Flap dynamics in pepsin-like aspartic proteases: a computational perspective using Plasmepsin-II and BACE-1 as model systems

Soumendranath Bhakat* and Pär Söderhjelm*

Division of Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, P.O. Box 124, SE-22100 Lund, Sweden

E-mail: bhakatsoumendranath@gmail.com; par.soderhjelm@bpc.lu.se

Phone: +46-(0)-46-2228161

Abstract

Flexibility of β hairpin structure known as *flap* plays a key role in catalytic activity and substrate intake in pepsin-like aspartic proteases. Most of these enzymes share structural and sequence similarity. Tyrosine is a conserved residue present in the flap region of pepsin-like aspartic proteases. In apo protease, tyrosine remains in a dynamic equilibrium between normal and flipped states due to rotation of χ_1 angle (distributions of the χ_1 angle centred around $+\frac{\pi}{3}$ *radian* or $-\frac{\pi}{3}$ *radian* are denoted as *normal*, whereas distribution centred around $\pm\pi$ *radian* is denoted as *flipped*). In this study, we have used apo Plm-II and BACE-1 as model systems. Independent MD simulations starting with Plm-II and BACE-1 remained stuck either in normal or flipped state. Metadynamics simulations using torsion angles (χ_1 and χ_2 of Tyr) as CVs sampled transition between normal and flipped states. Qualitatively, flipped and normal states predicted to be equally populated. These states were stabilised by H-bond interactions to tryptophan (normal) and catalytic aspartate (flipped) respectively.

Further, mutation of tyrosine to an amino-acid with smaller side-chain, such as alanine; reduced flap flexibility and resulted in a flap collapse. This is in accordance with previous experimental studies which showed that mutation to alanine resulted in loss of activity in pepsin-like aspartic proteases. Using apo plasmepsin-II and BACE-1 as model systems, we have hypothesised that the rotation of tyrosine side-chain is the key movement which governs the flap dynamics in all pepsin-like aspartic proteases.

Keywords: aspartic protease; flap; tyrosine; molecular dynamics; metadynamics

Introduction

Aspartic proteases play an important role in the life cycle of different pathogens and therefore emerged as one of the hot targets in the modern-day drug discovery. One of the main structural features of aspartic proteases is the presence of the β-hairpin-like conformation often termed as *flap*.[1] The flap is a highly flexible part which plays a critical role in the ligand binding by displaying a scissor-like motion necessary for ligand uptake (Figure 1). The conformational flexibility of aspartic proteases and the role of the flap in ligand binding were studied in several systems, e.g. HIV proteases,[2,3] cathepsin-D[4] and BACE-1.[5–7] Pepsin-like aspartic proteases[8] are a subset of aspartic proteases which has conserved *flap* and *coil* region (Figure 1). The flap acts a lid which covers the active site of these enzymes. Tyrosine (Tyr) is the conserved residue present in the *flap* of most of the pepsin-like aspartic proteases e.g. plasmepsin I, II, IV, human cathepsin-D, cathepsin-E, BACE-1, BACE-2 etc. It is believed that the rotation of Tyr side-chain plays an important role in the flap dynamics of these enzymes.

Plasmepsins (*Plm*) are a group of pepsin-like aspartic proteases present in *P. falciparum* (the parasite which causes Malaria) and expressed by ten different genes, *Plm* I, II, III, IV, V, VI, VII, IX, X and HAP. Due to the *neglected* status of Malaria, the conformational flexibility and structural biology of Plm has not generated much attention among structural and computational biologists when compared with HIV protease.[2,3,9,10] Inspired by the idea

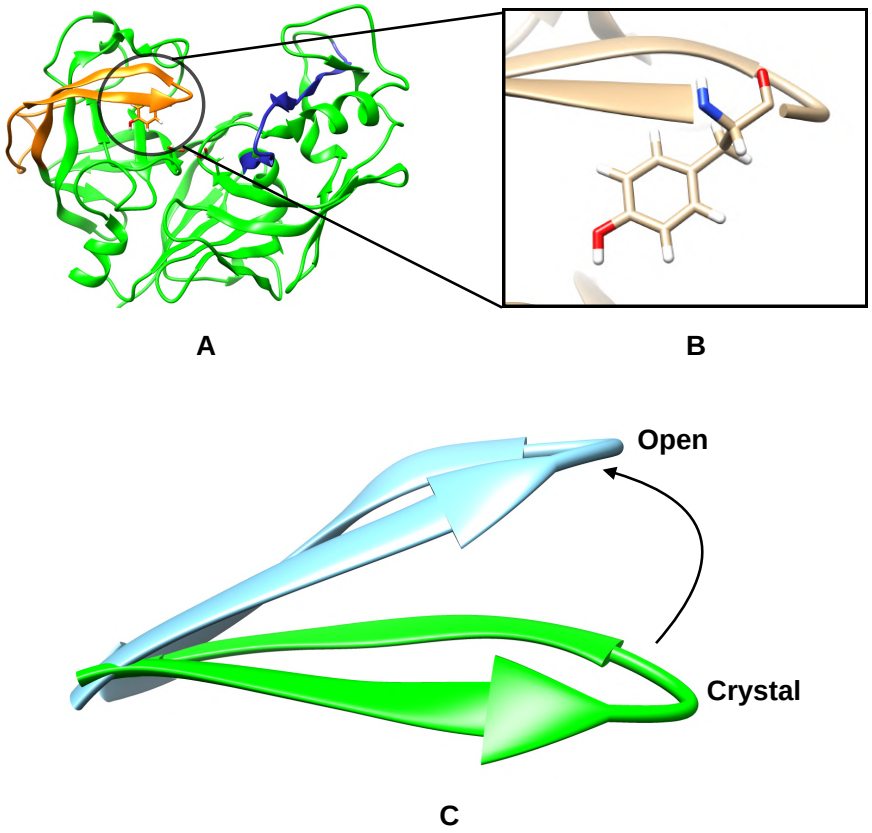

Figure 1: Flap (**Orange**) and coil (**Blue**) region of pepsin-like aspartic protease and location of the conserved tyrosine (Tyr) present in the flap (B). Difference in flap conformation in *crystal* and *open* state (C).

that the flap of HIV-protease spontaneously opens and close in MD simulation,[3] *Karubiu et al.*[11] used MD simulations and devised some distance parameters to understand the flap opening in Plm II. Further, *McGillewie et al.*[12,13] used the same distance parameters to understand the flap dynamics and conformational flexibility in Plm I, III, IV and V. *Friedman and Caflisch*[14] studied the flexibility of Plm II using MD simulation and hinted that the conserved tyrosine (Tyr) residue present at the flap region might play a role in flap flexibility. *Spronk and Carlson*[7] performed classical MD simulations on a homologous protein, BACE-1 and predicted that the side-chain fluctuation of Tyr governs different flap conformations. They reported three different conformations involving H-bond interactions between Tyr and three neighbouring amino acids (Trp, Ser and Asp). MD simulations also sampled occurrence

of a *self-inhibited* conformation where the Tyr points toward the catalytic region, known as the *binding site*. The presence of this orientation is not experimentally reported in any BACE-1 (also not in apo Plm-II) crystal structures to date. However, it was observed in crystal structures of other aspartic proteases such as chymosin and saccharopepsin ([15]). Due to lack of transitions among these conformations in MD simulation, it was difficult to estimate free energy difference among different conformations within reasonable time-scale.

Enhanced sampling methods such as metadynamics,[16–18] replica-exchange molecular dynamics (REMD),[19] umbrella sampling,[20] conformational flooding[21] etc. have been used regularly to overcome the sampling problem associated with MD simulation. Metadynamics is an enhanced sampling method,[16–18] which uses a repulsive bias potential along some reaction co-ordinates (also known as collective variables). Addition of bias potential pushes the system away from local free energy minima and accelerates sampling of conformational space .[22]

The aim of our study was to use MD and metadynamics simulations to understand conformational dynamics in pepsin-like aspartic proteases. Most of these proteases possesses structural and sequence similarity. Hence, understanding the conformational dynamics associated with a few of these enzymes can be applied to other homologous enzymes. We have used two homologous proteases, Plm-II (Plasmepsin-II) and BACE-1 (β secretase-1) as models to answer two fundamental questions: 1. what is role of the Tyr residue in flap dynamics? and 2. does rotation of χ_1 and χ_2 torsional angles of Tyr dictate different flap orientations? We have tested several different collective variables (CVs) within the metadynamics framework and discussed their effectiveness in sampling different flap conformations. We have also tested the effect of force fields and water models on the population of different flap conformations and suggested possible directions of future works.

Results

Flipping of the tyrosine side-chain

Based on earlier investigations, we expected the conserved tyrosine residue in the flap (Tyr-77 in Plm-II; Tyr-71 in BACE-1) to play an important role in the dynamics of the flap. Indeed, MD simulations of apo Plm-II and BACE-1 showed the hydroxyl group of this residue interchangeably involved in several hydrogen-bond interactions, and that slow conformational dynamics occurred around this residue, on a timescale of at least hundreds of nanoseconds. However, a closer analysis of the dynamics revealed that each particular hydrogen bond was not exceptionally strong. Instead, the high kinetic barriers appeared to occur between the rotational states of the χ_1 torsional angle of the tyrosine side-chain, with each state being able to accommodate several different hydrogen bonds. Thus, this torsional angle, possibly together with its neighbour χ_2, should be a good candidate for biasing in enhanced-sampling simulations.

Our notation for the χ_1 and χ_2 angles of the tyrosine residue is presented in Figure 2. The distributions of the χ_1 angle centred around $+\frac{\pi}{3}$ *radian* or $-\frac{\pi}{3}$ *radian* are denoted as *normal*, whereas the distribution centred around $\pm\pi$ *radian* is denoted as *flipped* (Figure 2). Crystal structures of Plm-II are always in the normal state, whereas crystal structures of BACE-1 have captured the residue in different states.

Plasmepsin-II

During MD simulations of apo Plm-II, Tyr-77 formed five major H-bond interactions: with Trp-41, Ser-37, Asp-34, Asn-39 and Gly-216, respectively (Figure 3). The distribution of the χ_1 angle of Tyr-77 showed that only one of the four independent MD simulations[1] sampled both the normal and flipped states (Figure 4). The normal state was stabilised by formation of interchanging H-bonds to Trp-41, Ser-37 and Asn-39, whereas the flipped state

[1]Only simulations with the TIP3P water model + FF14SB force-field are considered here, whereas other combinations are discussed in a later section

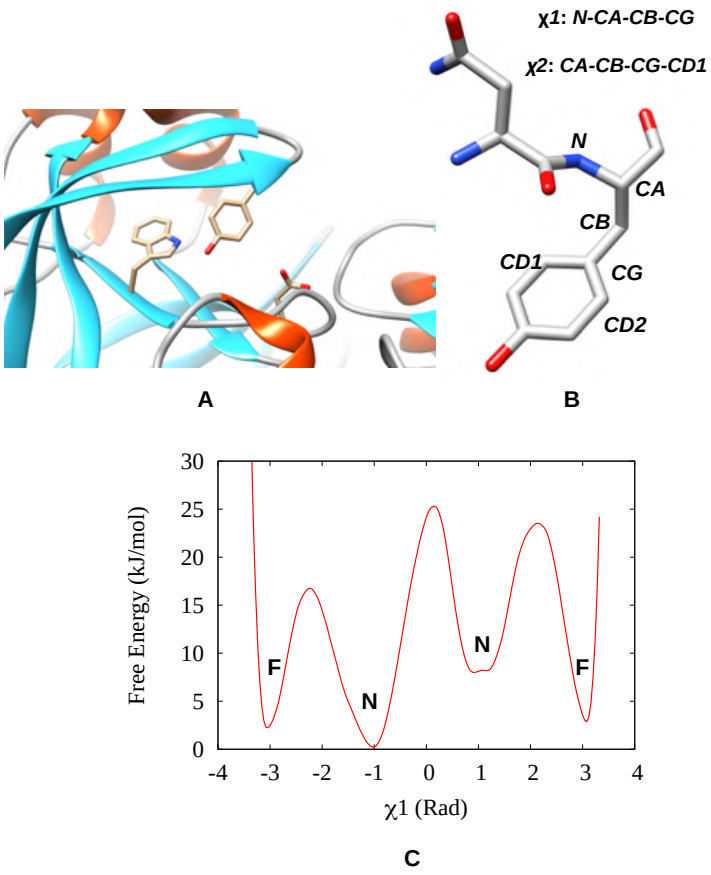

Figure 2: Location of the Tyr, Trp,Asp triad in a typical pepsin-like aspartic protease, Plm-II (A). χ_1 and χ_2 angles of Tyr (B) along with a typical distribution of χ_1 angle (C). **N** and **F** denotes normal and flipped states respectively.

was stabilised by H-bonds to Asp-34 (Figure 22 in the Supplementary Information). Both these states also sampled the *solvent-exposed* conformations, in which the Tyr-77 side-chain didn't form any H-bond interactions with neighbouring residues but only with the solvent. Lack of sampling and a great variation of population in the independent MD runs prevented us from predicting the free energy difference between different conformational states of Tyr.

Metadynamics simulations with biasing on the χ_1 and χ_2 angles (hereafter denoted Torsion-Metad) enhanced the sampling of the rotational degrees of freedom of the Tyr side-chain, which led to an enhanced transition between the normal and flipped states (Figure 4). Similarly to MD, the metadynamics simulations sampled formation of several interchanging H-bond interactions involving Tyr-77. Qualitatively, the flipped and normal states were

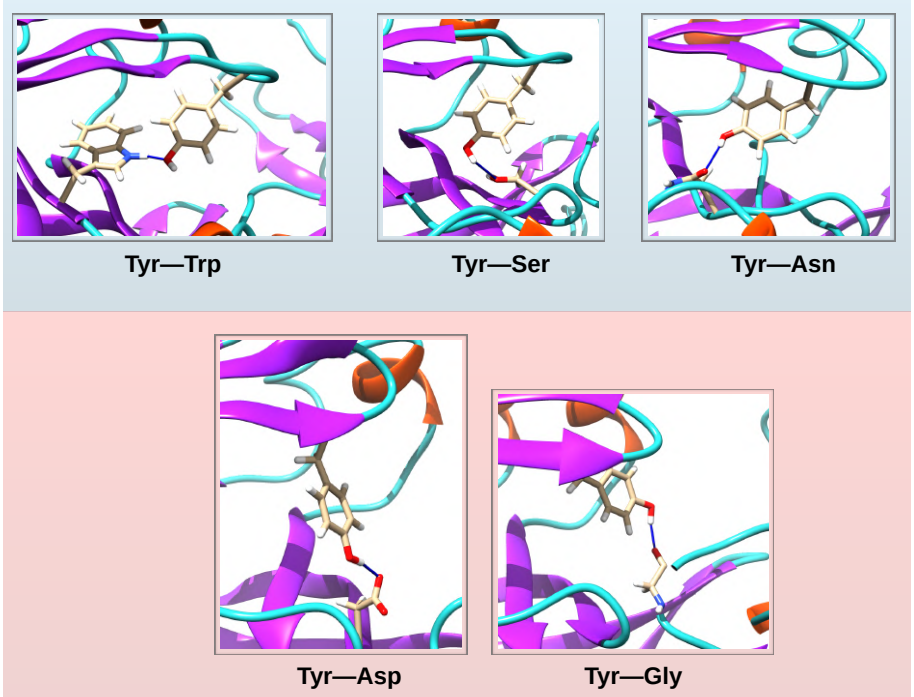

Figure 3: H-bond interactions involving Tyr-77 in Plm-II. *Blue* and *red* shadow indicates normal and flipped states respectively.

found to be equally populated (see subsection *Convergence*) with a barrier $> 10\ kJ/mol$. The normal state was stabilised by interchanging H-bond interactions involving Trp-41, Asn-39 and Ser-37. Whereas, the flipped state was stabilised by H-bonds to Asp-34 and Gly-216 (Figure 5 and Figure 24 in the Supplementary Information). The H-bonds to Trp-41 and Asp-34 were found to be the dominant interactions stabilising these states (Figure 5).

BACE-1

The MD simulations of BACE-1 were initiated using three different crystal structure, denoted SO, SN, and SSO. These structures differs in terms of orientation of Tyr-71 and the extent of flap opening (Figure 16 in the Supplementary Information). The MD simulations starting from the SO and SN conformations only sampled the normal state (Figure 6) of Tyr-71, which was stabilised by H-bond interaction to Trp-76 (comparable to Trp-41 in apo Plm-II). However, MD simulations started from the SSO conformation remained stuck in the flipped state (Figure 6), which was stabilised by formation of a H-bond to Asp-32 (comparable to

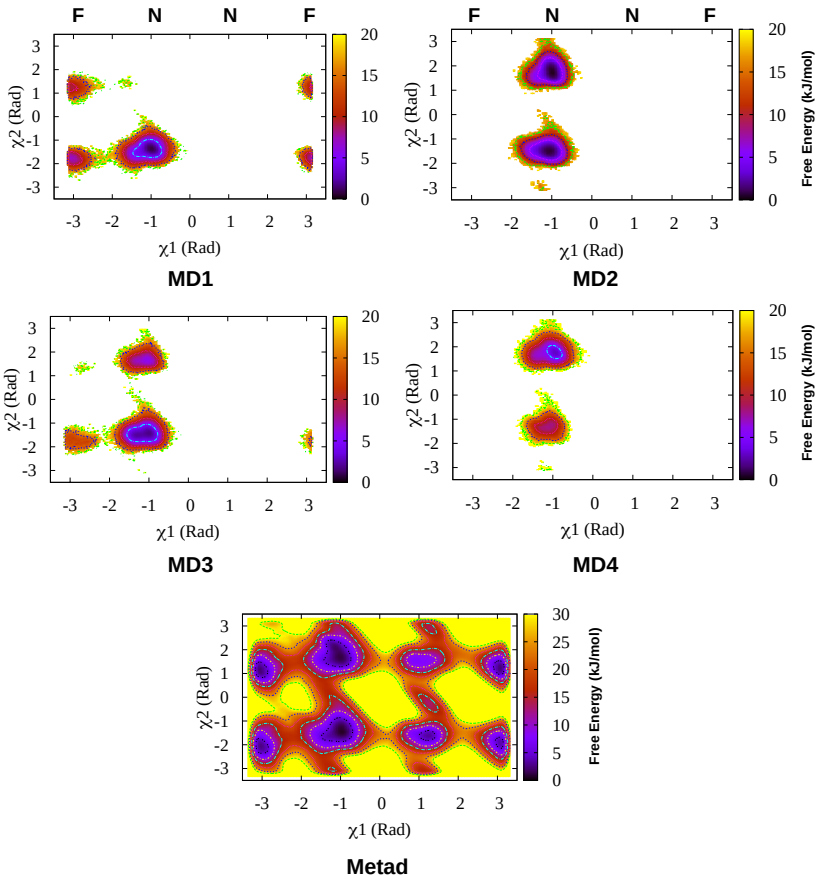

Figure 4: Sampling of χ_1 and χ_2 angles during four independent MD simulations (denoted as MD **1-4**) with TIP3P water model. FES reweighted on aforementioned dihedral angles in Torsion-Metad simulation shows sampling of normal (**N**) and flipped (**F**) states.

Asp-34 in Plm-II) (Figure 25 in the Supplementary Information).

Torsion-Metad simulations enhanced the transition between the normal and flipped states (Figure 13 and Figure 25 in the Supplementary Information). The sampling of these two states in BACE-1 is comparable with Plm-II. H-bond interactions to Trp-76 and Asp-32 dominate the normal and flipped states, respectively (Figure 7). Metadynamics simulations also sampled H-bond to Ser-35 and Lys-107. The H-bond to Lys is caused by an unusual backwards orientation of Tyr-71, which we defined as *Tyr-back* conformation. The free-energy difference between the H-bonds to Trp-76 and Ser-35 was predicted to be ~ 5 kJ/mol. The difference in free energy between the H-bonds to Trp-76 and Lys-107 was ~ 20 kJ/mol (Figure 7).

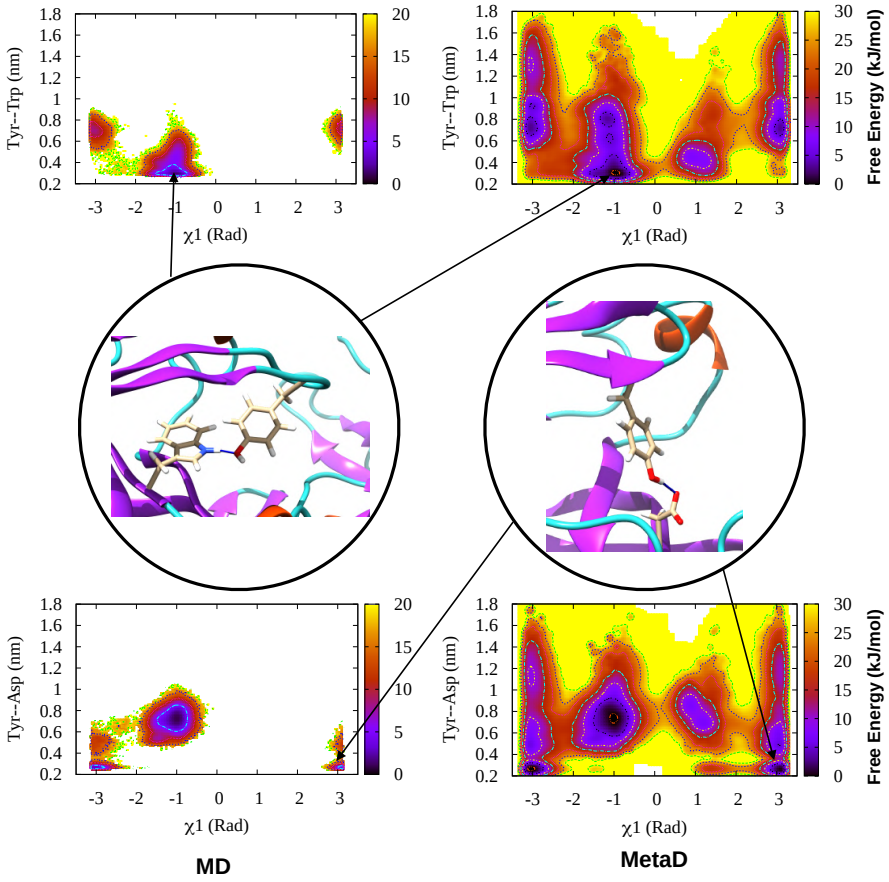

Figure 5: FES reweighted on χ_1 and H-bond distances (H-bond to Trp and Asp) in case of Plm-II. Conformations correspond to H-bond to Trp and Asp were also highlighted. Torsion-Metad simulation did a significantly better sampling of different H-bonded conformations.

Implications for flap opening

Plasmepsin-II

To describe large-scale motion of the flap, especially *opening* of the flap, we defined two distances as illustrated in Figure 8. For each MD simulation, the free energy surface was calculated from the corresponding probability distribution with respect to DIST2 and DIST3 as well as to DIST2 and χ_1. The lack of correlation between DIST2 and DIST3 indicates that the flap and the coil regions move largely independently. The extent of flap opening (DIST2) and the rotation of χ_1 angle of Tyr-77 are related as follows:

1. Normal state: A broader basin around DIST2 $\sim$1.2 nm is stabilised by interchanging

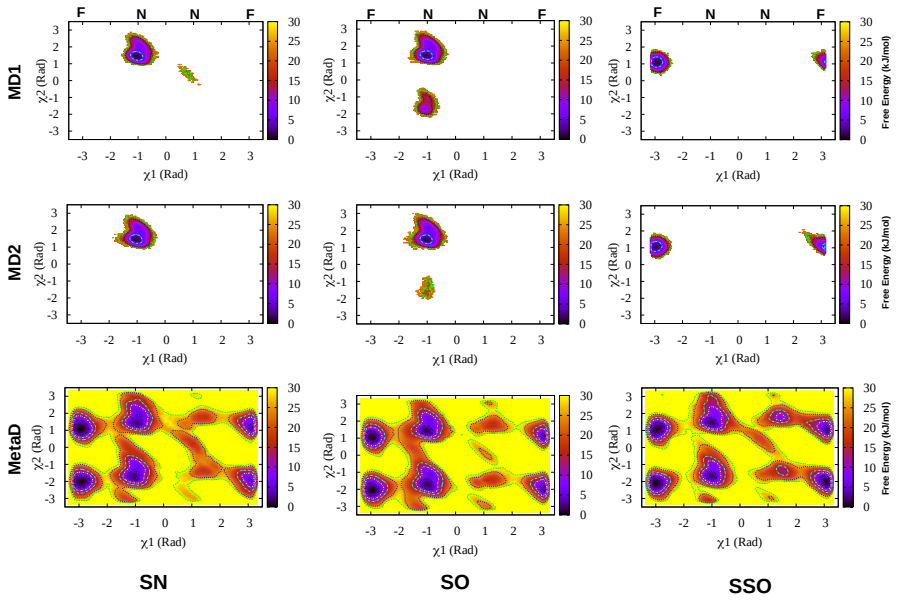

Figure 6: Sampling of normal (**N**) and flipped (**F**) states in MD and Torsion-Metad simulations starting with SN, SO and SSO conformations respectively. Independent MD simulations starting with SN and SO conformations only sampled the **N** state ($-\frac{\pi}{3}$ *radian*) whereas simulations starting with SSO conformation only sampled the **F** state. Torsion-Metad simulations starting from SN, SO and SSO conformations sampled both **N** and **F** states.

H-bonds to Trp-41, Ser-37 and Asn-39. The free-energy minimum centred around DIST2 ∼1.6 nm fluctuates between H-bond to Trp-43 and the solvent exposed conformations of Tyr-77.

2. Flipped state: The free-energy minimum centred around DIST2 ∼1.05 nm is stabilised by interchanging H-bonds to Asp-34 and Gly-216.

3. The flap adapts an *open* conformation around DIST2 ∼2.0 nm.

All four independent MD simulations with TIP3P water model sampled free-energy minima centred around DIST2 ∼1.2 nm and ∼1.6 nm respectively (Figure 8). In MD simulations, the free-energy basin around DIST2 ∼1.2 nm sampled only interchanging H-bonds to Trp-41 and Ser-37. Only one of the four independent simulations sampled DIST2 around ∼1.05 nm (stabilised by the H-bond to Asp-34). Moreover, one of the four independent MD simulations sampled the flap opening.

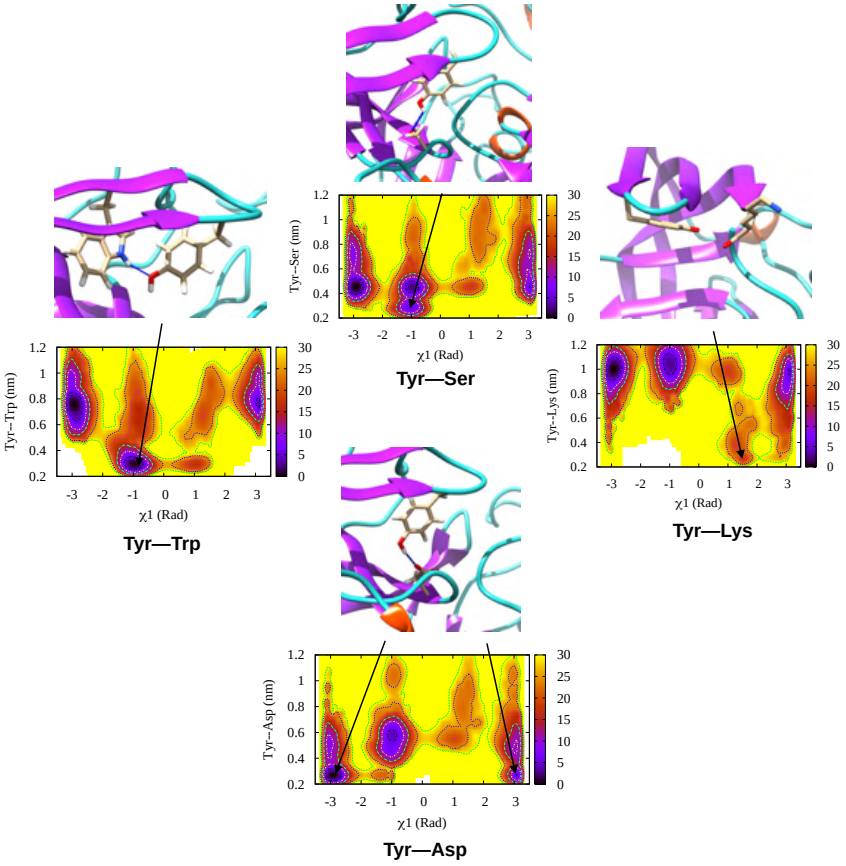

Figure 7: FES reweighted on χ_1 and different H-bond distances in apo BACE-1. 2D FES reweighted on H-bond distances (Trp, Ser, Asp and Lys) and χ_1 for each metadynamics simulations starting with SN, SO and SSO conformations can be found in Supplementary Informations.

Torsion-Metad simulations gave a better picture of flap dynamics in Plm-II (Figure 8). Besides sampling typical minima around DIST2 ∼1.05 nm, ∼1.2 nm and ∼1.6 nm; normal state of apo Plm-II sampled an additional free energy minimum centred around DIST2 ∼0.8 nm which was stabilised by a H-bond to Asn-39. The flap tip adapted an *inwardly* bended conformation in the aforementioned basin (DIST2 ∼0.8 nm). Free-energy minima centred around DIST2 ∼1.05 nm and ∼1.2 nm were found to be equally populated and the free energy cost of flap opening (transition from DIST2 of ∼ 1.2 nm to ∼ 2.0 nm) was predicted to be 5 kJ/mol.

The sampling of the open conformation is not directly influenced by the enhanced sam-

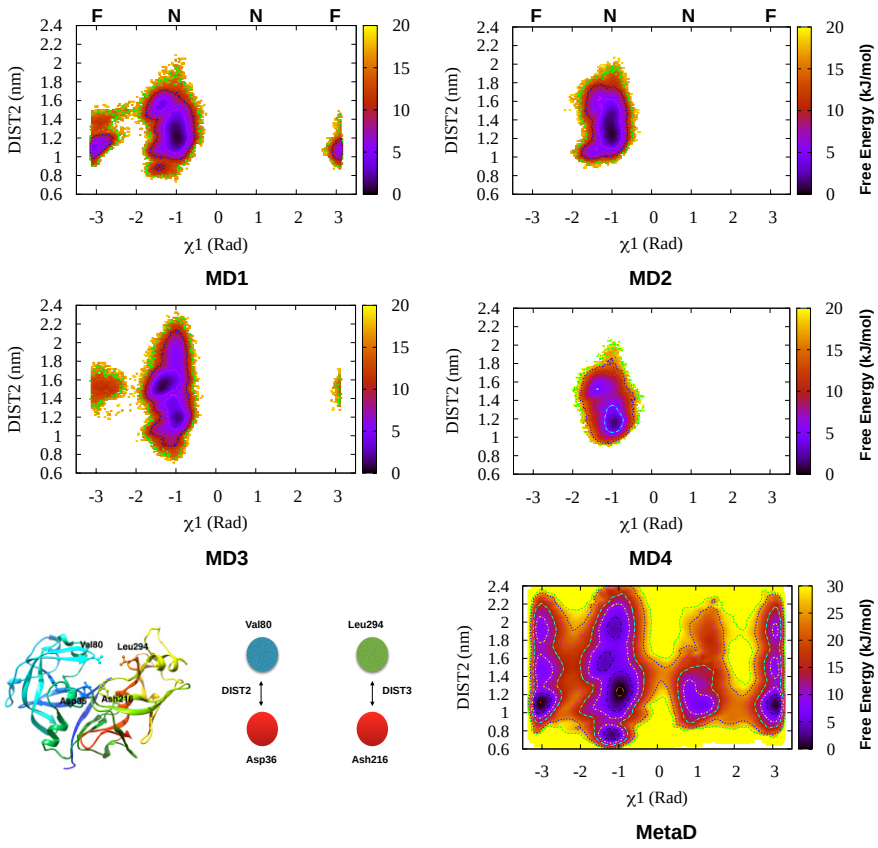

Figure 8: FES reweighted on χ_1 and DIST2 shows sampling of overall flap flexibility during MD (denoted as MD**1-4**) and Torsion-Metad simulations. **N** and **F** denotes normal and flipped states of Tyr-77 respectively.

pling of the Tyr-77 side-chains (Table 7 in Supplementary Informations). The only significance appears to be that one avoids being trapped in the normal or flipped states. Once the H-bond interactions between Tyr-77 and neighbouring residues are broken and it becomes solvent-exposed, the sampling of the flap opening is in fact similar among standard MD simulations and metadynamics simulations.

BACE-1

We have devised analogous DIST2 and DIST3 descriptors (see Supplementary Informations) for BACE-1. Flap opening (DIST2) and rotation of the χ_1 angle associated with Tyr-71 in the MD simulations are related as follows:

1. Normal state: The free-energy minimum centred around DIST2 $\sim$ 1.2 nm is mainly

stabilised by interchanging H-bonds to Trp-76, Ser-35 and Lys-107.

2. Flipped state: The free-energy basin centred around DIST2 ∼1.05 nm is stabilised by formation of a H-bond to Asp-32.

Figure 9 again shows that MD simulations started from the SO and SN conformations only sampled the normal state, whereas MD simulations started from the SSO conformation remained stuck in the flipped state. The Torsion-Metad simulations sampled both states. The free-energy basins centred around DIST2 of ∼1.2 nm and ∼1.05 nm were found to be equally populated, which showed that the flap remains in a dynamic equilibrium between the normal and flipped conformations (Figure 9). The largest flap opening (DIST2 ∼1.8 nm) was observed in the Torsion-Metad simulation started from SN conformation. The cost of flap opening was predicted to be ∼ 15 kJ/mol.

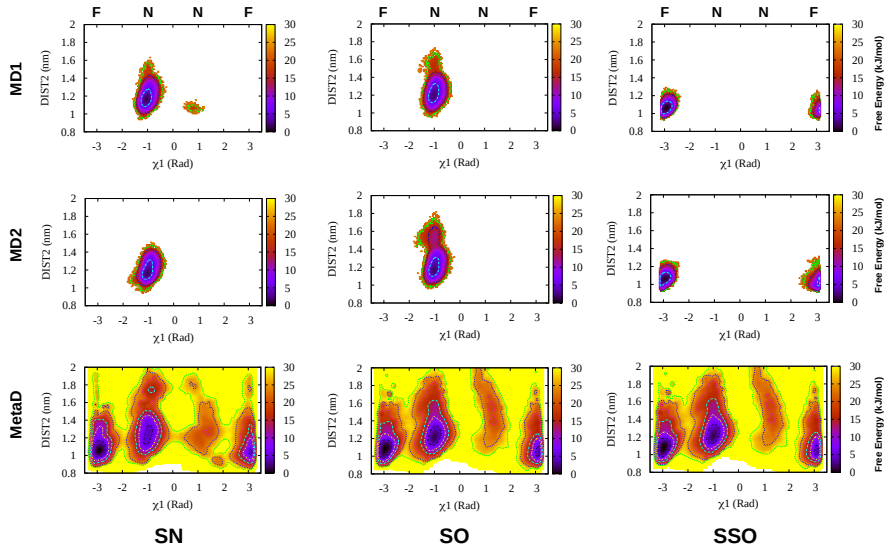

Figure 9: FES reweighted on χ_1 and DIST2 in MD and Torsion-Metad simulations starting with SN, SO and SSO conformations. **N** and **F** denotes normal and flipped states of Tyr-73.

Extent of opening needed for ligand release

In order to understand the extent of flap opening necessary to accommodate the ligand in the active site, we have performed well-tempered metadynamics simulation using a *Plm-ligand*

complex. The distance between the centre of mass of the active site residues and the ligand heavy atoms was selected as CV to capture ligand unbinding. The aim of this simulation was to gain a rough picture of the ligand unbinding process. This type of simulation will not necessarily follow a low-energy path. Plotting DIST2 as a function of time (or equivalently the stage of unbinding) indicated that a flap opening of ∼2.0 nm (corresponding to the open conformation in the apo simulations) is sufficient for ligand binding/unbinding (Fig. 10). Tyr-77 remained in the normal state during the unbinding event. It is important to note that the amount of flap opening might depend on the size of the ligand. In our case, we chose a small molecule inhibitor with hydroxylethylamine scaffold[23] as a ligand, keeping in mind the importance of such inhibitors in antimalarial drug discovery.

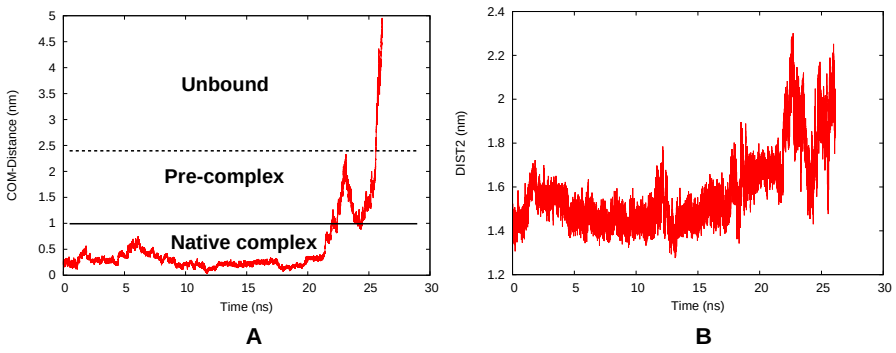

Figure 10: Time evolution of the unbinding CV (distance between the centre of mass of the active site residues and ligand heavy atoms) during metadynamics simulation (A). During the pre-complex stage, the flap opens up in order to facilitate ligand exit. The DIST2 as a function of simulation time shows the extent of flap opening during ligand exit (B).

Forcefield, watermodel and choice of CVs

All results presented up to this point were obtained using the FF14SB force field for the protein and the TIP3P water model. In order to test the effect of the force field and water model on the population of normal and flipped states, we used apo Plm-II as a model system.

Force field

We performed a Torsion-Metad simulation with the CHARMM36 force field and the TIP3P water model. Metadynamics sampled both the normal and flipped conformations of Tyr-77. The flipped state was predicted to be slightly more populated compared to the normal state. The free-energy difference between the flipped and normal state was predicted to be $\sim$6 kJ/mol (Figure 21 in the Supplementary Information).

Water model

We also carried out four independent MD simulations of apo Plm-II with TIP4P-Ew water model (and the FF14SB force field). Two of the four independent MD simulations sampled both normal and flipped states (Figure 23 in the Supplementary Information). Besides H-bond interactions to Trp-41, Ser-37, Asn-39 and Asp-34; these TIP4P simulations sampled an additional H-bond to Gly-216 (Figure 23 in the Supplementary Information).

Torsion-Metad simulations with TIP4P water model sampled both the normal and flipped states. The population of normal and flipped states as well as different H-bond interactions were comparable to the TIP3P results (Figure 11 and Figure 26 in the Supplementary Information).

Choice of CVs

We have also performed metadynamics simulations of apo Plm-II using three other types of collective variables, namely principle components (PCA), centre-of-mass distances (COM), and time-lagged independent components (TICA), as described in the Supplementary Information). Metadynamics simulations using PCA and COM CVs did a poor sampling of the normal and flipped states (Figure 11). This is due to the fact that the slow degrees of freedom (χ_1 and χ_2 angles of Tyr-77) were not explicitly biased during these simulations. Dimensionality reduction using TICA was able to capture rotational degrees of freedom associated with Tyr-77. The metadynamics simulation using the first and third TICA components as

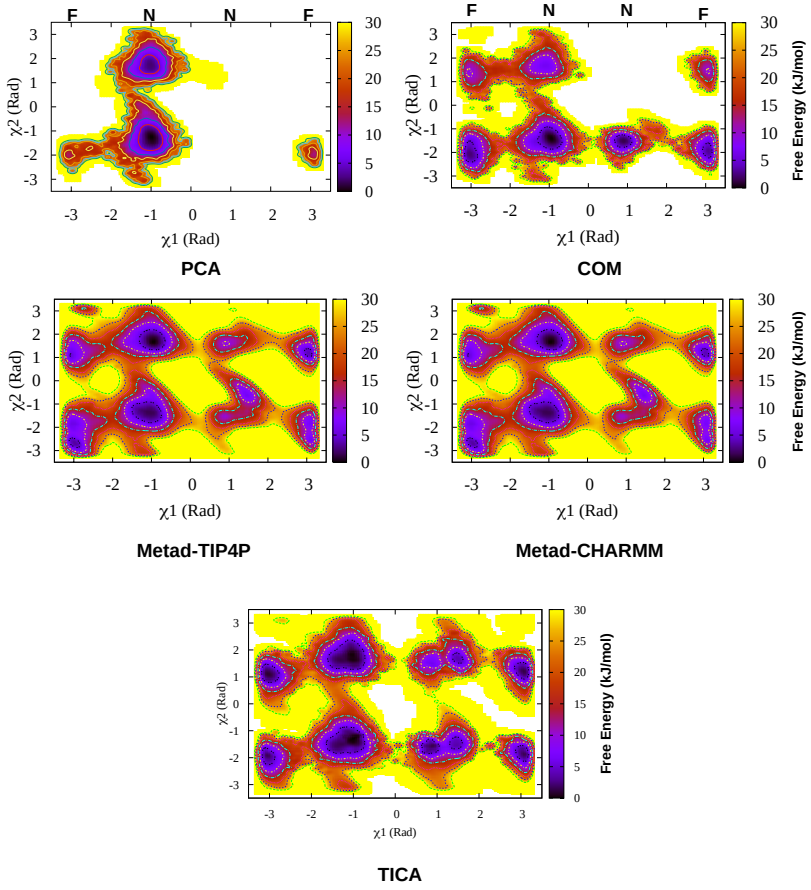

Figure 11: FES reweighted on χ_1 and χ_2 for metadynamics simulations using PCA, COM, torsion (denoted as Metad-TIP4P and Metad-CHARMM), TIC CVs.

CVs sampled both the normal and flipped states (Figure 11). Qualitatively, the sampling using TICA metadynamics and Torsion-Metad were comparable to eachother.

Convergence

Plasmepsin-II

The Torsion-Metad simulations[2] using TIP3P water model reached apparent convergence after around ∼600 ns. To assess the convergence, we have calculated the free-energy difference between the flipped and normal ($-\frac{\pi}{3}$ *radian*) state as a function of simulation time (Figure 12). The last ∼150 ns of each simulation was used for statistical analysis. The aver-

[2]We have performed two independent Torsion-Metad simulations using same CV

age free-energy difference between the flipped and normal state calculated to be 1.34 ± 1.86 and -0.31 ± 1.62 kJ/mol respectively (see the Supplementary Information for statistical analysis).

Torsion-Metad with TIP4P water model and CHARMM force-field reached apparent convergence around ∼600 ns and ∼500 ns respectively (Figure 21 in Supplementary Informations). The average free-energy difference between the flipped and normal state was predicted to be 2.3 ± 0.6 kJ/mol and -5.8 ± 1.9 kJ/mol for the TIP4P and CHARMM simulations, respectively.

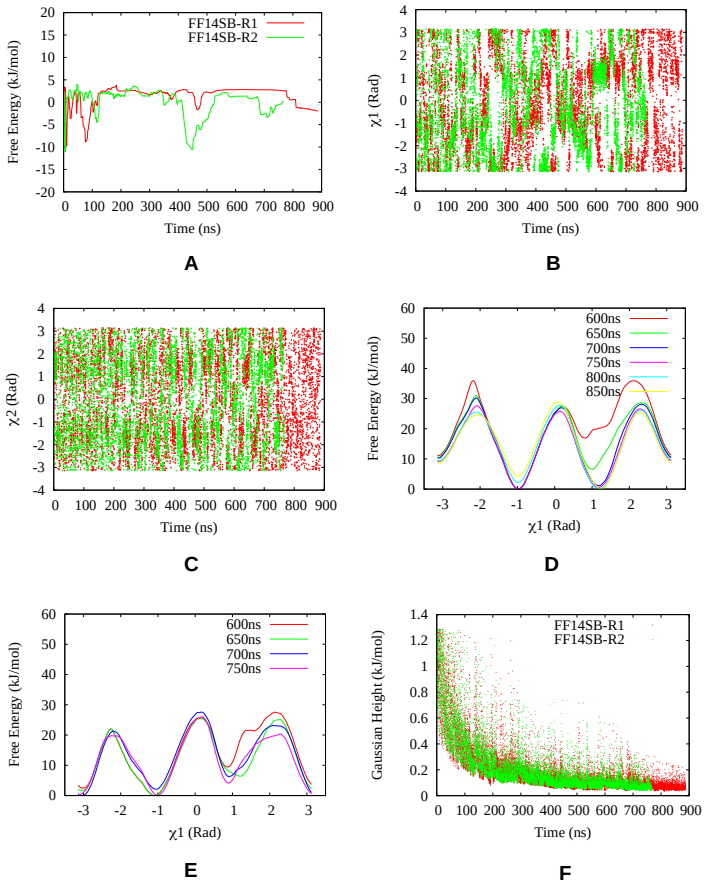

Figure 12: Free energy difference between flipped and normal state ($-\frac{\pi}{3}$ *radian*) as a function of simulation time (A). Sampling of χ_1 and χ_2 angles (B and C) during metadynamics simulations shows diffusive behavior of CVs space when the hill height quite small (F). In the last part (∼600 ns onward), the free energy profiles (D and E) look similar, apart from a constant offset. Using all these observations, we can say that the two independent metadynamics simulations reached apparent convergence.

BACE-1

Torsion-Metad simulations starting with SO, SN and SSO conformations reached convergence around ∼300 ns. Statistical analyses were carried out using the time interval of 300-550 ns (Figure 13). The hill heights were very small at this part of simulation. The average free energy difference between flipped and normal state ($-\frac{\pi}{3}$ *radian*) calculated to be -1.4 ± 2.6 kJ/mol, -2.8 ± 3.1 kJ/mol and -6.1 ± 4.0 kJ/mol for SO, SN and SSO respectively (see the Supplementary Information). In the presentations of results in the previous sections, we have only included the Torsion-Metad simulation started from the SN conformation.

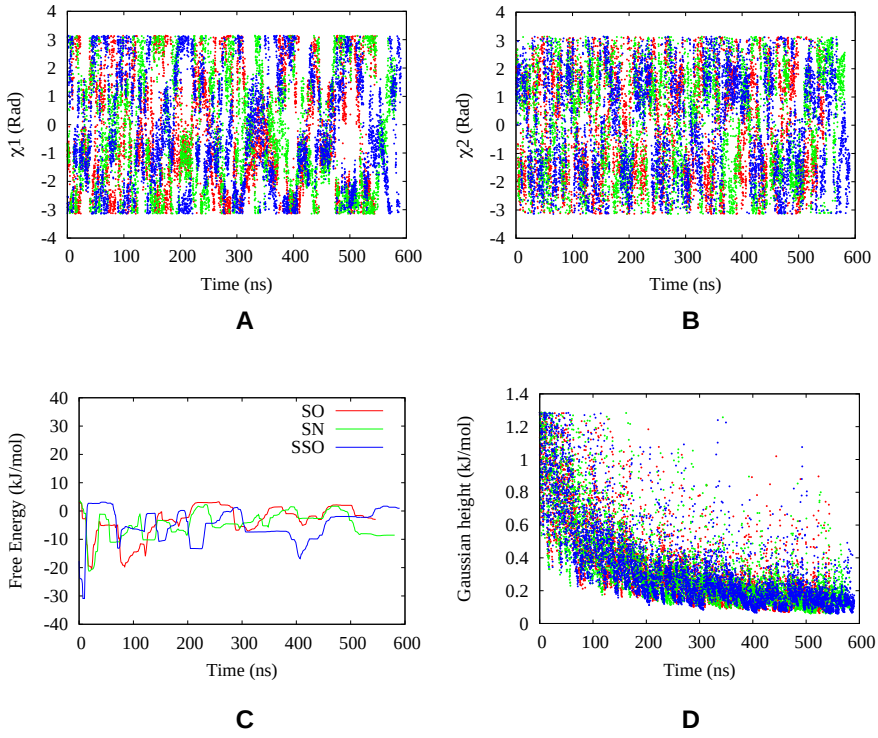

Figure 13: Fluctuation of χ_1 (A) and χ_2 (B) during Torsion-Metad with SO, SN and SSO conformations. The free-energy difference between the flipped and normal state as a function of simulation time shows apparent convergence of metadynamics simulations (C).

Mutational study

We computationally mutated the conserved Tyr to alanine (Ala) in apo Plm-II and BACE-1. Alanine doesn't possess a bulky side-chain like Tyr. The apparent free-energy surface on DIST2 and DIST3 shows a complete flap collapse in both Plm-II and BACE-1 (Figure 14).

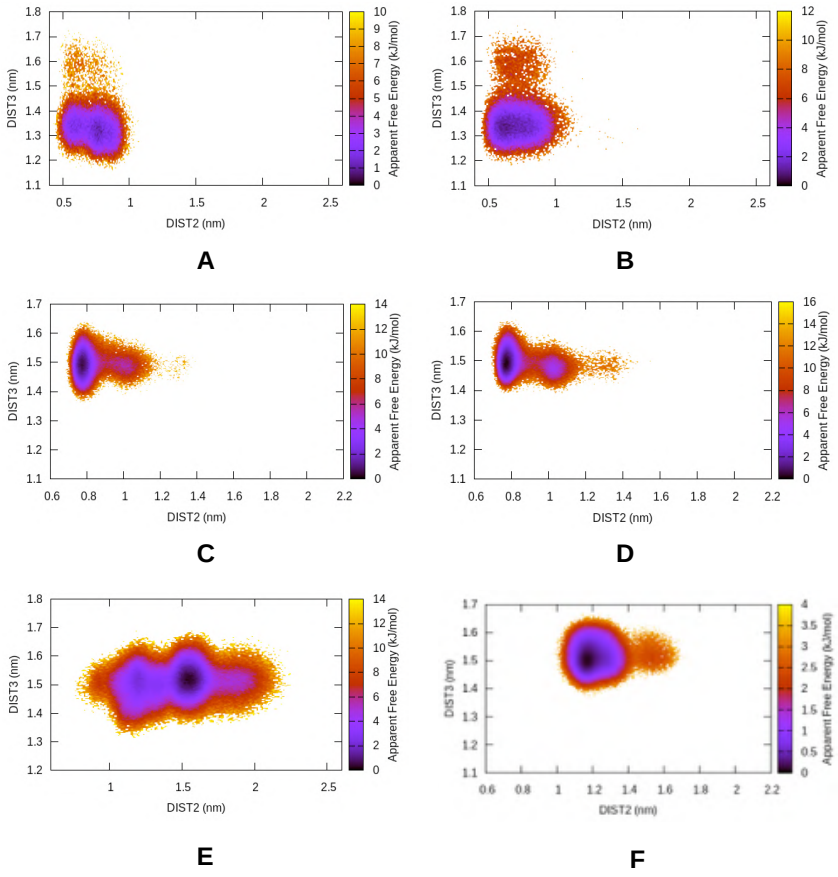

Figure 14: FES projected on DIST2 and DIST3 in case of apo Plm-II (A and B) and BACE-1 (C and D) with Ala mutation. This can be compared with wild type Plm-II (E) and BACE-1 (F). Projection on DIST2 highlights that the flap loses flexibility in mutant protein.

Discussion

In solution, the flap region of apo Plm-II and BACE-1 adapts different conformations due to rotation of χ_1 and χ_2 angles of the conserved tyrosine residue (Tyr-77 and Tyr-71 in case of Plm-II and BACE-1 respectively). Mutation of Tyr to Ala led to complete flap collapse in

Plm-II and BACE-1. Previous experimental studies highlighted that mutation of Tyr to Ala resulted in loss of activity.[24] Mutation of Tyr to other amino acids with smaller side-chains, e.g. Thr, Ile, Val in chymosin (another pepsin-like aspartic protease), also resulted in loss of activity. However, mutation of Tyr to Phe in pepsin didn't diminish the activity of the protein.[25,26] Like Tyr, Phe also possesses rotational degrees of freedom along the χ_1 and χ_2 angles. Hence, the flap of pepsin-like proteases with Phe (*Toxoplasma gondii* aspartic proteases, Plm IX, X and Plm V) can still remain dynamic in the apo conformation.[27,28]

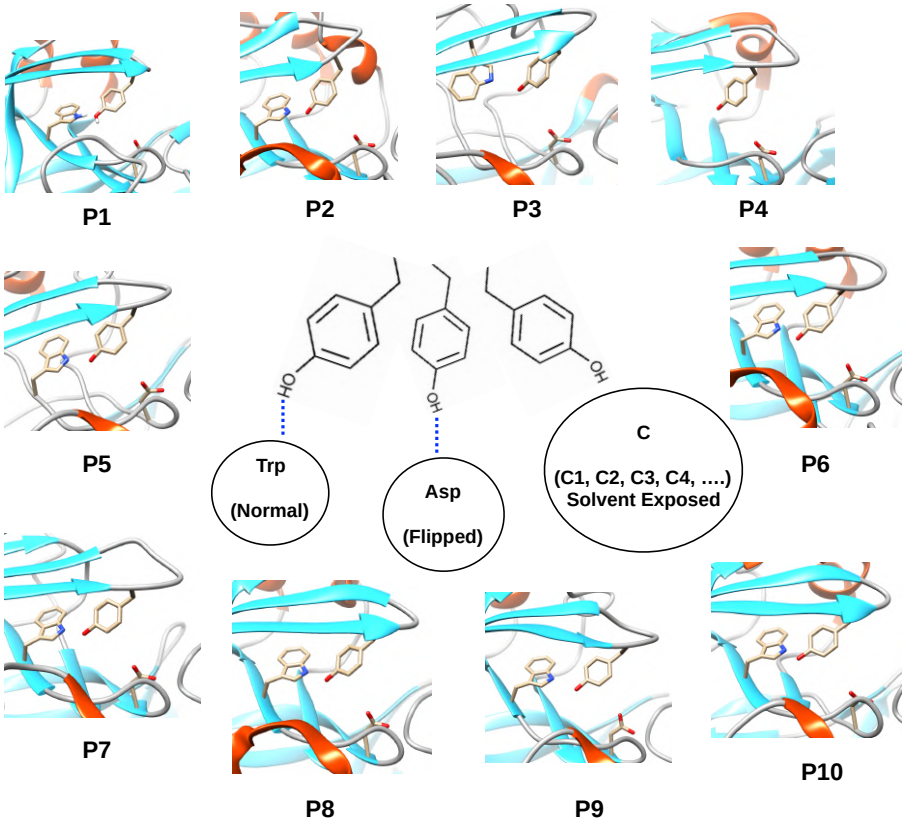

Figure 15: Three different conformational states associated with side-chain flexibility of Tyr in pepsin-like aspartic proteases. Normal and flipped states are stabilised by H-bond to Trp and Asp. C is the solvent exposed subspace which consists of several rotameric conformations (C1, C2, C3, C4, ... etc.) of Tyr. P1 to P10 denotes crystal structures of human cathepsin-D (PDB:1LYA), cathepsin-E (PDB:1TZS), BACE-2 (PDB:3ZKQ), Plm-V (PDB:4ZL4), bovine chymosin (PDB:4AUC), human pepsin (PDB:3UTL), candidapepsin (PDB:2QZW), human renin (PDB:5SY2), Plm-I (PDB:3QS1) and Plm-IV (PDB:1LS5) respectively. All these structures (except Plm-V) possessed conserved Tyr, Trp and Asp residues similar to Plm-II and BACE-1. Plm-V is missing the conserved Trp residue.

MD simulations sampled both normal and flipped states. However, low transition probability between normal and flipped states in MD simulation makes it a *rare* event. Rotation of the χ_1 angle of Tyr was found to be the slowest degree of freedom separating normal and flipped states. On the other hand, rotation of the χ_2 angle acts as an auxiliary collective variable. Metadynamics simulations with χ_1 and χ_2 angles as CVs sampled the normal and flipped states much more reliably. Both these states were stabilised by formation of interchanging H-bonds with neighbouring residues. A common pattern among these interactions is the formation of H-bonds to Trp and Asp (Figure 15). In *R. pusillus* protease, the normal state was stabilised by a hydrogen bond between Tyr-75 and Trp-39.[29] In a set of experiments, Park et al.[30] replaced Trp-39 with other residues in *R. pusillus* protease and observed a decreased activity, which suggests a role of Trp in stabilising the normal state of Tyr. Crystal structure of Plm-V from *P. vivax* (PDB: 6C4G,[31] 4ZL4[32]) doesn't possess the Trp residue but the flap region still remains dynamical due to the rotational degrees of freedom associated with Tyr .[33]

Spontaneous flap opening was also sampled during MD and metadynamics simulations. However, the extent of flap opening varied among Plm-II and BACE-1. The metadynamics-based unbinding study provided an overview of the extent of flap opening necessary for substrate entry in apo Plm-II. In that simulation, Tyr remained in the normal state during ligand unbinding. In the future, more extensive simulation studies are necessary to understand the actual role of Tyr in substrate binding.

Finally, we also want to address our thoughts regarding convergence of metadynamics and the problems associated with it. In the metadynamics simulations, the normal and flipped states of Tyr can access a wide range of conformations that are solvent exposed. If one imagines that the normal and flipped states of Tyr fluctuates between only two H-bonded states, A and B, then the Torsion-Metad simulations would have been much easier to converge. However, in our case, the longer we run the Torsion-Metad simulations, the flipped and normal states tend to sample new conformations in the solvent exposed subspace

(denoted by C), which makes it harder to converge within a reasonable time (Figure 15).

Conclusion

The dynamic nature of the flap in pepsin-like aspartic proteases is necessary for its catalytic activity and substrate binding. Most of these aspartic proteases possess a conserved triad of Tyr, Trp and Asp (Figure 15). We predict that the flap dynamics in almost all pepsin-like aspartic proteases is governed by the rotational degrees of freedom associated with the Tyr residue. The orientation of tyrosine directly influences the volume of the binding site. Drug designers may be able to exploit this property by designing inhibitors of varying size which can lock tyrosine at any desired conformation. We believe that our study will act as a starting point to perform experiments that can give structural and mechanistic insights into pepsin-like aspartic proteases.

Acknowledgement

PS thanks the Crafoordska Foundation and the Swedish Research Council for financial support. The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at LUNARC (Lund University) and HPC2N (Umeå University).

Supporting Information Available

Computational Methods

Starting structure: Plm II

The co-ordinates of the apo Plm II (PDB ID: *1LF4*)[34] was retrieved from protein data bank. Prior to simulation, the protonation state of the structure was adjusted using the *H++* server[35] and missing cysteine–cysteine sulphur–sulphur (S-S) bridges were created. The catalytic aspartic acid at position 214 was protonated while the other one (Asp34) remained negatively charged.

Starting structure: BACE-1

The co-ordinates of the apo BACE-1 were retrieved from PDB (PDB: 1W50 (SO), 3TPL (SN) and 1SGZ (SSO)). These structures differs in terms of orientation of tyrosine and extent of flap opening (Figure 16). The structures of BACE-1 were prepared similarly that of Plm-II.

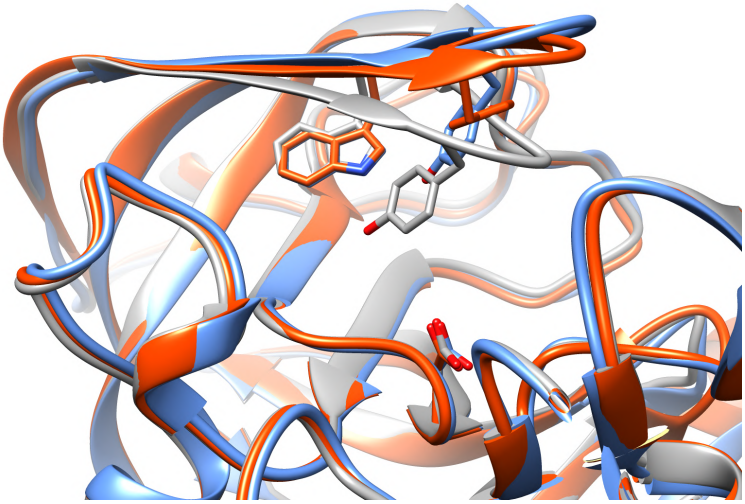

Figure 16: Difference in flap opening and orientation of Tyr in SO (PDB:1W50, *blue*), SN (PDB:3TPL, *grey*) and SSO (PDB:1SGZ, *orange*) conformations of apo BACE-1. In SSO conformation, Tyr points backwards.

MD simulations

TIP3P water model

The protein was treated with the FF14SB[36] force field and immersed into a truncated octahedron box with TIP3P[37] water molecules. The system was neutralised by 9 sodium ions. The box was set such that no protein atom was within 1.0 nm of the box edge. Initial restrained minimisation was performed with steepest descent algorithm for 750 steps followed by 1750 steps of conjugate gradients with a restrained potential of 40 kJ/mol applied on the C_α atoms. The restrained minimisation was followed by 200 steps of unrestrained minimisation. The minimisation was followed by gradual heating (from 0K to 300K) for 400 ps with a harmonic restrained of 40 kJ/mol applied on the C_α atoms and a Langevin thermostat with a collision frequency of $1ps^{-1}$ using NVT ensemble. The system was subsequently equilibrated at 300K in an NPT ensemble for 5 ns without restraint and a Berendsen barostat[38] used to maintain the pressure at 1 bar. Long-range electrostatics were treated using particle mesh Ewald (PME) method[39] with a grid spacing of 0.1 nm and van-der Waals (vdW) cut-off of 1.2 nm. Finally, independent production runs (using apo Plm-II and BACE-1) with time step of 2 fs were performed with different initial starting velocities and LINCS algorithm[40] was used to constrain the bonds of all hydrogen atoms during the simulation.

Table 1: Length (in ns) of MD simulations starting with TIP3P and TIP4P-Ew water models using apo Plm-II.

Simulation	Length
TIP3P	
MD1	528
MD2	562
MD3	442
MD4	584
TIP4P-Ew	
MD1	669
MD2	535
MD3	530
MD4	530

Table 2: Length (in ns) of MD simulations starting with SO, SSO and SN conformation of BACE-1. The simulations were carried out in TIP3P water model.

Simulation	Length
SO	
MD1	774
MD2	835
SSO	
MD1	518
MD2	513
SN	
MD1	790
MD2	791

TIP4P water model

We have also carried out independent MD simulations using the TIP4P-Ew water model (referred to as TIP4P from here on).[41] Simulation parameters were identical as in the TIP4P simulations. The system was equilibrated in the NPT ensemble using the Berendsen barostat for at least 5 ns. Independent production runs (using apo Plm-II) were performed with time step of 2 fs.

Additional simulations were performed also performed where apo Plm-II was described using CHARMM36 force-field ([42]). Other simulation parameters were kept identical as in TIP3P simulation. The system was equilibrated in NPT ensemble using Berendsen barostat for 5 ns before starting metadynamics simulations.

Metadynamics and Choice of CVs

In metadynamics simulation,[16] an external bias potential is constructed in the space of a few selected degrees of freedom, known as collective variables (CVs). This potential is built as a sum of Gaussians deposited along the trajectory in the CVs space. The applied bias potential in metadynamics pushes the system away from a local energy minimum into new regions of phase space.

In standard metadynamics simulation, gaussians of a pre-defined height are added dur-

ing the course of metadynamics. As a result, the system is eventually pushed to explore high free-energy regions. This problem is solved by applying the well-tempered metadynamics protocol,[43] where the height of the gaussian is decreased with simulation time which eventually leads to smooth convergence in a long time scale.

We have used principal component analysis (PCA), time-independent component analysis (TICA),[44] distances (COM) and torsion angles (Torsion-MetaD) as CVs during our metadynamics investigations.

Torsion Angles

We have used χ_1 and χ_2 angles of the Tyr residue in the flap (Tyr-77 in case of Plm-II and Tyr-71 in case of BACE-1) as CVs to perform well-tempered metadynamics simulations. All metadynamics simulations were performed at 300K with a gaussian height of 1.2 kJ/mol and a width of 0.05 *radian* deposited every 1 ps. The bias factor was set to 15. Three independent metadynamics simulations (two with TIP3P and one with TIP4P-Ew water model) were performed in case of apo Plm-II. In case of BACE-1, three different metadynamics simulations (using TIP3P water model) were performed with the same settings but starting from the SO, SN and SSO conformations, respectively.

For Plm-II, we have also performed metadynamics simulations with CHARMM36 forcefield[42] (with TIP3P water model) using torsion angles as CVs.

Distance

We have used the distance between the centre of mass (COM) of the catalytic aspartic residues (taking into account only C_α atoms of Asp34 and Ash214) and the COM of the flap region (taking into account all C_α atoms of residues 58-88) as the first collective variable (CV1). The distance between the COM of the catalytic aspartic residues and the COM of the coil region (taking into account all C_α atoms of residues 282-302) was used as the second collective variable (CV2) (Figure 17). CV1 and CV2 were used in a 2D WT-MetaD

simulation with the TIP3P water model. Metadynamics simulations were performed at 300K with a gaussian height of 1.2 kJ/mol and a width of 0.02 nm deposited every 20 ps, and a bias factor of 10. An upper and lower wall at 2.05 nm and 1.60 nm respectively, were placed along CV1, whereas in case of CV2, the upper and the lower wall were placed at 1.75 nm and 1.35 nm, respectively.

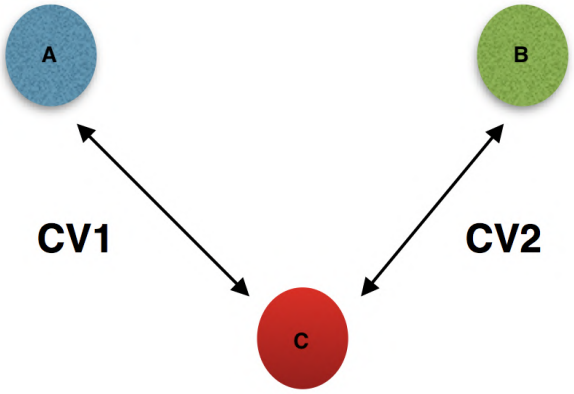

Figure 17: A simplified view of the COM CVs used in this study. **A**, **B** and **C** denotes centre of mass corresponds to flap (residue 58-88), coil (residue 282-302) and catalytic aspartic dyads respectively.

Principal Components

Principal component analysis (PCA) was performed in order to capture the most prominent motions from our MD simulations.[45] Trajectories from two independent MD runs in TIP4P water model (starting with apo Plm-II) were combined and PCA analysis was performed on C_α atoms of the protein (ignoring the tail part, Gly0–Asn3) using the *g_covar* tool integrated with *Gromacs*. We have used the first two eigenvectors (PC1 and PC2) as CVs in a 2D WT-MetaD with the TIP3P water model. The temperature was set at 300K with a Gaussian height and width of 1.0 kJ/mol and 0.00025 nm, respectively. The bias factor was 10. Walls were applied along the eigenvectors to restrict the sampling of regions of high free energy.

Time-independent component analysis

Time-independent component analysis[3] was performed using *MSM Builder 3.8*[46] using one representative MD simulation with TIP4P water model. Rotational and translational degrees of freedoms were removed from the MD trajectory before performing the dimensionality reduction. The dimensionality reduction started by transforming raw cartesian co-ordinates into a subset of important dihedral features using *DihedralFeaturiser*. We used χ_1 and χ_2 angles associated with the flap region (residues 73-83) of apo Plm II to generate the features (Figure 19). High-dimensional dihedral features were then reduced to 10 time-lagged independent components by performing TICA using a *lag time* of 10. The TICA features were transferred to a *Plumed* input file via a *Python* script. We were mainly focused on the first 5 TICs. *Plumed Driver* was used to extract the projection of the first 5 TICs from MD trajectories. Finally, well-tempered metadynamics was performed using TIC1 and TIC3 (Figure 18) as CVs with a Gaussian width and height of 0.06 and 1.2 kJ/mol at 300K. The bias factor was set at 15.

Unbinding using metadynamics: Plm-II-ligand complex

Prior to metadynamics simulation on Plm-ligand complex (PDB ID: *4Y6M*),[23] we used the following procedures to prepare the system for simulation. The ligand was extracted from the complex and was parameterised by the general amber force field (GAFF)[47] whereas the protein was treated using the FF14SB force field. The complex was then neutralised using 5 sodium ions and immersed into an truncated octahedral box with TIP3P water models such that no solute atom is within 1.0 nm of the box edge. Initial minimisation with steepest descent algorithm was performed for 5000 steps followed by restrained equilibration (a restrained of 500 kJ/mol was applied on protein C_α and ligand heavy atoms) for 2 ns in the NPT ensemble with Berendsen barostat used to maintain the pressure at 1 bar. The temperature was fixed at 300 K in both cases using the velocity rescaling algorithm. All

[3]using lag time 10

Table 3: Mean-free sin and cos transformed χ_1 and χ_2 angles and corresponding feature index.

Feature Index	Name
0	sin χ_1 74
1	sin χ_1 75
2	sin χ_1 76
3	sin χ_1 77
4	sin χ_1 78
5	sin χ_1 79
6	sin χ_1 81
7	sin χ_1 82
8	sin χ_1 83
9	cos χ_1 74
10	cos χ_1 75
11	cos χ_1 76
12	cos χ_1 77
13	cos χ_1 78
14	cos χ_1 79
15	cos χ_1 81
16	cos χ_1 82
17	cos χ_1 83
18	sin χ_2 74
19	sin χ_2 75
20	sin χ_2 76
21	sin χ_2 77
22	cos χ_2 78
23	cos χ_2 75
24	cos χ_2 76
25	cos χ_2 77

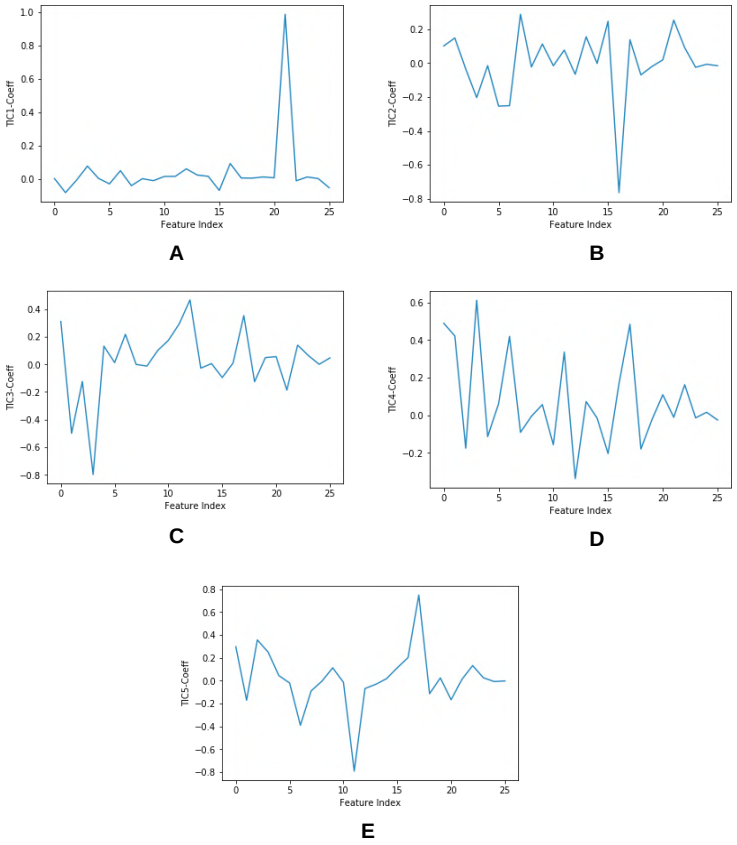

Figure 18: Coefficients corresponding first 5 TICs. The X axis corresponds to the features which are described in Table 3. One can see that TIC1 and TIC3 able to capture rotational degrees of freedom associated with Tyr side-chain.

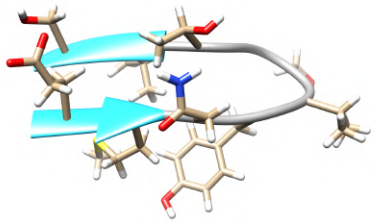

Figure 19: Structural of flap region (residues 73-83) used in generating TICs in case of apo Plm-II.

bonds lengths were constrained using the LINCS algorithm[40] to enable a time-step of 2 fs. A non-bonded cut-off of 1.2 nm was used, and long-range electrostatics were treated by PME using a grid spacing of 0.1 nm. The restrained equilibration was followed by an unrestrained equilibration for 5 ns in the NPT ensemble keeping all other simulation settings identical as

before. Finally a snapshot with a volume near the average of this ensemble was used as a starting point for the metadynamics simulation.

In order to estimate the motion necessary for ligand unbinding, we have performed well-tempered metadynamics using the distance between centre of mass of the active site residues[4] and the ligand heavy atoms as CV. This kind of CV was used recently by *Dodda et al.* in order to study unbinding of a small molecule inhibitor from HIV NNRTI.[48] The metadynamics simulation was performed at 300K with a gaussian height of 1.2 kJ/mol and a width at 0.011 nm, deposited every 1 ps. The simulation was stopped when CV distance reached $\sim$ 5 nm.

Reweighting with distance

To understand the conformational flexibility of flap and coil region region in apo Plm-II, we devised two distances (DIST2 and DIST3) which quantifies conformational flexibility across all simulations (Figure 20). In case of apo BACE-1, C_α distance between Asp-32 and Thr-72 is referred as DIST2 whereas the C_α distance between Ash-228 and Gln-326 is referred as DIST3.

The distances were not used as CVs. In our initial test calculations, we observed that a moving harmonic potential on the distances led to distorted flap conformations. However, we reweighted the free energy landscapes as a function of DIST2 and DIST3 as the distances gave a better quantitative view of the movement of the flap and the coil region. Reweighting the free energy surfaces on DIST2 and DIST3 also gives an understanding of which is the dominant movement between the flap and the coil region. Reweighting was also performed on several H-bond distances involving Tyr.

A reweighting algorithm has been used to calculate the unbiased probability distribution of variables in the well-tempered metadynamics simulations.[49] If a bias is acting on a system, it is constantly changing the biased probability distribution of the system. Using the reweighting algorithm, we can remove the effect of bias and recover the unbiased probability

[4]active site residues are: Tyr77, Asp34, Asp214, Ile123, Ser37, Ile32, Val78, Thr217.

distribution along chosen degrees of freedom.

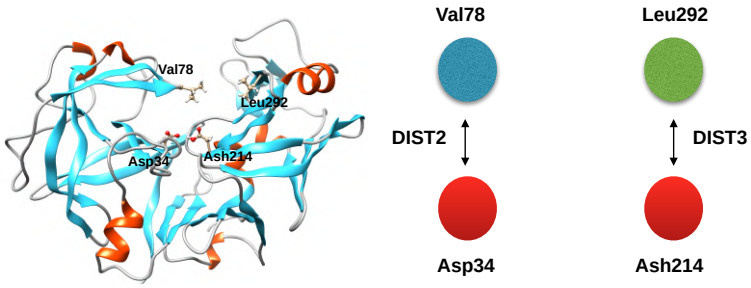

Figure 20: Representation of distance matrices, DIST2 and DIST3 used in this study to quantify movement of flap (DIST2) and coil region (DIST3) in apo Plm-II.

Computational tools

All simulations were performed using *Gromacs 5.0.2*[50] patched with *Plumed*.[51] The *tLeap* module of Amber was used to generate topology and co-ordinate files. *Acpype*[52] was used to convert the Amber topology and co-ordinates to a *Gromacs* compatible version.

Statistical analysis

Time evaluation of free energy difference between flipped and normal ($-\frac{\pi}{3}$ *radian*) state in Torsion-Metad simulations. We integrated multiple free-energy profiles in two basins defined by the following intervals on χ_1 space: flipped, $+2.5 < \chi_1 < -2.5$; normal, $-1.5 < \chi_1 < -0.5$.

The block averages are calculated over a period of 100 ns and the precision is approximated by standard deviation in each block.

Table 4: Block averaging performed on the free energy difference between flipped and normal state in two independent Torsion-Metad simulations with TIP3P water model. Final denotes averaging over time-interval 600-887 ns (R1) and 600-766 ns (R2) respectively.

	Average (kJ/mol)	Std. Dev
TIP3P-R1		
600-650 ns	2.82	0.01
650-700 ns	2.79	0.004
700-750 ns	2.75	0.03
750-800 ns	1.79	0.91
800-850 ns	-0.91	0.78
Final	1.34	1.86
TIP3P-R2		
600-650 ns	1.26	0.18
650-700 ns	-0.28	1.63
700-750 ns	-1.93	0.90
Final	-0.31	1.62

Table 5: Block averaging performed on the free energy difference between flipped and normal state in Torsion-Metad simulations with apo Plm-II.*Final* denotes the time interval of 600-700 ns and 500-650 ns for TIP4P and CHARMM respectively.

	Average (kJ/mol)	Std. Dev
TIP4P		
600-650 ns	2.59	0.07
650-700 ns	1.87	0.62
Final	2.35	0.58
CHARMM		
500-550 ns	-5.18	0.35
550-600 ns	-6.58	2.65
600-650 ns	-5.57	1.59
Final	-5.76	1.89

Table 6: Block averaging performed on the free energy difference between flipped and normal state in Torsion-Metad simulations starting with SO, SN and SSO conformations. *Final* denotes the time interval of 300-550 ns.

	Average (kJ/mol)	Std. Dev. (kJ/mol)
SO		
300-350 ns	-4.44	2.22
350-400 ns	0.25	1.12
400-450 ns	-2.79	1.17
450-500 ns	1.69	0.73
500-550 ns	-1.62	1.32
Final	-1.40	2.60
SN		
300-350 ns	-3.80	1.17
350-400 ns	-0.37	1.41
400-450 ns	-2.39	0.66
450-500 ns	0.29	1.34
500-550 ns	-7.47	1.49
Final	-2.76	3.06
SSO		
300-350 ns	-6.98	1.17
350-400 ns	-7.94	1.35
400-450 ns	-11.45	2.68
450-500 ns	-2.52	1.11
500-550 ns	-1.63	0.98
Final	-6.10	3.96

Figures and Tables

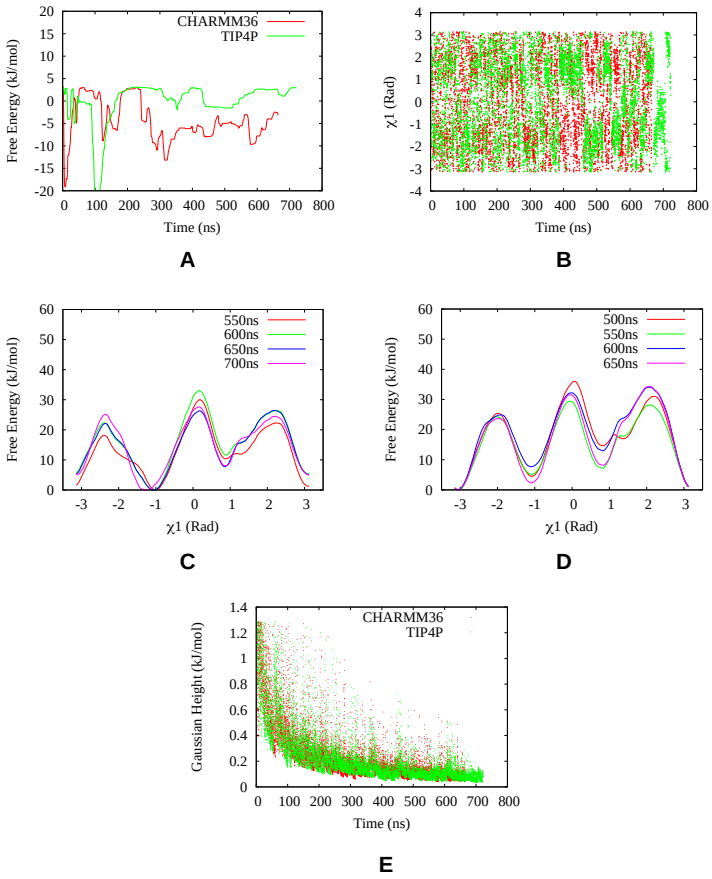

Figure 21: Free energy difference between flipped and normal state as a function of simulation time (A). Torsion-Metad with TIP4P water model converged around same value that of TIP3P (Figure 12). Fluctuation of χ_1 as a function of simulation time (B) shows diffusive behavior in the CVs space when the hill height quite small (E). Free energy profile of χ_1 from the last $\sim$150 ns of Torsion-Metad with TIP4P (C) and CHARMM (D) looks quite similar, apart from a constant offset.

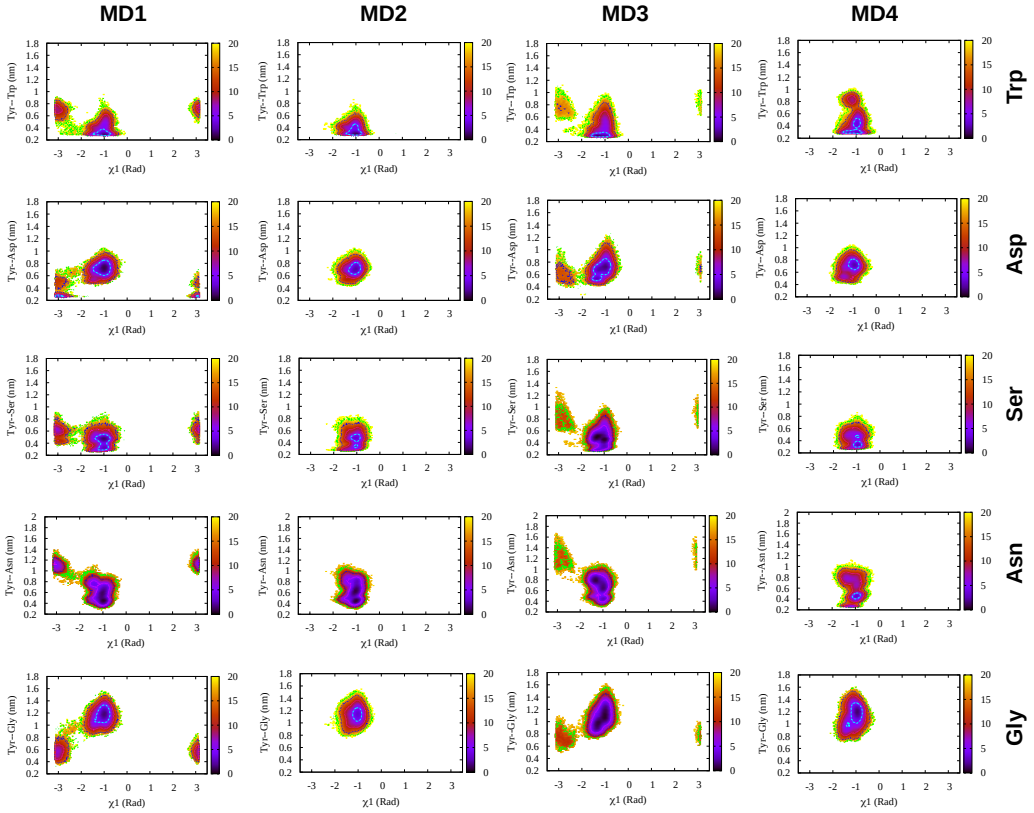

Figure 22: 2D FES projected on χ_1 and H-bond distances in case of independent MD simulations with TIP3P water model. **MD1** did a slight sampling of normal and flipped states. The flipped state was stabilised by formation of H-bond to Asp. In case of **MD1-MD4**, the normal state was stabilised by inter-changing H-bonds to Trp and Ser. **MD4** sampled additional H-bond to Asn.

Table 7: Sampling of open conformations (DIST2 > 1.8 nm.) in MD (TIP3P water model) and metadynamics simulations. Tor-Metad-1 and Tor-Metad-2 correspond to two independent metadynamics runs with torsion CVs.

Simulation	% Open
R1-T3P	8
R2-T3P	0.2
R3-T3P	0.2
R4-T3P	0.4
Tor-Metad-1	12
Tor-Metad-2	9

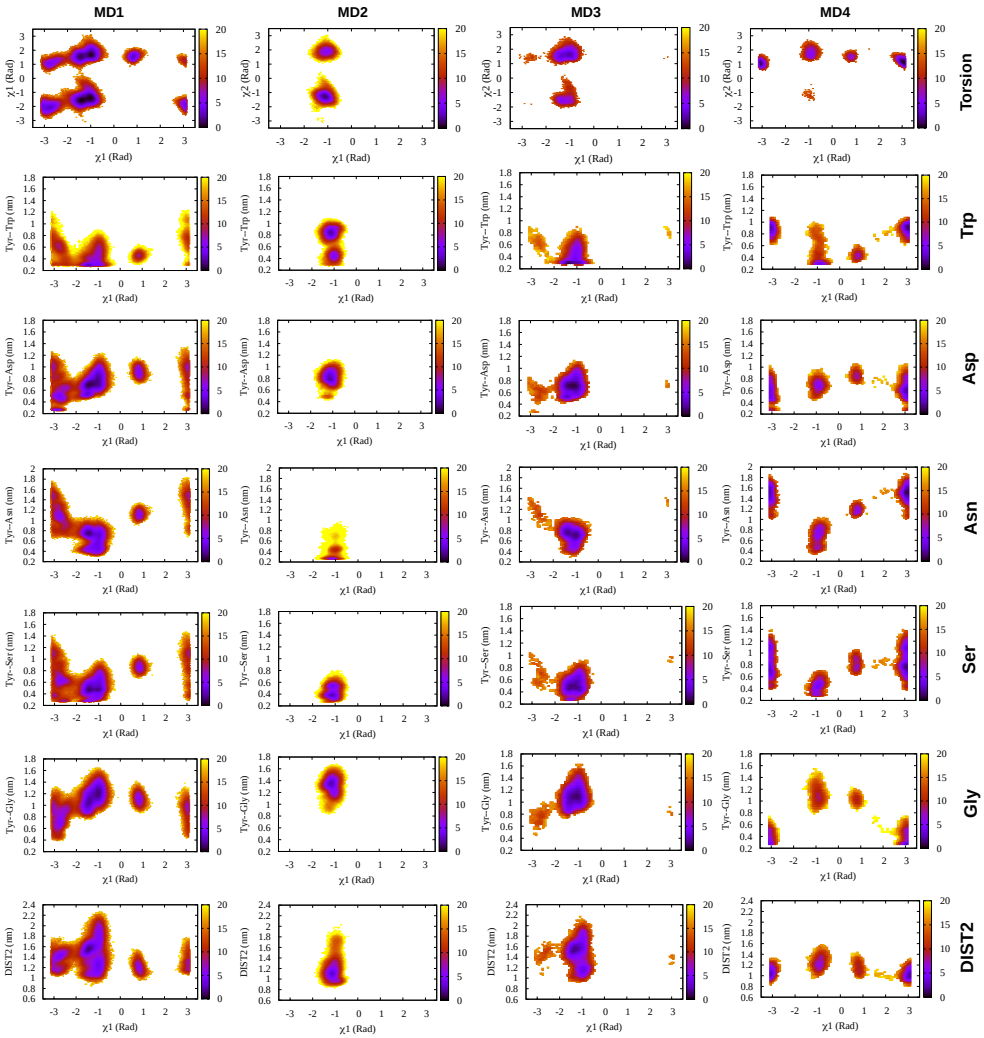

Figure 23: 2D FES projected on χ_1 and other variables (e.g. χ_2, H-bond distances (Trp, Asp, Ser, Asn and Gly) and DIST2) in case of four independent MD simulations with TIP4P water-model. **MD1** and **MD4** sampled both normal and flipped states. The flipped state in **MD4** was stabilised by inter-changing H-bond to Asp and Gly.

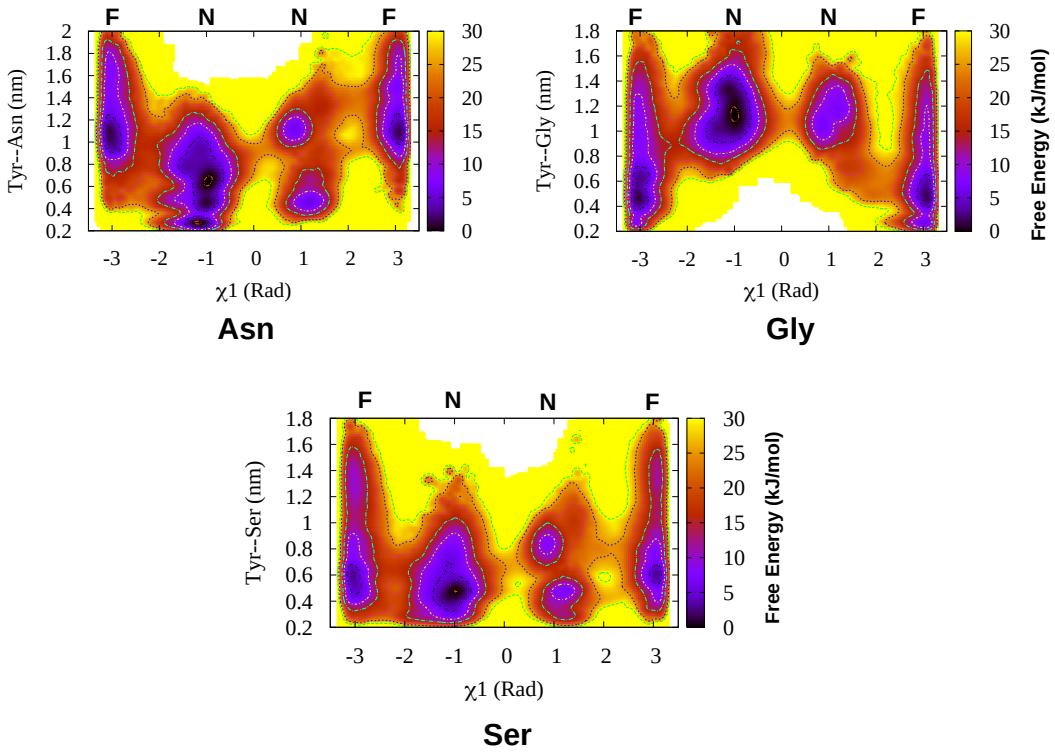

Figure 24: FES reweighted on χ_1 and H-bond distances (Asn, Ser and Gly) in case of Torsion-MetaD simulations with apo Plm-II. **F** and **N** denotes flipped and normal states respectively.

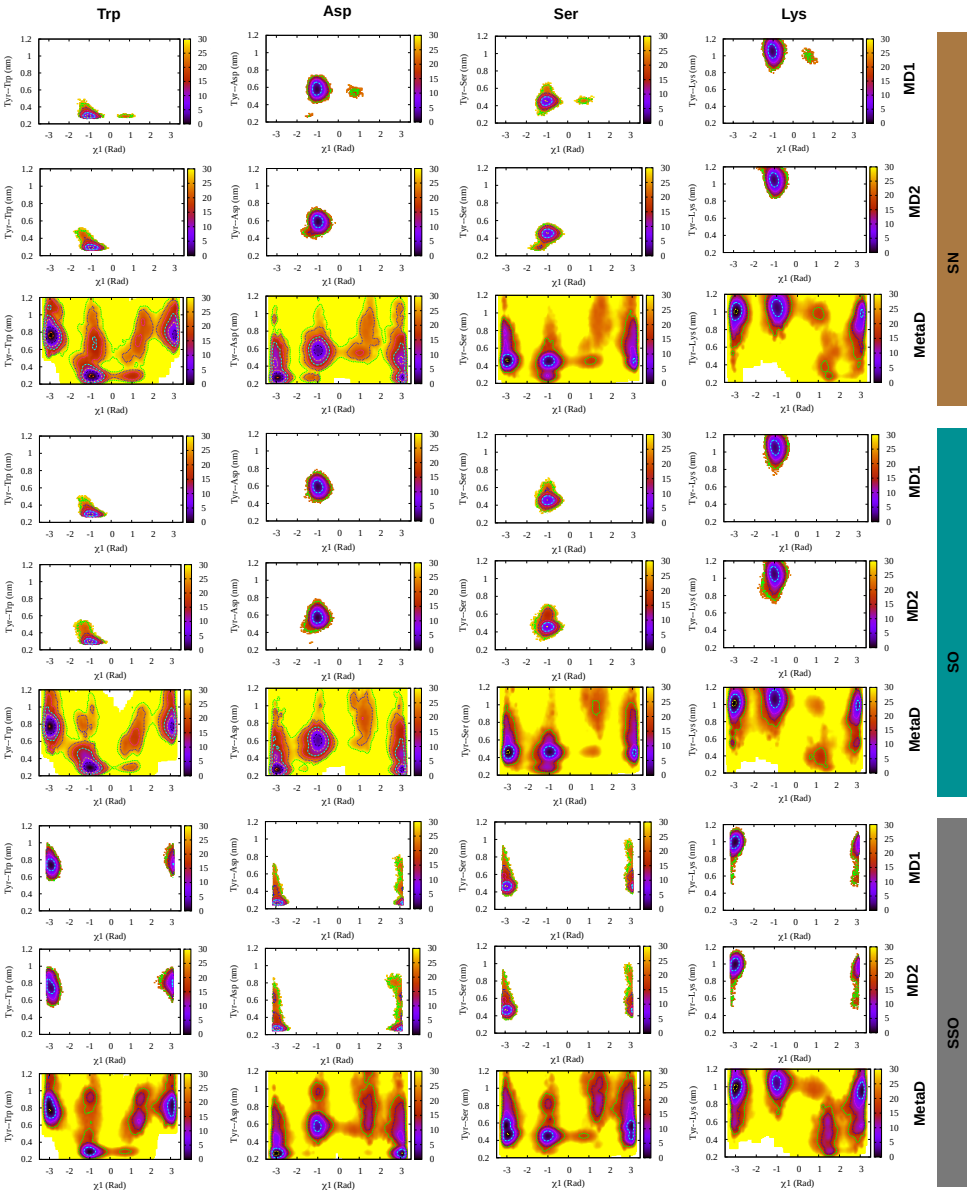

Figure 25: 2D FES projected on χ_1 and H-bond distances in case of MD and Torsion-Metad simulations starting with SN, SO and SSO conformations. MD simulations starting with SO and SN conformations only sampled the normal state. Whereas, simulations starting with SSO conformation remained stuck in the flipped state. Torsion-Metad starting with SO, SN and SSO conformations sampled both normal and flipped states.

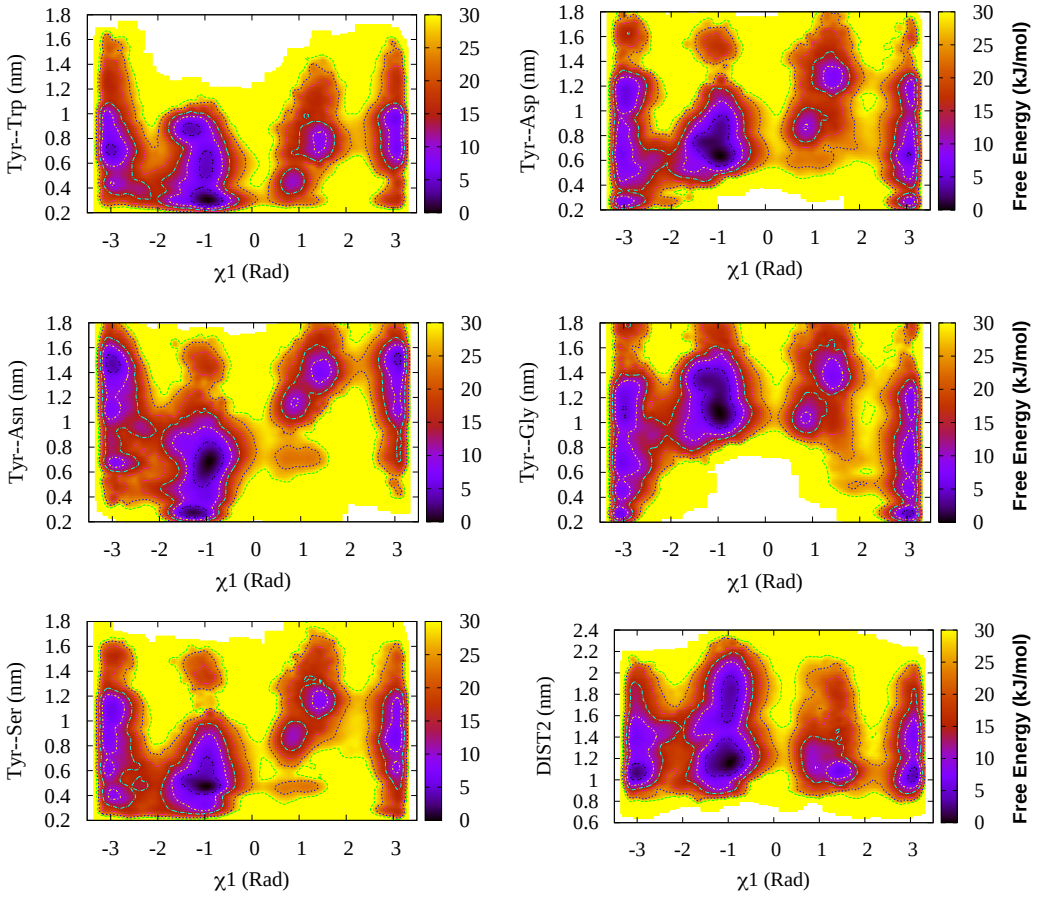

Figure 26: 2D FES reweighted on χ_1 and distances (H-bond to Trp, Asp, Ser, Asn, Gly and DIST2) in Torsion-Metad simulation with TIP4P water model in apo Plm-II.

References

(1) Mahanti, M.; Bhakat, S.; Nilsson, U. J.; Söderhjelm, P. Flap Dynamics in Aspartic Proteases: A Computational Perspective. *Chemical Biology and Drug Design* **2016**, *88*, 159–177.

(2) Sadiq, S. K.; De Fabritiis, G. Explicit solvent dynamics and energetics of HIV-1 protease flap opening and closing. *Proteins: Structure, Function, and Bioinformatics* **2010**, *78*, 2873–2885.

(3) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 915–920.

(4) Arodola, O. A.; Soliman, M. E. Molecular Dynamics Simulations of Ligand-Induced Flap Conformational Changes in Cathepsin-D—A Comparative Study. *Journal of Cellular Biochemistry* **2016**, *117*, 2643–2657.

(5) Kumalo, H. M.; Bhakat, S.; Soliman, M. E. Investigation of flap flexibility of β-secretase using molecular dynamic simulations. *Journal of Biomolecular Structure and Dynamics* **2016**, *34*, 1008–1019.

(6) Kumalo, H. M.; Soliman, M. E. A comparative molecular dynamics study on BACE1 and BACE2 flap flexibility. *Journal of Receptors and Signal Transduction* **2016**, *36*, 505–514.

(7) Spronk, S. A.; Carlson, H. A. The role of tyrosine 71 in modulating the flap conformations of BACE1. *Proteins: Structure, Function, and Bioinformatics* **2011**, *79*, 2247–2259.

(8) Dunn, B. M. Structure and Mechanism of the Pepsin-Like Family of Aspartic Peptidases. *Chemical Reviews* **2002**, *102*, 4431–4458.

(9) Dan, N.; Bhakat, S. New paradigm of an old target: An update on structural biology and current progress in drug design towards plasmepsin {II}. *European Journal of Medicinal Chemistry* **2015**, *95*, 324 – 348.

(10) Ersmark, K.; Samuelsson, B.; Hallberg, A. Plasmepsins as potential targets for new antimalarial therapy. *Medicinal Research Reviews* **2006**, *26*, 626–666.

(11) Karubiu, W.; Bhakat, S.; McGillewie, L.; Soliman, M. E. S. Flap dynamics of plasmepsin proteases: insight into proposed parameters and molecular dynamics. *Molecular BioSystems* **2015**, *11*, 1061–1066.

(12) McGillewie, L.; Soliman, M. E. The binding landscape of plasmepsin V and the implications for flap dynamics. *Molecular BioSystems* **2016**, *12*, 1457–1467.

(13) McGillewie, L.; Soliman, M. E. Flap flexibility amongst plasmepsins I, II, III, IV, and V: Sequence, structural, and molecular dynamics analyses. *Proteins: Structure, Function, and Bioinformatics* **2015**, *83*, 1693–1705.

(14) Friedman, R.; Caflisch, A. Pepsinogen-like activation intermediate of plasmepsin II revealed by molecular dynamics analysis. *Proteins: Structure, Function, and Bioinformatics* **2008**, *73*, 814–827.

(15) Andreeva, N.; Dill, J.; Gilliland, G. L. Can enzymes adopt a self-inhibited form? results of X-ray crystallographic studies of chymosin. *Biochemical and Biophysical Research Communications* **1992**, *184*, 1074 – 1081.

(16) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 826–843.

(17) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics* **2008**, *71*, 126601.

(18) Tiwary, P.; Parrinello, M. From Metadynamics to Dynamics. *Phys. Rev. Lett.* **2013**, *111*, 230602.

(19) Abrams, C.; Bussi, G. Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. *Entropy* **2014**, *16*, 163.

(20) Kastner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 932–942.

(21) Grubmuller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E* **1995**, *52*, 2893–2906.

(22) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annual Review of Physical Chemistry* **2016**, *67*, 159–184, PMID: 26980304.

(23) Recacha, R.; Leitans, J.; Akopjana, I.; Aprupe, L.; Trapencieris, P.; Jaudzems, K.; Jirgensons, A.; Tars, K. Structures of plasmepsin II from *Plasmodium falciparum* in complex with two hydroxyethylamine-based inhibitors. *Acta Crystallographica Section F* **2015**, *71*, 1531–1539.

(24) Suzuki, F.; Goto, K.; Shiratori, Y.; Inagami, T.; Murakami, K.; Nakamura, Y. Tyrosine 83 of renin has an important role in renin?angiotensinogen reaction. *Protein and Peptide Letters* **1996**, *3*, 45–49.

(25) Park, Y.-N.; Aikawa, J.-i.; Nishiyama, M.; Horinouchi, S.; Beppu, T. Involvement of a residue at position 75 in the catalytic mechanism of a fungal aspartic proteinase, Rhizomucor pusilus pepsin. Replacement of tyrosine 75 on the flap by asparagine enhances catalytic efficiency. *Protein Engineering, Design and Selection* **1996**, *9*, 869–875.

(26) Suzuki, J.; Sasaki, K.; Sasao, Y.; Hamu, A.; Kawasaki, H.; Nishiyama, M.; Hori-

nouchi, S.; Beppu, T. Alteration of catalytic properties of chymosin by site-directed mutagenesis. *Protein Engineering, Design and Selection* **1989**, *2*, 563–569.

(27) Munsamy, G.; Ramharack, P.; Soliman, M. E. S. Egress and invasion machinery of malaria: an in-depth look into the structural and functional features of the flap dynamics of plasmepsin IX and X. *RSC Adv.* **2018**, *8*, 21829–21840.

(28) Mukherjee, B.; Tessaro, F.; Vahokoski, J.; Kursula, I.; Marq, J.-B.; Scapozza, L.; Soldati-Favre, D. Modeling and resistant alleles explain the selectivity of antimalarial compound 49c towards apicomplexan aspartyl proteases. *The EMBO Journal* **2018**, *37*, e98047.

(29) Park, Y. N.; Aikawa, J.; Nishiyama, M.; Horinouchi, S.; Beppu, T. Involvement of a residue at position 75 in the catalytic mechanism of a fungal aspartic proteinase, Rhizomucor pusillus pepsin. Replacement of tyrosine 75 on the flap by asparagine enhances catalytic efficiency. *Protein Eng* **1996**, *9*, 869–875.

(30) Park, Y. N.; Aikawa, J.; Nishiyama, M.; Horinouchi, S.; Beppu, T. Site-directed mutagenesis of conserved Trp39 in Rhizomucor pusillus pepsin: possible role of Trp39 in maintaining Tyr75 in the correct orientation for maximizing catalytic activity. *J Biochem* **1997**, *121*, 118–121.

(31) Nguyen, W.; Hodder, A. N.; de Lezongard, R. B.; Czabotar, P. E.; Jarman, K. E.; O'Neill, M. T.; Thompson, J. K.; Sabroux, H. J.; Cowman, A. F.; Boddey, J. A.; Sleebs, B. E. Enhanced antimalarial activity of plasmepsin V inhibitors by modification of the P2 position of PEXEL peptidomimetics. *European Journal of Medicinal Chemistry* **2018**, *154*, 182 – 198.

(32) Hodder, A. N.; Sleebs, B. E.; Czabotar, P. E.; Gazdik, M.; Xu, Y.; O'Neill, M. T.; Lopaticki, S.; Nebl, T.; Triglia, T.; Smith, B. J.; Lowes, K.; Boddey, J. A.; Cowman, A. F.

Structural basis for plasmepsin V inhibition that blocks export of malaria proteins to human erythrocytes. *Nature Structural & Molecular Biology* **2015**, *22*, 590–596.

(33) Bhaumik, P.; Gustchina, A.; Wlodawer, A. Structural studies of vacuolar plasmepsins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2012**, *1824*, 207 – 223, Proteolysis 50 years after the discovery of lysosome.

(34) Asojo, O. A.; Gulnik, S. V.; Afonina, E.; Yu, B.; Ellman, J. A.; Haque, T. S.; Silva, A. M. Novel Uncomplexed and Complexed Structures of Plasmepsin II, an Aspartic Protease from Plasmodium falciparum. *Journal of Molecular Biology* **2003**, *327*, 173 – 181.

(35) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research* **2012**, *40*, W537–W541.

(36) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **2015**, *11*, 3696–3713.

(37) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.

(38) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*, 3684–3690.

(39) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.

(40) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.

(41) Alejandre, J.; Chapela, G. A. The surface tension of TIP4P/2005 water model using the Ewald sums for the dispersion interactions. *The Journal of Chemical Physics* **2010**, *132*, 014701.

(42) Huang, J.; MacKerell Jr, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry* **2013**, *34*, 2135–2145.

(43) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Letter* **2008**, *100*, 20603.

(44) M. Sultan, M.; Pande, V. S. tICA-Metadynamics: Accelerating Metadynamics by Using Kinetically Selected Collective Variables. *Journal of Chemical Theory and Computation* **2017**, *13*, 2440–2447.

(45) David, C. C.; Jacobs, D. J. *Protein Dynamics: Methods and Protocols*; Humana Press: Totowa, NJ, 2014; pp 193–226.

(46) Harrigan, M. P.; Sultan, M. M.; Hernández, C. X.; Husic, B. E.; Eastman, P.; Schwantes, C. R.; Beauchamp, K. A.; McGibbon, R. T.; Pande, V. S. MSMBuilder: Statistical Models for Biomolecular Dynamics. *Biophysical Journal* **2017**, *112*, 10 – 15.

(47) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.

(48) Dodda, L. S.; Tirado-Rives, J.; Jorgensen, W. L. Unbinding Dynamics of Non-

Nucleoside Inhibitors from HIV-1 Reverse Transcriptase. *The Journal of Physical Chemistry B* **2019**, *123*, 1741–1748.

(49) Tiwary, P.; Parrinello, M. A Time-Independent Free Energy Estimator for Metadynamics. *The Journal of Physical Chemistry B* **2015**, *119*, 736–742, PMID: 25046020.

(50) Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19 – 25.

(51) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **2009**, *180*, 1961 – 1972.

(52) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2012**, *5*, 367.