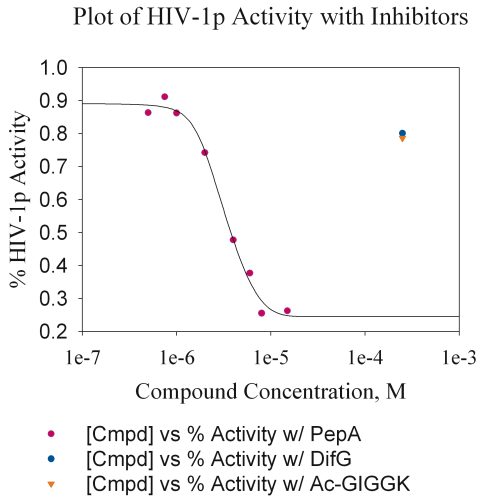

Figure 5-10: The activity of HIVp in the presence of 250 μM peptides is given: DifG (blue circle) and the model peptide for GIGG (Ac-GIGGK, orange triangle). The IC50 curve of the control PepA is also shown (purple circles).

5.4 Conclusion

As demonstrated by both unrestrained and restrained simulations of protease-ligand complexes, the contacts seen in the 1TW7 crystal structure do not exemplify allosteric control. The peptide-HIVp complexes were unstable and freely dissociated. The interactions appear too weak, even when modified to improve the contacts. Despite restraints to maintain contacts to the elbow region, their association with HIVp had no significant affect on flap mobility. Perryman *et al*. have used restraints within the elbow region to control flap dynamics[10], but it appears that maintaining only the elbow contacts in 1TW7 is not able to force this control. Furthermore, experimental testing showed no inhibitory activity by small peptides representing those crystal contacts. This study further refines the conclusions of Layten *et al*[253] to show that the altered conformation in

1TW7 is solely the result of packing effects and not the result of a symmetric environment which fortuitously presents allosteric contacts.

It must be emphasized that this study does not refute the possibility of allosteric control via the elbow region, but it does indicate that peptide-based molecules may be less appropriate for these efforts. Hornak *et al*. showed that it was possible for the small inhibitor XK263 to correct itself during LD sampling after initial improper placement.[255] That result does not mean that all ligands can correct poor contacts during an LD simulation, but we may have expected at least one of the 12 peptides to correct their placement into a stable alternative location if peptide ligands were appropriate. Our results could explain why no allosteric inhibitors have been found serendipitously, despite a significant effort to develop competitive inhibitors using peptide scaffolds.[264]

It is possible that effective allosteric inhibitors will require more contacts between the structural features of HIVp. The contacts in 1TW7 placed the flap tip in contact with only the elbow and cantilever (residues 59-75) regions. As such, these were the only contacts maintained in our simulations. However, the nearby "fulcrum" region (residues 11-22) has also been shown to be correlated with flap motion.[190,205,265,266] It is possible that a small molecule will have to contact all three regions to gain adequate allosteric control of a region as large as the flaps.

This work was published as: Lexa KW, Damm KL, Quintero JJ, Gestwicki JE, Carlson HA. Clarifying Allosteric Control of Flap Conformations in the 1TW7 Crystal Structure of HIV-1 Protease. *Proteins* **2009**, *74*, 872-80.

Appendix

A.1 Exploring the Potential for Allostery at the Elbow Region of HIV-1 Protease

A.1.1. Introduction

Studies have shown that the conformational flexibility of these flaps is highly anti-correlated to other regions on the protease, particularly the elbow (residues 35-42/35'-42').[190,196,204,205,253,256,265-271] Perryman *et al*. performed a series of harmonic restraint coarse-grained and MD simulations that held a pair of atoms in the elbow region a set distance apart.[190,205,265] These studies demonstrate that constraining movement at the elbow region impacts the conformational sampling of the flaps. Although this research is limited in that distance restraints were used rather than a small molecule, it laid the groundwork for the theory of allosteric regulation of HIVp activity via the elbow region. To demonstrate that allosteric control through the elbow site is a valid therapeutic path to pursue for controlling the flexibility and therefore activity of HIVp, further studies were necessary to examine the impact of a small molecule bound at the elbow region.

Using the MPS technique[101,272,273], we flooded the elbow region of two different HIVp conformations, using MD snapshots of semi-open protease (PDB ID 1HHP[274]) as well as NMR structures of closed protease (PDB ID 1BVE[275]). Pharmacophore models for each of these conformations were generated with MPS, however the models consisted mostly of hydrophobic elements, demonstrating a lack of specific contacts available in the elbow region necessary for realistic drug design. The pharmacophore model resulting from the semi-open structure displayed only four elements between the elbow and cantilever. The closed conformation pharmacophore displayed a larger binding site, with five available contact elements, including a hydrogen bonding element. While the semi-open model highlighted contacts between the elbow and cantilever, the closed protease model illustrated the importance of contacts between the elbow, cantilever, and fulcrum (Fig. 3-1). This closed flap pharmacophore model implied that small molecules with affinity for the elbow region should not only form interactions the elbow and cantilever residues, but also the fulcrum residues. These additional contact regions should improve specificity for the elbow site and also contribute to allosteric regulation of the flap region.

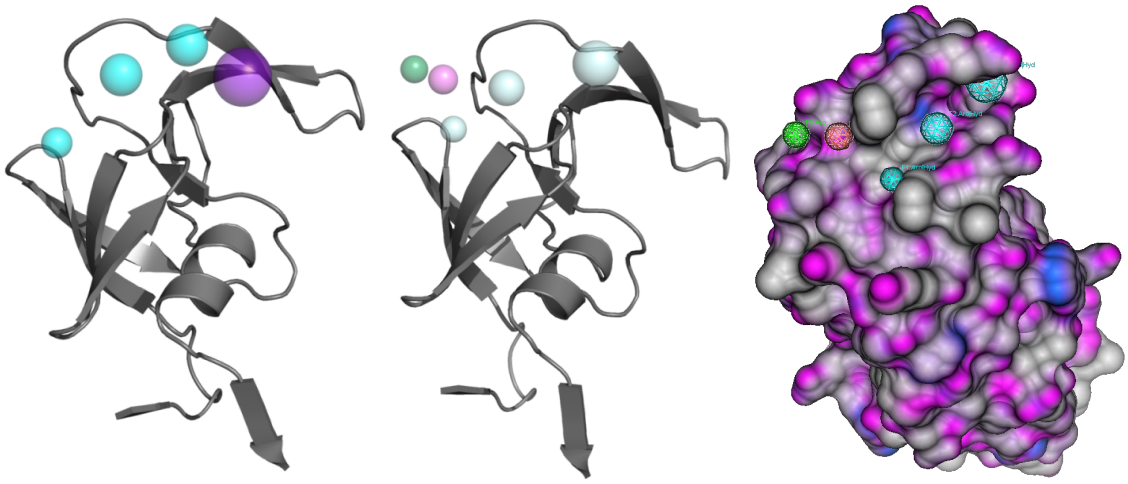

Figure A-1: The pharmacophore models of HIVp based on MD snapshots of the semi-open 1HHP structure is shown to the far left. In the middle is the MPS model based on NMR structures (1BVE) of closed HIVp showing the V-shaped binding pocket in the elbow region. In these figures, cyan represents an aromatic element and green an aromatic/hydrophobic element. The image to the right demonstrates the character of the elbow region binding site with an analytic Connolly surface: hydrogen bonding regions are shown in pink, hydrophobic in gray and mildly polar in dark blue.

We set out to elucidate this region of space at the elbow site through MD studies. Distance restraints as well as small molecule mimics were implemented to better understand the link between the cantilever-elbow-fulcrum triangle and the flaps. To further our understanding of the elbow region as an allosteric site with therapeutic potential, it was necessary to identify the appropriate conformational space that must be controlled in order to induce a shift in the flap position. To this end, we utilized distance restraints and ligand mimics at the elbow region.

To further clarify the nature of allosteric regulation, we also performed LD and MD simulations of HIVp tethered at the β-sheet region. Both an apo and bound tethered HIVp were crystallized by Pillai[276] and Bhat[277] respectively. The tether attaches the C-terminus of the first monomer to the N-terminus for the second monomer (Figure A-). Both structures were shown to adopt the closed conformation. These studies demonstrated another possible target for allosteric control of HIV-1p. While these simulations examined the importance of contacts at the dimer interface, they simultaneously provided a mechanism for exploring correlations between flexible residues. The β-sheet linker enabled us to study how allosteric regulation at the dimer interface might correlate with motion at the elbow as well as the flaps.

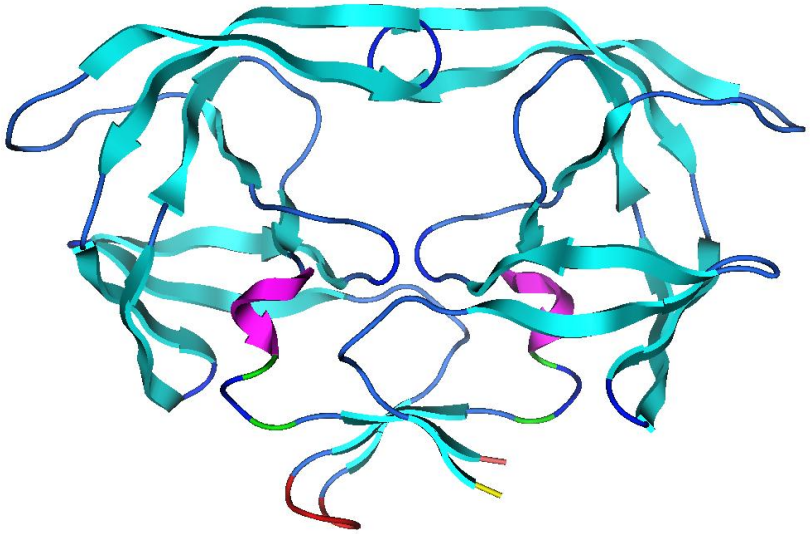

Figure A-2: The tethered HIVp, with the linker shown at the base of the dimer in red.

Methods

MPS

The MPS methodology originally developed by Carlson *et al*. was employed to create pharmacophore models of the allosteric sites. The 1BVE structural ensemble was downloaded from the PDB[278] and the 1HHP structural ensemble were generated through MD in AMBER8[259]. These files were prepared in PyMol[257,279], MOE2006.08[258] and tLEaP. Hydrogens were minimized in AMBER8 while constraining the movement of all other atoms.

Each structure file underwent an initial multi-unit search for interacting conformers (MUSIC)[280] run using BOSS[280,281] in preparation for flooding. Once prepared, these files were imported into PyMol and the region for flooding was centered at the elbow (Gly41) with a radius of 15 Å. The files were then flooded with small molecule probes three separate times: once with 500 benzenes, once with 500 ethanes, and once with 500 methanols. The resulting files were then run through MUSIC a second time to simultaneously minimize all 500 probes of a single type to the binding surface. The probe-probe interaction energy was ignored, allowing for clustering of the small molecules into groups of probes that represented potential binding interactions. Probes were clustered into groups by a Jarvis-Patrick algorithm developed in the Carlson lab.[282] All clusters with at least eight members were retained. The probe with the lowest energy

from each cluster was chosen to represent the cluster as a single parent probe. Then the coordinate files were aligned using the wRMSD code from the Carlson lab[283] and examined in PyMol. For the 28 coordinate files containing parent probes of benzene, clusters of parent probes with at least 14 probes were retained. The same was done for the coordinate files containing ethane and methanol. A pharmacophore table was created based on the clusters of parent probes, resulting in a final pharmacophore model of important features at the binding site.

AutoDock

AutoDock4.0[284] was used to perform grid-based docking of the ligand mimics into the allosteric binding pocket. To calculate partial charges on the atoms, AutoDockTools was used to assign Gasteiger (PEOE) partial charges. Lamarckian GA sampling[285] was implemented for searching the conformational space.

Amber Paramterization of Ligand Mimics

Force field parameters were created for the ligand mimics consisting of three or five artificial atoms by developing special libraries through antechamber and xLeAP using the gaff force field[245]. The atomic weight, radius, bond length, bond stretch, torsional angles, ε value and σ value were used to describe the artificial atom (Eq. 1-2). The depth of the Lennard-Jones well potential is signified by ε and the well width by σ. These atoms were created as vdW spheres with radii of 1.9 (SML), 2.2 (MD), or 3.55 (LRG) Å, linked by a 3.5 Å bond, and docked to the elbow region of apo 1HHP or apo 1PRO. The size of the effective van der Waals radius (r_{min}) was varied to fully explore the effect a small molecule could have upon flap movement through the elbow region. Additionally, we chose to use very loose force constants for bond stretching, angle bending and torsions to allow the ligand mimics to fully sample the area available for binding at the elbow region. The atoms were uncharged, but the well depth (ε) was set to five times that of oxygen, causing the surrounding atoms to preferentially interact with these artificial atoms (Table A-1).

Table A-1: Parameters used for artificial atoms in ligand mimics simulations compared to values for aromatic C.

Replacement parameter set for model beads compared to aromatic C				
MASS	Atomic Mass	Polariz-ability		
A	16.00	0.000		
CA	12.01	0.360		
BOND	Harmonic Force Constant	r_{eq}		
A-A	100.0	3.500		
CA-CA	469.0	1.400		
ANGLE	Harmonic Force Constant	θeq		
A-A-A	10.0	160.0		
CA-CA-CA	63.0	120.0		
DIHE	Div. Factor	Barrier Height (V_n)	Phase (γ)	Periodicity (n)
X-A-A-X	0.0	0.0	0.0	1.0
X-CA-CA-X	4.0	29.0	180.0	2.0
NONB	$\sigma^*2^{1/6}$	ε		
A	1.9/2.2/3.55	1.00		
C*	1.9080	0.0860		

Elbow MD Simulations

The restrained elbow simulations were initiated from semi-open (1HHP) and closed (1PRO) conformations following preparation in MOE2006.08[258] and PyMol[257,279]. MOE was used to check partial charges and protonation states of Asp, Glu, and His residues. PyMol was used to propagate symmetry contacts, remove crystallographic ligands, and ensure the neutralizing ions and ligand mimics were appropriately positioned. All of the explicit simulations followed the procedures outlined by Meagher et al.[243].

Hydrogens and neutralizing ions were added in the tLEaP module in AMBER and simulations were performed in AMBER8[245] according to the PME[232] method with the FF99SB force field[256] and TIP3P waters[215]. The waters were added as a truncated octagonal box extending 12 Å beyond the protein and the vdW cut-off was 10 Å. A timestep of 2 fs was employed and bonds to hydrogen were restrained with SHAKE[4]. The hydrogen atoms and water molecules were equilibrated before the protein was allowed to move to prevent system collapse.[243] The total simulation length was 35 ns per run at a constant temperature of 310K.

A series of six equilibration phases were performed for each simulation; during the first two, the temperature of the system was increased over 20 ps from 110K to 310K.

The temperature was then held constant for the duration. Restraints were placed on the heavy atoms and then gradually removed over the course of equilibration, from an initial force constant of 2.0 kcal/mol* Å ^{2} to 0.1 kcal/mol*Å^2 during the fourth phase. During the fifth phase, restraints were only placed on the backbone atoms, with force constants of 0.1 kcal/mol* Å^2. After the fifth phase, restraints were removed and the system was able to sample freely. Phases one through three were each 10 ps, while phases four and five were both 50 ps. Phase six was the final step before production, allowing the system to freely equilibrate over 2ns.

NMR Distance Restraints

The NMR restraint option in AMBER was implemented to restrict distances across the elbow cleft. The distances across Gly40Cα-Gln61Cα and Gly17Cα -Pro39Cα were held stable at 7.7 Å (open) or 10.5 Å (closed). The distance across Gly17Cα-Gln61Cα was held at 14.8 Å (closed) or 17.3 Å (open). The energy penalty for exceeding the upper or lower bound was set at 32 kcal/mol*Å or 60 kcal/mol*Å. These simulations are hereafter referred to as Closed1, Closed2, Open1, and Open2. These systems were minimized in sander to relax bad contacts prior to application of restraints across the elbow. Following phase five of equilibration (3.2.4), the NMR force restraints were applied with a weight multiplier of 0.1 to allow the system to adjustment to the restraints. This prevented system blow-up and any major structural violations. The weight multiplier was increased to 1.0 over 5000 steps and then maintained at 1.0 for the rest of the simulation.

This method of ramping up the weight of force constants worked for opening the protease, but we found that closing restraints required a more gradual application. In these cases, we began applying weights following the inital 800 ps of equilibration step six, starting with a value of 0.1 and increasing to 0.8 by incrementing plus 0.1 every 200 ps. The weight was further increased from 0.8 by 0.1 every 2000 ps until a constant of 1.0 was reached. The force constants were then maintained at a weight constant of 1.0 for the duration of the simulation.

Tethered β-sheet MD

Since the tethered apo structure (PDB ID 1G6L[276]) was disordered at the linker region, the starting structure was generated from the tethered holo HIVp crystal structure

(PDB ID 1HVC[277]) following ligand removal. All MD simulations were performed in AMBER8 with FF99SB. Hydrogens were added in the tleap module and minimized over 10,000 steps. The minimized structure was solvated by a 15 Å octahedral box of TIP3P waters, the long-range interaction cut-off was set to 10 Å, and chloride ions were added based on the electrostatic surface potential[248] to neutralize the charge. Equilibration was performed as described in *3.2.4*. The production run was carried out over 18 ns.

Calculated Properties

Success for each of our studies was defined as modified flap motion. To identify modified flap motion, several analytical methods were employed. *Ptraj*[10,249] was used to analyze and cluster simulation trajectories to give a picture of the flexibility inherent in our various protein-ligand and protein-restraint simulations. Commonly used metrics in the HIVp literature were calculated in order to quantify the extent of flap opening, flap tip curling, core stability, duration of the protein-ligand mimic bound state, and RMSF. The distances for the flap opening were calculated based on the distance between Asp25/25'Cα and Ile50/50'Cα. However, this metric only portrays the opening of one side of the protease and may not accurately reflect whether the active site is available for ligand binding. Therefore, an additional metric was used based on the center of mass (COM) between the two flap tips Ile50/50'Cα and catalytic Asp25Cα. Flap tip curling was measured by evaluating the angle between residues Gly48-Gly49-Ile50. Stability was assessed from the RMSD of the protease core from the starting structure. RMSF was calculated by *ptraj* using the atomicfluct command, where $(8/3)\pi^2*RMSF^2$ = B-factor. Clustering was based on the *means* algorithm available through *ptraj*. The ideal number of families to use for clustering our trajectories was based on the criteria for DBI, pSF, and SSR/SST as described by Shao *et al*.[249] For our tethered HIVp simulations, we generated cross-correlation plots with *ptraj* and Carlson group scripts to describe correlated motion over time.

Discussion

The Elbow

The triangle of space at the elbow region where the elbow, cantilever and fulcrum come together is mostly hydrophobic in nature. From our MPS model of closed HIVp, we

found that this binding site features one aromatic element, one donector element, and three aromatic/hydrophobic elements (Fig. 3-1). Due to the lack of specific contacts available in this region, performing VS to find novel leads that target this site may not be the best approach. The highly hydrophobic character of the elbow region would lead to the identification of many false positives. Instead, we performed a more complete exploration of the space by expanding our understanding of how a small molecule bound at the elbow might impact flap movement with the future goal of designing novel leads.

Elbow Restraint Simulations

By performing a series of simulations with distance restraints across the elbow-cantilever-fulcrum region (Figure A), it was possible to demonstrate where the most important interactions in the elbow region were occurring in order to modulate flap movement. This allowed the allosteric mechanism of action to be explored in-depth prior to performing virtual screening and structural biology experiments.

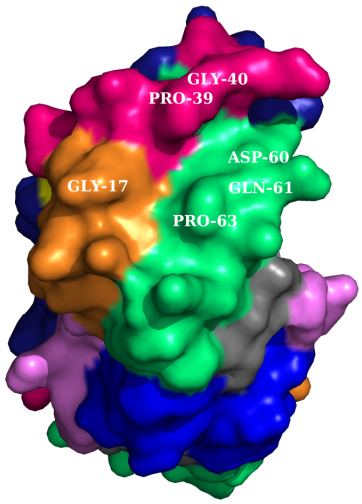

Figure A-3: Terminology for the topology of HIV-1 protease. Dark blue: flaps, residues 43-58. Yellow: flap tips, residues 49-52. Orange: fulcrum, residues 11-22. Lime green: cantilever, residues 59-72. Cyan: eye (active site), residues 23-30. Blue: whiskers (dimer interface), residues 1-5, 95-99. Hot pink: elbow, residues 35-42. Light blue: helix, residues 86-90. Red: wall turn, residues 79-84. Violet: nose, residues 6-10.

The effect of a combination of distance restraints at the elbow region was examined over a series of simulations. The distances used for each pairwise restraint were based on a comparison of the crystallographic distances seen between the four atoms used to

define the elbow cleft (Gly17, Pro39, Gly40, Gln61) as well as the distances used by Perryman *et al*.[204] We tried several alternate combinations of distance restraints because of the diverse distances observed in crystallographic structures of similar HIVp states. Additionally, the range of observed distances had overlap for the closed and semi-open conformations. The well width of the force restraints was also varied to examine the impact of a loose range of distances versus a more tightly controlled distance sampling. Combinations of one, two, or three restraints across the elbow region were used to explore the impact of constraining these secondary structural elements upon flap movement.

Eight LD simulations were performed over 15 ns to examine the effect of a combination of distance restraints at the elbow region. Although LD required less computational power, the difference between open and closed restraints was not significant (Figure A-). The most promising results were run in explicit solvent MD for 35 ns.

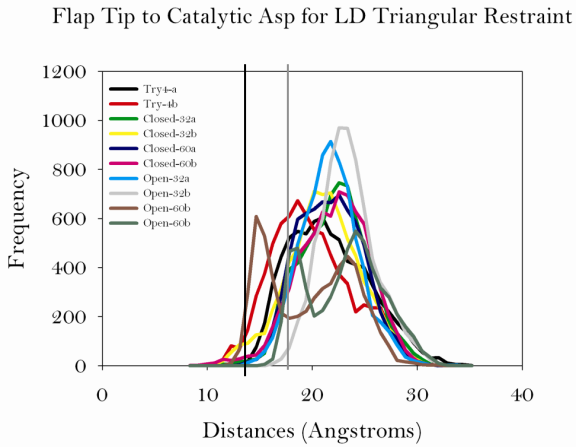

Figure A-4: Distances from the base of the active site to the flap tips as a measure of flap opening in the LD simulations. The typical distance for a closed conformation is shown as a black line, and the typical distance for a semi-open conformation is shown as a gray line.

Using a combination of two restraints, one across either the elbow-fulcrum or elbow-cantilever region, and another across the fulcrum-cantilever, led to a skewed conformation. Using two restraints across the elbow-fulcrum and elbow-cantilever did allow for flap closure. This indicated that the small fulcrum-cantilever cleft offers less important interactions when compared to the larger cleft between the elbow-fulcrum or the elbow-cantilever. Our MPS models of HIVp in the closed and semi-open state

reinforce this idea. Neither pharmacophore model displayed an element that took advantage of the small cleft available between the fulcrum and cantilever.

A full triangle of restraints led to normal closed, semi-open, and open conformations. We found it necessary to have all three distances restricted to a narrow well width to achieve significant flap closure. In addition, we found that the closed conformation was most frequently sampled and more stable when restraints across the elbow region were gradually applied. On the other hand, it was considerably easier to open the flaps, even without applying gradual restraints and without such a narrow well width. Furthermore, a heavier force constant of 60 kcal/mol*Å allowed for better flap closure than a lighter force constant of 32 kcal/mol*Å. The lower force constant led to more frequent sampling of semi-open flap positions. However, the lighter force constant did not perturb the flap conformation, whereas the higher force constant caused skewed-open flap conformations when used with opening restraints.

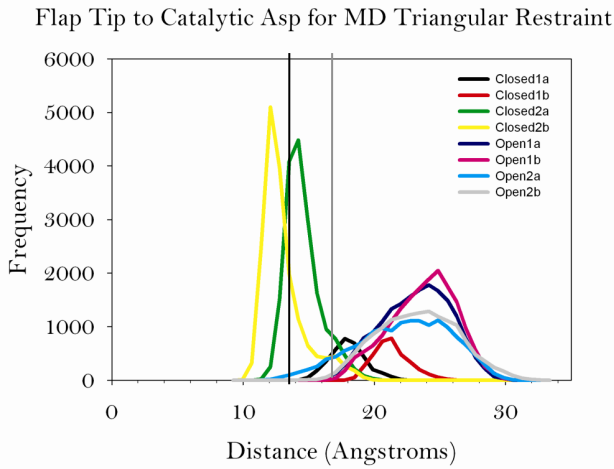

Figure A-5: Distances from the base of the active site to the flap tips as a measure of flap opening. The typical distance for a closed conformation is shown as a black line, and the typical distance for a semi-open conformation is shown as a gray line.

We found that Open1 and Open2 promoted full flap opening (Figure A-5). Both of the flap tip-to-catalytic site metrics quantitatively showed that the open restraints effectively pushed the semi-open conformation from 1HHP into an even more open conformation. The closed restraints shifted the ensemble toward the closed conformation when heavier force restraints were used. These metrics illustrated that our tighter closed restraints shifted the populated ensemble of conformations from the semi-open to the

closed conformation. We were able to successfully promote flap opening and flap closing from the semi-open conformation of 1HHP with three distance restraints. We also ran simulations from the closed conformation to illustrate forcing the closed flaps open.

Table A-2: In the semi-open crystal structure the Ile50-Asp25 distance was 17.8 Å. In the closed crystal structure the distance was 14.1 Å. In the semi-open structure of 1HHP, this COM distance measures 17.31 Å and in the closed structure of 1PRO it measures 13.1 Å.

Simulation	Distance (Å) Asp25-Ile50	Distance (Å) Asp25'-Ile50'	Distance (Å) Asp25– COM Ile50/50'	Distance (Å) Asp25'-COM Ile50/50'
Open1	23 ± 2.64	23.59 ± 2.51	22.89 ± 1.94	23.19 ± 1.95
Open2	22.15± 3.58	22.97± 3.04	22.22 ± 2.26	22.02 ± 2.37
Closed1	17.39 ± 1.43	20.66 ± 1.48	19.37 ± 1.52	18.17 ± 1.50
Closed2	13.93 ± 1.35	12.20 ± 1.63	14.06 ± 1.24	12.51 ± 1.5

RMSD and RMSF values were calculated to ensure that the distance restraints did not warp the structure of the protease (Figure A-6). We found that the core of the protease (all residues excluding the flaps) remained stable in all simulations with an average of 1.41 Å for the closed restraint simulations and 2.29 Å for the open restraint simulations. The average RMSD of the core for a 35 ns explicit MD simulation of apo 1HHP was 1.92 Å. We also found that the RMSF values were quite low with the exception of the flap residues. As would be expected, the simulations with closing restraints demonstrated far less fluctuation of flap residues.

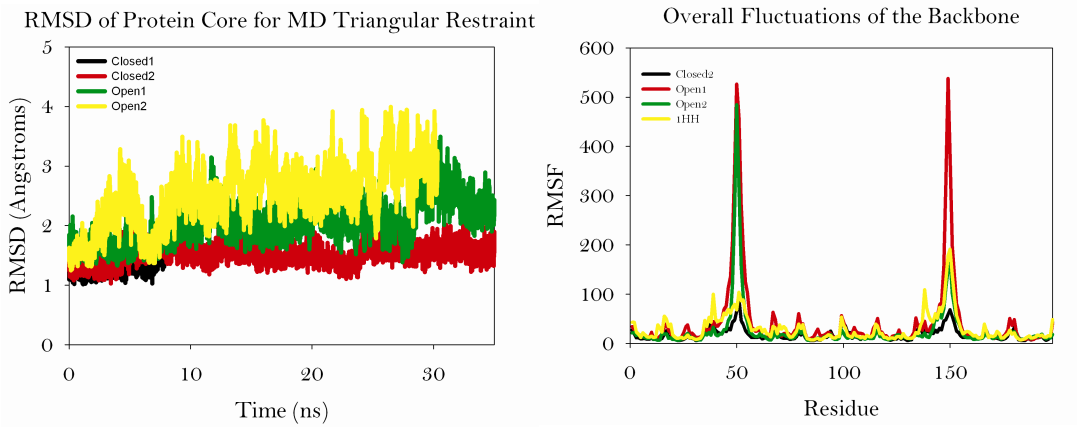

Figure A-6: The overall RMSD and RMSF calculated for the core (all residues but the flaps 43-58/43'-58') of HIVp over the length of the production run. The mean RMSD of the HIVp core for a simulation of apo 1HHP over 35ns was 1.92 ± 0.43Å. The core remains stable for the duration of the trajecories.

Our MD simulations were clustered into 15 families based on the last 10 ns of each 35 ns trajectory. The 85% most populated families were chosen to illustrate the most frequently occupied conformations present (Figure A-7). By the last 10 ns of the trajectory, the simulations were completely equilibrated and indicated the effect a small molecule bound between the cantilever-elbow-fulcrum regions may have on flap mobility. The combination of three distance restraints effectively kept the flaps in an open state or a closed state. The stronger force constant simulations demonstrated a more significant shift towards the desired flap conformation, particularly in the case of the closing restraint simulations.

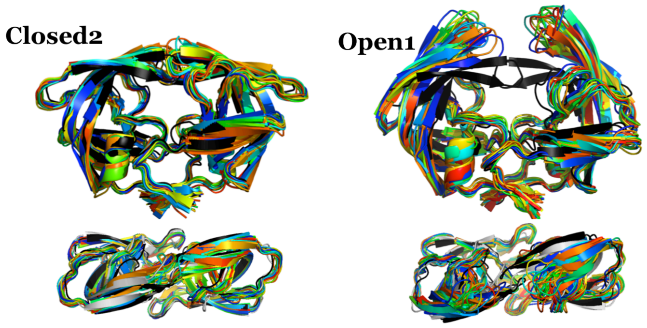

Figure A-7: The last 10 ns of each 35 ns production trajectory were clustered into 15 representative families based on the means algorithm in ptraj. These images show the front and top view of all 15 families from the closed restraint2 and open restraint1 MD simulations. These structures are overlaid with the crystal structure of a semi-open conformation (1HHP) in light gray and a closed conformation (1PRO) in black.

Ligand Mimic Simulations

We furthered our understanding of controlling flap motion through the elbow by constructing a model system with ligand mimics, allowing a small molecule to freely sample the elbow region, and testing for control over flap mobility. The ligand mimics SML, MED, LRG were introduced via AutoDock4 to the elbow region of 1HHP and 1PRO, according to our MPS model of closed HIVp. The SML and MED beads were meant to illustrate the effect of an allosteric flap opener[204] by causing the elbow region to contract and forcing the flaps to open, thereby elongating the time required for substrate binding. We hypothesized that the LRG beads would push the secondary structural features at the elbow site apart, similarly to the closed distance restraints, causing flap closing. The incorporation of explicit TIP3P waters allowed for realistic competition between solvent, ligand, and protein for forming favorable interactions.

Table A-3: The mean distances found for flap movement over the duration of our 35ns simulations initiated from the semi-open (1HHP) and closed (1PRO) conformations with ligand mimics LRG_5, MED_5 or SML_5 bound at each elbow region. The typical distance for Asp25/25' - Ile50/50' in a 35-ns simulation started from apo 1HHP was 23.32±1.03 and 23.56±0.71 Å. For Asp25/25' - COM Ile50&50' the average was 22.93±0.98 and 23.85±1.06 Å.

Simulation	Distance Asp25-Ile50 (Å)	Distance Asp25'-Ile50' (Å)	Distance Asp25-COM Ile50/50' (Å)	Distance Asp25'-COM Ile50/50' (Å)
1HHP LRG_5	13.99±1.95	13.52±1.93	14.28±1.53	13.62±1.58
1HHP MED_5	14.34±2.84	15.36±1.95	15.37±2.10	15.59±2.06
1HHP SML_5	19.83±4.10	18.32±3.47	18.44±2.85	20.59±3.59
1PRO LRG_5	13.23±1.50	13.25±0.96	12.20±1.27	13.02±1.13
1PRO MED_5	13.92±0.80	14.31±0.99	12.96±0.71	12.97±0.79
1PRO SML_5	14.83±0.79	15.45±1.20	14.34±0.80	13.28±0.84

For all ligand mimic simulations, the RMSD of the core region was compared to the position of the starting structure in order to ensure stability over the course of the simulation (Figure A-8). Our analysis showed that, excluding the flap region, there was relatively little deviation over the 35ns simulation time from the starting conformation. Distances from the COM of the flap tips to the catalytic Asp25/25' as well as the distance from flap tip Ile50/50' to catalytic Asp25/25' were also calculated and are presented in Table A-3. The differences between the extent of flap opening seen in simulations with 1PRO and 1HHP may indicate that MD initiation from the closed structure requires

longer simulation time for full exploration of the conformational space compared to simulations beginning from the semi-open structure.

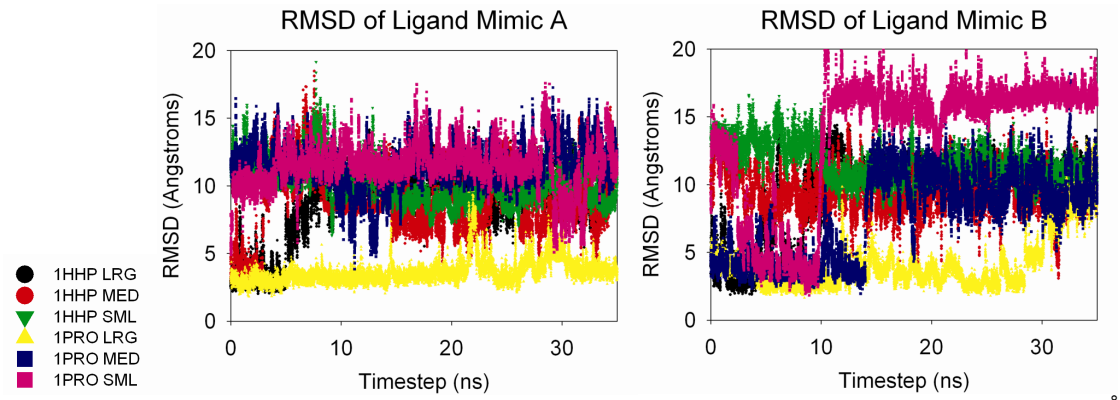

Figure A-8: Dissociation of ligand mimics from the bound position. An RMSD of approximately 11 Å or greater can indicate either dissociation from the elbow region or movement into a different area of the pocket. At an RMSD of about 15 Å, the ligand has dissociated from the elbow region.

Clustering of the last 10 ns from each 35 ns trajectory showed that the SML mimics had mostly dissociated from the elbow (Figure A-9). In the simulation of 1HHP-SML, the ligand mimics from both elbows dissociated and the structural symmetry was no longer preserved. Oddly, although the mimics have mostly dissociated in the 1PRO-SML simulations, the flaps are still closed. This could be explained by the position of one of the ligand mimics on the top of a flap. A crystallography study of darunavir identified a secondary binding site for the PI that the authors hypothesized could contribute to efficacy against MDR HIVp. One SML mimic was bound at that site, potentially explaining the continued flap closure.

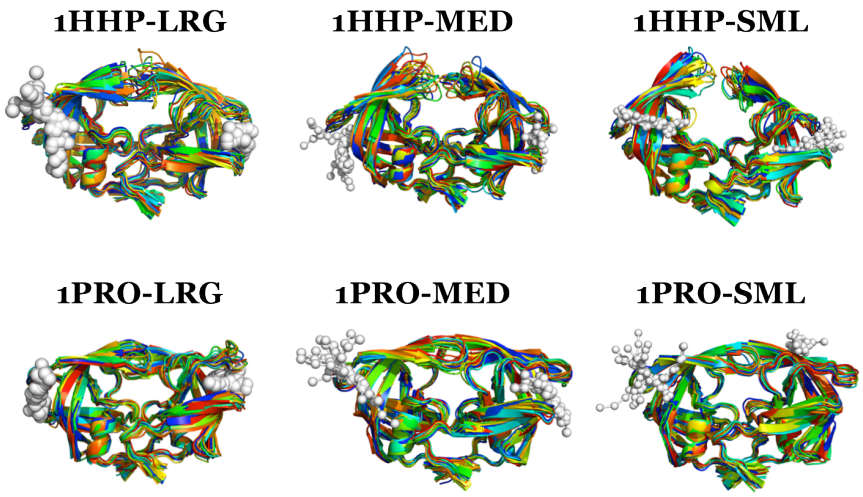

Figure A-9: Representative structures of the ligand mimic simulations obtained by clustering the trajectories of each run into 15 separate families to describe the conformations adopted by the protease when a ligand mimic is placed in the elbow region. Clustering was performed in ptraj according to the means algorithm on the final 10ns of the 35ns production run. The ligand mimics are displayed in white.

In our simulation of 1HHP-MED, the ligand mimics were still bound at the elbow region, although they shifted in position to occupy the space between the elbow and fulcrum. However, the flaps of this structure sampled the semi-open and open conformations and exhibited flap tip curling. For the 1PRO-MED system, on one side the ligand mimics have dissociated from the elbow, and on the other they have moved to occupy the cleft between the elbow and fulcrum. The flaps of this structure are closed.

The 1HHP-LRG and 1PRO-LRG runs showed greater binding stability at the elbow site. The flaps sampled a reduced number of semi-open conformations relative to the flexibility of the flaps seen in a simulation of 1HHP in explicit solvent. The simulation of 1PRO-LRG demonstrates higher stability of the ligand mimics and flaps than the 1HHP-LRG run. This could be because the simulation commenced in the closed structure, so the necessary space for binding was already available at the elbow region whereas starting from a semi-open conformation (narrower elbow cleft) required extra time for the protein to sample the space needed.

The SML and MED mimics were less likely to remain associated with the elbow region. Additionally, they were less capable of forcing flap opening. RMSD values for each ligand mimic were calculated compared to the starting docked conformation. However, these values did not accurately reflect the distance of the ligand mimic from

the binding site, as RMSDs of 8-13 Å could either indicate complete dissociation of the mimic from the elbow or a shift from occupying the V-shaped cleft between the cantilever-elbow-fulcrum to the cleft between the elbow-cantilever or elbow-fulcrum. Instead, the percent accessible surface area was calculated with NACCESS2.1.1. The accessible surface area for each family representative was calculated and then compared to the initial accessible surface area for the starting structure to gain an understanding of how our ligand mimic had moved in the elbow region.

These studies showed that it seems more tractable to keep the flaps closed with a ligand mimic inserted into the elbow region than to keep them open. These mimics did remain bound to the elbow region for a longer period; however, typically only one ligand mimic remained bound in the elbow region over time, while the other dissociated. This may indicate that the elbow region is an asymmetrical binding site.

Tethered Protease Simulations

The tethered HIVp form was the first evidence for conformational change of the protease flaps that was not driven by crystal packing. Therefore, sampling of the structure's energy landscape offered another approach for studying behavior modification of the flaps. As a result of the tether, the two monomers are no longer equivalent; each has a difference relationship to the tether. Behavior of monomer A reflects the constraints imposed by restraining the C-terminus of the monomer, while the behavior of monomer B reflects the effects of restraining the N-terminus of the monomer.

We found that throughout the entire 18 ns explicit MD simulation, tethered protease remained in its closed form, despite being unliganded. To further understand the potential flexibility of this system, we initiated several shorter (10 ns) implicit LD simulations. In our LD simulations, one flap covered the entire active site with its tip interacting at the eye site. The other flap sampled randomly. No open or semi-open flap conformations were seen in the trajectory snapshots from LD or explicit MD simulation.

Based on these results, we sought the corresponding distances that were the key for modifying flap behavior. To relate flexibility and conformation, cross-correlation plots were constructed from 500 ps – 1000 ps snapshots of the trajectory. These plots clearly illustrated the link between the β-sheets, the elbow region, and the flaps. Unfortunately,

how the conformational change demonstrated by these studies relates to inhibition is unclear given that the tethered mutant retains activity (~60%).

Conclusion

Our studies show that there is a V-shaped cleft made up of the space between three secondary structural elements: the cantilever, fulcrum and elbow. This region has been previously studied by applying one distance restraint to a pair of atoms at this region in all-atom MD and coarse-grained MD. We have expanded upon this work to demonstrate the relative importance of associations amongst the three secondary structures features at the proposed allosteric site. We found that distance restraints could both force flap closure as well as flap opening. This was the first time these three structural elements have been evaluated together to examine the relative significance of each sub-site within this region. We also showed that the small cleft between the fulcrum and cantilever is not as important for granting allosteric control over the flaps as the regions between the elbow and cantilever and the elbow and fulcrum.

We were the first to definitively show with a small molecule that flap closing is easier to achieve than flap opening. We found that flap mobility was affected by the range of flexibility available to the elbow region. From our studies with ligand mimics, and based on previous work in the Carlson lab, we believe the elbow region may be an asymmetrical binding site. Our ligand mimic simulations indicated that flap closing could be better controlled by a bigger ligand at the elbow site than flap opening could be by a small ligand. This makes intuitive sense, as it is generally easier to push two regions apart than encourage them to come closer together. After this research was completed, a similar study was published using carbon nanotube beads.[286]

Further work still needs to be performed to translate these studies into pharmaceutically-relevant small molecules that allosterically target the protease. Although the conformational restriction imposed by the tethered protease is quite interesting, the retained activity of the enzyme limits the applicability of our studies. However, we have shown that bulky small molecules do restrict flap movement when bound at the elbow region. Our development of simple 5-element van der Waals spheres

and our elucidation of the necessary elements to control at the elbow site will lead to more fruitful studies using specific small molecules to target the elbow region.

Bibliography

1. Leach AR. Molecular Modelling: Principles and Applications. Second ed: Prentice Hall; 2001.
2. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J Am Chem Soc. 1995;117(19):5179-97.
3. Verlet L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. Phys Rev. 1967;159(1):98.
4. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. J Comp Phys. 1977;23(3):327-41.
5. Birkhoff GD. Proof of the Ergodic Theorem. Proc Natl Acad Sci U S A. 1931;17(12):656-60. PMCID: PMC1076138.
6. Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, Merz KM, Jr., Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. J Comput Chem. 2005;26(16):1668-88.
7. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins. 2006;65(3):712-25.
8. Edinger SR, Cortis C, Shenkin PS, Friesner RA. Solvation free energies of peptides: Comparison of approximate continuum solvation models with accurate solution of the Poisson-Boltzmann equation. J Phys Chem B. 1997;101(7):1190-7.
9. Onufriev A, Bashford D, Case DA. Modification of the generalized Born model suitable for macromolecules. J Phys Chem B. 2000;104(15):3712-20.
10. Case DA, Darden TA, Cheatham I, T.E., Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA. AMBER 10. San Francisco.: University of California; 2008.
11. Jorgensen WL. The many roles of computation in drug discovery. Science (New York, NY. 2004;303(5665):1813-8.
12. Taft CA, Da Silva VB, Da Silva CH. Current topics in computer-aided drug design. J Pharm Sci. 2008;97(3):1089-98.
13. Talele TT, Khedkar SA, Rigby AC. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. Curr Top Med Chem. 2010;10(1):127-41.
14. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS. A critical assessment of docking programs and scoring functions. J Med Chem. 2006;49(20):5912-31.
15. Plewczynski D, Lazniewski M, Augustyniak R, Ginalski K. Can We Trust Docking Results? Evaluation of Seven Commonly Used Programs on PDBbind Database. J Comp Chem. 2011;32(4):742-55.
16. Leach AR, Shoichet BK, Peishoff CE. Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. J Med Chem. 2006;49(20):5851-5.
17. Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. Br J Pharmacol. 2008;153 Suppl 1:S7-26. PMCID: 2268060.
18. Englebienne P, Moitessier N. Docking ligands into flexible and solvated macromolecules. 4. Are popular scoring functions accurate for this class of proteins? J Chem Inf Model. 2009;49(6):1568-80.
19. Mukherjee S, Balius TE, Rizzo RC. Docking Validation Resources: Protein Family and Ligand Flexibility Experiments. J Chem Inf Model. 2010;50(11):1986-2000.
20. Huang SY, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. PCCP. 2010;12(40):12899-908.

21. Klebe G. Virtual ligand screening: strategies, perspectives and limitations. Drug Discov Today. 2006;11(13-14):580-94.
22. Zentgraf M, Steuber H, Koch C, La Motta C, Sartini S, Sotriffer CA, Klebe G. How reliable are current docking approaches for structure-based drug design? Lessons from aldose reductase. Angewandte Chemie. 2007;46(19):3575-8.
23. Velec HF, Gohlke H, Klebe G. DrugScore(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. J Med Chem. 2005;48(20):6296-303.
24. Teague SJ. Implications of protein flexibility for drug discovery. Nat Rev Drug Discov. 2003;2(7):527-41.
25. Teodoro ML, Kavraki LE. Conformational flexibility models for the receptor in structure based drug design. Curr Pharm Des. 2003;9(20):1635-48.
26. Sotriffer CA. Accounting for Induced-Fit Effects in Docking: What is Possible and What is Not? Curr Top Med Chem. 2011;11(2):179-91.
27. Alonso H, Bliznyuk AA, Gready JE. Combining docking and molecular dynamic simulations in drug design. Med Res Rev. 2006;26(5):531-68.
28. Sousa SF, Fernandes PA, Ramos MJ. Protein-ligand docking: current status and future challenges. Proteins. 2006;65(1):15-26.
29. Totrov M, Abagyan R. Flexible ligand docking to multiple receptor conformations: a practical alternative. Curr Opin Struct Biol. 2008;18(2):178-84. PMCID: 2396190.
30. Rao CB, Subramanian J, Sharma SD. Managing protein flexibility in docking and its applications. Drug Discov Today. 2009;14(7-8):394-400.
31. Durrant JD, McCammon JA. Computer-aided drug-discovery techniques that account for receptor flexibility. Curr Opin Pharma. 2010;10(6):770-4.
32. Fuentes G, Dastidar SG, Madhumalar A, Verma CS. Role of Protein Flexibility in the Discovery of New Drugs. Drug Dev Res. 2011;72(1):26-35.
33. Henzler AM, Rarey M. In Pursuit of Fully Flexible Protein-Ligand Docking: Modeling the Bilateral Mechanism of Binding. Molec Info. 2010;29(3):164-73.
34. Bottegoni G. Protein-ligand docking. Frontiers Biosci. 2011;16:2289-306.
35. Fischer E. Synthese des Traubenzuckers. Berichte der deutschen chemischen Gesellschaft. 1890;23(1):799-805.
36. Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. Proc Natl Acad Sci U S A. 1958;44(2):98-104. PMCID: 335371.
37. Tsai CJ, Ma B, Sham YY, Kumar S, Nussinov R. Structured disorder and conformational selection. Proteins. 2001;44(4):418-27.
38. Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R. Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci. 2000;9(1):10-9. PMCID: 2144430.
39. Tsai CJ, Ma B, Nussinov R. Folding and binding cascades: shifts in energy landscapes. Proc Natl Acad Sci U S A. 1999;96(18):9970-2. PMCID: 33715.
40. Berger C, Weber-Bornhauser S, Eggenberger J, Hanes J, Pluckthun A, Bosshard HR. Antigen recognition by conformational selection. FEBS Lett. 1999;450(1-2):149-53.
41. Foote J, Milstein C. Conformational isomerism and the diversity of antibodies. Proc Natl Acad Sci U S A. 1994;91(22):10370-4. PMCID: 45021.
42. Tsai CJ, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. Protein Sci. 1999;8(6):1181-90. PMCID: 2144348.
43. Sullivan SM, Holyoak T. Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. Proc Natl Acad Sci U S A. 2008;105(37):13829-34. PMCID: 2544539.
44. Weikl TR, von Deuster C. Selected-fit versus induced-fit protein binding: kinetic differences and mutational analysis. Proteins. 2009;75(1):104-10.
45. Hammes GG, Chang YC, Oas TG. Conformational selection or induced fit: a flux description of reaction mechanism. Proc Natl Acad Sci U S A. 2009;106(33):13737-41. PMCID: 2728963.
46. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Proteins. 2004;57(3):433-43.
47. Mobley DL, Dill KA. Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". Structure. 2009;17(4):489-98.

48. Murray CW, Baxter CA, Frenkel AD. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. J Comput Aided Mol Des. 1999;13(6):547-62.
49. Najmanovich R, Kuttner J, Sobolev V, Edelman M. Side-chain flexibility in proteins upon ligand binding. Proteins. 2000;39(3):261-8.
50. Smith LJ, Redfield C, Smith RA, Dobson CM, Clore GM, Gronenborn AM, Walter MR, Naganbushan TL, Wlodawer A. Comparison of four independently determined structures of human recombinant interleukin-4. Nat Struct Biol. 1994;1(5):301-10.
51. Eyal E, Gerzon S, Potapov V, Edelman M, Sobolev V. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. J Mol Biol. 2005;351(2):431-42.
52. Chopra G, Summa CM, Levitt M. Solvent dramatically affects protein structure refinement. Proc Natl Acad Sci U S A. 2008;105(51):20239-44. PMCID: 2600579.
53. Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J. The Protein Data Bank and the challenge of structural genomics. Nat Struct Biol. 2000;7 Suppl:957-9.
54. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct. 2000;29:291-325.
55. Schafferhans A, Klebe G. Docking ligands onto binding site representations derived from proteins built by homology modelling. J Mol Biol. 2001;307(1):407-27.
56. Philippopoulos M, Lim C. Exploring the dynamic information content of a protein NMR structure: comparison of a molecular dynamics simulation with the NMR and X-ray structures of Escherichia coli ribonuclease HI. Proteins. 1999;36(1):87-110.
57. Damm KL, Carlson HA. Exploring experimental sources of multiple protein conformations in structure-based drug design. J Am Chem Soc. 2007;129(26):8225-35.
58. Yang LW, Eyal E, Chennubhotla C, Jee J, Gronenborn AM, Bahar I. Insights into equilibrium dynamics of proteins from comparison of NMR and X-ray data with computational predictions. Structure. 2007;15(6):741-9. PMCID: 2760440.
59. Bolstad ES, Anderson AC. In pursuit of virtual lead optimization: the role of the receptor structure and ensembles in accurate docking. Proteins. 2008;73(3):566-80.
60. Laughton CA, Orozco M, Vranken W. COCO: a simple tool to enrich the representation of conformational variability in NMR structures. Proteins. 2009;75(1):206-16.
61. Jiang F, Kim SH. "Soft docking": matching of molecular surface cubes. J Mol Biol. 1991;219(1):79-102.
62. Ferrari AM, Wei BQ, Costantino L, Shoichet BK. Soft docking and multiple receptor conformations in virtual screening. J Med Chem. 2004;47(21):5076-84. PMCID: 1413506.
63. Taylor RD, Jewsbury PJ, Essex JW. FDS: flexible ligand and receptor docking with a continuum solvent model and soft-core energy function. J Comput Chem. 2003;24(13):1637-56.
64. Dunbrack RL, Jr., Karplus M. Backbone-dependent rotamer library for proteins. Application to side-chain prediction. J Mol Biol. 1993;230(2):543-74.
65. Leach AR. Ligand docking to proteins with discrete side-chain flexibility. J Mol Biol. 1994;235(1):345-56.
66. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009.
67. Kallblad P, Dean PM. Efficient conformational sampling of local side-chain flexibility. J Mol Biol. 2003;326(5):1651-65.
68. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol. 1996;261(3):470-89.
69. Abagyan R, Totrov M, Kuznetsov DA. ICM - A New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation. J Comp Chem. 1994;15(5):488-506.
70. Kairys V, Gilson MK. Enhanced docking with the mining minima optimizer: Acceleration and side-chain flexibility. J Comp Chem. 2002;23(16):1656-70.
71. Alberts IL, Todorov NP, Dean PM. Receptor flexibility in de novo ligand design and docking. J Med Chem. 2005;48(21):6585-96.
72. Alberts IL, Todorov NP, Kallblad P, Dean PM. Ligand docking and design in a flexible receptor site. QSAR Comb Sci. 2005;24(4):503-7.
73. Schnecke V, Kuhn LA. Virtual screening with solvation and ligand-induced complementarity. Perspect Drug Discov Des. 2000;20(1):171-90.

74. Zavodszky MI, Kuhn LA. Side-chain flexibility in protein-ligand binding: the minimal rotation hypothesis. Protein Sci. 2005;14(4):1104-14. PMCID: 2253453.
75. Zavodszky MI, Sanschagrin PC, Korde RS, Kuhn LA. Distilling the essential features of a protein surface for improving protein-ligand docking, scoring, and virtual screening. J Comp Mol Des. 2002;16(12):883-902.
76. Anderson AC, O'Neil RH, Surti TS, Stroud RM. Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. Chemistry & Biology. 2001;8(5):445-57.
77. Huang SY, Zou X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins. 2007;66(2):399-421.
78. Cavasotto CN, Abagyan RA. Protein flexibility in ligand docking and virtual screening to protein kinases. J Mol Biol. 2004;337(1):209-25.
79. Claussen H, Buning C, Rarey M, Lengauer T. FlexE: efficient molecular docking considering protein structure variations. J Mol Biol. 2001;308(2):377-95.
80. Bottegoni G, Kufareva I, Totrov M, Abagyan R. Four-Dimensional Docking: A Fast and Accurate Account of Discrete Receptor Flexibility in Ligand Docking. J Med Chem. 2009;52(2):397-406.
81. Corbeil CR, Englebienne P, Moitessier N. Docking ligands into flexible and solvated macromolecules. 1. Development and validation of FITTED 1.0. J Chem Inf Model. 2007;47(2):435-49.
82. Corbeil CR, Englebienne P, Yannopoulos CG, Chan L, Das SK, Bilimoria D, L'Heureux L, Moitessier N. Docking Ligands into flexible and solvated macromolecules. 2. Development and application of FITTED 1.5 to the virtual screening of potential HCV polymerase inhibitors. J Chem Inf Model. 2008;48(4):902-9.
83. Corbeil CR, Moitessier N. Docking ligands into flexible and solvated macromolecules. 3. Impact of input ligand conformation, protein flexibility, and water molecules on the accuracy of docking programs. J Chem Inf Model. 2009;49(4):997-1009.
84. Sherman W, Beard HS, Farid R. Use of an induced fit receptor structure in virtual screening. Chemical Biol & Drug Des. 2006;67(1):83-4.
85. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem. 2006;49(2):534-53.
86. Zacharias M. Rapid protein-ligand docking using soft modes from molecular dynamics simulations to account for protein deformability: binding of FK506 to FKBP. Proteins. 2004;54(4):759-67.
87. May A, Zacharias M. Protein-ligand docking accounting for receptor side chain and global flexibility in normal modes: evaluation on kinase inhibitor cross docking. J Med Chem. 2008;51(12):3499-506.
88. Wong S, Jacobson MP. Conformational selection in silico: loop latching motions and ligand binding in enzymes. Proteins. 2008;71(1):153-64.
89. Bottegoni G, Kufareva I, Totrov M, Abagyan R. A new method for ligand docking to flexible receptors by dual alanine scanning and refinement (SCARE). J Comput-Aid Mol Des. 2008;22(5):311-25.
90. Huang SY, Zou X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins. 2007;66(2):399-421.
91. Knegtel RM, Kuntz ID, Oshiro CM. Molecular docking to ensembles of protein structures. J Mol Biol. 1997;266(2):424-40.
92. Miranker A, Karplus M. Functionality maps of binding sites: a multiple copy simultaneous search method. Proteins. 1991;11(1):29-34.
93. Stultz CM, Karplus M. MCSS functionality maps for a flexible protein. Proteins. 1999;37(4):512-29.
94. Schubert CR, Stultz CM. The multi-copy simultaneous search methodology: a fundamental tool for structure-based drug design. J Comp Mol Des. 2009;23(8):475-89.
95. Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. Proteins. 2002;46(1):34-40.
96. Broughton HB. A method for including protein flexibility in protein-ligand docking: improving tools for database mining and virtual screening. J Mol Graphics Model. 2000;18(3):247-57, 302-4.
97. Kastenholz MA, Pastor M, Cruciani G, Haaksma EE, Fox T. GRID/CPCA: a new computational tool to design selective ligands. J Med Chem. 2000;43(16):3033-44.
98. Zentgraf M, Fokkens J, Sotriffer CA. Addressing protein flexibility and ligand selectivity by "in situ cross-docking". ChemMedChem. 2006;1(12):1355-9.
99. van Westen GJP, Wegner JK, Bender A, Ijzerman AP, van Vlijmen HWT. Mining protein dynamics from sets of crystal structures using "consensus structures". Protein Science. 2010;19(4):742-52.

100. Carlson HA, Masukawa KM, McCammon JA. Method for Including the Dynamic Fluctuations of a Protein in Computer-Aided Drug Design. J Phys Chem A. 1999;103:10213-9.
101. Carlson HA, Masukawa KM, Rubins K, Bushman FD, Jorgensen WL, Lins RD, Briggs JM, McCammon JA. Developing a dynamic pharmacophore model for HIV-1 integrase. J Med Chem. 2000;43(11):2100-14.
102. Bowman AL, Nikolovska-Coleska Z, Zhong H, Wang S, Carlson HA. Small molecule inhibitors of the MDM2-p53 interaction discovered by ensemble-based receptor models. J Am Chem Soc. 2007;129(42):12809-14.
103. Damm KL, Ung PM, Quintero JJ, Gestwicki JE, Carlson HA. A poke in the eye: Inhibiting HIV-1 protease through its flap-recognition pocket. Biopolymers. 2008;89(8):643-52.
104. Lin JH, Perryman AL, Schames JR, McCammon JA. Computational drug design accommodating receptor flexibility: The relaxed complex scheme. J Am Chem Soc. 2002;124(20):5632-3.
105. Lin JH, Perryman AL, Schames JR, McCammon JA. The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. Biopolymers. 2003;68(1):47-62.
106. McCammon JA. Target flexibility in molecular recognition. Biochim Biophys Acta. 2005;1754(1-2):221-4.
107. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE. Calculating Structures and Free Energies of Complex Molecules:‚Äâ Combining Molecular Mechanics and Continuum Models. Acc Chem Res. 2000;33(12):889-97.
108. Amaro RE, Baron R, McCammon JA. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. J Comp Mol Des. 2008;22(9):693-705.
109. Perryman AL, Lin JH, McCammon JA. Optimization and computational evaluation of a series of potential active site inhibitors of the V82F/I84V drug-resistant mutant of HIV-1 protease: An application of the relaxed complex method of structure-based drug design. Chem Biol & Drug Des. 2006;67(5):336-45.
110. Cheng LS, Amaro RE, Xu D, Li WW, Arzberger PW, McCammon JA. Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. J Med Chem. 2008;51(13):3878-94.
111. Babakhani A, Talley TT, Taylor P, McCammon JA. A virtual screening study of the acetylcholine binding protein using a relaxed-complex approach. Comput Biol and Chem. 2009;33(2):160-70.
112. Barakat K, Tuszynski J. Relaxed complex scheme suggests novel inhibitors for the lyase activity of DNA polymerase beta. J Mol Graph. 2010;29(5):702-16.
113. Barakat K, Mane J, Friesen D, Tuszynski J. Ensemble-based virtual screening reveals dual-inhibitors for the p53-MDM2/MDMX interactions. J Mol Graph. 2010;28(6):555-68.
114. Durrant JD, Keranen H, Wilson BA, McCammon JA. Computational Identification of Uncharacterized Cruzain Binding Sites. PLoS Neglected Tropical Diseases. 2010;4(5):-.
115. Meiler J, Baker D. ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. Proteins. 2006;65(3):538-48.
116. Davis IW, Baker D. RosettaLigand docking with full ligand and receptor flexibility. J Mol Biol. 2009;385(2):381-92.
117. Davis IW, Raha K, Head MS, Baker D. Blind docking of pharmaceutically relevant compounds using RosettaLigand. Protein Sci. 2009.
118. Rueda M, Bottegoni G, Abagyan R. Recipes for the selection of experimental protein conformations for virtual screening. J Chem Inf Model. 2010;50(1):186-93. PMCID: 2811216.
119. Barril X, Morley SD. Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. J Med Chem. 2005;48(13):4432-43.
120. May A, Sieker F, Zacharias M. How to efficiently include receptor flexibility during computational docking. Curr Comput-Aid Drug Des. 2008;4(2):143-53.
121. May A, Zacharias M. Accounting for global protein deformability during protein-protein and protein-ligand docking. Biochim Biophys Acta. 2005;1754(1-2):225-31.
122. Zacharias M, Sklenar H. Harmonic modes as variables to approximately account for receptor flexibility in ligand-receptor docking simulations: Application to DNA minor groove ligand complex. J Comput Chem. 1999;20(3):287-300.
123. Keseru GM, Kolossvary I. Fully flexible low-mode docking: application to induced fit in HIV integrase. J Am Chem Soc. 2001;123(50):12708-9.
124. Carcache LM, Rodriguez J, Rein KS. The structural basis for kainoid selectivity at AMPA receptors revealed by low-mode docking calculations. Bioorg Med Chem. 2003;11(4):551-9.

125. Cavasotto CN, Kovacs JA, Abagyan RA. Representing receptor flexibility in ligand docking through relevant normal modes. J Am Chem Soc. 2005;127(26):9632-40.
126. Rueda M, Bottegoni G, Abagyan R. Consistent improvement of cross-docking results using binding site ensembles generated with elastic network normal modes. J Chem Inf Model. 2009;49(3):716-25.
127. Sperandio O, Mouawad L, Pinto E, Villoutreix BO, Perahia D, Miteva MA. How to choose relevant multiple receptor conformations for virtual screening: a test case of Cdk2 and normal mode analysis. Eur Biophys J. 2010;39(9):1365-72.
128. Wei BQ, Weaver LH, Ferrari AM, Matthews BW, Shoichet BK. Testing a flexible-receptor docking algorithm in a model binding site. J Mol Biol. 2004;337(5):1161-82.
129. Craig IR, Essex JW, Spiegel K. Ensemble Docking into Multiple Crystallographically Derived Protein Structures: An Evaluation Based on the Statistical Analysis of Enrichments. J Chem Inf Model. 2010;50(4):511-24.
130. Frembgen-Kesner T, Elcock AH. Computational sampling of a cryptic drug binding site in a protein receptor: explicit solvent molecular dynamics and inhibitor docking to p38 MAP kinase. J Mol Biol. 2006;359(1):202-14.
131. Huang ZA, Wong CF. Docking Flexible Peptide to Flexible Protein by Molecular Dynamics Using Two Implicit-Solvent Models: An Evaluation in Protein Kinase and Phosphatase Systems. J Phys Chem B. 2009;113(43):14343-54.
132. Huang ZN, Wong CF, Wheeler RA. Flexible protein-flexible ligand docking with disrupted velocity simulated annealing. Proteins. 2008;71(1):440-54.
133. Bisson WH, Cheltsov AV, Bruey-Sedano N, Lin B, Chen J, Goldberger N, May LT, Christopoulos A, Dalton JT, Sexton PM, Zhang XK, Abagyan R. Discovery of antiandrogen activity of nonsteroidal scaffolds of marketed drugs. Proc Natl Acad Sci U S A. 2007;104(29):11927-32. PMCID: PMC1924583.
134. Bolstad ES, Anderson AC. In pursuit of virtual lead optimization: pruning ensembles of receptor structures for increased efficiency and accuracy during docking. Proteins. 2009;75(1):62-74. PMCID: 2649978.
135. Zhu J, Fan H, Liu H, Shi Y. Structure-based ligand design for flexible proteins: application of new F-DycoBlock. J Comp Mol Des. 2001;15(11):979-96.
136. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF. Protein flexibility predictions using graph theory. Proteins. 2001;44(2):150-65.
137. Zavodszky MI, Lei M, Thorpe MF, Day AR, Kuhn LA. Modeling correlated main-chain motions in proteins for flexible molecular recognition. Proteins. 2004;57(2):243-61.
138. Nabuurs SB, Wagener M, De Vlieg J. A flexible approach to induced fit docking. J Med Chem. 2007;50(26):6507-18.
139. Claussen H, Buning C, Rarey M, Lengauer T. Molecular docking into the flexible active site of aldose reductase using FLEXE. Holtje HD, Sippl W, editors2001.
140. Zhao Y, Sanner MF. FLIPDock: Docking flexible ligands into flexible receptors. Proteins. 2007;68(3):726-37.
141. Zhao Y, Sanner MF. Protein-ligand docking with multiple flexible side chains. J Comp Mol Des. 2008;22(9):673-9.
142. Popov VM, Yee WA, Anderson AC. Towards in silico lead optimization: scores from ensembles of protein/ligand conformations reliably correlate with biological activity. Proteins. 2007;66(2):375-87.
143. Ota N, Agard DA. Binding mode prediction for a flexible ligand in a flexible pocket using multi-conformation simulated annealing pseudo crystallographic refinement. J Mol Biol. 2001;314(3):607-17.
144. Mustata GI, Soares TA, Briggs JM. Molecular dynamics studies of alanine racemase: A structural model for drug design. Biopolymers. 2003;70(2):186-200.
145. Mustata GI, Briggs JM. A structure-based design approach for the identification of novel inhibitors: application to an alanine racemase. J Comput Aided Mol Des. 2002;16(12):935-53.
146. Leong MK. A novel approach using pharmacophore ensemble/support vector machine (PhE/SVM) for prediction of hERG liability. Chem Res Toxicol. 2007;20(2):217-26.
147. Leong MK, Chen TH. Prediction of cytochrome P450 2B6-substrate interactions using pharmacophore ensemble/support vector machine (PhE/SVM) approach. Med Chem. 2008;4(4):396-406.
148. Leong MK, Chen YM, Chen HB, Chen PH. Development of a new predictive model for interactions with human cytochrome P450 2A6 using pharmacophore ensemble/support vector machine (PhE/SVM) approach. Pharm Res. 2009;26(4):987-1000.
149. Subramanian J, Sharma S, B-Rao C. A novel computational analysis of ligand-induced conformational changes in the ATP binding sites of cyclin dependent kinases. J Med Chem. 2006;49(18):5434-41.

150. Subramanian J, Sharma S, B-Rao C. Modeling and selection of flexible proteins for structure-based drug design: Backbone and side chain movements in p38 MAPK. ChemMedChem. 2008;3(2):336-44.
151. Landon MR, Amaro RE, Baron R, Ngan CH, Ozonoff D, McCammon JA, Vajda S. Novel druggable hot spots in avian influenza neuraminidase H5N1 revealed by computational solvent mapping of a reduced and representative receptor ensemble. Chem Biol & Drug Des. 2008;71(2):106-16.
152. Firth-Clark S, Willems HM, Williams A, Harris W. Generation and selection of novel estrogen receptor ligands using the de novo structure-based design tool, SkelGen. J Chem Inf Model. 2006;46(2):642-7.
153. Jain AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. J Med Chem. 2003;46(4):499-511.
154. Jain AN. Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. J Comput Aided Mol Des. 2007;21(5):281-306.
155. Jain AN. Effects of protein conformation in docking: improved pose prediction through protein pocket adaptation. J Comp Mol Des. 2009;23(6):355-74.
156. Erlanson DA, McDowell RS, O'Brien T. Fragment-based drug discovery. J Med Chem. 2004;47(14):3463-82.
157. Mattos C, Ringe D. Locating and characterizing binding sites on proteins. Nat Biotechnol. 1996;14(5):595-9.
158. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. Science. 1996;274(5292):1531-4.
159. Wiesmann C, Barr KJ, Kung J, Zhu J, Erlanson DA, Shen W, Fahr BJ, Zhong M, Taylor L, Randal M, McDowell RS, Hansen SK. Allosteric inhibition of protein tyrosine phosphatase 1B. Nat Struct Mol Biol. 2004;11(8):730-7.
160. Howard S, Berdini V, Boulstridge JA, Carr MG, Cross DM, Curry J, Devine LA, Early TR, Fazal L, Gill AL, Heathcote M, Maman S, Matthews JE, McMenamin RL, Navarro EF, O'Brien MA, O'Reilly M, Rees DC, Reule M, Tisi D, Williams G, Vinkovic M, Wyatt PG. Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent aurora kinase activity. J Med Chem. 2009;52(2):379-88.
161. Murray CW, Rees DC. The rise of fragment-based drug discovery. Nat Chem. 2009;1(3):187-92.
162. English AC, Done SH, Caves LSD, Groom CR, Hubbard RE. Locating interaction sites on proteins: The crystal structure of thermolysin soaked in 2% to 100% isopropanol. Proteins. 1999;37(4):628-40.
163. English AC, Groom CR, Hubbard RE. Experimental and computational mapping of the binding surface of a crystalline protein. Protein Eng. 2001;14(1):47-59.
164. Schmitke JL, Stern LJ, Klibanov AM. The crystal structure of subtilisin Carlsberg in anhydrous dioxane and its comparison with those in water and acetonitrile. Proc Natl Acad Sci U S A. 1997;94(9):4250-5.
165. Schmitke JL, Stern LJ, Klibanov AM. Comparison of x-ray crystal structures of an acyl-enzyme intermediate of subtilisin Carlsberg formed in anhydrous acetonitrile and in water. Proc Natl Acad Sci U S A. 1998;95(22):12918-23.
166. Fitzpatrick PA, Steinmetz AC, Ringe D, Klibanov AM. Enzyme crystal structure in a neat organic solvent. Proc Natl Acad Sci U S A. 1993;90(18):8653-7. PMCID: 47416.
167. Dechene M, Wink G, Smith M, Swartz P, Mattos C. Multiple solvent crystal structures of ribonuclease A: an assessment of the method. Proteins. 2009;76(4):861-81.
168. Fedorov AA, Joseph-McCarthy D, Fedorov E, Sirakova D, Graf I, Almo SC. Ionic interactions in crystalline bovine pancreatic ribonuclease A. Biochemistry. 1996;35(50):15962-79.
169. Ho WC, Luo C, Zhao KH, Chai XM, Fitzgerald MX, Marmorstein R. High-resolution structure of the p53 core domain: implications for binding small-molecule stabilizing compounds. Acta Crystallogr Sect D-Biol Crystallogr. 2006;62:1484-93.
170. Wang Z, Zhu G, Huang Q, Qian M, Shao M, Jia Y, Tang Y. X-ray studies on cross-linked lysozyme crystals in acetonitrile-water mixture. Biochimica et biophysica acta. 1998;1384(2):335-44.
171. Buhrman G, de Serrano V, Mattos C. Organic solvents order the dynamic switch II in Ras crystals. Structure. 2003;11(7):747-51.
172. JosephMcCarthy D, Fedorov AA, Almo SC. Comparison of experimental and computational functional group mapping of an RNase A structure: Implications for computer-aided drug design. Protein Eng. 1996;9(9):773-80.
173. Clark M, Guarnieri F, Shkurko I, Wiseman J. Grand canonical Monte Carlo simulation of ligand-protein binding. J Chem Info Model. 2006;46(1):231-42.

174. Moore WR, Jr. Maximizing discovery efficiency with a computationally driven fragment approach. Curr Opin Drug Disc Devel. 2005;8(3):355-64.
175. Clark M, Meshkat S, Wiseman JS. Grand canonical free-energy calculations of protein-ligand binding. J Chem Info Model. 2009;49(4):934-43.
176. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, Landon MR, Mattos C, Vajda S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. Bioinformatics. 2009;25(5):621-7. PMCID: 2647826.
177. Seco J, Luque FJ, Barril X. Binding site detection and druggability index from first principles. J Med Chem. 2009;52(8):2363-71.
178. Yang C, Wang S. Computational Analysis of Protein Hotspots. ACS Med Chem Lett. 2010;1:125-9.
179. Guvench O, MacKerell AD, Jr. Computational fragment-based binding site identification by ligand competitive saturation. PLoS Comput Biol. 2009;5(7):e1000435. PMCID: 2700966.
180. Raman EP, Yu W, Guvench O, Mackerell AD. Reproducing Crystal Binding Modes of Ligand Functional Groups Using Site-Identification by Ligand Competitive Saturation (SILCS) Simulations. J Chem Inf Model. 2011;51(4):877-96. PMCID: PMC3090225.
181. Lexa KW, Carlson HA. Full Protein Flexibility Is Essential for Proper Hot-Spot Mapping. J Am Chem Soc. 2011;133:200-2. PMCID: PMC3081398.
182. Gallo RC. A reflection on HIV/AIDS research after 25 years. Retrovirology. 2006;3:72.
183. World Health Organization U. AIDS epidemic update. 2009.
184. UNAIDS. UNAIDS Report on the Global AIDS Epidemic: UNAIDS; 2010.
185. Kohl NE, Emini EA, Schleif WA, Davis LJ, Heimbach JC, Dixon RA, Scolnick EM, Sigal IS. Active human immunodeficiency virus protease is required for viral infectivity. Proc Natl Acad Sci U S A. 1988;85(13):4686-90.
186. Seelmeier S, Schmidt H, Turk V, von der Helm K. Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. Proc Natl Acad Sci U S A. 1988;85(18):6612-6.
187. Eder J, Hommel U, Cumin F, Martoglio B, Gerhartz B. Aspartic proteases in drug discovery. Curr Pharm Des. 2007;13(3):271-85.
188. McKeage K, Perry CM, Keam SJ. Darunavir: a review of its use in the management of HIV infection in adults. Drugs. 2009;69(4):477-503.
189. Mehellou Y, De Clercq E. Twenty-six years of anti-HIV drug discovery: where do we stand and where do we go? J Med Chem.53(2):521-38.
190. Perryman AL, Lin JH, McCammon JA. Restrained molecular dynamics simulations of HIV-1 protease: the first step in validating a new target for drug design. Biopolymers. 2006;82(3):272-84.
191. Vega S, Kang LW, Velazquez-Campoy A, Kiso Y, Amzel LM, Freire E. A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease. Proteins. 2004;55(3):594-602.
192. Muzammil S, Armstrong AA, Kang LW, Jakalian A, Bonneau PR, Schmelmer V, Amzel LM, Freire E. Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations. J Virol. 2007;81(10):5144-54.
193. Lapatto R, Blundell T, Hemmings A, Overington J, Wilderspin A, Wood S, Merson JR, Whittle PJ, Danley DE, Geoghegan KF, et al. X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes. Nature. 1989;342(6247):299-302.
194. Navia MA, Fitzgerald PM, McKeever BM, Leu CT, Heimbach JC, Herber WK, Sigal IS, Darke PL, Springer JP. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature. 1989;337(6208):615-20.
195. Wlodawer A, Miller M, Jaskolski M, Sathyanarayana BK, Baldwin E, Weber IT, Selk LM, Clawson L, Schneider J, Kent SB. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science. 1989;245(4918):616-21.
196. Hornak V, Simmerling C. Targeting structural flexibility in HIV-1 protease inhibitor binding. Drug Discov Today. 2007;12(3-4):132-8.
197. Freedberg DI, Ishima R, Jacob J, Wang YX, Kustanovich I, Louis JM, Torchia DA. Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamics calculations. Protein Sci. 2002;11(2):221-32.
198. Ishima R, Freedberg DI, Wang YX, Louis JM, Torchia DA. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. Structure. 1999;7(9):1047-55.

199. Nicholson LK, Yamazaki T, Torchia DA, Grzesiek S, Bax A, Stahl SJ, Kaufman JD, Wingfield PT, Lam PY, Jadhav PK, et al. Flexibility and function in HIV-1 protease. Nat Struct Biol. 1995;2(4):274-80.
200. Toth G, Borics A. Closing of the flaps of HIV-1 protease induced by substrate binding: a model of a flap closing mechanism in retroviral aspartic proteases. Biochem. 2006;45(21):6606-14.
201. Toth G, Borics A. Flap opening mechanism of HIV-1 protease. J Mol Graph Model. 2006;24(6):465-74.
202. Ding F, Layten M, Simmerling C. Solution Structure of HIV-1 Protease Flaps Probed by Comparison of Molecular Dynamics Simulation Ensembles and EPR Experiments. J Am Chem Soc. 2008.
203. Wlodawer A, Erickson JW. Structure-based inhibitors of HIV-1 protease. Annu Rev Biochem. 1993;62:543-85.
204. Perryman AL, Lin JH, McCammon JA. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs. Protein Sci. 2004;13(4):1108-23.
205. Trylska J, Tozzini V, Chang CE, McCammon JA. HIV-1 protease substrate binding and product release pathways explored with coarse-grained molecular dynamics. Biophys J. 2007;92(12):4179-87.
206. Chang CE, Trylska J, Tozzini V, McCammon JA. Binding pathways of ligands to HIV-1 protease: coarse-grained and atomistic simulations. Chem Biol Drug Des. 2007;69(1):5-13.
207. Rick SW, Erickson JW, Burt SK. Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease. Proteins. 1998;32(1):7-16.
208. Muzammil S, Ross P, Freire E. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. Biochemistry. 2003;42(3):631-8.
209. Weber IT. Comparison of the crystal structures and intersubunit interactions of human immunodeficiency and Rous sarcoma virus proteases. J Biol Chem. 1990;265(18):10492-6.
210. Zhang ZY, Poorman RA, Maggiora LL, Heinrikson RL, Kezdy FJ. Dissociative inhibition of dimeric enzymes. Kinetic characterization of the inhibition of HIV-1 protease by its COOH-terminal tetrapeptide. J Biol Chem. 1991;266(24):15591-4.
211. Meagher KL, Carlson HA. Incorporating protein flexibility in structure-based drug discovery: using HIV-1 protease as a test case. J Am Chem Soc. 2004;126(41):13276-81.
212. Bowman AL, Lerner MG, Carlson HA. Protein flexibility and species specificity in structure-based drug discovery: dihydrofolate reductase as a test system. J Am Chem Soc. 2007;129(12):3634-40.
213. Meagher KL, Lerner MG, Carlson HA. Refining the multiple protein structure pharmacophore method: consistency across three independent HIV-1 protease models. J Med Chem. 2006;49(12):3478-84.
214. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235-42.
215. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983;79(2):926-35.
216. Grabuleda X, Jaime C, Kollman PA. Molecular dynamics simulation studies of liquid acetonitrile: New six-site model. J Comp Chem. 2000;21(10):901-8.
217. Andrea TA, Swope WC, Andersen HC. The role of long ranged forces in determining the structure and properties of liquid water. J Chem Phys. 1983;79:4576.
218. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. J Comp Chem. 2004;25(13):1605-12.
219. Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. Acta Cryst. 2003;D59:1131-7.
220. Aburi M, Smith PE. A Combined Simulation and Kirkwood-Buff Approach to Quantify Cosolvent Effects on the Conformational Preferences of Peptides in Solution. J Phys Chem B. 2004;108(22):7382-8.
221. Allen KN, Bellamacina CR, Ding X, Jeffery CJ, Mattos C, Petsko GA, Ringe D. An Experimental Approach to Mapping the Binding Surfaces of Crystalline Proteins. J Phys Chem. 1996;100(7):2605-11.
222. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. Science. 1995;267(5196):383-6.
223. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol. 1998;280(1):1-9.
224. Murray CW, Rees DC. The rise of fragment-based drug discovery. Nat Chem. 2009;1(3):187-92.
225. Howard S, Berdini V, Boulstridge JA, Carr MG, Cross DM, Curry J, Devine LA, Early TR, Fazal L, Gill AL, Heathcote M, Maman S, Matthews JE, McMenamin RL, Navarro EF, O'Brien MA, O'Reilly M, Rees DC, Reule M, Tisi D, Williams G, Vinkovic M, Wyatt PG. Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent aurora kinase activity. J Med Chem. 2009;52(2):379-88.

226. English AC, Done SH, Caves LS, Groom CR, Hubbard RE. Locating interaction sites on proteins: the crystal structure of thermolysin soaked in 2% to 100% isopropanol. Proteins. 1999;37(4):628-40.
227. Mattos C, Bellamacina CR, Peisach E, Pereira A, Vitkup D, Petsko GA, Ringe D. Multiple solvent crystal structures: probing binding sites, plasticity and hydration. J Mol Biol. 2006;357(5):1471-82.
228. Deshpande A, Nimsadkar S, Mande SC. Effect of alcohols on protein hydration: crystallographic analysis of hen egg-white lysozyme in the presence of alcohols. Acta Crystallogr D Biol Crystallogr. 2005;61(Pt 7):1005-8.
229. Hobrath JV, Wang SM. Computational elucidation of the structural basis of ligand binding to the dopamine 3 receptor through docking and homology modeling. J Med Chem. 2006;49(15):4470-6.
230. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, 3rd, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 2007;35(Web Server issue):W375-83. PMCID: 1933162.
231. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J Am Chem Soc. 1996;118:11225-36.
232. Darden T, York D, Pedersen L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. J Chem Phys. 1993;98(12):10089-92.
233. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):235-42.
234. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vonrhein C, Womack TO. Buster. 1.6.0 ed. Cambridge, United Kingdom: Global Phasing Ltd.; 2011.
235. Kleywegt GJ, Harris MR, Zou J-y, Taylor TC, Wahlby A, Jones TA. The Uppsala Electron-Density Server. Acta Crystallogr D Biol Crystallogr. 2004;60(12 Part 1):2240-9.
236. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501. PMCID: PMC2852313.
237. DeLano WL. The PyMOL Molecular Graphics System. Palo Alto, CA: DeLano Scientific; 2008.
238. Englert L, Silber K, Steuber H, Brass S, Over B, Gerber H-D, Heine A, Diederich WE, Klebe G. Fragment-Based Lead Discovery: Screening and Optimizing Fragments for Thermolysin Inhibition. ChemMedChem.5(6):930-40.
239. Wang SM, Yang CY. Computational Analysis of Protein Hotspots. ACS Med Chem Lett. 2010;1(3):125-9.
240. Arnold U, Leich F, Neumann P, Lilie H, Ulbrich-Hofmann R. Crystal structure of RNase A tandem enzymes and their interaction with the cytosolic ribonuclease inhibitor. FEBS J. 2011;278(2):331-40.
241. Böttcher J, Blum A, Dörr S, Heine A, Diederich WE, Klebe G. Targeting the open-flap conformation of HIV-1 protease with pyrrolidine-based inhibitors. ChemMedChem. 2008;3(9):1337-44.
242. Perryman AL, Zhang Q, Soutter HH, Rosenfield R, McRee DE, Olson AJ, Elder JE, Stout CD. Fragment-Based Screen against HIV Protease. Chem Biol Drug Des. 2010;75(3):257-68.
243. Meagher KL, Carlson HA. Solvation influences flap collapse in HIV-1 protease. Proteins. 2005;58(1):119-25.
244. MOE. Montreal, Canada: Chemical Computing Group, Inc.; 2008.
245. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem. 2004;25(9):1157-74.
246. Jakalian A, Bush BL, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method. J Comput Chem. 2000;21(2):132-46.
247. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc Natl Acad Sci U S A. 2001;98(18):10037-41.
248. Lerner M. APBS plugin. Ann Arbor, MI2004.
249. Shao J, Tanner SW, Thompson N, Cheatham TE. Clustering molecular dynamics trajectories: I. Characterizing the performance of different clustering algorithms. J Chem Theory Comp. 2007;3:2312-34.
250. Davies DL, Bouldin DW. A Cluster Separation Measure. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 1979;PAMI-1(2):224-7.
251. Calinski T, Harabasz J. A dendrite method for cluster analysis. Comm in Stat. 1974;3(1):1-27.
252. Martin P, Vickrey JF, Proteasa G, Jimenez YL, Wawrzak Z, Winters MA, Merigan TC, Kovari LC. "Wide-open" 1.3 A structure of a multidrug-resistant HIV-1 protease as a drug target. Structure. 2005;13(12):1887-95.

253. Layten M, Hornak V, Simmerling C. The Open Structure of a Multi-Drug-Resistant HIV-1 Protease is Stabilized by Crystal Packing Contacts. J Am Chem Soc. 2006;128(41):13360-1.
254. Hubbard SJ, Thornton JM. NACCESS. 2.1.1 ed. London: Department of Biochemistry and Molecular Biology, Department of Biochemistry and Molecular Biology. 1993. Calculates solvent accessible surface area.
255. Hornak V, Okur A, Rizzo RC, Simmerling C. HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state. J Am Chem Soc. 2006;128(9):2812-3.
256. Hornak V, Okur A, Rizzo RC, Simmerling C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. Proc Natl Acad Sci U S A. 2006;103(4):915-20.
257. DeLano WL. The PyMOL Molecular Graphics System. San Carlos, CA, USA.: DeLano Scientific; 2002.
258. MOE (Molecular Operating Environment). 2005.06 ed. Montreal, Canada: Chemical Computing Group; 2005.
259. Case DA, Darden TA, Cheatham TE, III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister C, Caldwell J, Ross W, Kollman P. AMBER 8. version 8 ed. San Francisco, CA2004.
260. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins. 2006;65(3):712-25.
261. Feig M, Karanicolas J, Brooks CL, 3rd. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graph Model. 2004;22(5):377-95.
262. Matayoshi ED, Wang GT, Krafft GA, Erickson J. Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. Science. 1990;247(4945):954-8.
263. Toth MV, Marshall GR. A simple, continuous fluorometric assay for HIV protease. Int J Pept Protein Res. 1990;36(6):544-50.
264. Babine RE, Bender SL. Molecular Recognition of Protein-Ligand Complexes: Applications to Drug Design. Chem Rev. 1997;97(5):1359-472.
265. Tozzini V, McCammon JA. A coarse grained model for the dynamics of flap opening in HIV-1 protease. Chem Phys Lett. 2005;413:123-8.
266. Tozzini V, Trylska J, Chang CE, McCammon JA. Flap opening dynamics in HIV-1 protease explored with a coarse-grained model. J Struct Biol. 2006.
267. Harte WE, Jr., Beveridge DL. Probing structure-function relationships in human immunodeficiency virus type 1 protease via molecular dynamics simulation. Method Enz. 1994;241:178-95.
268. Harte WE, Jr., Swaminathan S, Beveridge DL. Molecular dynamics of HIV-1 protease. Proteins. 1992;13(3):175-94.
269. Harte WE, Jr., Swaminathan S, Mansuri MM, Martin JC, Rosenberg IE, Beveridge DL. Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. Proc Natl Acad Sci U S A. 1990;87(22):8864-8.
270. Swaminathan S, Harte WE, Jr., Beveridge DL. Investigation of Domain Structure in Proteins via Molecular Dynamics Simulation: Application to HIV-1 Protease Dimer. J Am Chem Soc. 1991;113:2717-21.
271. Piana S, Carloni P, Parrinello M. Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease. J Mol Biol. 2002;319(2):567-83.
272. Carlson HA, McCammon JA. Accommodating protein flexibility in computational drug design. Mol Pharmacol. 2000;57(2):213-8.
273. Carlson HA, Masukawa KM, McCammon JA. Method for including the dynamic fluctuations of a protein in computer-aided drug design. J Phys Chem A. 1999;103:10213-19.
274. Spinelli S, Liu QZ, Alzari PM, Hirel PH, Poljak RJ. The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. Biochimie. 1991;73(11):1391-6.
275. Yamazaki T, Hinck AP, Wang YX, Nicholson LK, Torchia DA, Wingfield P, Stahl SJ, Kaufman JD, Chang CH, Domaille PJ, Lam PY. Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy. Protein Sci. 1996;5(3):495-506.
276. Pillai B, Kannan KK, Hosur MV. 1.9 angstrom X-ray study shows closed flap conformation in crystals of tethered HIV-1PR. Proteins. 2001;43(1):57-64.
277. Bhat TN, Baldwin ET, Liu B, Cheng YS, Erickson JW. Crystal structure of a tethered dimer of HIV-1 proteinase complexed with an inhibitor. Nat Struct Biol. 1994;1(8):552-6.

278. Berman H, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. Nat Struct Biol. 2003;10(12):980.
279. DeLano WL. The PyMOL Molecular Graphics System. 0.99 ed2006.
280. Jorgensen WL. BOSS. New Haven: Yale University Press; 1997.
281. Jorgensen W. BOSS. 4.2 ed. New Haven, CT: Yale University; 2000.
282. Lerner MG, Meagher KL, Carlson HA. Automated clustering of probe molecules from solvent mapping of protein surfaces: new algorithms applied to hot-spot mapping and structure-based drug design. J Comput Aided Mol Des. 2008;22(10):727-36.
283. Damm KL, Carlson HA. Gaussian-weighted RMSD superposition of proteins: a structural comparison for flexible proteins and predicted protein structures. Biophys J. 2006;90(12):4558-73.
284. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. J Comp Chem. 1998;19:1639-62.
285. Turner GW, Tedesco E, Harris KDM, Johnston RL, Kariuki BM. Implementation of Lamarckian concepts in a Genetic Algorithm for structure solution from powder diffraction data. Chem Phys Lett. 2000;321:183-90.
286. Cheng Y, Li D, Ji B, Shi X, Gao H. Structure-based design of carbon nanotubes as HIV-1 protease inhibitors: Atomistic and coarse-grained simulations. J Mol Graph Model. 2010;29(2):171-7.