J Mol Model (2007) 13:1151–1156
DOI 10.1007/s00894-007-0242-3

ORIGINAL PAPER

Molecular dynamics studies on HIV-1 protease: a comparison of the flap motions between wild type protease and the M46I/G51D double mutant

Antonino Lauria · Mario Ippolito · Anna Maria Almerico

Received: 1 July 2007 / Accepted: 8 August 2007 / Published online: 6 September 2007

Abstract The emergence of drug-resistant mutants of HIV-1 is a tragic effect associated with conventional long-treatment therapies against acquired immunodeficiency syndrome. These mutations frequently involve the aspartic protease encoded by the virus; knowledge of the molecular mechanisms underlying the conformational changes of HIV-1 protease mutants may be useful in developing more effective and longer lasting treatment regimes. The flap regions of the protease are the target of a particular type of mutations occurring far from the active site. These mutations modify the affinity for both substrate and ligands, thus conferring resistance. In this work, molecular dynamics simulations were performed on a native wild type HIV-1 protease and on the drug-resistant M46I/G51D double mutant. The simulation was carried out for a time of 3.5 ns using the GROMOS96 force field, with implementation of the SPC216 explicit solvation model. The results show that the flaps may exist in an ensemble of conformations between a "closed" and an "open" conformation. The behaviour of the flap tips during simulations is different between the native enzyme and the mutant. The mutation pattern leads to stabilization of the flaps in a semi-open configuration.

Keywords HIV-1 protease · Molecular dynamics simulations · Flap motions · HIV-1 drug-resistant mutants · Gromacs 3.2

A. Lauria (✉) · M. Ippolito · A. M. Almerico
Dipartimento Farmacochimico, Tossicologico e Biologico,
Università di Palermo,
Via Archirafi 32,
90123 Palermo, Italy
e-mail: lauria@unipa.it

Introduction

The pandemic diffusion of acquired immunodeficiency syndrome (AIDS) has promoted an unprecedented series of efforts in order to understand and combat this lethal disease: the design, synthesis and evaluation of potential new therapeutic agents still represents a significant challenge in this research field. Intensive studies has focussed on the viral protease (PR), one of the three enzymes playing a key role in the viral lifecycle of HIV-1, the etiological agent of AIDS. Since 1995, PR inhibitors have been fundamental components in the chemotherapy of HIV-1 infection. Unfortunately, in addition to issues of patient compliance, the long-term efficacy of this antiretroviral therapy is hampered by the emergence of viral strains that exhibit resistance to PR inhibitors.

HIV-1 PR is an enzyme that cleaves the virally encoded Gag and Gag-Pol polyproteins [1, 2]. This process is essential for the production of mature virions. PR is a homodimer formed by two chains of 99 amino acid residues, and belongs to the aspartic proteinases class, which also includes human enzymes like pepsin and rennin. Aspartic proteinases are characterised by the presence of a highly conserved catalytic triad Asp25(25′)-Thr26(26′)-Gly27(27′), which is usually located at the base of the substrate binding site; in particular, the Asp25(25′) residues are responsible for the cleavage of the substrate's scissile peptide bond. Two β-hairpins, termed "flaps" located above the active site cavity control access for both substrate and inhibitors. Large structural differences exist between the bound and free states of PR. In ligand-bound form, the flaps are pulled in toward the bottom of the active site ("closed" form); when the enzyme is in ligand-unbound form, a "semi-open" conformation seems to be preferred, where the flap tips are shifted away from the catalytic site, but are still

substantially closed and in contact with each other. Recently, a very interesting study [3] succeeded in reproducing the dynamic behaviour of HIV-1 PR using molecular dynamics with a coarse-grained model. This model was able to correctly detect and predict the main features of the native enzyme from both a thermodynamic and a kinetic point of view; it was also able to predict HIV-1 PR enzyme conformations that are inaccessible experimentally.

Unfortunately, our understanding of the details of the mechanism by which substrate processing is achieved remains incomplete. Many mechanisms have been proposed, falling into two general classes: general acid–base catalysis, and catalysis via a covalently bound intermediate. Molecular and quantum mechanic studies [4–6] accounted for the acid–base catalysis mechanism; according to this mechanism, substrate peptide bond hybridisation is disrupted through a nucleophilic attack on its carbonyl group by a water molecule coordinated to the active site aspartic residues. The *gem*-diol intermediate is subsequently cleaved, yielding the carboxylic and amide groups.

The emergence of drug-resistant mutants is a serious side effect associated with current anti-HIV treatment regimes. HIV-1 PR is one such target for mutations; PR variants are able to acquire resistance through complementary strategies involving both active-site and non active-site mutations that may reduce the affinity of the enzyme for its substrate and inhibitors. In addition, there are many mutations, known as "compensatory mutations", that are usually located far from the active site and can act through long-range electrostatic interactions. Understanding the molecular mechanisms governing HIV-1 PR flap mobility has deep implications for the design of new therapeutic agents, such as allosteric inhibitors intended to interfere with flap opening and subsequent enzymatic function.

Here we investigate the conformational changes occurring in the flap region when the well known M46I mutation [7] is artificially associated with the G51D mutation. In fact, previous studies [8] have shown that residue G51 is directly involved in flap curling, thus its mutation contributes to drug resistance. The results obtained show that drug resistance can be related to flap mobility and to the changes in charge distribution and in steric hindrance due to the introduced mutations. Moreover, the simulation permitted observation of stabilisation of the dimeric structure in the mutant enzyme.

Materials and methods

The structure of native HIV-1 PR was obtained from the Brookhaven Protein Data Bank (PDB code: 1HHP) [9]. The M46I/G51D double mutant was obtained by a manual amino acid substitution using Deep View [10] followed by an energy minimisation of the side chains using a GROMOS96 force field [11].

The GROMACS 3.2 package [12, 13] was used to perform molecular dynamics (MD) simulations. This package is a collection of programs and libraries for the MD simulations and the subsequent analysis of trajectory data. Simulations were performed using a general triclinic cell geometry. Pressure and temperature coupling is implemented for all types of simulation cells; in this study the Berendsen's weak coupling algorithm scheme was used for both pressure and temperature [14].

Before starting MD, the protonation state of Asp25 of every macromolecule was adjusted as protonated according to experimental observations predicting that, under physiological conditions, only one of the two active site aspartates is protonated while the other remains in the carboxylated form.

The two HIV-1 PR molecules were solvated in a rectangular box using periodic boundary conditions and the SPC water model [15]; the distance between the grid box and the protein was set to 1.0 nm. The systems, each containing a dimer of PR with water molecules, were submitted to 500 steps of steepest descent minimisation converging to a value of 2,000 kJ mol^{-1} nm^{-1}, applying the Particle-Mesh Ewald method [16] at 310 K. Further, a 100 ps position-restrained MD simulation was performed by keeping the protein coordinates fixed, and allowing the water molecules to equilibrate themselves. Finally, a 3,500 ps MD was performed on each of the two systems at 310 K temperature. A leap-frog algorithm was used for integrating Newton equations of motion for 1,750,000 simulation steps, with a time step of 0.002 ps.

The trajectory files were analysed through the use of g_rms and g_rmsf GROMACS utilities in order to obtain the root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) values. Moreover VEGA [17] and VMD [18] programs were used for trajectory analysis and for the manipulation of the simulation snapshot structures. Linux OpenOffice was used for numeric data manipulation and graphical presentation.

Results and discussion

One of the major problems associated with treatment regimes targetted against HIV-1 PR is the acquisition of molecular resistance. Protease inhibitors tend to lose their potency and effectiveness against their molecular target when rapid point mutations alter the viral genome. The flaps of the PR molecule are the target for a particular type of mutations occurring distal to the active site that modify the affinity of the enzyme for both substrate and ligands.

Many studies have reported data indicating that "compensatory mutations" affect the dynamics of conformational changes, thus causing a decrease in ligand affinity. In particular, recent work [19, 20] has shown that flap curling could be altered by some mutations that cause increased flexibility of the flaps. These alterations affect the equilibrium between the closed, semi-open and open conformations of the protease. Moreover, these mutations can result in decreased affinity for both substrate and inhibitors, thus providing a plausible explanation for the ability to increase the level of resistance.

Many HIV-1 PR mutation patterns have been collected and characterised [19]. Nanosecond-scale MD simulations of the M46I mutant [21, 22] indicated that this mutation induces subtle differences in the dynamics of the molecule, through stabilisation of the flaps and a decrease in their flexibility. Here, we combined M46I with the G51D mutation, which occurs in a region known to be responsible for flap curling. To investigate the flap mobility of the double mutant compared to wild type PR, we performed 3.5 ns MD simulations (Fig. 1). In our experiments an unbound HIV-1 PR structure (pdb code: 1HHP) was used in which the flap tips are present in the "semi-open" conformation, which seems to be the conformation preferentially adopted by ligand-free enzyme. Graphical representations of the simulations show that the flaps may exist in an ensemble of conformations ranging between a "closed" and an "open" conformation. Moreover, movement of the flaps during simulations differs between the wild type enzyme and the mutant; in fact, in the case of WT enzyme an irreversible opening of the flaps was observed, whereas the mutant enzyme exhibited an incomplete opening that led to the stabilisation of the semi-open conformation.

Examination of the trajectories obtained from MD simulations performed for wild-type and double-mutant HIV-1 PRs shows that, after a 1,200 ps period of relaxation, all systems deviated in a similar fashion from their starting structures, resulting in a backbone RMSD of approximately 0.4–0.6 Å (Fig. 2a) during the simulations. Specifically, statistical analysis of the RMSD values obtained after the relaxation period (1,200–3,500 ps) resulted in an RMSD range of 0.63÷0.42 Å for the wild-type enzyme and 0.38÷0.57 Å for the double mutant, with average values of 0.51± 0.033 Å for the wild type and 0.46±0.036 Å for the mutant. This magnitude of fluctuations, together with the very small difference between the average RMSD values after the relaxation period (0.050 Å), leads to the conclusion that the simulations produced stable trajectories, thus providing a suitable basis for further analyses. With the aim of determining whether the mutation affects the dynamic behaviour of the residues, the RMSF values of the HIV-1 PR residues was calculated (Fig. 2b). Analysis of the fluctuations revealed that the greatest degree of flexibility is contained in the flap region (between residues 40 and 60) for both wild type and mutant enzymes; moreover, the mutant exhibited a decreased RMSF in the flap region (0.48 Å vs 0.61 Å), confirming that the occurrence of the mutation leads to a more rigid protein form.

According to many reported simulations, the opening process of the entire flap region is triggered by curling of the flap tips; thus, a comparison of the distance between the flap tips (the distance between residues 50 and 50′) can be useful in understanding the differences in flap opening between wild type HIV-1 PR and the mutant (Fig. 3a,b). Our simulation revealed substantial differences between the mutant and the wild type both in the variation of distance versus time and in the distance distribution. In the case of the wild type PR, a period of tip curling occurred until 600 ps; after 620 ps the opening process becomes irreversible and the distance between the tips of the flaps grows continually until 3,000 ps. After this period, the distance increases only a little due to the partial winding of the flap tip. The mutant enzyme exhibited a minor period of flap curling (a few hundred picoseconds), after which complex flap motion occurs leading, after 1,000 ps, to the semi-open form of HIV-1 PR, which is stable until the end of the simulation. Moreover, monitoring of the distribution of flap tip distances (Fig. 3b) showed a predominance of medium-range distances (10–20 Å) for the mutant enzyme, while in the case of the wild type protein, long-range distances (>24 Å) seem to be favoured. The average distance between the flap tips (Ile50 and 50′) and catalytic aspartates (Asp25 and 25′) was also measured from the

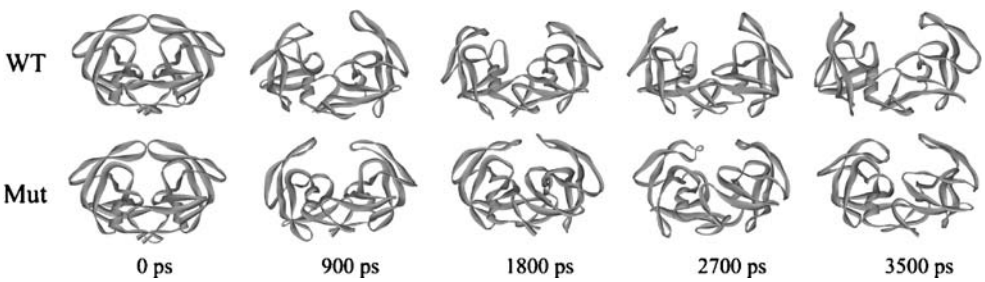

Fig. 1 Snapshots of backbone conformations of HIV-1 protease (PR) throughout the trajectories

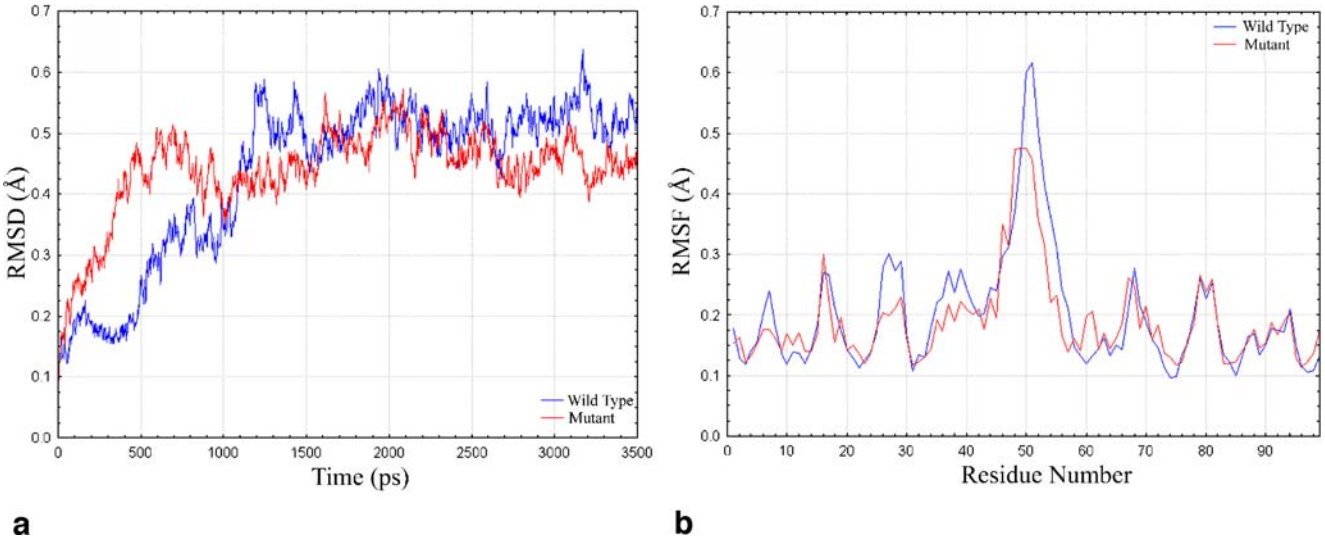

Fig. 2 **a** Root-mean-square deviation (RMSD) values for backbone atoms along the 3.5 ns trajectory. **b** C_α atom root-mean-square fluctuation (RMSF) values given for each residue number in the two HIV-1 PRs

simulation and their distribution is shown in Fig. 4a. The results confirm that the distance between flap tips and catalytic aspartates in clearly different for WT and mutant HIV-1 PR. The broadest peak of WT is centred between 22 and 24 Å, while the broadest peak for the mutant lies between 18 and 21 Å. A minimum overlap is also observed; moreover it should be emphasized that the wild type protein presents a larger distribution of short distances due to the longer period of flap curling that occurs before complete opening. Previous studies [23, 24] have demonstrated that the flap separation process is preceded by the formation of a hydrophobic cluster between the side-chains of Ile50 and Pro81 of the same monomer, which destabilises the symmetrical flap–flap interactions. The G51D mutation probably interferes with this clustering due to the insertion of the charged carboxylic group. This substitution not only makes flap curling more difficult but can also dramatically change the electrostatic potential that the substrate polypeptide or the inhibitor would encounter upon binding.

Comparison between the WT and the mutant PR structures in the closed and open form (Fig. 5) can help to better understand the results. At 0 ps the two proteins are very similar. The only observable difference is the presence of the two Asp groups deriving from mutation of Gly51. At 3,500 ps the WT shows a large flap opening, which exposes the active site, where the catalytic Asp25 and Asp25′ residues are visible as a red region corresponding to their carboxylic groups. In the case of the M46I/G51D double mutant, in the 3,500 ps snapshot the flaps are in a semi-open conformation, which keeps the active site more

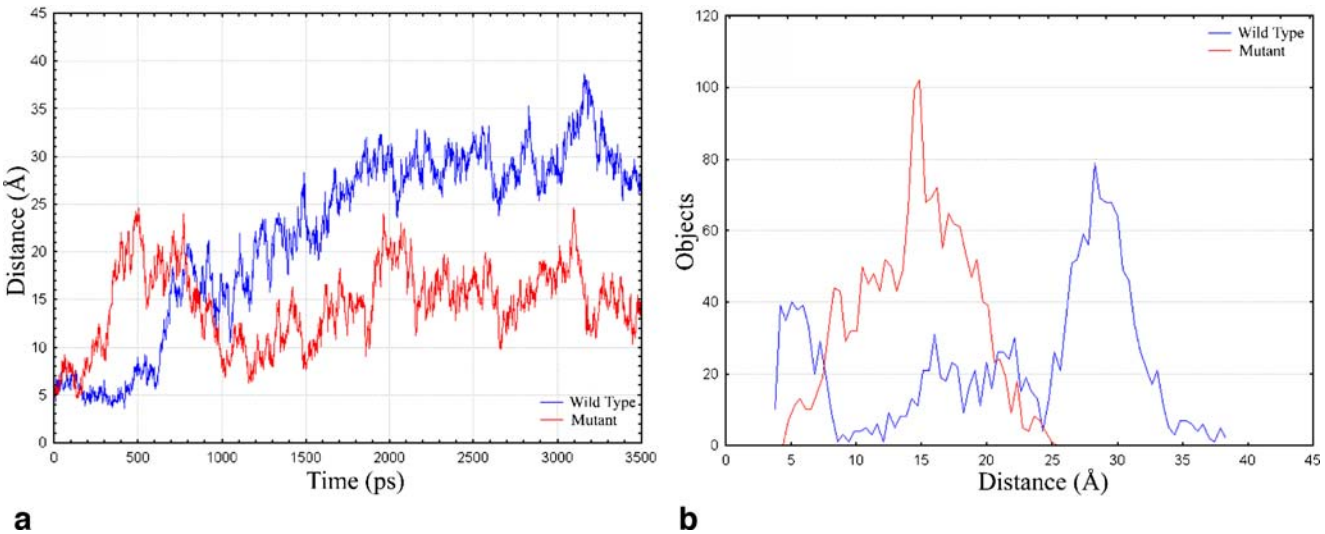

Fig. 3 **a** Distances between the tips of the flaps (residues Ile50–Ile50′) over the entire simulation. **b** Distribution of flap tip distances

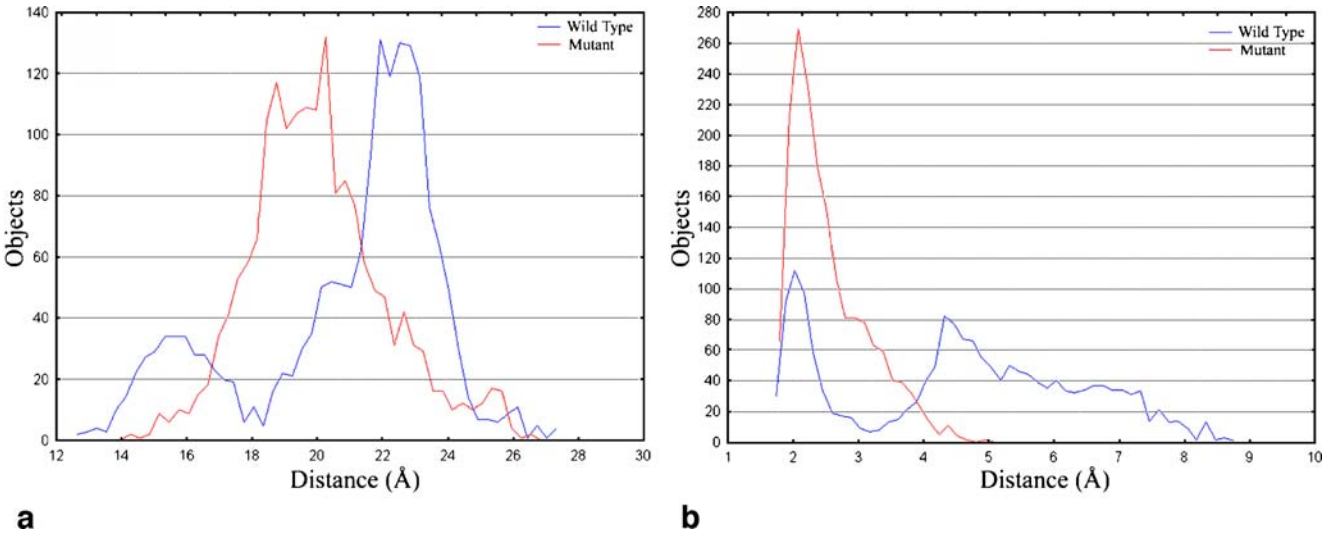

Fig. 4 **a** Distribution of the average distances between Ile50/Ile50′ and Asp25/Asp25′. **b** Distribution of the distances between the "Fireman's Grip" (residue Thr26–Thr26′)

closed; the catalytic Asp25 and Asp25′ are kept buried in the active site, and the whole protein seems to adopt a more closed conformation.

Since the mutant HIV-1 PR gives more closed dimer conformation during the simulation, the dimer stability of the two proteins was compared. Two main structural groups are responsible for dimer stability in the aspartic proteinase class: the N and C terminals intertwined in an antiparallel β-sheet, and a hydrogen bond supporting the active site (called "the fireman's grip"). This bond is formed by the interaction between the side chain oxygen of Thr26 and the NH_2 group of Thr26′. Since the fireman's grip has been shown to be indispensable for dimer stability of HIV-1 PR [25], we examined the distribution of distances between Thr26 and Thr26′ (Fig. 4b).

The results revealed a partial overlap between the two structures, with a peak occurring between 1.8 and 2.5 Å, which is completely compatible with the existence of a hydrogen bond. Despite this partial overlap, it emerged that in part of the wild type conformation this interaction does not occur, suggesting that during flap opening stabilisation due to the fireman's grip is lost and thus the only contribution to dimer stability is that afforded by the interwinding of the N and C terminals. The distances measured for the mutant enzyme confirm the role of the mutation in closing the dimer form of the protein and

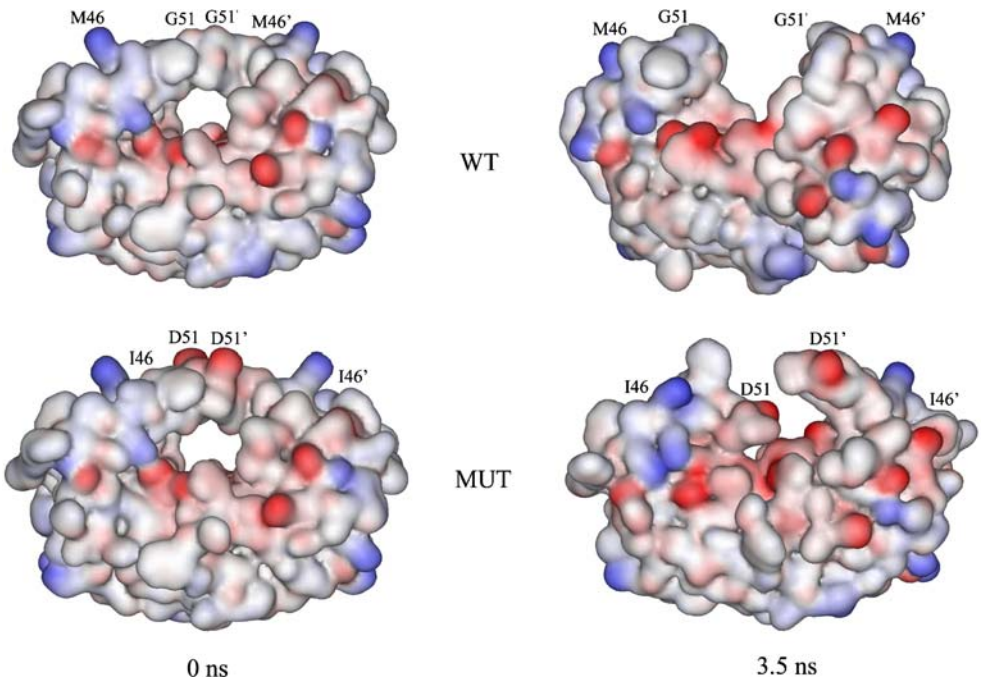

Fig. 5 Molecular electrostatic potential of wild-type and mutant HIV-1 PR at 0 and 3,500 ps, respectively

reducing access to the active site. The difference between the mutant and the wild type is in the zone between 4.0 and 5.5 Å. In fact, while in the mutant enzyme the major distance is 4.2 Å, the wild type structure exhibits distances of up to 5.3 Å. This result accounts for the fact that the double mutation affects not only flap motion, but can also contribute to the stability of the HIV-1 PR dimer.

Conclusions

In this study, we used 3,500 ps MD simulations in order to compare the flap opening mechanism between a native wild type HIV-1 PR and a mutant PR that has shown clinical resistance to long-term conventional treatments and contains an additional mutation on a residue involved in flap mobility. The simulation was performed in the presence of a flexible water model in order to mimic the effects of the solvent on protein structure. A position-restrained MD was useful in optimising the position of the solvent molecules with respect to the protein atoms.

The results obtained show that the double mutation can delay flap opening and can stabilise a series of semi-open flap conformations, thus enhancing the steric hindrance that a possible substrate or inhibitor would encounter in order to enter the active site. In particular, while the M46I mutation contributes to make the flap tips less flexible, the G51D substitution dramatically changes the charge distribution in the proximity of the region responsible for flap–flap interactions. The global effect of these mutations is to interfere with the network of weak polar interactions that make flap opening a key mechanism for regulating access to the active site.

In agreement with previous studies [8], our results also support an irreversible mechanism of flap opening. The predominance of semi-open conformations rather that irreversible opening has recently been observed [26] under continuous solvent conditions, also confirming that reversible flap motion occurs after several nanoseconds of simulation. Efforts to extend our experiments to this larger timescale are in progress, in order to check if our mutant yields similar results. Nevertheless, the presence of different solvation models or the use of several steps of position restraining could underlie the observed differences.

References

1. Seelmeier S, Schmidt H, Turk V, Von der Helm K (1988) Proc Natl Acad Sci USA 85:6612–6616
2. Huff JR (1991) J Med Chem 34:2305–2314
3. Tozzini V, Trylska J, Chang C, McCammon JA (2007) J Struct Biol 157:606–615
4. Chatfield DC, Brooks BR (1995) J Am Chem Soc 117:5561–5572
5. Silva AM, Cachau RE, Sham HL, Erickson JW (1996) J Mol Biol 255:321–340
6. Chatfield DC, Eurenius KP, Brooks BR (1998) J Mol Struct 423:79–92
7. Ohtaka H, Schon A, Freire E (2003) Biochemistry 42:13659–13666
8. Scott WRP, Schiffer CA (2000) Structure 8:1259–1265
9. Spinelli S, Liu QZ, Alzari PM, Hirel PH, Poljak RJ (1991) Biochimie 73:1391–1393
10. http://www.expasy.org/spdbv/
11. Oostenbrink C, Villa A, Mark AE, Van Gusteren WF (2004) J Comput Chem 25:1656–1676
12. Berendsen HJC, Van der Spoel D, Van Drunen R (1995) Comput Phys Commun 91:43–56
13. Lindahl E, Hess B, Van der Spoel D (2001) J Mol Model 7:306–317
14. Berendsen HJC, Postma JPM, DiNola A, Haak JR (1984) J Chem Phys 81:3684–3690
15. Ferguson DM (1995) J Comp Chem 16:501–511
16. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577–8593
17. Pedretti A, Villa L, Vistoli, G (2002) J Mol Graph 21:47–49
18. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38
19. Perryman AL, Lin J-H, McCammon A (2004) Protein Sci 13:1108–1123
20. Meiselbach H, Horn AHC, Harrer T, Sticht H (2007) J Mol Model 13:297–304
21. Maschera B, Darby G, Palu G, Wright LL, Tisdale M, Myers R, Blair ED, Fufine ES (1996) J Biol Chem 271:33231–33235
22. Piana S, Carloni P, Rothlingsberger U (2002) Protein Sci 11:2393–2402
23. Wu TD, Schiffer CA, Gonzales MJ, Taylor J, Kantor R, Chou S, Israeliski D, Zolopa AR, Fessel WJ, Shafer RW (2003) J Virol 77:4836–4847
24. Tòth G, Borics A (2006) J Mol Graph Model 24:465–474
25. Ingr M, Uhlìkovà T, Strisovsky K, Majerovà E, Konvalinka J (2003) Protein Sci 12:2173–2182
26. Hornak V, Okur A, Rizzo RC, Simmerling C (2006) Proc Natl Acad Sci USA 103:915–920