CHAPTER 5
CONCLUSION

5.1 Full-length model

The structure of the full-length HIV-1 IN containing all three domains, N-terminal, core and C-terminal domain was successfully built up using the available experimental data of the two-domain structures, CORE-N and CORE-C. The structure is comparable to the previous model structure proposed by Luca *et al.* The missing residues were modeled using the homology modeling. The model regions show high variety as their flexibility when compare to the CORE only and the other two-domain, CORE-N and CORE-C structures. The orientation of the two-end domain, N-terminal and C-terminal domain, were placed in the right position based on the crystal data of the CORE-N and CORE-C. The core and N-terminal portions are in a compact form while the C-terminal portion is rather extended. This is different from what found in the other full-length model proposed by Podtelezhnikov *et al.*, in which the two-end domains, N-terminal and C-terminal, were not connected.

The dimer and tetramer forms of the HIV-1 IN were also proposed herein in the same manner. The missing residues were completely modeled. The dimer structure is comparable to the Luca *et al.*'s model. However, the tetramer model differs from Podtelezhnikov *et al.*'s model especially for the orientation of the N-terminal and the C-terminal domain.

The full-length HIV-1 IN modeled here aims to propose another alternative complete model in which theoretically synthesized based on the available experimental structure. The structure was fully optimized and solvated to investigate the structural and dynamical properties and to obtain the information at the mean time that the full-length structure has not been able to resolve using experimental technique.

5.2 Structural and dynamical properties

The crystallographic structures of the various forms of the HIV-1 IN were considered in this study. The missing residues present in the crystal structure were constructed using homology modeling method. By comparison to the structure after modeling of the HIV-1 IN, the results reveal that the modeled regions of various systems, residues 47 – 55 in the N-terminal and core connection region and residues 140 – 148 in the core region, were significantly different. These results are in agreement with the experimental data in which these regions cannot be resolved. The linkage between the central core domain and the two-end terminal domains, the N-terminal and the C-terminal are significantly different. This supports the previous study which states that these connections are flexible.

The structural and dynamical properties were explored using MD simulations. The 2-ns time scale was chosen and compared to the previous MD studies. All systems were under equilibrium before the properties were analyzed. Among the three domains, the core domain has the least flexibility while the two end terminal domains have larger fluctuation; particularly, the C-terminal domain has the largest fluctuation with the highest RMSD values. Such fluctuation in the two end domain regions is due to the flexibility of the elbow linkage between the domains. Such flexibility may play important role in binding to DNA during the integration process. Divalent metal ion, Mg^{2+}, binds to the catalytic residues in an octahedral fashion, by which three coordination from HIV-1 IN (two from Asp64 and one from Asp116), thus, allow three water to come in.

The absence of the metal ion, Mg^{2+}, in the active site of the core domain has slightly effect on the conformation of the catalytic residues. Binding of the Mg^{2+} to the Asp64 and Asp116 forces these residues to have only one preferential torsion angle in comparing to two preferential angles found in the FULL system. The RMSD over the 2-ns simulation shows that the FULL+ION has high mobility than that of the FULL one.

5.3 Effect of the divalent metal ion

The divalent metal ion, Mg^{2+}, plays essential role in the integration process in which being catalyzed by the HIV-1 IN. By coordination with three catalytic residues, Asp64, Asp116 and Glu152, in the active site of the catalytic core domain, the Mg^{2+} was proposed to join the polynucleotidyl transfer.

In our study, the effect of the divalent metal ion in the active site of the catalytic core domain was investigated by the MD studies of the two full-length HIV-1 IN systems, the FULL and the FULL+ION. The results reveal that the metal binding in the active site affect the conformation of the catalytic triad. The fluctuations of each catalytic residue are smaller when Mg^{2+} is present in the active site. In addition, the two conformational preferences of the Asp64 in the FULL system were changed, *i.e.* exist only one preference in the FULL+ION system. The active site binding pocket formed by three residues is altered by which the FULL has larger in size than that of the FULL+ION.

5.4 Effect of the two-end terminal domains

The addition of the other parts to the core domain seems to have an effect on the structural and dynamical properties of the core system but not significant. The overall structure of the core domain is rather conserved.

Even the core domain only can carry out the disintegration reaction; the previous studies reveal that all three domains are essential for the integration process. Considering the missing region, residues 140 – 148, it can be seen that the model residues were significantly different in all structure. The result from MD simulation also shows the high fluctuation of this region. Another region in the core domain, residues 188 – 192, also shows high fluctuation. These observations are well agreed with the experimental data. For the second missing region, residues 47 – 55, which is the connection between the core domain and the C-terminal domain, the model structure was significantly different. The MD studies show the same result.

The structure and the mobility of the core domain seem to be affected by the presence of the second and third parts. In addition, the system composing of two domains are different from the complete full-length structure. This implies that the addition of the third part affect the structural and dynamical behavior of the HIV-1 IN structure.

5.5 Dimer and tetramer full-length HIV-1 IN complexed with DNA

The wild type full-length HIV-1 IN for both dimer and tetramer forms complexed with DNA was also proposed in this study. The results were compared to the two available theoretical models proposed by Luca *et al.* and Podtelezhnikov *et al.* The DNA binding in our model differs from Podtelezhnikov *et al.* model in the C-terminal region.

The host DNA binds to HIV-1 IN in the interface region between the core domain of subunits B and D in our model. The viral DNA has a contact with the C-terminal domain. Some key residues observed to interact with the DNA are as follows: Lys156, Lys159, Lys160, Lys186, Arg187 and Lys188 in the core region. Residues Gly245 – Pro261 were found to contact with the viral DNA.

We proposed an alternative theoretical model of wild type full-length HIV-1 IN structure complexed with the viral and the host DNA in the presence of the divalent metal ion, two atoms of Mg^{2+} in the catalytic core domain and a Zn^{2+} in the N-terminal domain. The structural investigation shows that the result is comparable to the available model.

5.6 Suggestion for further works

In this study, we have proposed the full-length HIV-1 IN model based on the two-domain crystal structures, CORE-N and CORE-C. The MD simulations of various systems, FULL, FULL+ION, CORE-N, CORE-C and CORE were carried out. The divalent metal ion, Zn^{2+}, was placed in the N-terminal region, while, the Mg^{2+} was located in the active site of the catalytic core domain.

a. Since we pay much attention to the core domain which is the active site that binds inhibitor, the coordination of the Zn^{2+} and the two histidine and two cysteine moieties were kept fix for the entire simulation. We suggest performing simulation with non-constraint in this region in order to obtain more realistic information in this region. However, the Zn^{2+}-enzyme as well as Zn^{2+}-water force-field potentials have to be re-parameterized.

b. Another Mg^{2+} was suggested to bind at the second position, between Asp64 and Glu152. However the position of this ion has not been identified in the HIV-1 IN crystal structure. It was proposed that this ion play role in the binding with DNA. For the next simulation, we would like to propose to add the second Mg^{2+} in the core domain region.

c. The MD simulations for the systems including inhibitors and both viral and host DNA in various domain structures, CORE, CORE-N, CORE-C and the FULL+ION should be performed to clarify the effect of the end terminal domains on the binding of the inhibitors and DNA.

d. The interaction energy between the HIV-1 IN, particularly in the active site pocket, and various inhibitors should be studied using high level of calculation. The reaction mechanism is interesting to figure out for the HIV-1 IN.