Insights into a mutation-assisted lateral drug escape mechanism from the HIV-1 protease active site

Supporting Information

S. Kashif Sadiq , Shunzhou Wan and Peter V. Coveney*

Centre for Computational Science, Department of Chemistry, University College London, Christopher Ingold Laboratories, 20 Gordon Street, London, WC1H 0AJ.

August 29, 2007

Structural flexibility

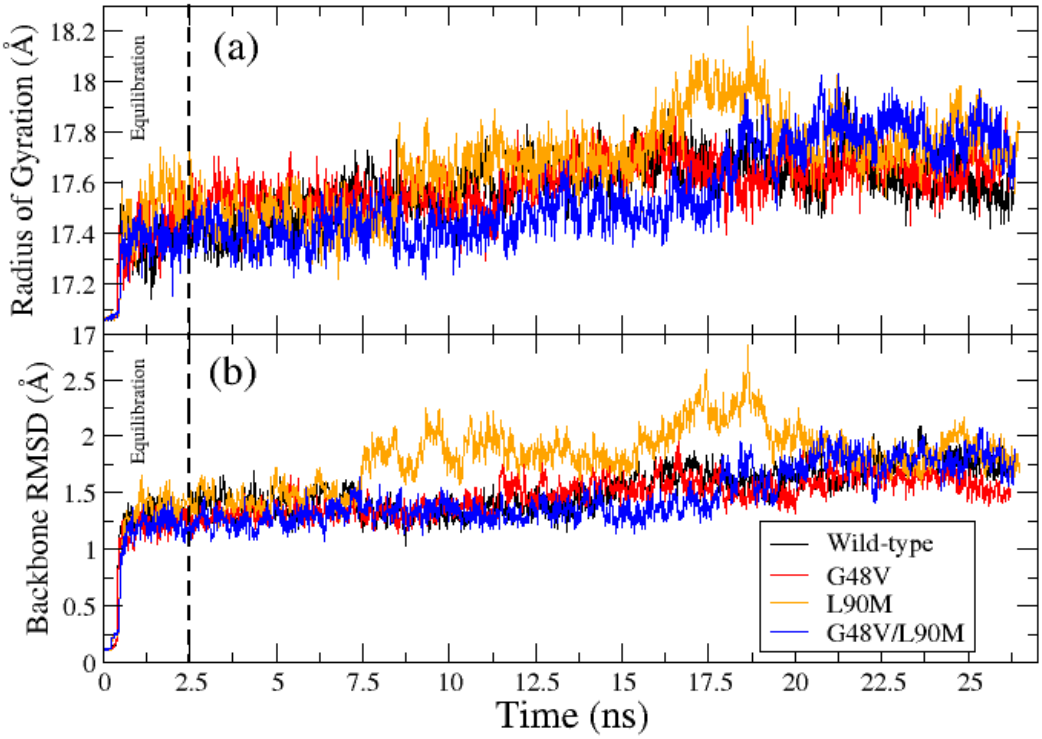

Figure 1: (a) Radius of gyration and (b) RMSD of backbone atoms of HIV-1 protease relative to the crystal structure, for the wildtype (black), G48V (red), L90M (orange) and G48V/L90M (blue) systems. The L90M mutant shows slightly greater flexibility than the other systems across the longer timescale of 25 ns.

*p.v.coveney@ucl.ac.uk

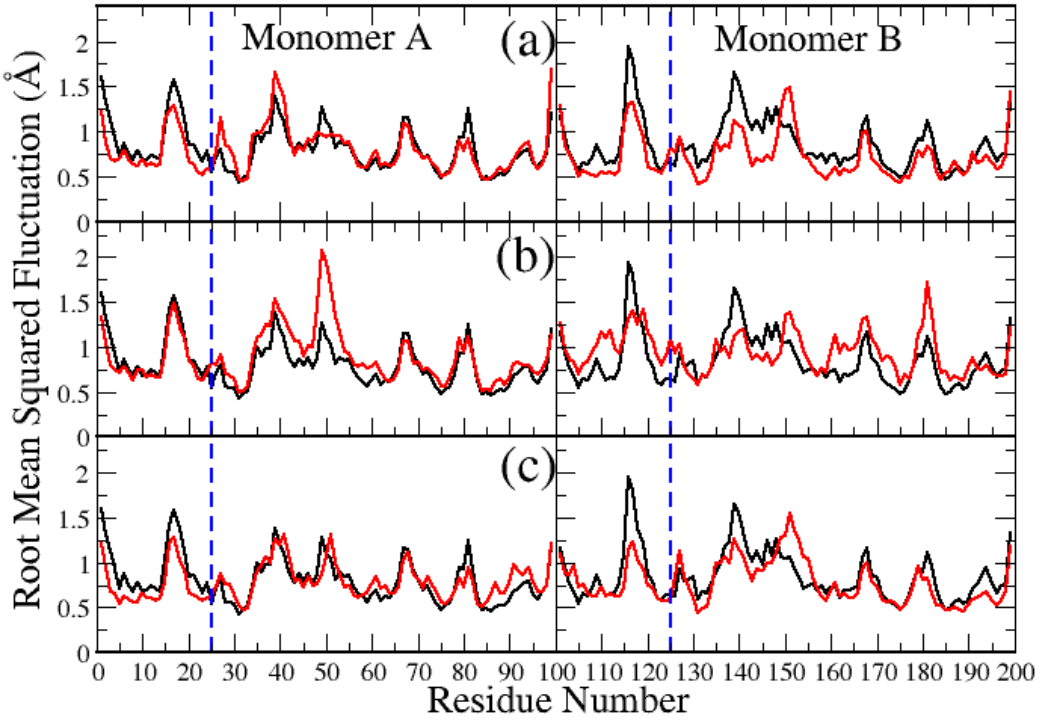

Figure 2: Comparative profiles versus amino acid residue number of the RMSF of backbone atoms of HIV-1 protease, relative to the average structure after equilibration. The (a) G48V, (b) L90M and (c) G48V/L90M mutant systems (red) are compared against the wildtype (black). The L90M system shows increased flexibility of the catalytic aspartic residues (distinguished by the dashed blue lines) as compared to all other systems. The G48V system also shows decreased flexibility of the inner flap residues of monomer B (143-149). This is explained by the increased association with the P2 subsite of saquinavir.

Coupled flap and inhibitor dynamics

Distance Cross-Correlation	WT	G48V	L90M	G48V/L90M
$Cc(\|\mathbf{r}_{FA}\|:\|\mathbf{r}_{FS}\|)$	**0.85**	0.39	0.25	0.41
$Cc(\|\mathbf{r}_{FA}\|:\|\mathbf{r}_{SA}\|)$	0.41	**0.91**	0.13	**0.90**
$Cc(\|\mathbf{r}_{FS}\|:\|\mathbf{r}_{SA}\|)$	0.01	0.44	0.45	0.22

Table 1: Cross-correlation coefficients between three distance metrics, Flap-Asp ($|\mathbf{r}_{FA}|$), Flap-Saq ($|\mathbf{r}_{FS}|$) and Saq-Asp ($|\mathbf{r}_{SA}|$), across all four protease systems. A high value of $Cc(|\mathbf{r}_{FA}|:|\mathbf{r}_{FS}|)$ indicates tight drug-coupling to the aspartic acid dyad, whilst a high value of $Cc(|\mathbf{r}_{FA}|:|\mathbf{r}_{SA}|)$ indicates tight drug-coupling to the flaps. Only in the wildtype is the drug significantly coupled to the aspartic acid dyad. The G48V-containing systems both exhibit tight coupling of the drug to the flaps. $Cc(|\mathbf{r}_{FS}|:|\mathbf{r}_{SA}|)$ is not significant in any system as the point 'Saq' does not necessarily lie on the Flap-Asp vector.

Conformational changes of saquinavir

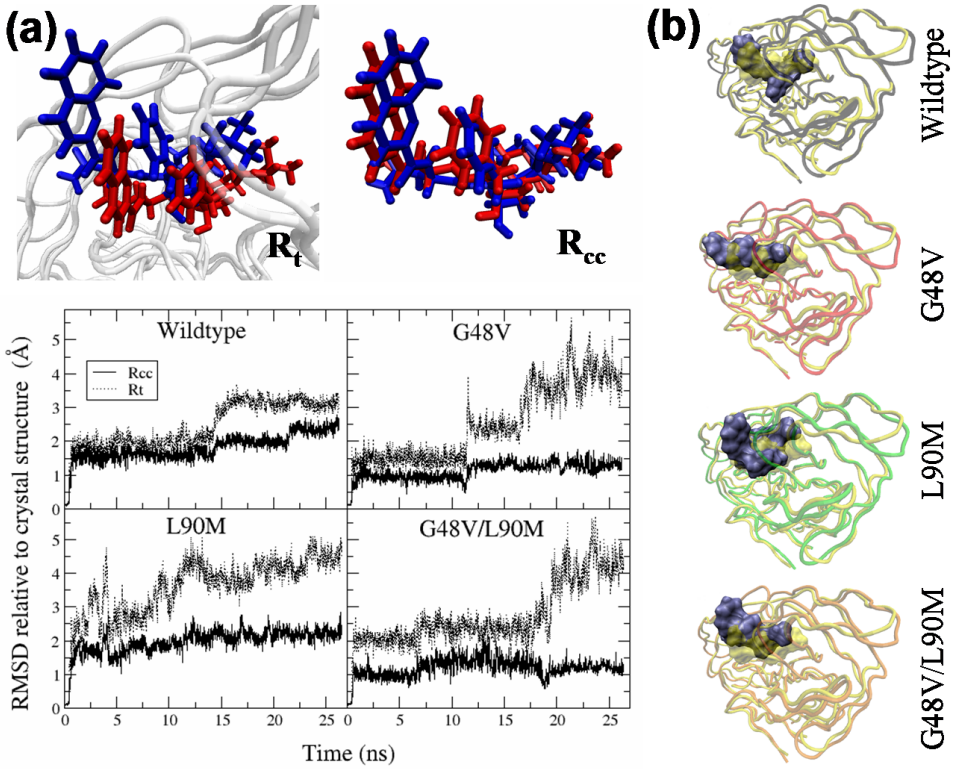

Figure 3: (a) Differences in the RMSD of saquinavir relative to its crystal structure in all four systems, for two different alignment protocols. R_t (dotted line) shows the RMSD of saquinavir atoms after alignment of the protein backbones, R_{cc} (solid line) after alignment to the heavy atoms of saquinavir. (b) Comparison of solvent accessible surface area plot (probe radius 1.4 Å) of saquinavir from the wildtype crystal structure 1HXB (transparent yellow) with the averaged structure over the last 1 ns of simulation (opaque ice-blue) for all systems. The backbones of the proteases are depicted in black, red, green and orange for wildtype, G48V, L90M and G48V/L90M mutant systems respectively and are aligned (R_t) before comparing the structures of the inhibitor.

Conformational changes of the entire saquinavir molecule were determined by taking root-mean-squared deviation (RMSD) measurements of the drug relative to its crystal structure for two different alignment protocols (Figure 3(a)). In the first protocol (R_t), prior to calculation of the RMSD at each step, the protein backbones of the two relevant systems were aligned. In the second protocol (R_{cc}), the heavy atoms of saquinavir in both systems were aligned to each other. R_t therefore measures the total motion of saquinavir from its original position, incorporating translational and rotational motion of the centre of mass of the molecule as well as net conformational changes, whilst R_{cc} measures solely net conformational change. The difference between R_t and R_{cc} is an indication of the degree of bulk translational and rotational motion which we term the 'bulk' motion. By the end of the simulations, R_t (dashed line) is larger in all mutants as compared to the wildtype. The largest R_{cc} (solid line) is observed for the L90M system, whilst saquinavir changes conformation little in the G48V and G48V/L90M systems. However, whilst this observation corresponds to lateral motion away from the active site centre in the L90M system, the change in R_{cc} in the wildtype corresponds to motion towards the catalytic centre. In the wildtype R_t and R_{cc} are within 1 Å of each other at the end of the simulation, implying little 'bulk' motion of the molecule. The increased separation of R_t from R_{cc} for all mutant systems demonstrates significant 'bulk' motion of saquinavir within these systems from its initial position and up to a 3 Å deviation from the active site centre by the end of the simulations. Solvent accessible surface area plots of the averaged structure across the last 1 ns of simulation show elongation of saquinavir in the G48V and G48V/L90M mutants through rotation of the quinoline and phenyl moieties into a conformation where their planes are parallel to each other, and demonstrate the differential position of the drug in each of the four systems compared to the starting structure (Figure 3(b)). Lifting of the inhibitor towards the flap as well as significant translation out of the site are also observable. The drug undergoes conformational changes to sit lower in the wildtype active site than when in the crystal structure, although bulk motion is not significant. Large conformational change is apparent in the L90M system with the P1′ subsite of the drug rising into the flaps. The larger degree of 'bulk' motion of all mutant systems compared with the wildtype together with the conformational changes associated with the wildtype are also confirmed by direct visual inspection.

Water coordination in the active site

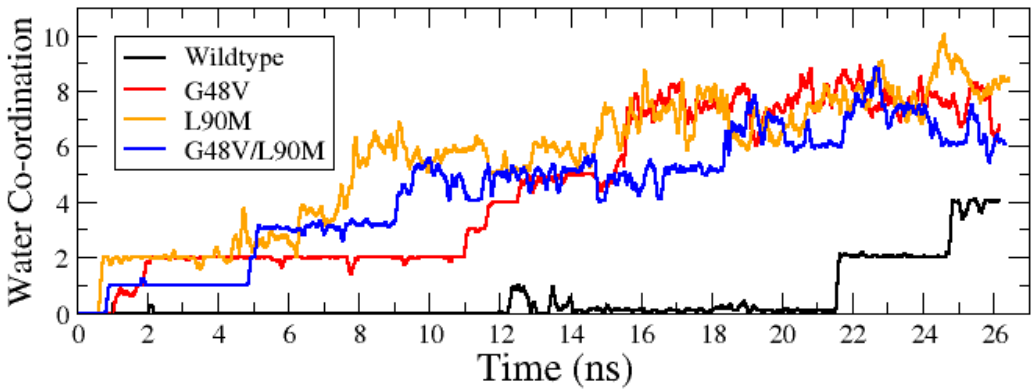

Figure 4: Water coordination within a 3 Å radius of the catalytic dyad, using a 100 ps running average window. Water ingress into this region occurs more for all mutants than for the wildtype. Furthermore, loss of hydrogen bonds with the aspartic acid dyad coincides with further water ingress in the G48V and G48V/L90M systems.

Simulation and analyses of saquinavir bound to mono-protonated HIV-1 proteases

As mentioned in the main manuscript, there are two likely protonation states for saquinavir bound to HIV-1 protease. The mono-protonated state has been shown in previous studies to be thermodynamically favourable, whilst a di-anionic state for the apo-protease exists under physiological conditions. Previous studies have also shown that any protonation is likely to occur upon or after ligand binding and so a di-anionic dyad with saquinavir bound is likely to exist prior to such protonation. In the main manuscript, we have studied the dynamics of saquinavir bound to such a di-anionic protease for the wildtype and G48V, L90M and G48V/L90M mutant proteases and observed a lateral dissociation mechanism induced by the G48V mutation. A dissociation mechanism has not been observed in the mono-protonated state in the literature and it is thus likely that the G48V mutation takes advantage of the di-anionic state to induce resistance via the mechanism described in the main manuscript. To support this claim and in order to understand better the effects of mono-protonation of the catalytic dyad on the dynamics of saquinavir in the active site of HIV-1 protease, we have implemented molecular dynamics simulations for saquinavir bound to the same proteases, but with the catalytic dyad in a mono-protonated state, over a timescale of 10 ns for each system.

Methods

The initial preparation, minimisation and equilibration protocol for these systems was similar to that implemented for the di-anionic study. We describe the differences here. The 1HXB crystal structure (resolution 2.3 Å) was used for the wildtype and G48V systems, whilst the 1FB7 structure (resolution 2.6 Å) was used for the L90M and G48V/L90M systems. The Leap module in the AMBER9 software package was used to protonate residue ASP25 of the protease. The assignment of residue numbers across monomers in this study was in a reverse order to that of the di-anionic study. ASP25 in the mono-protonated systems therefore corresponded to ASP125 in the di-anionic systems. Five Cl^- counter-ions were added to electrically neutralise each system, which was then solvated using atomistic TIP3P water in a cubic box with at least 14 Å distance around the complex. The size of each prepared system was 45314, 45338, 42670 and 42676 atoms for the wildtype, G48V, L90M and G48V/L90M systems respectively.

Minimisation was conducted for 2000 steps for each system owing to a comparatively larger system size. Restrained atoms were held with a force constant of 4 kcal/mol/Å^2. Each system was annealed from 50 K to 300 K over a period of 50 ps, then maintained at 300 K for a further 200 ps in the NPT ensemble. For each mutation, the relaxation protocol (described in the main manuscript) was implemented sequentially (50ps/mutation). The relaxed region was again constrained before moving on to the next mutation. This was followed by gradual constraint relaxation on the drug atoms followed by relaxation on the protease atoms over a further 350 ps. Finally, unrestrained equilibration in the NPT ensemble was conducted up to a total equilibration time of 2 ns for each system. Each system was then simulated for 10 ns in the production phase.

Results

We determined the same Flap-Asp (black line), Flap-Saq (red line) and Saq-Asp (orange line) distances as described in the di-anionic study (see Figure 5). The Flap-Asp distance remained stable at around 12 Å for all systems. Furthermore, the G48V, L90M and G48V/L90M systems exhibited almost identical Flap-Saq and Saq-Asp distances across the whole trajectory, remaining at 5 Å and 8 Å respectively. The wildtype exhibited conformational readjustments for about 6 ns into the production phase, followed by stabilisation of the Flap-Saq and Saq-Asp distances at 6 Å and 7 Å respectively. Therefore, in comparison to the 14.9 Å defined as the semi-open flap conformation, all systems remained in the closed conformation for the whole duration of the simulation. Furthermore, there was no evidence of coupled motion between the Flap-Asp and either the Flap-Saq or Saq-Asp vectors, the magnitudes of cross-correlation coefficients for all

distance pairs were all less than 0.7, and no G48V-related motion of the inhibitor towards the flaps was observed.

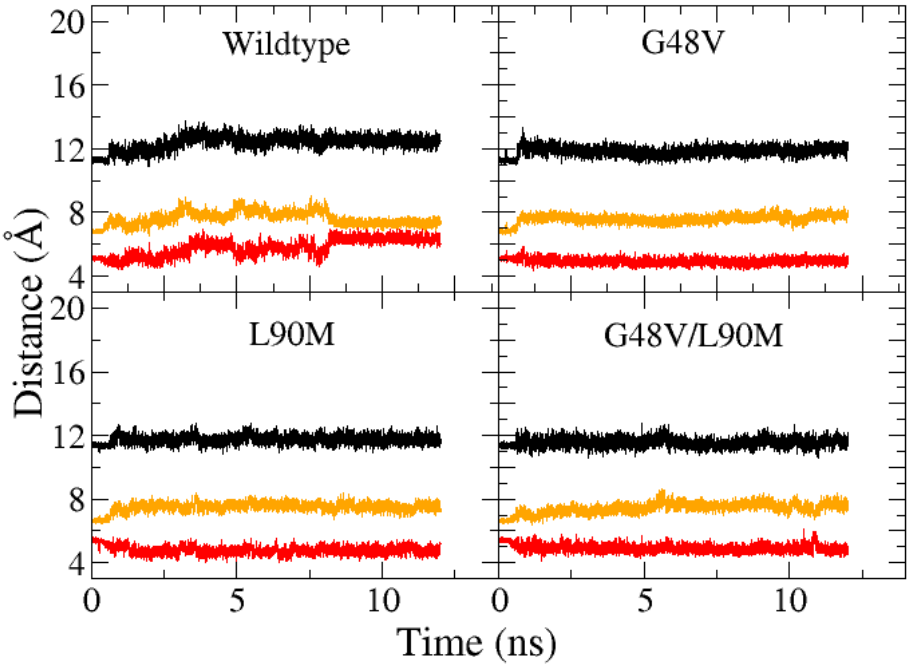

Figure 5: Time evolution of the Flap-Asp, Flap-Saq and Saq-Asp vectors over 10 ns for each mono-protonated protease system. Unlike the di-anionic proteases, the flaps remain in a closed conformation with a Flap-Asp distance (black line) of approxiamtely 12 Å. Furthermore, there is no consistent G48V-related motion of the inhibitor towards the flaps.

The lateral motion of the inhibitor was also determined (see Figure 6). In each system, the drug exhibited little lateral motion away from its original position. Indeed the wildtype moved the most (∼1Å) out of all the systems and each system exhibited stable fluctuations in between -0.5 Å and 1 Å from the original positon of $d_L \sim 2$ Å. This contrasted with the lateral motion observed in the di-anionic study, in which lateral motion up to 4 Å was observed from the original position.

Conformational changes of the saquinavir molecule as compared to net motion of the drug were also determined using RMSD analysis (see Figure 7) across two different alignment protocols, R_t (alignment of the protease backbone) and R_{cc} (alignment of the heavy atoms of saquinavir). The wildtype exhibited significant 'bulk' deviation in the first 6 ns post-equilibration, but subsequently stabilised with a total RMSD (R_t) of just under 3 Å, with almost no separation of R_{cc} from R_t. All mutant systems exhibited no significant change in RMSD post-equilibration. Furthermore, there was very little separation of R_t from R_{cc} in these systems, confirming that little 'bulk' motion of the inhibitor occured from its original position.

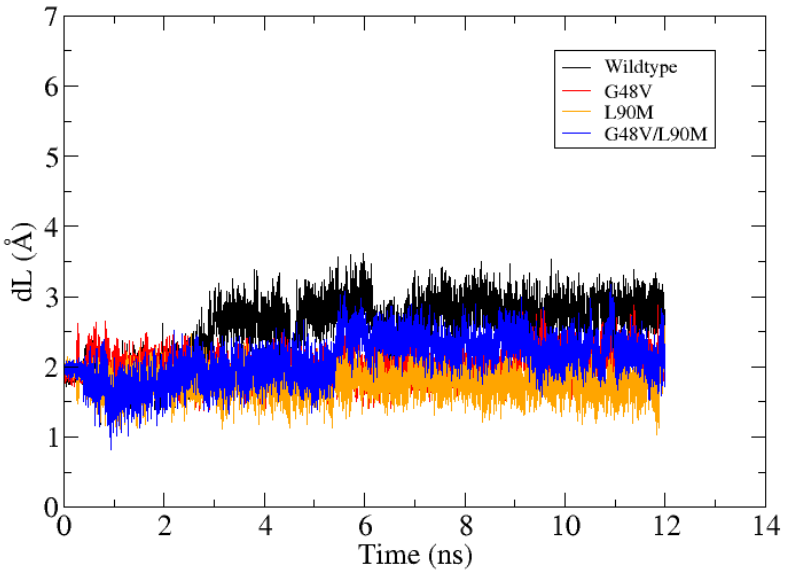

Figure 6: Lateral motion of saquinavir within the active site as measured by the time evolution of the lateral vector d_L (the perpendicular distance of the centre of mass of saquinavir (Saq) from the Flap-Asp vector). The lateral motion in the wildtype protease is the largest with ∼1 Å deviation of the drug from its original position.

Discussion

Our study on the dynamics of saquinavir bound to the mono-protonated HIV-1 protease for the wildtype and G48V, L90M and G48V/L90M mutant systems confirms the stabilising effect of protonation on the binding event. The flaps in all systems remain closed and the drug remains bound in the active site with a maximum deviation of 3 Å observed for the wildtype. Furthermore, the lateral motion is less than 1 Å away from the original position while the G48V-related lateral dissociation mechanism is not observed in the mono-protonated systems. As the mono-protonated system has been reported to be energetically favourable when saquinavir is bound and, from our study, exhibits stable binding in the closed conformation, the G48V mutation is unlikely to be able to take advantage of this state to confer resistance. By contrast, the di-anionic state exhibits a consistent lateral dissociation mechanism observed in the G48V and G48V/L90M systems, which is directly due to the additional flap-coupling induced by the G48V mutation. As protonation occurs upon or after ligand binding, at physiological pH, the G48V mutation is likely to confer resistance whilst the drug is bound to a di-anionic dyad by inducing the first stages of the lateral dissociation, prior to protonation.

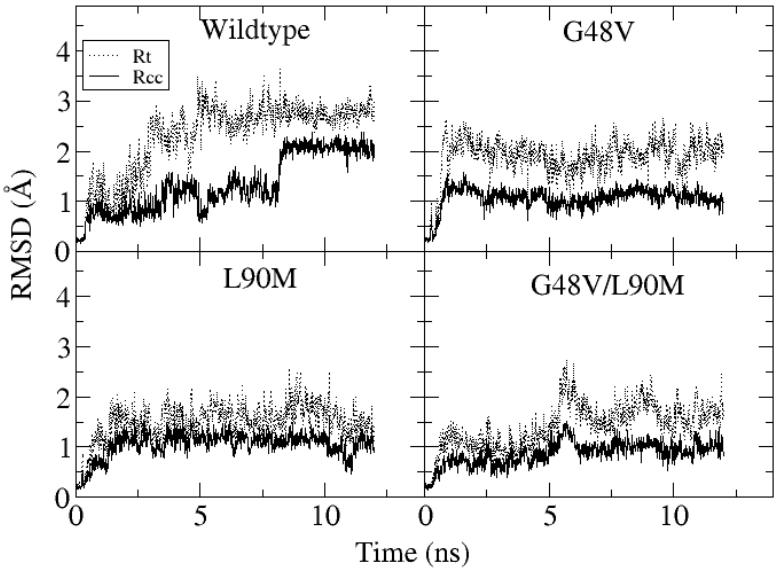

Figure 7: Differences in the RMSD of saquinavir relative to its crystal structure in all four mono-protonated protease systems, for two different alignment protocols. R_t (dotted line) shows the RMSD of saquinavir atoms after alignment of the protein backbones, R_{cc} (solid line) after alignment to the heavy atoms of saquinavir. After initial changes in the position of the drug, which in the wildtype lasts for several nanoseconds, all systems exhibit stable inhibitor RMSDs. Unlike the situation pertaining in the di-anionic study, there is no observed separation of R_t and R_{cc} here, indicating insignificant 'bulk' motion of the drug.