CHAPTER 4

PREDICTION HIV-1 PR MUTANT DUE TO 6 FDA-APPROVED DRUGS

4.1 Introduction

As known and given in details in chapter 1, the HIV-1 protease (PR) is an essential enzyme for the production of the infectious virions. This enzyme is one of the promising targets. Various protease inhibitors are now in clinical used. HIV-1 PR composes of two identical polypeptides consisted of 99 amino acid residues each, which Asp-Ser/Thr-Gly[161-164] catalytic residues are conserved. During the past decade, six protease inhibitors (PIs) have been approved by FDA and successfully used to decrease the death rate due to AIDS. Recently, 4 drugs were additionally approved. The major problem associated with the existing drug against HIV-1 PR is molecular resistance. Their effectiveness was lost against HIV-1 PR due to the rapid point mutations of the genome of HIV. Mutation V82A was associated with the initial loss of antiviral activity. Subsequently, the following mutations were observed most frequently: I54V, A71V, and M36I. Later, proteases containing mutations I84V, K20R, M46I, L33F, and L90M appeared. Primary mutations directly confer resistance to one or more protease inhibitors, whereas secondary mutations only contribute to resistance and often occur together with primary ones or in synergistic form with other secondary mutations. Active-site mutations are exclusively primary ones, but not all primary mutations are necessarily limited to the active-site (e.g., the nonactive site mutations at sequence positions 46, 88 and 90 can also directly confer resistance). The mechanism of active-site mutations can frequently be rationalized in structural

terms in which resistance is directly associated with a change of the contacts and thus the interaction energy between drug and target. In contrast, the mechanism of nonactive-site mutations, the so-called compensatory mutation, influencing binding from distal locations, is not satisfyingly understood. In general, the common mutation found for protease inhibitors were given in Table 4.1, and additional mutated residues were summarized in Table 4.2 for the specific FDA-approved drugs corresponding with the high and the intermediate resistant level[165].

The present work have shown the ability of prediction of the HIV-1 PR mutation due to 6 FDA-approved drugs as the high and intermediate resistant levels (Table 4.2) using decomposition (DC) free energy. Although, the protonation state for each drug was already proposed in chapter 3, it remains interesting whether the different protonation state will effect the prediction of the HIV-1 PR mutation. In order to understand this suspected point, the prediction of HIV-1 PR mutation due to 6 drugs was examined for the three protonation states (D25, D25', D25,25'). In addition, the proposed prediction process was also applied for the influenza virus enzyme, neuraminidase subtypes N1, N2, and N9 complexed with the oseltamivir.

Table 4.1 The frequent mutated residues classified by their the locations in the HIV-1 PR.

Mutation site	**Res.no.**
Substrate cleft	8, 30, 32, 48, 50, 82, 84
Protease flap	46, 47, 53, 54
Nearby the conserved residues	24, 33, 73, 88, 90
Contribution to resistance	10, 20, 36, 63, 71, 77, 93

Table 4.2 The mutated residues of specific FDA-drugs which lead to the high and intermediate resistant levels as the clinical data report.

Drug	High resistant level	Intermediate resistant level
LPV	82	46, 47, 50, 54, 84, 90
RTV	82, 84	32, 46, 47, 50, 53, 54, 90
SQV	48, 84, 90	53, 54, 82
IDV	82, 84	32, 46, 47, 53, 54, 90
APV	50, 84	32, 46, 47, 54, 82, 90
NFV	30, 84, 90	46, 48, 54, 82, 88

4.2 Calculation Details

MM-GBSA has been used to predict the binding free energy between individual residue and inhibitor (ΔG_{total}), following equation 4.1, using decomposition energy (DC) module in AMBER 8. GB model of Onufriev *et al.*[166] has been used in this work.

$$\Delta G_{res} = \left\langle G^{res+inhibitor} \right\rangle - \left\langle G^{res} \right\rangle - \left\langle G^{inhibitor} \right\rangle \qquad (4.1)$$

The $\left\langle G^{res+inhibitor} \right\rangle$ is the average binding free energy between individual residue located on the protein and the inhibitor, while $\left\langle G^{res} \right\rangle$ and $\left\langle G^{inhibitor} \right\rangle$ are the average binding free energy of individual residue and inhibitor, respectively.

$$\left\langle G^{x} \right\rangle = H^{x}_{gas} + G^{x}_{solvation} - TS^{x} \qquad (4.2)$$

$\langle G^x \rangle$ contains internal gas-phase energies, solvation free energies, and entropic terms with the superscript x represents the free energies of the three terms shown in equation 4.1, the complex of individual residue and inhibitor, individual residue, inhibitor. The furthermore details show in Gohlke, H. *et al.*[167]

For DC binding free energy calculations, 100 sampling configurations were taken from explicit MD simulations of HIV-1 PR bound to 6 drugs, computed for 3 possible protonation states. Calculation were additionally examined for the Neuraminidase subtypes N1, N2, and N9 where the MD trajectories were taken from Aruksakunwong, O., *et al.*[168].

4.3 Results and Discussions

4.3.1 Prediction of mutation-pattern for HIV-1 PR complexed with six FDA-approved drugs.

The DC binding free energy (ΔG_{res}) according to equation 4.1 was computed for the HIV-1 PR complexed with 6 FDA-approved drugs (LPV, RTV, SQV, IDV, APV, and NFV) at the 3 protonation states (D25, D25', and D25,25'). The results were summarized in Figure 4.1. The mutated residues of high and intermediate resistant levels according to the clinical data were also given for comparison.

The DC energy patterns for each chain show three minimum regions taken place at catalytic (residues 23-33), flaps (residues 47-53), and C-terminal (residues 77-84) regions. This observation is true for both chains, all 6 drugs and all protonation states. Additional peak was found only at the residues 8-10 of chain B of the IDV complex.

Surprisingly, exclude the catalytic region where the residues are usually conserve and display tight enzyme-inhibitor binding, the high and intermediate levels of mutated residues were found to locate within the other two low DC energy regions, flap and C-terminal. Such direct selection does not depend on protonation states and type of inhibitors.

4.3.2 Prediction of mutation for the influenza Neraminidase

Special attention turns into the influenza neuraminidase (NA),. This enzyme plays a critical role in the releasing of virus after replication. This is to examined the free energy decomposition model in predicting the mutation. Therefore, three subtypes of NA, N1, N2, and N9, complexed with oseltamivir (OTV) have been chosen. The MD trajectories were taken from Aruksakunwong, O., *et al.*[168]. The DC free energies for the three complexes were plotted in Figure 4.2 in which the residues that interact strongly, both repulsion and attraction, were labeled. Note that the NA consists of 449, 388, and 387 residues for N1, N2, and N9, respectively, where 118, 292, and 371 are the catalytic residues. The mutated residues specific to each subtype of NA which propose clinically were also given in the plot.

The plots for the three NA subtype show similar DC free energy patterns. The strong attraction DC energy for the OTV-N2 complex take place ate the 119, 151, 292 and 371 residues. The corresponding set for the OTV-N3 complex is 118, 150, 292 and 370. Interestingly, the two high resistant residues, 119 and 292, are also included in the strong attraction regions. The prediction is not valid for the OTV-N1 complex. This can be due to the following reasons *(i)* the homology model was used as starting geometry of MD simulations for the OTV-N1 complex because the x-ray structure was not available at that time. and *(ii)* The OTV does not match to the

N1 catalytic cavity. It was investigated in details in reference of Aruksakunwong, O., *et al.*[165] and it was found that due to strong repulsion of the OTV's bulky sidechain.

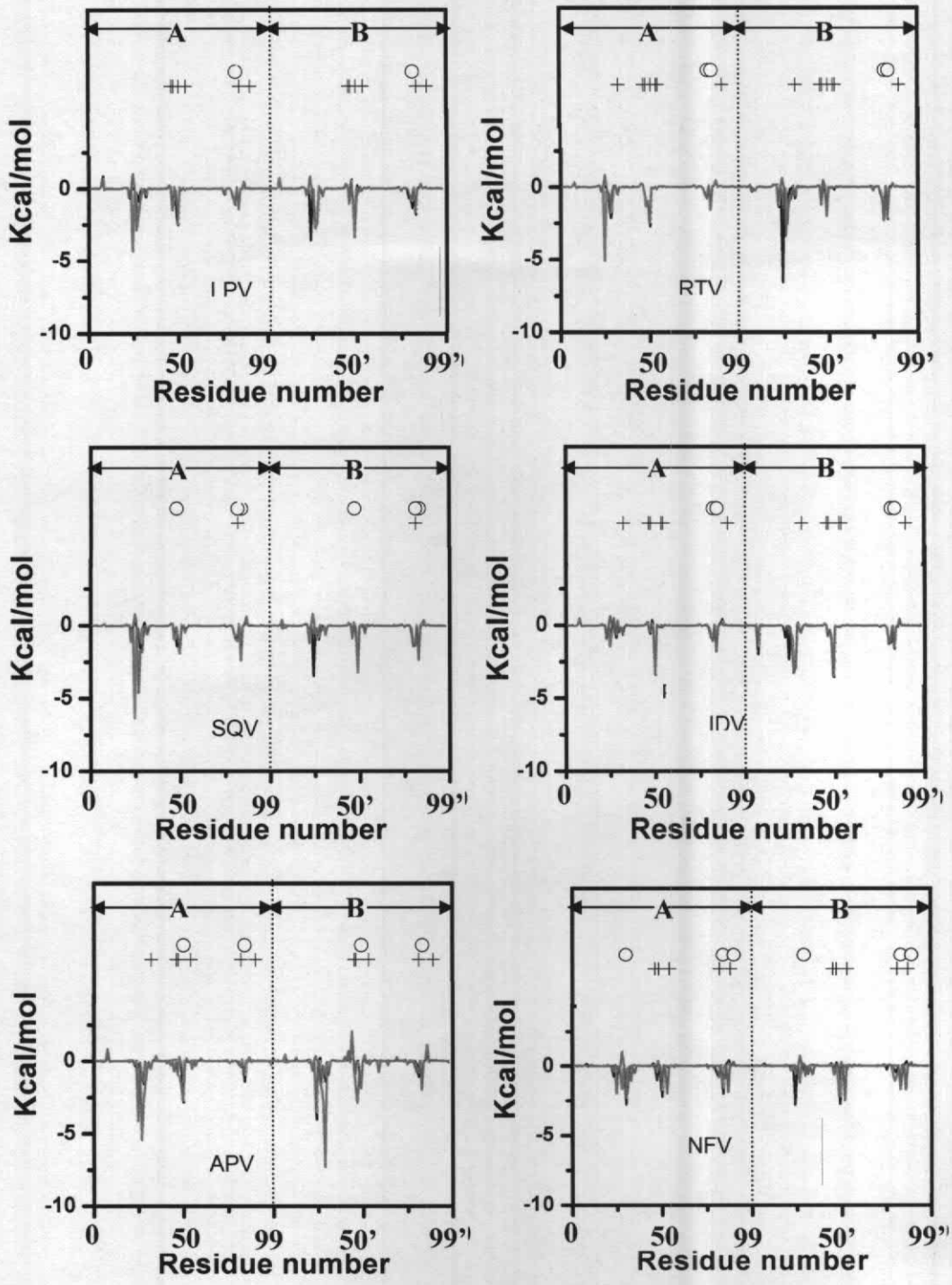

Figure 4.1 The decomposition energy for both chains (A and B) of the HIV-1 PR complexed with 6 drugs (LPV, RTV, SQV, IDV, APV, and NFV) for the D25 (black), D25' (red) and D25,25 (green) protonation caculated using MMGB/SA approach.

representing in black, red and green lines, respectively. The mutated residues of high (o) and intermediate (+) resistant levels according to clinical data were also given for comparison.

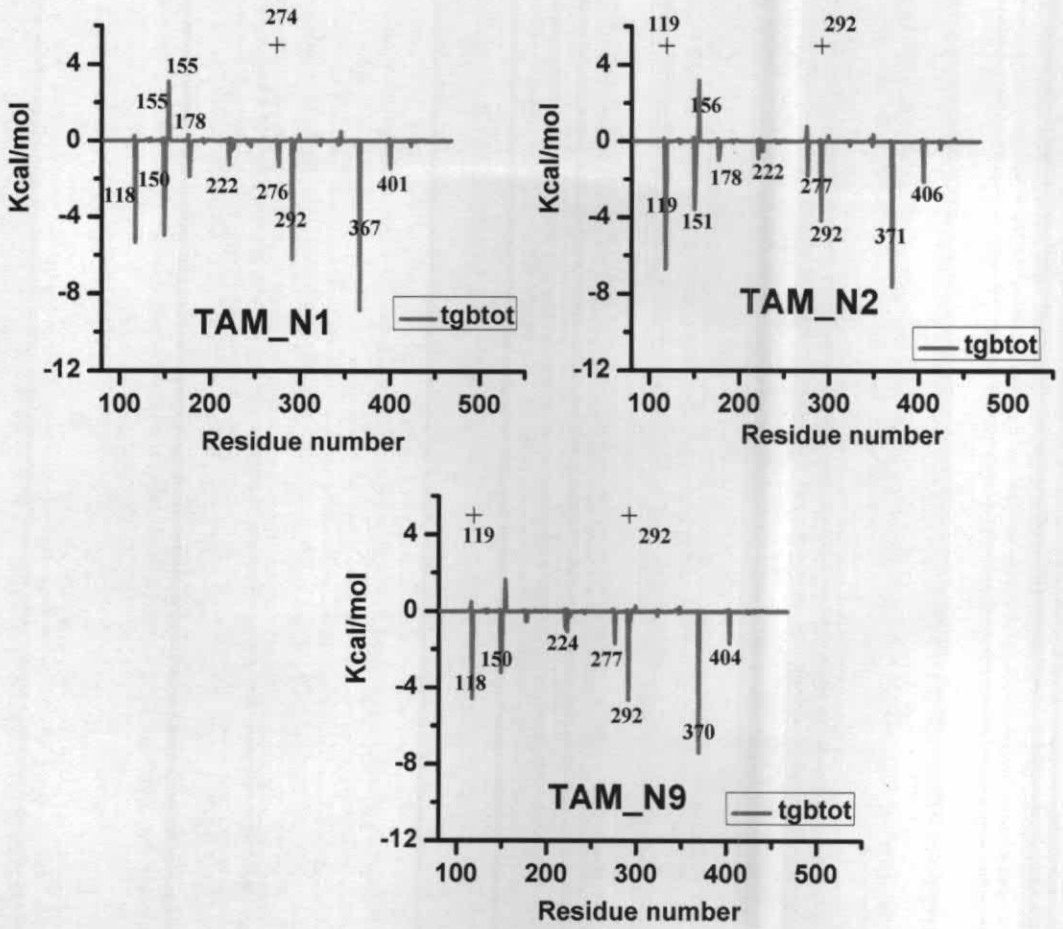

Figure 4.2 The decomposition energy of the three subtypes of Neraminidase (N1, N2, and N9) complexed with oseltamivir, OTV, with the clinical mutated residues (label as +).

4.3.3 Insight into the source of Gly48 and Ile84 mutation using SQV complex as a case study

To understand detailed information, we now turn our attention to investigate molecular properties around the inhibitor. Attention is focused on the flexibility of the residues lying in the flap region and those located within 3 Å from all atoms of saquinavir for three protonation states, D25, D25' and D25,25'. The RMSD for those residues with respect to the X-ray structure for three states were evaluated and plotted in Figure 4.3a-4.3b.

It is interesting to note that within a spherical radius of 3 Å from the saquinavir highest and lowest displacements were observed at Ile84 and Gly48 residues of chain A (Figure 4.3a-4.3b), respectively (Gly48 is not within 3 Å limit for D25,25'). As it was found that Ile84 and Gly48 are two of the most frequent residues where mutation toward a protease inhibitor[166] takes place. The data lead us to conclude that a conformational change is a primary source of mutation of Ile84. In contrast, Gly48 mutation can be described in terms of indirect displacement of Phe53 in the flap region. As can be seen in Figure 4.4b that the RMSDs of Phe53 of the D25 chain A and of D25,25' chain B are significantly higher than those of the other two states. A clear description of the mutation of Gly48 was given in the next paragraph.

To explore the conformational changes due to complexation, subtraction between the RMSDs for the residues of the free enzyme and the complex of D25 was taken into account because the previous results indicating D25 is the most favorable state. These subtracted values of those residues, lying in the flap region and those located within 3 Å from the atoms of saquinavir were plotted in Figure 4.4a-4.4b, i.e., a positive value indicates higher rigidity of the complex than the free enzyme. Dramatic changes were detected on the residues Ile50', Phe53' (Figure 4.4b)

4.3.3 Insight into the source of Gly48 and Ile84 mutation using SQV complex as a case study

To understand detailed information, we now turn our attention to investigate molecular properties around the inhibitor. Attention is focused on the flexibility of the residues lying in the flap region and those located within 3 Å from all atoms of saquinavir for three protonation states, D25, D25' and D25,25'. The RMSD for those residues with respect to the X-ray structure for three states were evaluated and plotted in Figure 4.3a-4.3b.

It is interesting to note that within a spherical radius of 3 Å from the saquinavir highest and lowest displacements were observed at Ile84 and Gly48 residues of chain A (Figure 4.3a-4.3b), respectively (Gly48 is not within 3 Å limit for D25,25'). As it was found that Ile84 and Gly48 are two of the most frequent residues where mutation toward a protease inhibitor[169] takes place. The data lead us to conclude that a conformational change is a primary source of mutation of Ile84. In contrast, Gly48 mutation can be described in terms of indirect displacement of Phe53 in the flap region. As can be seen in Figure 4.4b that the RMSDs of Phe53 of the D25 chain A and of D25,25' chain B are significantly higher than those of the other two states. A clear description of the mutation of Gly48 was given in the next paragraph.

To explore the conformational changes due to complexation, subtraction between the RMSDs for the residues of the free enzyme and the complex of D25 was taken into account because the previous results indicating D25 is the most favorable state. These subtracted values of those residues, lying in the flap region and those located within 3 Å from the atoms of saquinavir were plotted in Figure 4.4a-4.4b, i.e., a positive value indicates higher rigidity of the complex than the free enzyme. Dramatic changes were detected on the residues Ile50', Phe53' (Figure 4.4b)

and Asp25 (Figure 4.4a) where the RMSD of free enzyme are much higher than those of the complex. This observation can be easily explained in terms of binding between enzyme and inhibitor which leads directly to rigidity of the complex. Surprisingly, residue Phe53 (Figure 4.4b) in the flap region in the complex form is less rigid than the free form. This is in contrast to what is generally known and reported by Zhongwei Zhu et al.[170] using free energy techniques showing that a barrier to flap opening exists in the presence of an inhibitor while this evidence disappears for the free enzyme. This observation confirms the lower rigidity of Phe53 chain A of D25 state in the complex than the free forms (Figure 4.4b). To understand the reason for the mutation due to the indirect displacement, superimposition between the enzyme structures in the free and complex forms was examined, and the molecular structure in an area close to Phe53 of both chains was displayed in Figure 4.5. It can be clearly seen from the plot that the presence of saquinavir leads to dramatic changes of the conformations of Phe53 and Phe53'.

Gly48, on one side, was found to form strong hydrogen bonds with saquinavir (Figure 4.6), 100% occupied throughout the simulation, while it was, on the other side, repelled by the hydrophobic Phe53 residue. This leads directly to a very high rigidity of Gly48 (Figure 4.3b, chain A) as well as a high displacement of Phe53. Such unfavorable conditions were supposed to facilitate the mutation of Gly48. Due to the unsymmetric character of the saquinavir molecule, Gly48' was not detected to bind to the other end of saquinavir. Therefore, Gly48' is less interfered by the presence of saquinavir in comparison to Gly48. This can be also a reason for the lower RMSD of Phe53' than Phe53 (Figure 4.3b, D25) and the higher degree of drug resistance of Gly48 than Gly48'. This evidence is fully supported by the X-ray crystallographic analysis and molecular modeling experiments.[169] A clear

conclusion is that in addition to the I84V mutation due to high displacement of Ile84 in the presence of saquinavir (Figure 4.3a), cooperation effect can be also due to the limited space (Figure 4.5) in the HIV-1 PR pocket, especially for the G48V where saquinavir requires more space on the Gly48 than the Gly48'sides.

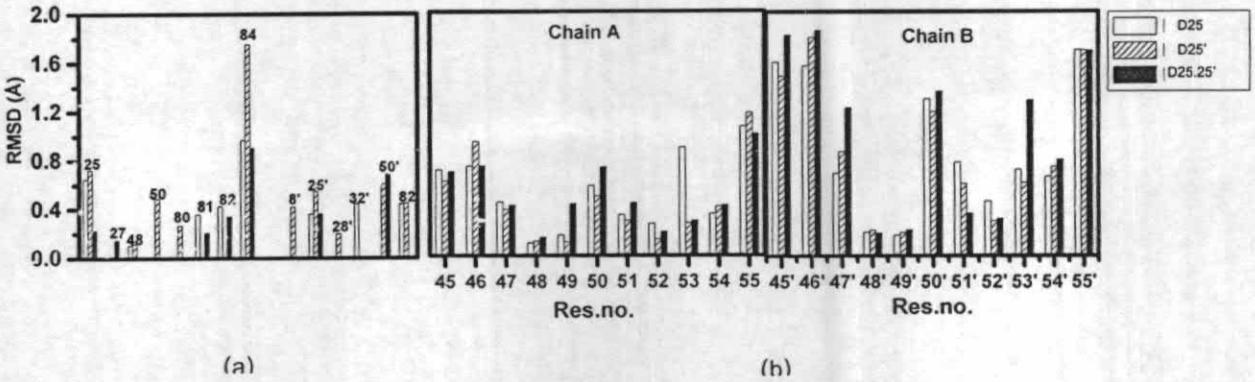

Figure 4.3 Comparison of RMSD for the D25, D25' and D25,25' with respect to their average structure for the residues located within 3 Å around saquinavir (a) and in the flap region, residues 45-55 (b).

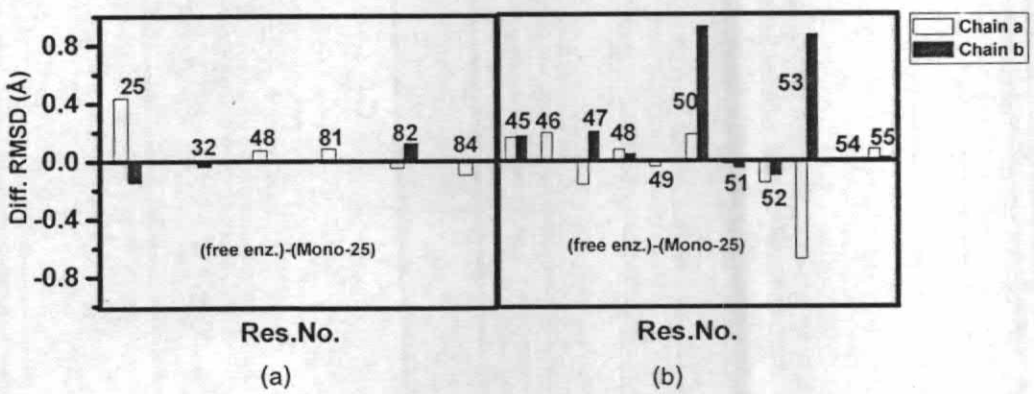

Figure 4.4 Subtraction between the RMSDs of the free enzyme and complex for the Mono-25 state lying within 3 Å from the atoms of saquinavir (a) and the flap region, residues 45-55 (b).

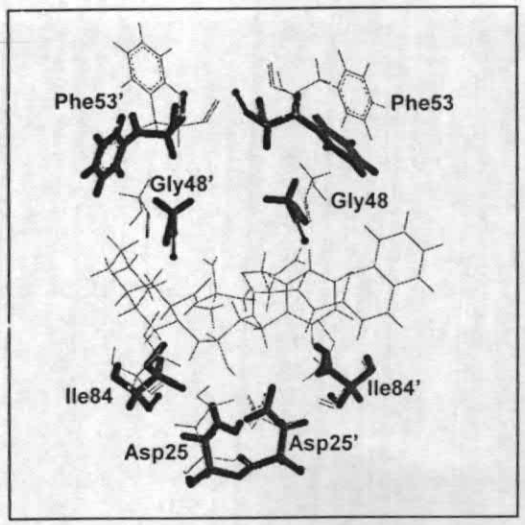

Figure 4.5 Superimposition between the average structures of HIV-1 PR in free (stick) and complex forms (line) where only selected residues are displayed.

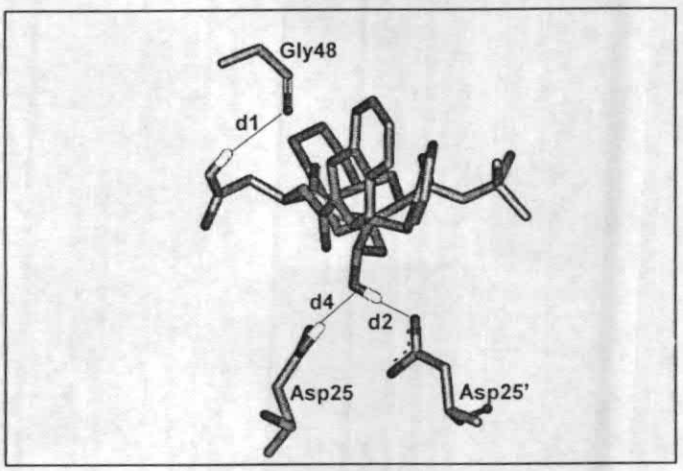

Figure 4.6 Binding between saquinavir and HIV-1 PR in the D25 states where solid denote hydrogen bonds which were detected 100% of the total configurations after equilibration, respectively.

4.4 Conclusion

Free energy Decomposition, implemented in AMBER, is an attractive approach to predict the possible mutated residues in terms of high and intermediate for HIV-1 PR complexed with inhibitor/drug. The different protonation states of HIV-1 PR bound to 6 FDA-approved drugs are not significant effect on prediction of HIV-1 PR/inhibitor using binding free energy decomposition. In addition to the capable prediction of HIV-1 PR mutation, this method is valid for the prediction of three types of neuraminidase/oseltamivir complex. Study of Molecular dynamic simulation is able to investigate insight into the details of the source of mutation. For instance, the source of Gly48 and Ile84 mutation for HIV-1 PR complexed with SQV is revealed that conformational change is a primary source of mutation for Ile84. In contrast, Gly48 mutation can be described in terms of indirect displacement of Phe53 in the flap regions.