Journal of Bioinformatics and Computational Biology
Vol. 10, No. 6 (2012) 1250018 (19 pages)

DOI: 10.1142/S0219720012500187

Imperial College Press
www.icpress.co.uk

CONFORMATIONAL DYNAMICS OF HIV-1 PROTEASE: A COMPARATIVE MOLECULAR DYNAMICS SIMULATION STUDY WITH MULTIPLE AMBER FORCE FIELDS

BISWA RANJAN MEHER[*,‡,¶], MATTAPARTHI VENKATA SATISH KUMAR[*,§,‖], SMRITI SHARMA[†,**] and PRADIPTA BANDYOPADHYAY[†,††]

[*]*Computational Biology Research Laboratory*
Department of Biotechnology, Indian Institute of Technology
Guwahati, Assam 781039, India

[†]*Centre for Computational Biology and Bioinformatics*
School of Computational and Integrative Sciences
Jawaharlal Nehru University, New Delhi 110067, India

[‡]*Computational Chemistry Laboratory*
Department of Natural Sciences, Albany State University, Albany
Georgia 31705, USA

[§]*Centre for Condensed Matter Theory*
Department of Physics, Indian Institute of Science, Bangalore 560012, India
[¶]*brmeher@gmail.com*
[‖]*mvenkatasatishkumar@gmail.com*
[**]*ssmriti@gmail.com*
[††]*praban07@gmail.com*

Received 6 February 2012
Revised 1 June 2012
Accepted 27 June 2012
Published 31 July 2012

Flap dynamics of HIV-1 protease (HIV-pr) controls the entry of inhibitors and substrates to the active site. Dynamical models from previous simulations are not all consistent with each other and not all are supported by the NMR results. In the present work, the effect of force field on the dynamics of HIV-pr is investigated by MD simulations using three AMBER force fields ff99, ff99SB, and ff03. The generalized order parameters for amide backbone are calculated from the three force fields and compared with the NMR S2 values. We found that the ff99SB and ff03 force field calculated order parameters agree reasonably well with the NMR S2 values, whereas ff99 calculated values deviate most from the NMR order parameters. Stereochemical geometry of protein models from each force field also agrees well with the remarks from NMR S2 values. However, between ff99SB and ff03, there are several differences, most notably in the loop regions. It is found that these loops are, in general, more flexible in the ff03 force field. This results in a larger active site cavity in the simulation with the ff03 force field. The effect of this difference in computer-aided drug design against flexible receptors is discussed.

Keywords: Molecular dynamics; force field; S2 order parameter; flap movement; flexible binding site.

[¶]Corresponding author.

1. Introduction

The development of better drugs against various Human Immunodeficiency Virus (HIV) proteins is an important goal to combat the HIV virus.[1] HIV-pr, one of the most indispensable enzymes for HIV replication, is an important target for drug design. It acts at the late stage of infection by cleaving the Gag and Gag−Pol polyproteins to yield mature infectious virions for the continuation of the viral life cycle. Several drugs targeting HIV-pr have been developed and approved by the Food and Drug Administration (FDA) for the treatment of HIV infection (www.UNAIDS.org). Unfortunately, resistance to these drugs has built up quickly in the form of mutations in the protease. The recent 10 protease inhibitors in clinical use[2] suffer from taxing side effects, poor pharmacokinetic properties, and the development of drug resistance.[3,4] So there is a great demand for the drugs that are less susceptible to mutations. A detailed knowledge of HIV-pr structure, dynamics, and its interaction with ligands and substrates is required to design effective and potent drugs.

Structurally, HIV-pr is a homodimeric aspartyl protease whose active site is capped by two identical flexible, glycine-rich ß-hairpins, or flaps that restrict access to the active site (shown in Fig. 1). With the functional Asp residues located at the dimer interface, the protease contains 198 residues with each subunit containing 99 residues. The residues of HIV-pr are numbered as 1−99 and 100−198 for chain A and

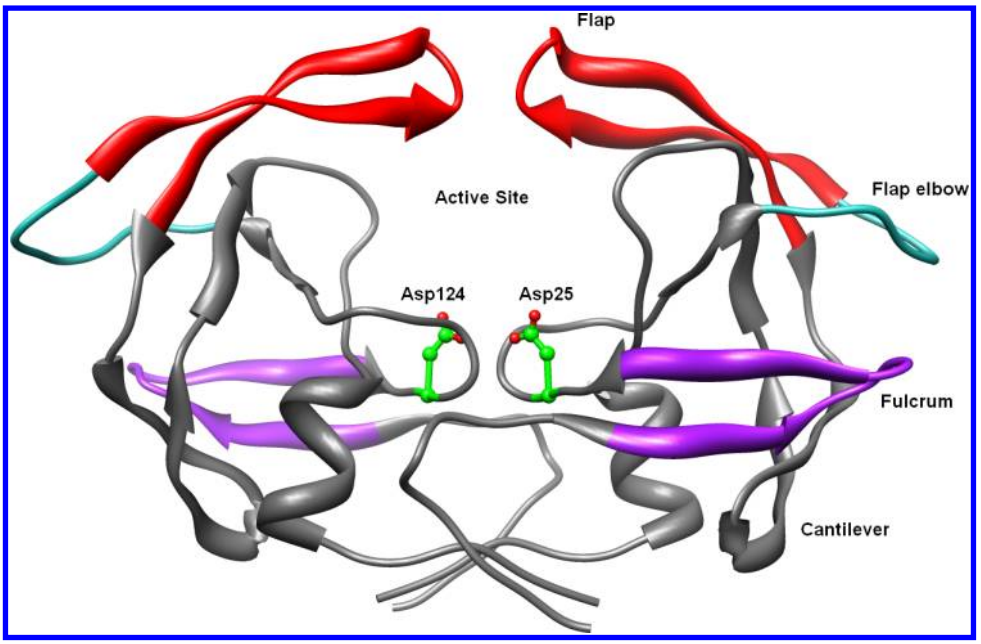


Fig. 1. Structure of HIV-pr. Catalytic residues (Asp25 in chain-A and Asp124 in chain-B) are represented by ball and stick. Different regions of the protein are colored in different colors like flap (red), flap elbow (cyan), and fulcrum (purple). For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

B, respectively. The residue index for different regions of HIV-pr are shown in brackets, flap (43−58 and 142−157), flap elbow (35−42 and 134−141), fulcrum (11−22 and 110−121), cantilever (59−75 and 158−174), and active site regions (23−30, 78−82 and 122−129, 177−181) for both the chains A and B.

NMR experiments propose a model for HIV-pr dynamics, where the flaps are described as an ensemble of semi-open structures with rapid movement of the flap tips. It has been suggested that the flap dynamics has two different time scales. The flap tip motion is rapid (in the nanosecond range), while the full flap movement is much slower (in the micro- to millisecond range).[5] Computer simulations of HIV-pr have been highly valuable because of their ability to observe the protein at atomic resolution, which is difficult for experiments. There have been a large number of computer simulation studies to understand the HIV-pr dynamics primarily using molecular dynamics (MD) simulation.[6−14] Our previous work on the MD simulation study of I47V mutation of HIV-pr showed that, along with a subtle change in the flap dynamics, a change in a specific interaction between the side chain of the residue 47 in chain B and the ligand was likely the main reason for drug resistance.[15] However, computer simulation studies provide various plausible models for the flap dynamics. Several simulations[7,10] suggested one interesting model where the flaps are in curled conformation (flaps move inward, toward the active site) not corroborated by the NMR results. A recent NMR study on HIV-pr by Ishima *et al.* suggests that the flaps weakly interact with each other even in the absence of ligands or substrates.[16] They also concluded that their experimental findings can be explained without involving any curled conformation. Also, some simulations observed rapid flap opening (in nanoseconds), contradicting the NMR model.[7] It is likely that some of the observations gained from the computer simulations may be influenced by the details of the simulation protocols followed. The initial solvent distribution is already discussed by Carlson's group. There are multiple studies in the recent past involving comparison of dynamics of certain molecules using AMBER (Assisted Model Building with Energy Refinement) force fields with NMR results.[17−21] The studies by different authors on different platforms have diverse opinions on the force field accuracy. However, to the best of our knowledge, there is no systematic study on the effect of force field on the conformational dynamics of HIV-pr, a kind of protein with outstanding therapeutic importance and a larger protein as compared to the previously published protein models. Further, the results from such a study will be useful for future investigation of proteins with flexible binding sites. Significant differences in MD trajectories may arise from several factors, such as the energy function used in the simulation (force field), the underlying integrator for time propagation, initial solvent distribution, initial velocity distribution, and convergence of the trajectories, to name a few.

In this study, we mainly focused on the force field effects on the conformational dynamics, especially the flap dynamics of HIV-pr by using three different AMBER force fields namely ff99,[22] ff03,[23] and ff99SB.[24] The TIP3P[25] water model was used with all the force fields. All other simulation protocols remained the same for all three trajectories. An MD simulation for 10 ns was performed initially for each of the force

fields. Afterward, simulations with ff99SB and ff03 force fields were continued for up to 30 ns. Another trajectory of 30 ns was run for ff99SB and ff03 starting from a different initial velocity distribution. The ff99SB force field differs from the ff99 force field by having better backbone dihedral parameters. In the ff03 force field, the charges are determined by performing quantum chemical calculation of peptides in condensed phase environment and the main chain torsion parameters are also different from ff99 and ff99SB. While there are several structural and dynamical quantities that can be calculated from MD simulation, it is not easy to validate those without any experimental results. In the present work, we have focused primarily on the N−H S2 order parameters calculated from the simulations and compared them with the available experimental NMR order parameters.[26] Various structural parameters were also examined to understand the force field dependence of the conformational dynamics. One major source of discrepancy between NMR S2 and calculated S2 values is the extent of conformational sampling in the MD simulation. It is difficult to obtain a full convergence of conformation sampling of large biomolecular systems using an explicit solvent.[27] Thus no attempt is made in the current work to achieve full convergence. Rather the focus of this work is to observe the differences in result from MD simulation with different force fields. Two separate trajectories were run in order to assess the quality of the results.

The results of the simulation indicate that the N−H S2 order parameters calculated from the ff99SB and ff03 are reasonably close to the NMR S2 ($S2_{NMR}$) values for most of the residues. The ff99 values show maximum deviation from the $S2_{NMR}$ values. As a result, we stopped MD simulation using ff99 after 10 ns. Although there are some similarities between ff99SB and ff03, significant differences remain for several residues. These are the residues in the loop region, many of these being glycine. A similar finding on a different protein, GB3, is reported.[28] A major consequence of the different fluctuations of these loop residues is the larger active site cavity observed in the ff03 simulation. Examination of hydrogen bonding (H-bonding) of the loop residues show that the H-bonding pattern differs for some of the residues. It appears that multiple factors, such as H-bonding description, different torsion parameters in the two force fields, and the interplay between different terms in the force field are contributing to the differential mobility of the loop residues. This work shows that the complex dynamics of HIV-pr can be quite sensitive to the force field difference. This partly explains the discrepancy seen in previous simulations of HIV-pr.

One important consequence of the sensitivity of the binding site flexibility to the simulation parameters is how the conformational dynamics of wild-type and different mutant forms of HIV-pr depend on the simulation details. Another consequence of our result is in the ensemble-based docking approach against flexible receptors. In this approach, potential inhibitors are docked to an ensemble of protein structure, very often obtained from MD simulation.[29] However, the binding of ligands may be influenced by the variation in the average geometry and flexibility of the binding site between simulations with different parameters. More generally, many of the conclusions regarding biological function (such as binding, enzymatic action) are drawn

from MD simulations by observing the differences in dynamics. However, the current work shows that careful examination with different simulation parameters is required before making such kind of conclusion.

2. Computational Methods

2.1. *System setups*

All the three MD simulations were started using a crystal structure of HIV-pr (pdb ID: 1HHP)[30] of resolution 2.7 Å with semi-open conformation. The LEaP module of the AMBER 9 program package[31] was used to prepare the system for simulation. The TIP3P[25] water model was used in all the simulations due to its better structural and thermodynamic properties. The system was immersed in a water box of size $83.5 \times 62.5 \times 68.9\,\text{Å}^3$ containing more than 8000 water molecules. A cutoff of 10 Å was used along the three axes to discard any water molecule if it is farther than the cutoff from any solute molecule. A default cutoff of 8.0 Å was used for Lennard-Jones interactions. The net positive charge on the system was neutralized through the addition of seven chloride ions. The protonation states of catalytic aspartates (Asp-25 and Asp-124) vary depending on the binding of inhibitors or substrates. As the starting model in the current simulation study is a ligand-free structure, the protonation states of both the catalytic aspartates were kept deprotonated in accordance to the finding in Ref. 32 that the carboxyl of Asp25 was protonated only when it is in complexed form. The electrostatic interactions were calculated with the particle mesh ewald (PME) method.[33] Constant temperature and pressure conditions in the simulation were achieved by coupling the system to a Berendsen's thermostat and barostat.[34] Bonds involving the hydrogen atoms were constrained to their equilibrium position with the SHAKE algorithm. The generalized order parameters (S2) were calculated from a plateau region of the N−H internuclear vector autocorrelation function using the *ptraj* module. Especially, the autocorrelation function was calculated up to time of half of trajectory length and the mean of last 5 ns was taken as S2.[24] The correlation curves obtained were fit to the Lipari−Szabo "model-free" equation.[35]

2.2. *Molecular dynamics simulations*

The system was minimized in two phases. In the first phase, the system was minimized giving restraints ($30\,\text{kcal/mol/Å}^2$) to protein and crystallographic waters for 500 steps with subsequent second phase minimization of the whole system. Then the system was heated from 0−300 K with a gap of 50 K over 30 ps with a 1 fs time step. The protein atoms were restrained with force constant of $30\,\text{kcal/mol/Å}^2$ at the NVT ensemble. After that, the force constant was reduced by $10\,\text{kcal/mol/Å}^2$ in each step to reach the unrestrained structure in three steps of 10 ps each. The system was then switched over to the NPT ensemble and equilibrated without any restraints for 180 ps. The system was equilibrated in total of 240 ps. The time step for MD simulation for the production run was 2 fs. All the three trajectories were initially run for 10 ns. The ff99SB and ff03 simulations were later extended to 30 ns. One more

trajectory of 30 ns was run with ff99SB and ff03 force fields with a different initial velocity distribution to check the biasness of a single trajectory. The convergence of energies, temperature, pressure, and global RMSD was used to verify the stability of the systems. The hydrogen bonds were analyzed using the *ptraj* module of AMBER program. Formation of the H-bonds depends on the distance and angle cutoff as follows: (a) distance between proton donor and acceptor atoms were $\leq$3.5 Å, and (b) the angle between donor-H … acceptor was $\geq 120°$. Graphic visualization, making of the videos, and presentation of protein structures were done using Chimera.[36] Calculation and analysis of the Ramachandran plot for the protein structures were done using the PROCHECK[37] program.

3. Results and Discussions

3.1. *RMSD of the Cα atoms*

Figure 2 shows the RMSD values of the Cα atoms relative to the initial equilibrated structure calculated from the three simulations, which ensure stable trajectories. The ff99 simulation data is up to 10 ns, while the ff99SB and ff03 simulations are up to 30 ns. The ff99 force field shows the maximum RMSD among the three force fields

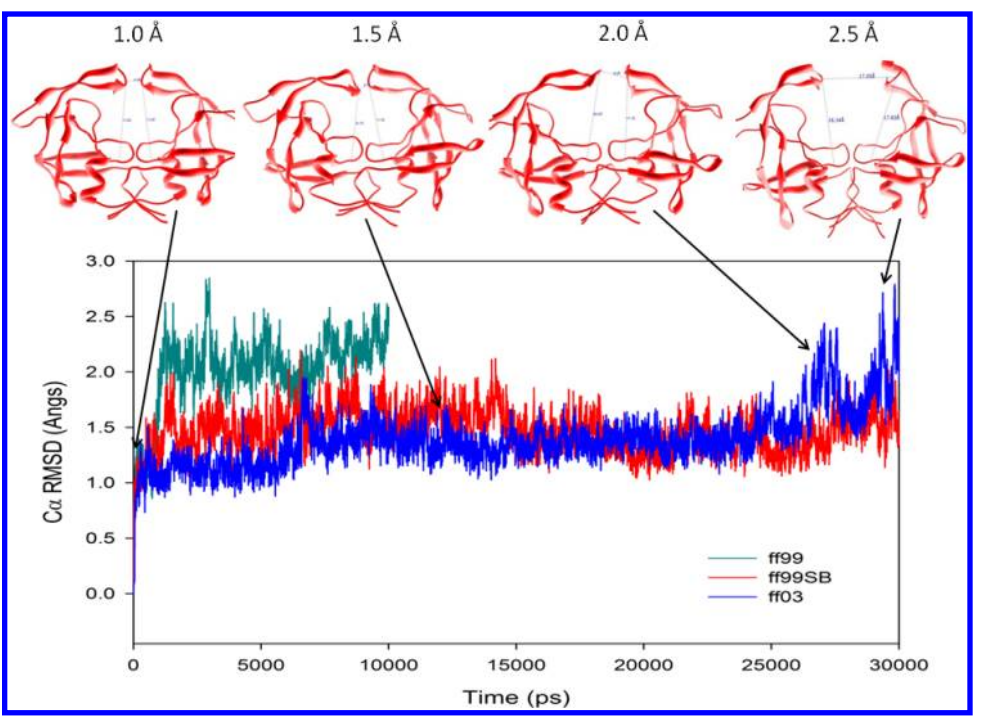
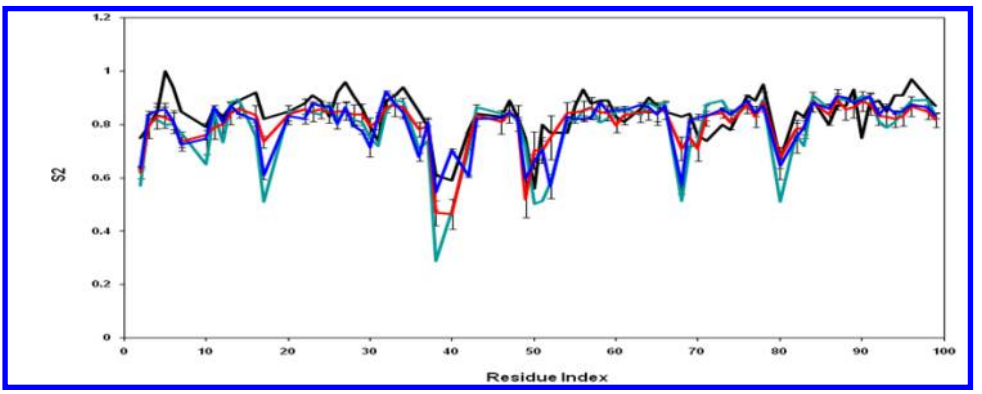


Fig. 2. RMSD values for the Cα atoms for the trajectories with force fields ff99, ff99SB, and ff03. All the simulations started from a semi-open crystal structure (1HHP). Snapshots (cartoon diagrams, side view) along the trajectory show the structures with RMSD values 1.0, 1.5, 2.0, and 2.5 Å. Large flap openings are sampled (structure with 2.5 Å RMSD) with flap tip (Ile50Cα-Ile149Cα) distances reaching $\geq$17.0 Å.

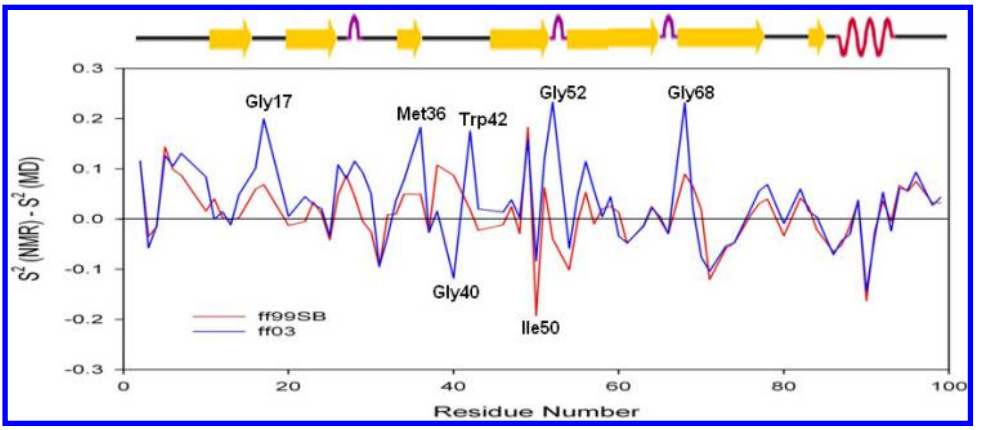
with values fluctuating mostly between 2.0 and 2.5 Å. RMSD values fluctuate generally below 2 Å for ff99SB and ff03 force fields. There is a sharp increase of RMSD in the ff03 trajectory in the last 3 ns of the MD simulation, suggesting a major change in HIV-pr conformation. This change in conformation leads to a flap opening distance between the flap tips (Ile50Cα-Ile149Cα) up to $\geq$17.0 Å in case of ff03 force field.

3.2. *Generalized N−H order parameter (S2)*

The calculated S2 values ($S2_{MD}$) for the N−H bond vector averaged over the two chains of HIV-pr from the three force fields are compared with the experimental $S2_{NMR}$.[26] In general, most of the $S2_{MD}$ values are lower than the $S2_{NMR}$ values as found by the previous workers.[24] Figure 3(a) shows the comparison of $S2_{MD}$



(a)



(b)

Fig. 3. (a) Comparison of experimental (black line) NMR S2 N−H generalized-order parameter values with the calculated values from ff99 (cyan line), ff99SB (red line) and ff03 (blue line) force field simulations. Order parameters (S2) averaged for both monomers, with error bars reflecting the difference in case of ff99SB and ff03. (b) The difference between ff99SB and ff03 calculated order parameters from the NMR S2 values.

calculated with three force fields with $S2_{NMR}$. It is clearly seen that the ff99 values differ most from the $S2_{NMR}$. The ff99 calculated S2 values are lowest and hence makes HIV-pr most flexible among the three force fields considered. It shows six regions of HIV-pr highly flexible (low values of S2), not corroborated by the NMR results. The ff99SB values show similar trend as the $S2_{NMR}$ with maximum differences comes in the residues 38−40 (flap elbow) and in the flap tip region. The maximum deviation between the ff99 and ff99SB results is observed in the residues Leu10, Gly16-17 (fulcrum region), Thr31, flap region (49−52), Gly68, and Thr80. The experimental values are almost always higher than the ff99SB values except for the residues Thr31, Ile50, Ile54, Ala71, and Leu90.

The difference between the S2 calculated using ff99SB and ff03 is depicted in Fig. 3(b), which shows the difference between $S2_{MD}$ and $S2_{NMR}$ for these two force fields. The ff99SB values are, in general, higher than the ff03 values with the exception of residues 40 and 70. This is similar to the findings of Trbovic and coworkers.[28] The major difference between the ff99SB and ff03 values come for the residues Gly17, Met36, Gly40, Trp42, Ile50, Gly52, and Gly68. All of these residues are part of different loop regions. Gly17 is in the bend of the beta-hairpin of the fulcrum. Met36, Gly40 and Trp42 belong to the flap elbow loop. Residues Ile50 and Gly52 are part of the flap and Gly68 belongs to the part of the mobile loop 68−71 in the cantilever region (see Fig. 1). Among the glycines, Gly68 is a special case, since this residue is part of the mobile loop 68−71, with histidine in position 71, which is likely to be affected by the proton exchange, not captured in our simulation.[8] This in turn could change the dynamics of Gly68. One of the most important results from the present work is the higher mobility of many of the loop residues in the ff03 force field simulation. Except for the active site region (containing the catalytic triad and the active site wall), all the loops show difference in S2 calculated by ff99SB and ff03 force fields. It is to be noted that the $S2_{NMR}$ values were derived from an auto-catalysis-resistant mutant of HIV pr, where the mutations like Q7K, L33I, and L63I were carried out. This could one of the reasons for the difference between $S2_{NMR}$ and $S2_{MD}$. Chemical exchange is found to be important for the terminal residues (4−6, 97−99) and the flap residues in the NMR experiments.[5] Although the time scale of chemical exchange is not accounted for in our MD simulation, the differences that are observed between ff99SB and ff03 are found for other loops, where chemical exchange is not significant. Thus the difference between ff99SB and ff03 cannot be due to the omission of chemical exchange in the MD simulation. To cross-check the results, $S2_{MD}$ values are calculated from the second trajectory of 30 ns for both ff03 and ff99SB. In the second set of trajectories (Fig. S1, Supplementary Material section), the $S2_{MD}$ of all residues converge well with the residues from first set of trajectories except for few of the flap residues. Moreover, for the loop residues, ff03 derived S2 values are mostly lower as seen in the first trajectory. Therefore, this suggests that the conclusions drawn from the first set of trajectories are generalizable and not simulation specific.

Table 1. Ramachandran plot calculation and comparative analysis of the 1000 protein models from the three force fields (ff99, ff99SB and ff03) computed with the PROCHECK program.[a]

	ff99		ff99SB		ff03	
	Number of residues	%	Number of residues	%	Number of residues	%
Residues in most favored regions [A,B,L]	135	85.4	145	91.8	142	89.9
Residues in additional allowed regions [a,b,l,p]	21	13.3	12	7.6	15	9.5
Residues in generously allowed regions [∼a, ∼b, ∼l, ∼p]	2	1.3	0	0.0	1	0.6
Residues in disallowed regions	0	0.0	1	0.6	0	0.0
Number of non-glycine and non-proline residues	158	100	158	100	158	100
Number of end-residues (excl. Gly and Pro)	2	—	2	—	2	—
Number of glycine residues (shown as triangles)	26	—	26	—	26	—
Number of proline residues	12	—	12	—	12	—
Total number of residues	198	—	198	—	198	—

[a]Based on an analysis of 118 structures of resolution of at least 2.0 Å and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favored regions. Residues in the most favored regions are in [A,B,L], additional allowed regions are in [a,b,l,p] and generously allowed regions are in [∼a, ∼b, ∼l, ∼p] [from PROCHECK].

3.3. *Ramachandran Plot*

To diagnose the stereochemical geometry of each residue, we analyzed the Ramachandran Plot for the 1000 protein structures obtained from the initial 10 ns of the three trajectories. The data point (dihedral angles) on the plot for the calculated structures shows ample differences, according to which residues in most favored regions for ff99, ff99SB, and ff03 are 85.4%, 91.8%, and 89.9% respectively, that ensures the geometrically acceptable quality of ff99SB is greater than the ff03 and ff99 (Fig. S2 in the Supplementary Material section and Table 1).

It is to be noted that, the ff99 force field describes the protein dynamics poorly as compare to the ff99SB and ff03 force fields in the current study which is evident from the Cα RMSD values, S2 order parameters and the percentage of occupancy of residues in the most favored regions in Ramachandran plot. Hence, for the rest of this manuscript; we will be reporting the difference between the ff99SB and ff03 force fields only.

The differences in the Ramachandran angles of the interesting residues (Gly17, Met36, Gly40, Trp42, Ile50, Gly52, and Gly68) were also checked and found there are considerable differences in the Φ and Ψ angles. As for example the Ramachandran plot of residues Gly17 and Gly68 are shown in Fig. 4. It can be seen that the ff03 simulation samples two regions of Φ, Ψ space (Φ−180 to −45 and again +45 to +180), while ff99SB samples only one region of the Φ, Ψ

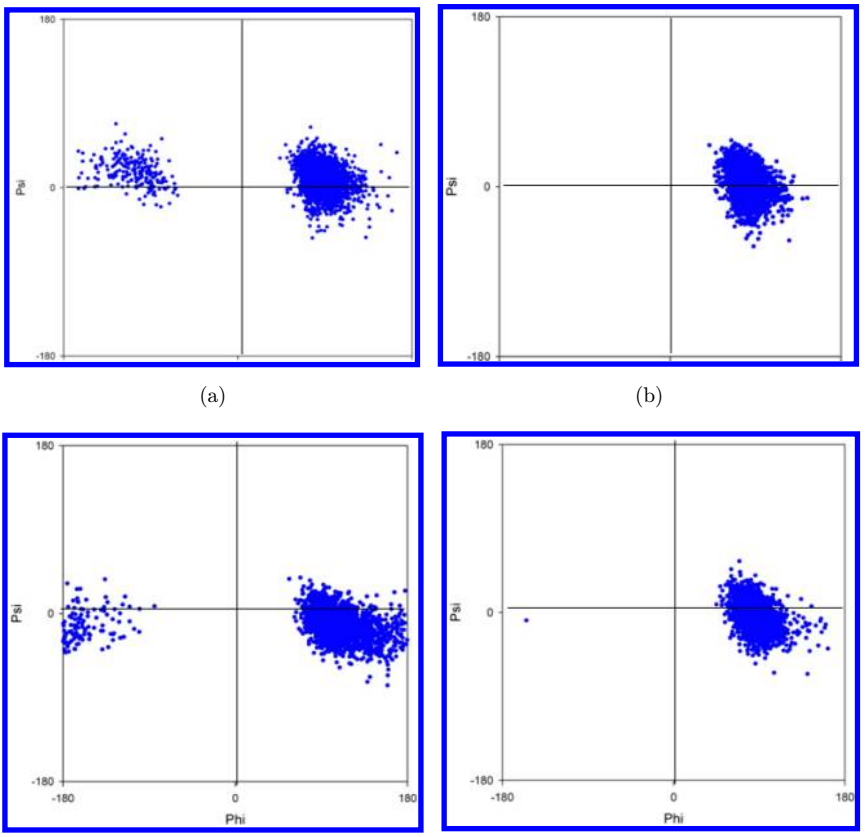


Fig. 4. Ramachandran plot for Glycine17 for (a) ff03 (b) ff99SB force fields and for Glycine68 for (c) ff03 (d) ff99SB force fields simulation. X-axis indicates the values for the angle Phi (Φ) and Y-axis indicates for angle Psi (Ψ). Protein structures were extracted from the whole 30 ns simulation trajectories.

space. Examination of Ramachandran plots of other five residues also shows that the sampling of secondary structure by ff99SB and ff03 are different (Fig. S3, Supplementary Material section).

3.4. *B-factors comparison*

The temperature factor (B-factor) of the protein with the ff99SB and ff03 force fields were calculated and compared with the experimental X-ray structure B-factor as shown in Fig. 5. The calculated B-factors are much higher than that for the X-ray structure, although the broad outline is similar with few exceptions. It is well known that for high-resolution crystal structures, the calculated B-factors can be much higher than that for experimental structures.

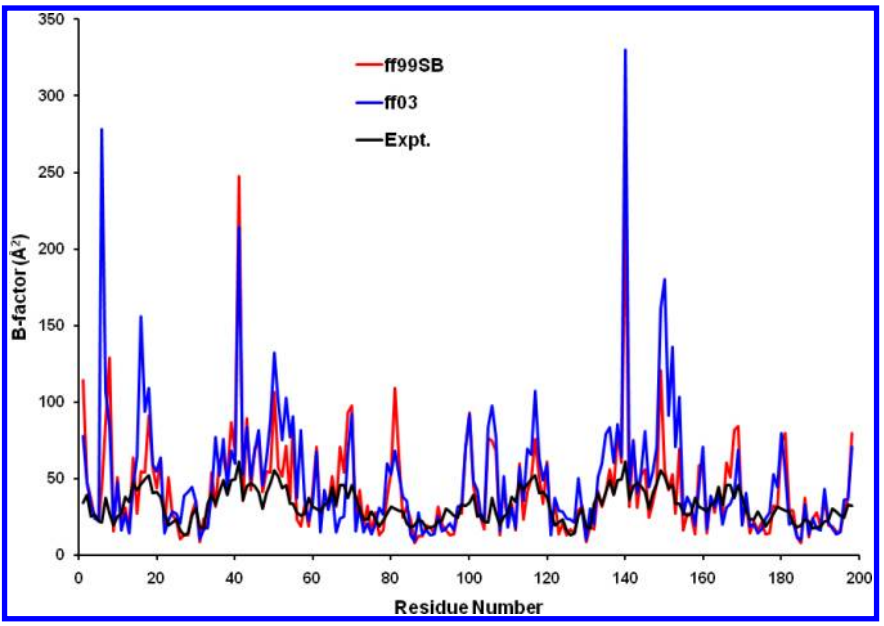


Fig. 5. Comparison of the calculated B-factor for ff99SB and ff03 force fields with that of X-ray crystallographic experimental B-factor (obtained from PDB id: 1HHP).

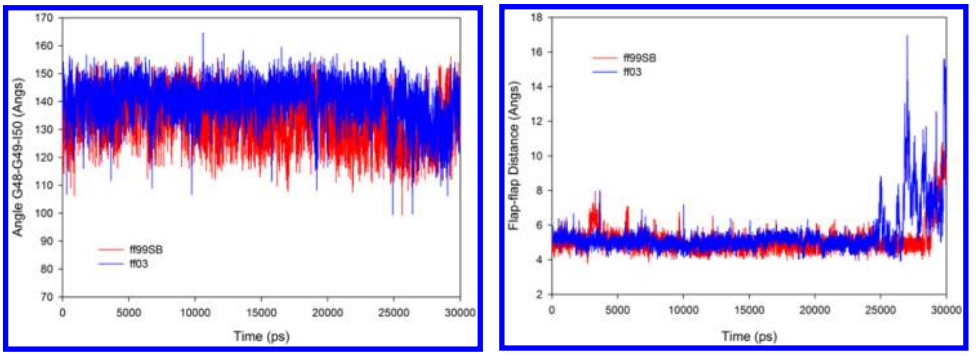
This finding may be correlated to the fact that high-resolution crystals are classically highly packed and such properties are not included in most simulations (including those performed in this study). Therefore, the ability of these force field parameters to simulate the true dynamics of the protein can be evaluated highly.

3.5. *Flap movement*

Understanding the flap movement is an important issue pertaining to the conformational dynamics of HIV-pr. In the present work, flap movement was examined in detail for the ff99SB and ff03 trajectories.

3.5.1. *Flap curling and flap−flap distance*

The examination of the relevant angles involving the residues in the flap region shows that the flap curling is not significant in either of the ff03 and or the ff99SB trajectories (Fig. 6(a)). The flap−flap distance (i.e. the distance between Ile50Cα-Ile149Cα) is found to be mostly overlapping between the two trajectories except around 30 ns, indicating that the flap−flap distance is not sensitive to the differences between ff99SB and ff03 force fields (Fig. 6(b)).

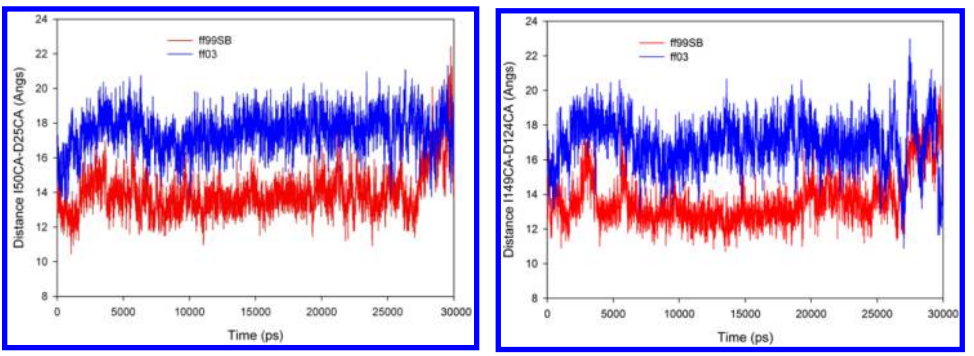


(a) (b)

Fig. 6. (a) TriCa angle in the flap region for the residues Gly48-Gly49-Ile50 in chain A of HIV-pr. (b) Time series plot for the flap−flap distance from the ff99SB and ff03 force field simulations.

3.6. *Flap−active site distance*

The flap−active site distance (the distance between Ile50Cα-Asp25Cα and the corresponding distance in chain B) shows significant difference between these two force fields. The flap−active site distance for both chains A and B of HIV-pr are shown in Fig. 7 for the two force fields. For the chain A (Fig. 7(a)), it can be seen that, the ff03 distance is fluctuating between 15 to 20 Å, while the ff99SB distance is fluctuating between 12 and 16 Å except for the last 3 ns of the simulation. The average value for this distance is 17.5 Å with standard deviation (SD) of 1.1 Å for the ff03 force field. For the ff99SB force field the average distance is 14.0 Å with standard deviation of 1.3 Å. The chain B distance (Fig. 7(b)) also show similar features. The average flap−active site distance in chain B is found to be 13.7 Å (1.4 Å as SD) for ff99SB and 16.9 Å (1.4 Å as SD) for ff03. The flap−active site distance is a measure of flap opening. Since, this distance on average is 3 Å longer



(a) (b)

Fig. 7. Flap−active site distance for (a) Chain A and (b) Chain B from the ff99SB and ff03 force field simulations.

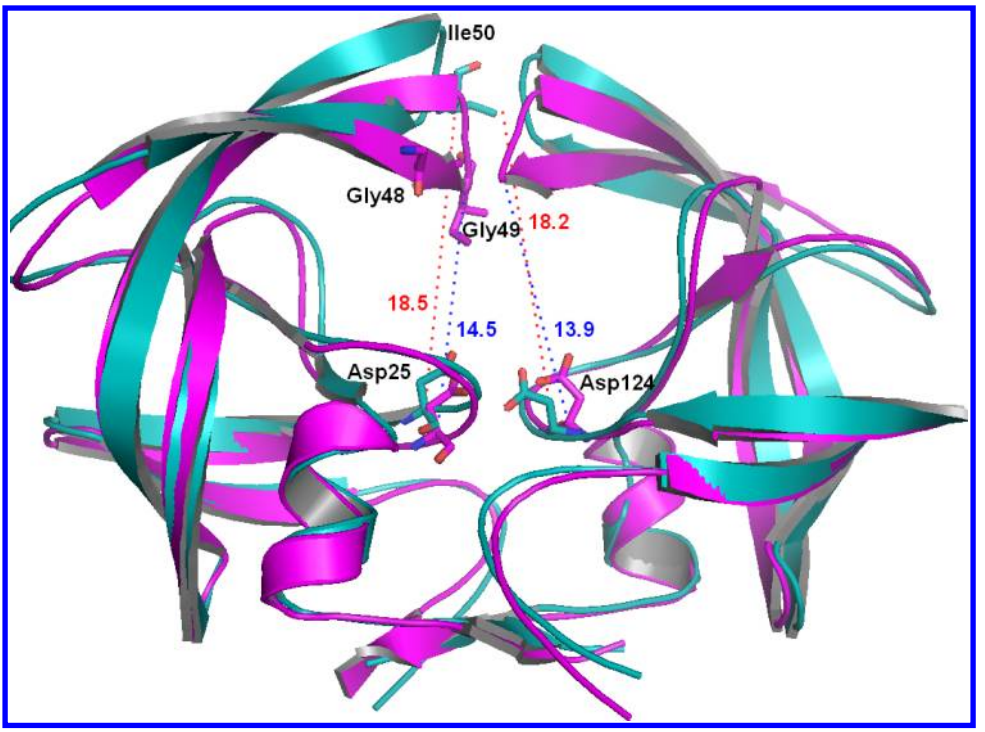


Fig. 8. Superimposition of two representative structures obtained from ff03 (cyan) and ff99SB (magenta) simulations. Residues (Asp25 (124), Gly48, Gly49 and Ile50) involved in the distance and angle calculations are labeled in the figure. However, for clarity reason we have not shown the residues for flap region in case of chain-B. Structure with ff03 force fields shows larger cavity size compared to ff99SB. The relative distances between the active site to the flap {Asp25 (124) Cα−Ile50 (149) Cα} is more in case of ff03 (18.5 and 18.2) compared to ff99SB (14.5 and 13.9) force field. The red dashed line and numbers indicates the distance for ff03 whereas blue is for ff99SB.

for both chains in ff03 simulation, the opening of the cavity is more according to the ff03 simulation. Figure 8 shows superimposition of two representative structures from ff03 and ff99SB force fields showing the larger cavity size in the ff03 simulation.

A closer inspection of the movement of different loops reveals that the differential motion of flap elbow (containing residue 17) and the fulcrum (containing residue 40) are contributing most to the difference seen in the active site-flap distance. Flap elbow and fulcrum movement is pushing the two flaps upwards, thus the flap−flap distance does not change but the flap−active site distance changes significantly. Please check the corresponding video files for ff99SB and ff03 force fields in the URL https://www.dropbox.com/home/JBCB-423%20Video%20files#:::99814589.

We have investigated possible reasons for the differential mobility in the loop regions. The ff03 and ff99SB differ both in their charges as well in their dihedral parameters. Both these parameters can influence the relative fluctuation. Also, the

interaction with TIP3P water molecules may also contribute to their differences. The difference in H-bonding patterns involving the residues of interest and interaction of water molecules with these seven loop residues have been investigated systematically. The H-bonding with water molecules for these seven loop residues did not show any significant difference. The intramolecular H-bonding of these residues shows that for Met36 CO and Leu38 NH H-bond, there is 11% occupancy according to the ff03 force field, while according to the ff99SB force field, the occupancy is 19%. For Met36 NH and Glu34 CO H-bond, the occupancy is 1% according to ff99SB and 13% according to ff03. The other major difference was in the H-bond involving Gly52 CO and Gly49 NH. Here the occupancy is 53% in ff99SB and 38% in ff03 force fields, while for Gly52 NH and Gly49 CO H-bonds, the occupancies are 6% and 21% for ff99SB and ff03, respectively. Gly68 NH and Ile66 CO H-bond has occupancies 1% and 12% for the ff99SB and ff03 trajectories, respectively. Table 2 shows all the observed H-bonds involving these interesting residues (and Fig. S4 in the Supplementary Material section shows the corresponding residues involved in the H-bonding).

3.7. *Analysis of interactions*

From Table 2, it can be said that although H-bonding difference exists between ff03 and ff99SB for some residues, intramolecular H-bonding alone is not sufficient to explain the differential fluctuation of the loop residues. It is likely that torsion

Table 2. Hydrogen bond distances involving the interesting residues Gly17, Met36, Gly40, Trp42, Ile50, Gly52, Gly68, His69, and their adjacent residues (for both the chains) with the percentage of occupancy in the ff99SB and ff03 force fields.[a,b]

	ff99SB		ff03	
Residue@Atom (Acceptor...Donor)	Occupancy (%)	Distance (Å)	Occupancy (%)	Distance (Å)
Met36-O...Leu38-N	19.0	2.87	11.8	2.89
Met36-N...Glu34-N	1.0	2.90	13.0	2.89
Met135-O...Leu137-N	12.4	2.89	37.3	2.87
Gly40-N...Leu38-O	5.5	2.88	0.05	2.94
Gly139-N...Leu137-O	2.0	2.88	0.1	2.85
Ile50-O...Gly150-N	—	—	2.2	2.88
Gly52-O...Gly49-N	53.5	2.86	38.4	2.88
Gly52-N...Gly49-O	6.4	2.91	21.2	2.90
Gly151-O...Gly148-N	38.1	2.86	28.4	2.88
Gly151-N...Gly148-O	11.5	2.91	6.2	2.91
Gly68-N...Ile66-O	1.1	2.78	12.4	2.81
Gly68-O...Ile66-N	0.01	2.85	0.13	2.83
Gly167-N...Ile165-O	1.7	2.77	15.3	2.80
His69-N...Ile66-O	58.5	2.88	45.1	2.89
His168-N...Ile165-O	54.4	2.88	45.4	2.89

[a]The Hydrogen bonds are determined by the donor acceptor atom distance of ≤ 3.5 Å and acceptor...H-donor angle of $\geq 120°$.

[b]Interesting residues with no hydrogen bonds to their adjacent residues are not listed in the above table.

parameters are also contributing to the differences, since most differences are in the highly flexible glycines. Although it is not easy to determine the exact reason for the differences between these force fields, it seems that a combination of several factors, including different descriptions of H-bonding, different charge sets, and different torsion parameters, contributes to the difference. One caveat to our results, as mentioned before, is the insufficient sampling in the MD simulations. However, it is encouraging that qualitatively most of the $S2_{MD}$ values follow the trend seen in the $S2_{NMR}$ values. Also, the difference between ff99SB and ff03 follows the same trend as seen by the other authors.[28] Our results showed that, among the tested AMBER force fields in the current study, ff99SB was found to produce the structural and dynamics data in comparably good agreement with the experiment. Moreover, our results also show the visible differences in the mobility of the loop residues with different force fields.

This work shows how the subtle differences in the force fields can influence the protein dynamics. Also, the observation that active site size variation is directly linked to the fluctuation of the flap elbow and fulcrum can be used to design allosteric ligands. This is also found by McCammon group, with their simulation of HIV-pr with ff99 force field.[10] The sensitivity of loop dynamics with force field is likely to be important not only for HIV-pr but also for other flexible proteins such as protein kinases. The outcome of this work would be important for proteins with flexible binding sites, which may show different fluctuation depending on the different force field parameters. One important case where this is a major problem is in the ensemble based docking approach for proteins with flexible binding site, where ligands are docked to the snapshots of MD simulation.[29] If the average geometry and fluctuation of the binding site differ significantly on simulation parameters, then this may as well influence the ligand binding and the relative ranking of ligands. Another major issue in HIV-pr dynamics is whether the different mutants have similar conformational dynamics as the wild type. It is suggested that the mutations away from the active site modulate the conformational dynamics in such a way that it reduces inhibitor binding affinity. Already there are several published works on the difference of dynamics between wild type and various mutants of HIV-pr using MD simulation. It would be interesting to check the effect of force field on the mutant flexibility as compared to the wild type for some of these.

4. Summary

The current work describes a comparison of three AMBER force fields, ff99, ff99SB, and ff03 by performing MD simulation with these force fields to understand the flap dynamics of HIV-pr, a protein of outstanding therapeutic importance. The effect of force field difference on the flexibility of HIV-pr and its significance in *in silico* drug design against HIV-pr is reported in this work. However, in a more general sense, this work also shed light on the difficulty in modeling dynamics of proteins with flexible binding site and *in silico* drug design against flexible receptors. Our simulation results show that ff99 force field's description of protein dynamics to be poor. The

ff99SB and ff03 force fields' overall description is similar but they vary in dynamics of various loop residues. Many of these residues are glycines present in the loops. The calculated S2 values for these residues (except one) are significantly lower in the ff03 compared to ff99SB (and deviate more from the NMR S2 values) indicating greater flexibility. Two different 30 ns MD trajectories were used to check the ff03 and ff99SB S2 values. These fluctuations lead to a larger active site cavity in the ff03 force field as a consequence of longer active site-flap distance ($\sim$3 Å for both chains). The interactions determining the loop fluctuation indicate that it is likely that a combination of several factors like H-bonding, electrostatics, and torsion parameters are contributing to this difference. The difficulty of modeling loops in MD simulation can be important for proteins with flexible binding sites. We expect that the current work can provide some helpful insights to the effects of force fields and simulation protocols on the nature of conformational dynamics for HIV-pr and aid the future design of potent inhibitors for the flexible binding site of HIV-pr.

Acknowledgments

BRM thanks AICTE, Govt. of India for the financial support in the form of National Doctoral Fellowship (NDF). The authors thank IIT Guwahati for providing the computational facility. This work is funded by a Department of Science and Technology (DST) fast-track grant (SR/FT/L-43/2004) awarded to PB. The authors thank anonymous reviewers for the constructive suggestions.

References

1. Ghosh AK, Chapsal BD, Weber IT Mitsuya H, Design of HIV protease inhibitors targeting protein backbone: An effective strategy for combating drug resistance, *Acc Chem Res* **41**:78−86, 2008.
2. Eder J, Hommel U, Cumin F, Martoglio B Gerhartz B, Aspartic proteases in drug discovery, *Curr Pharm Des* **13**:271−285, 2007.
3. McKeage K, Perry CM, Keam SJ, Darunavir: A review of its use in the management of HIV infection in adults, *Drugs* **69**:477−503, 2009.
4. 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**:521−538, 2010.
5. 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* **7**:1047−1055, 1999.
6. Collins JR, Burt SK, Erickson JW, Flap opening in HIV-1 protease simulated by 'activated' molecular dynamics, *Nat Struct Biol* **2**:334−338, 1995.
7. Scott WR, Schiffer CA, Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance, *Structure* **8**:1259−1265, 2000.
8. Piana S, Carloni P, Rothlisberger U, Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations, *Protein Sci* **11**:2393−2402, 2002.
9. Piana S, Carloni P, Parrinello M, Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease, *J Mol Biol* **319**:567−583, 2002.
10. 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* **13**:1108−1123, 2004.

11. Toth G, Borics A, Flap opening mechanism of HIV-1 protease, *J Mol Graph Model* **24**:465−474, 2006.
12. Ode H, Neya S, Hata M, Sugiura W, Hoshino T, Computational simulations of HIV-1 proteases−multi-drug resistance due to nonactive site mutation L90M, *J Am Chem Soc* **128**:7887−7895, 2006.
13. Hornak V, Okur A, Rizzo RC, Simmerling C, HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, *Proc Natl Acad Sci USA* **103**:915−920, 2006.
14. Meagher KL, Carlson HA, Solvation influences flap collapse in HIV-1 protease, *Proteins: Struct Funct Bioinf* **58**:119−125, 2005.
15. Bandyopadhyay P, Meher BR, Drug resistance of HIV-1 protease against JE-2147: I47V mutation investigated by molecular dynamics simulation, *Chem Biol Drug Des* **67**:155−161, 2006.
16. Ishima R, Louis JM, A diverse view of protein dynamics from NMR studies of HIV-1 protease flaps, *Proteins: Struct Funct Bioinf* **70**:1408−1415, 2008.
17. Yildirim I, Stern HA, Kennedy SD, Tubbs JD, Turner DH, Reparameterization of RNA chi torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine, *J Chem Theory Comput* **6**:1520−1531, 2010.
18. Hayre NR, Singh RR, Cox DL, Evaluating force field accuracy with long-time simulations of a beta-hairpin tryptophan zipper peptide, *J Chem Phys* **134**:0351031−0351038, 2011.
19. Wickstrom L, Okur A, Simmerling C, Evaluating the performance of the ff99SB force field based on NMR scalar coupling data, *Biophys J* **97**:853−856, 2009.
20. Voelz VA, Dill KA, Chorny I, Peptoid conformational free energy landscapes from implicit-solvent molecular simulations in AMBER. *Biopolymers* **96**:639−650, 2011.
21. Tong Y, Ji CG, Mei Y, Zhang JZ, Simulation of NMR data reveals that proteins' local structures are stabilized by electronic polarization, *J Am Chem Soc* **131**:8636−8641, 2009.
22. Wang J, Cieplak P, Kollman P, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* **21**:1049−1074, 2000.
23. Duan Y *et al.*, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, *J Comput Chem* **24**:1999−2012, 2003.
24. 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: Struct Funct Bioinf* **65**:712−725, 2006.
25. Jorgensen WL, Chandrasekhar J, Madura JD, Roger WI, Klein ML, Comparison of simple potential functions for simulating liquid water, *J Chem Phys* **79**:926−935, 1983.
26. 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* **11**:221−232, 2002.
27. Lyman E, Zuckerman DM, Ensemble-based convergence analysis of biomolecular trajectories, *Biophys J* **91**:164−172, 2006.
28. Trbovic N, Kim B, Friesner RA, Palmer AG, 3rd, Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation, *Proteins: Struct Funct Bioinf* **71**:684−694, 2008.
29. Perryman AL, Lin JH, Andrew McCammon J, 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* **67**:336−345, 2006.

30. 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* **73**:1391−1396, 1991.
31. Case DA *et al.*, AMBER 9, University of California, San Francisco, 2006.
32. Wang YX, Freedberg DI, Yamazaki T, Wingfield PT, Stahl SJ, Kaufman JD, Kiso Y, Torchia DA, Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272, *Biochemistry* **35**:9945−9950, 1996.
33. Essmann U, Perera L, Berkowitz ML, Darden TA, Lee H, Pedersen LG, A smooth particle mesh Ewald method, *J Chem Phys* **103**:8577−8593, 1995.
34. Berendsen HJC, Postma JPM, Van Gunsteren WF, DiNola A, Haak JR, Molecular dynamics with coupling to an external bath, *J Chem Phys* **81**:3684−3690, 1984.
35. Lipari G, Szabo A, Model free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J Am Chem Soc* **104**:4546−4559, 1982.
36. 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 Comput Chem* **25**:1605−1612, 2004.
37. Laskowski RA, MacArthur MW, Moss DS, Thornton JM, PROCHECK — a program to check the stereochemical quality of protein structures, *J App Cryst* **26**:283−291, 1993.

Biswa Ranjan Meher received his Ph.D. in Biotechnology from Indian Institute of Technology, Guwahati, India, in 2009. In the same year, he joined as a Research Assistant in the Center for Bioinformatics, University of Kansas, Lawrence, KS, USA. He is currently working as a Research Associate at the Computational Chemistry Laboratory, Albany State University, Albany, GA, USA. His research interests includes computer modeling and simulations of complex biomolecular systems, structure-based drug design, and nano-bioapplications in medicine.

Mattaparthi Venkata Satish Kumar has worked as an Associate Professor (1999−2006) in the Department of Chemical Engineering, Vignan University, Andhra Pradesh, India. He received his Ph.D. in Biotechnology from Indian Institute of Technology, Guwahati, India, in 2010. He is currently a Postdoctoral Research Fellow at the Centre for Condensed Matter Physics, Indian Institute of Science, Bangalore, India. His research interests include modeling of nanoscaled advanced materials : nanoparticle growth as well as DNA-assisted self-assembly of nanoparticles using molecular simulation; structure, function and dynamics of intrinsically disordered proteins; protein aggregation; and enzyme kinetics.

Smriti Sharma received her B.Pharm degree from GGSIP University, Delhi, India. Currently she is working as a Ph.D. Research Scholar at the Centre for Computational Biology and Bioinformatics, School of Computational and Integrative Scinces, Jawaharlal Nehru University, New Delhi. Her research interest includes understanding the reaction mechanism of β-lactamase protein using electrostatics, molecular dynamics, and quantum mehanical calculations.

Pradipta Bandyopadhyay is an Associate Professor at the School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi. He received his M.Sc. degree from Indian Institute of Technology, Kanpur and completed his Ph.D. from Institute of Molecular Science, Okazaki, Japan. He worked as a post-doctoral researcher at Iowa State University and University of California, San Francisco. He also served as an Assistant Professor in the Department of Biotechnology at the Indian Institute of Technology, Guwahati, India from 2004 to 2007 and was a Visiting Associate Professor at the University of California, San Francisco in 2008. His research interests include development and application of advanced Monte Carlo simulation techniques, understanding the energy landscape of large molecules, and application of computational tools to solve biological problems.