

โครงสร้าง พลศาสตร์และการยึดจับของเอชไอวี-1 โปรทีเอสที่ติดต่อยา
โดยการกลายพันธุ์ในตำแหน่งหลัก

นางสาวอรจิรา อารักษ์สกุลวงศ์

สถาบันวิทยบริการ

จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาเคมี ภาควิชาเคมี


คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2549

ISBN 974-14-2971-1

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

STRUCTURE, DYNAMICS AND BINDING OF DRUG RESISTANCE HIV-1 PROTEASE
WITH MAJOR MUTATIONS



MISS ORNJIRA ARUKSAKUNWONG

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Chemistry

Department of Chemistry

Faculty of Science

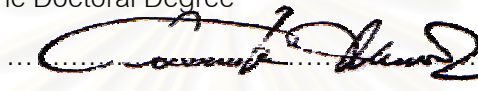
Chulalongkorn University

Academic year 2006

ISBN 974-14-2971-1

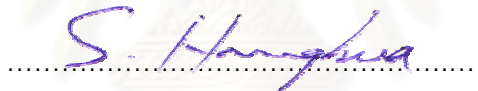
Thesis Title Structure, Dynamics and Binding of Drug Resistance HIV-1
Protease with Major Mutations
By Miss Ornjira Aruksakunwong
Filed of Study Chemistry
Thesis Advisor Professor Supot Hannongbua, Ph. D.
Thesis Co-advisor Associate Professor Pornthep Sompornpisut, Ph. D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of
the Requirements for the Doctoral Degree

.....Dean of the Faculty of Science
(Professor Piamsak Menasveta, Ph. D.)

COMMITTEE


.....Chairman
(Associate Professor Sirirat Kokpol, Ph. D.)

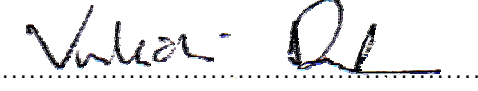
.....Thesis Advisor
(Professor Supot Hannongbua, Ph. D.)

.....Thesis Co-advisor
(Associate Professor Pornthep Sompornpisut, Ph. D.)

.....Member
(Associate Professor Wasun Chantrastita, Ph. D.)

.....Member
(Associate Professor Surapong Pinitglang, Ph. D.)

.....Member
(Professor Kritsana Sagarik, Ph. D.)

.....Member
(Associate Professor Vudhichai Parasuk, Ph. D.)

อรจิรา อารักษ์สกุลวงศ์ : โครงสร้าง พลศาสตร์และการยึดจับของเอชไอวี-1 โปรตีเอสที่ติดต่อยา โดยการกลายพันธุ์ในตำแหน่งหลัก. (STRUCTURE, DYNAMICS AND BINDING OF DRUG RESISTANCE HIV-1 PROTEASE WITH MAJOR MUTATIONS) อ. ที่ปรึกษา : ศ.ดร. สุพจน์ หารหนองบัว, อ.ที่ปรึกษาร่วม : ผศ.ดร. พรเทพ สมพรพิสุทธิ์ 200 หน้า. ISBN 974-14-2971-1.

เอนไซม์เอชไอวี - 1 โปรตีเอส เป็นหนึ่งในเอนไซม์เป้าหมายที่สำคัญสำหรับการพัฒนายาต้านไวรัส เอดส์ ถึงปัจจุบันมีตัวยับยั้งที่ออกฤทธิ์กับโปรตีเอสที่ใช้อยู่ในทางการแพทย์หลายชนิด อย่างไรก็ตาม อุปสรรคสำคัญในการรักษาผู้ติดเชื้อเอชไอวีด้วยยาต้านเอนไซม์โปรตีเอสคือการดื้อยา เพื่อความเข้าใจ ข้อมูลพื้นฐานทางโครงสร้างของเอนไซม์ที่เกิดการกลายพันธุ์ งานวิจัยนี้จึงทำการศึกษาระบบเชิงซ้อนของโปรตีเอสที่เกิดและไม่เกิดการกลายพันธุ์ ด้วยวิธีจำลองกลศาสตร์เชิงพลวัต โดยทำการศึกษากับ ยาซาควินาเวียและยาริทนาเวีย ในขั้นตอนแรกได้ทำการศึกษาเพื่อตอบคำถามเกี่ยวกับสภาพการรับ โปรตอนของบริเวณเร่งซึ่งได้แก่เรสิดิว Asp25 และ Asp25' โดยศึกษาทั้งสามสภาพการรับโปรตอนที่เป็นไปได้ ผลจากการศึกษาพบว่าสถานะที่เหมาะสมคือการรับโปรตอนตำแหน่ง Asp25 จากนั้นจึงใช้วิธีจำลองกลศาสตร์เชิงพลวัตศึกษาเอนไซม์ที่ไม่เกิดการกลายพันธุ์และที่เกิดการกลายพันธุ์ในตำแหน่งหลัก คือ G48V, L90M และ G48V/L90M สำหรับยาซาควินาเวีย และ V82F, I84V และ V82F/I84V สำหรับยาริทนาเวีย ผลการศึกษาแสดงให้เห็นว่าการกลายพันธุ์ในสองตำแหน่งหลักมีผลทำให้เกิดการเปลี่ยนแปลงทางโครงสร้างของเรสิดิวที่สำคัญในบริเวณเร่ง และโครงสร้างของยาขณะที่อยู่ในบริเวณเร่ง ซึ่งทำให้ประสิทธิภาพของการยึดจับระหว่างยากับเอนไซม์ลดลง ในขั้นตอนสุดท้ายได้ทำการคำนวณพลังงานเสรีและการยึดจับกันระหว่างตัวยากับเอนไซม์โปรตีเอส โดยใช้วิธี MM/PBSA และ ONIOM3 ผลที่ได้จากการคำนวณทั้งสองวิธีสอดคล้องกับผลที่ได้จากการทดลองเป็นอย่างดี ซึ่งการเข้าใจถึงข้อมูล พื้นฐานดังกล่าวจำเป็นอย่างยิ่งสำหรับการคัดกรองและการออกแบบยาเพื่อใช้ในการต่อต้านเอนไซม์เอชไอวี-1 โปรตีเอสที่เกิดการกลายพันธุ์

ภาควิชา.....เคมี.....ลายมือชื่อนิสิต.....
 สาขาวิชา.....เคมี.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
 ปีการศึกษา.....2549.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4573851323 : MAJOR CHEMISTRY

KEY WORD: HIV-1 PROTEASE / SAQUINAVIR / RITONAVIR / DRUG RESISTANCE

ORNJIRA ARUKSAKUNWONG : STRUCTURE, DYNAMICS AND BINDING OF DRUG RESISTANCE HIV-1 PROTEASE WITH MAJOR MUTATIONS. THESIS ADVISOR : PROF. SUPOT HANNONGBUA, Ph.D., THESIS COADVISOR : ASSOC. PROF. PORNTHEP SOMPORNPISTUT, Ph.D., 200 pp. ISBN 974-14-2971-1.

HIV-1 protease is one of the importance targets for drug development. To date, many protease inhibitors are in clinical used, however, a major problem in the treatment of the HIV infection with protease inhibitor is the development of drug resistance. To understand the molecular basic of drug resistance, the molecular dynamic (MD) simulation of wild-type and mutant proteases complexed with inhibitors, saquinavir and ritonavir, were carried out. The protonation states of the two aspartic residues, Asp25 and Asp25', has been examined by simulating four possible protonation states, non-protonation, mono-protonation and di-protonation of the active site residues and evaluated by means of properties of structure and dynamics, quantum chemical and free energy calculations. The results lead to a conclusion that the appropriate condition is the monoproteination on Asp25. The MD simulations of a wild-type, G48V, L90M and G48V/L90M of HIV-1 protease with saquinavir and those of a wild-type, V82F, I84V, V82F/I84V of HIV-1 protease with ritonavir have been carried out. It was found that the double mutation of both complexes has the largest impact on the bindings. The conformation changes in double mutant complexes consequently reduce the interaction between drug and protease. The results for both WT and mutants are in a good agreement with the experimental inhibition constant and the resistance fold. Fundamental understanding of the above properties is known to be prerequisite for the structure-based drug screening and design, both for wild and mutant types.

Department.....Chemistry.....Student's Signature.....*Ornjira Anuksakunwong*.....
 Field of Study.....Chemistry.....Advisor's Signature.....*S. Hannongbua*.....
 Academic Year.....2006.....Co-advisor's Signature.....*P. Sompornpistut*.....

ACKNOWLEDGEMENTS

First, all affectionate gratitude is acknowledged to Aruksakunwong family, my parents for their understandings, supports and encouragement throughout.

This thesis was completely finished with the excellent helps from my thesis advisor, Prof. Dr. Supot Hannongbua, who always gives me his guiding, advising and encouraging for all my study in Austria and Thailand. This is also to my co-advisor Assoc. Prof. Dr. Pornthep Sompornpisut. I would like to thank Prof. Dr. Peter Wolschann at University of Vienna for his suggestion and support during my stay in Austria.

I would like also to express my thanks to Assoc. Prof. Dr. Vudhichai Parasuk and Assist. Prof. Dr. Somsak Tonmunphean, who give me the basic computer skill and very useful suggestions, also the basic in computational chemistry.

I am very appreciating to Assoc. Prof. Dr. Sirirat Kokpol, Assoc. Prof. Dr. Wasun Chantratitra, Assist. Prof. Dr. Surapong Pinitglang and Prof. Dr. Kritsana Sagarik for their substantial as thesis committee.

Grateful thanks should be given to Computational Chemistry Unit Cell (CCUC) at Department of Chemistry, Chulalongkorn University and the ZID of the University of Vienna for providing the computation time and all facilities.

Finally, this thesis was completely finished with the funding from the Thailand Research Fund (TRF) and the ÖAD within the Asea-Unet and ZID of the University of Vienna.

CONTENTS

	Pages
ABSTRACT IN THAI.....	1
ABSTRACT IN ENGLISH.....	2
ACKNOWLEDGMENT.....	3
CONTENTS.....	4
LIST OF TABLES.....	8
LIST OF FIGURES.....	10
NOTATIONS.....	19
CHAPTER 1 INTRODUCTION.....	22
1.1 Research rationale.....	22
1.2 Acquired immunodeficiency syndrome (AIDS).....	23
1.3 Human immunodeficiency virus (HIV).....	26
1.3.1 The clinical course of HIV-1 infection.....	27
1.3.1.1 Primary infection.....	27
1.3.1.2 Clinical latency	28
1.3.1.3 The declaration of AIDS.....	28
1.3.2 Structure of HIV.....	28
1.3.3 Replication of HIV.....	31
1.3.3.1 Virus entry.....	31
1.3.3.2 Uncoating the capsid core	31
1.3.3.3 Reverse transcription	32
1.3.3.4 Nuclear import	33
1.3.3.5 Integration	33
1.3.3.6 Transcription and translation	33
1.3.3.7 Production of a new virion	34
1.3.3.8 Virion maturation	34
1.3.3.9 External factors	34
1.3.4 Treatment.....	35

	Pages
1.4 HIV-1 Protease.....	36
1.4.1 Functions.....	36
1.4.2 Structure.....	37
1.4.2.1 The terminal domains.....	38
1.4.2.2 The core domains	38
1.4.2.3 The flap domains	40
1.4.3 Hydrolysis mechanism of HIV-1 PR.....	40
1.5 HIV-1 PR inhibitor.....	43
1.6 HIV-1 Drug resistance	47
1.6.1 Evolution of HIV-1 drug resistance	47
1.6.2 HIV-1 protease drug resistance.....	48
1.7 Saquinavir resistance	52
1.8 Ritonavir resistance	53
1.9 Research objectives.....	56
 CHAPTER 2 THEORY.....	 58
2.1 Potential energy functions.....	58
2.1.1 Intramolecular interactions.....	59
2.1.2 Non-bonded interactions.....	60
2.2 Energy minimization.....	63
2.2.1 Steepest descent method.....	64
2.2.2 Conjugate gradient method.....	65
2.2.3 Newton-Raphson method.....	66
2.3 Molecular dynamics simulation.....	67
2.3.1 Basic theory of molecular dynamics.....	68
2.3.2 Integration algorithms.....	71
2.3.3 The basic steps in MD simulation.....	74
2.3.4 Periodic boundary condition.....	75

	Pages
2.3.5 Potential cut-off for non-bonded interaction.....	76
2.3.6 Treatment of long range interaction.....	76
2.4 The molecular mechanics Poisson-Boltzmann surface area (MM/PBSA)	79
2.4.1 Poisson-Boltzmann (PB) model.....	81
2.4.2 Solvent accessible surface area (SASA).....	83
 CHAPTER 3 MD SIMULATIONS OF DIFFERENT PROTONATION STATES OF G48V HIV-1 PR COMPLEXED WITH SAQUINAVIR IN AQUEOUS SOLUTION.....	 84
3.1 Computational methods.....	86
3.1.1 Atomic charges of saquinavir.....	86
3.1.2 The pK_a calculations of HIV-1 PR.....	87
3.1.3 Preparation of the initial structures.....	89
3.1.4 Molecular dynamics simulations	91
3.1.5 Quantum chemical calculations	92
3.2 Results and discussion.....	93
3.2.1 Reliability of the simulations.....	93
3.2.2 Flexibility of enzyme.....	96
3.2.3 Flexibility of the saquinavir inhibitor.....	98
3.2.4 Enzyme – saquinavir Interaction.....	103
3.3 Conclusion.....	108
 CHAPTER 4 MD SIMULATIONS OF WILD-TYPE AND MUTANTS HIV-1 PR COMPLEXED WITH SAQUINAVIR IN AQUEOUS SOLUTION.....	 109
4.1 Computational methods.....	110
4.1.1 Preparation of the initial structures.....	110
4.1.2 Molecular dynamics simulations	111
4.1.3 Calculations of the Binding free energy.....	112
4.2 Results and discussion.....	113

	Pages
4.2.1 Reliability of the simulations	113
4.2.2 Flexibility and conformational changes of SQV subsites.....	117
4.2.3 Flexibility and conformational changes of HIV-1 PR.....	121
4.2.4 Hydrogen bonding in binding site.....	125
4.2.5 Specific interaction energy at the catalytic site of the HIV-1 PR....	127
4.3 Conclusion.....	132
 CHAPTER 5 MD SIMULATIONS OF WILD-TYPE AND MUTANTS HIV-1 PR COMPLEXED WITH RITONAVIR IN AQUEOUS SOLUTION	133
5.1 Computational methods.....	134
5.1.1 Atomic charges of ritonavir.....	134
5.1.2 Preparation of the initial structures.....	136
5.1.3 Molecular dynamics simulations	137
5.1.4 Calculations binding free energy.....	137
5.2 Results and discussion.....	138
5.2.1 Reliability of the simulations.....	138
5.2.2 Structural features of the HIV-1 protease.....	141
5.2.3 Conformational flexibility of ritonavir	143
5.2.4 Ritonavir-HIV-1 protease binding energy.....	148
5.2.5 Effect of mutation to the binding in the catalytic domain	149
5.3 Conclusion.....	150
 CHAPTER 6 CONCLUSION.....	151
 REFERENCES.....	152
APPENDICES.....	165
VITAE.....	200

LIST OF TABLES

	pages
Table 1.1 Estimates of the global HIV & AIDS epidemic at the end of 2005 (million).....	25
Table 1.2 HIV-gene functions.....	29
Table 1.3 Inhibition of HIV-1 PR.....	47
Table 1.4 HIV-1 resistances to protease inhibitors.....	50
Table 1.5 Comparison of K_i values with association and dissociation rate constants for WT and mutant protease with saquinavir.....	53
Table 3.1 The proposed protonation states of HIV-1 PR from the previous studies.....	85
Table 3.2 RESP charges of saquinavir molecule obtained from single point HF/6-31G* calculation on the HF/6-31G** optimized geometry.....	88
Table 3.3 Number of water molecules, counter ions and total atoms together with the dimension of simulation boxes of the three simulations, Mono25, Mono25' and Dipro.....	90
Table 3.4 Mean values of total, potential and kinetic energy as well as temperature.....	96
Table 3.5 Interaction between saquinavir (SQV) and the catalytic residues, Asp25 and Asp25', of HIV-1 PR for the three protonated states yielded from <i>ab initio</i> calculations (see text for more details).....	106
Table 3.6 Average RMSDs as well as their deviations for the ASP25 and Asp25' of the Mono-25, Mono-25' and Di-Pro systems evaluated after equilibration, 400-1000 ps.....	106
Table 4.1 Number of water molecules, counter ions and total atoms together with the dimension of simulation boxes of the three simulations, WT, G48V, L90M and G48V/L90M.....	111

LIST OF TABLES (cont.)

	pages
Table 4.2 Mean values of total, potential and kinetic energies as well as temperature for the four simulated systems.....	117
Table 4.3 The interaction energy calculated by ONIOM3 and the component of decomposition.....	131
Table 5.1 RESP charges of ritonavir molecule obtained from single point HF/6-31G* calculation on the HF/6-31G** optimized geometry.....	135
Table 5.2 Number of water molecules, counter ions and total atoms together with the dimension of simulation boxes of the four simulations, WT, V82F, I84V and V82F/I84V.....	136
Table 5.3 Mean values of total, potential and kinetic energy as well as temperature.....	141
Table 5.4 Calculated volumes in the binding cavity of the protease complexes.	147
Table 5.5 Averaged energy contributions (kcal/mol) to the binding free energy for wild-type and mutant HIV-1 protease-ritonavir complexes.....	149

LIST OF FIGURES

10

	pages
Figure 1.1	The latest statistics on the world epidemic of AIDS & HIV were published by UNAIDS/WHO in May 2006, and refer to the end of 2005..... 24
Figure 1.2	Estimated global numbers of people living with HIV, 2001-2005.... 25
Figure 1.3	Graph showing HIV virus and CD4 levels over the course of an untreated infection..... 27
Figure 1.4	Structure of HIV..... 29
Figure 1.5	Drawing of the mature HIV virion surrounded by ribbon representations of the structurally characterized viral proteins and protein fragments. The protein structures have been drawn to the same scale..... 30
Figure 1.6	Schematic representation of the replication cycle of HIV..... 32
Figure 1.7	The translational products of the HIV <i>gag-pol</i> gene and the sites at which the gene product is cleaved by the virus-encoded protease. <i>p17</i> denotes capsid protein (CA), <i>p24</i> matrix protein (MA), and <i>p7</i> nucleocapsid (NC); <i>p2</i> , <i>p1</i> , and <i>p6</i> are small proteins with unknown functions. The arrows denote cleavage events catalyzed by the HIV-specific protease..... 37
Figure 1.8	Structure of HIV-1 PR..... 38
Figure 1.9	The 'fireman's grip', a stereotypical rigid network structure involving the Asp-Thr-Gly signature sequence in the retroviral proteases. Amino acids are identified by three-letter codes. The catalytic aspartic acid residue (Asp) is hydrogen-bonded to the backbone NH group of the glycine (Gly) two amino acids further along in the sequence. In addition, the OH groups of threonine (Thr) are hydrogen bonded to two points on the opposite domain or monomer, the backbone NH group (blue) of the threonine and to the carbonyl oxygen (red) of the residue before the catalytic

LIST OF FIGURES (cont.)

11

		pages
	aspartic acid.....	39
Figure 1.10	Ribbon drawings of (a) apo- and (b) inhibited HIV protease showing the relatively open and closed position of the flap (top of the images). These images were produced with MolScript and Raster3D on the PDB files 3HVP and 1AJX.....	40
Figure 1.11	Standard nomenclature $P_1...P_n$, $P'_1...P'_n$ is used to designate amino acid residues of peptide substrates. The corresponding binding sites on the protease are referred to as $S_1...S_n$, $S'_1...S'_n$ subsites.....	41
Figure 1.12	Mechanism of HIV-1 PR One of the proposed reaction mechanisms. 1) Enzyme reaction centre: catalytic aspartates (25 and 125), peptide bond and lytic water molecule. 2) Tetrahedral intermediate following HW2 transfer between water molecule and OD2 of Asp125, nucleophilic attack of the resulting hydroxy anion (OW-HW1) on the peptide carbon (C), and HD2 transfer between Asp25 and the carbonyl oxygen of the peptide bond. 3) An <i>anti-gauche</i> to <i>gauche</i> transition of the C-N bond allowing the nitrogen lone pair to accept HW2 proton of Asp125. 4) C-N bond rupture and hydrogen transfer between one of the hydroxyl groups of the tetrahedral intermediate onto Asp25. Balls represent carbon atoms bonded to other fragments of the protein or the substrate.....	42
Figure 1.13	Schematic representations of two flaps of HIV PR and their hydrogen bonds with the water molecule. The same water molecule makes another two hydrogen bonds with the peptidyl inhibitor or substrate.....	43
Figure 1.14	Clinical inhibitors of HIV-1 PR.....	46

LIST OF FIGURES (cont.)

12

		pages
Figure 1.15	Mutations in the protease gene associated with resistance to protease inhibitor.....	51
Figure 1.16	Computer-generated model of HIV-1 protease-ritonavir complexes of wild-type HIV-1 _{NL4-3} and the V82F/I84V double mutant. (a) Interaction between ritonavir and the protease at the S1' subsite; (b) the same interactions at the S1' binding subsite. I84V decreases the interaction with the C _β group of the benzyl side chain of ritonavir, whereas V82F results in a severe spatial overlap with the phenyl ring of the inhibitor at the P1' site. Note the effects of the double mutant on the van der Waals interactions between the enzyme and the inhibitor.....	55
Figure 2.1	Geometry of a simple chain molecule, illustrating the definition of interatomic distance r_{23} , bend angle θ_{234} , and torsion angle ϕ_{1234} .	59
Figure 2.2	Lennard-Jones pair potential showing the r^{-12} and r^{-6} contributions. Also shown is the WCA (Water Conservation Area) shifted repulsive part of the potential.....	61
Figure 2.3	Empirical potential energy function and their components.....	62
Figure 2.4	The method of steepest descents.....	65
Figure 2.5	The method of conjugate gradient.....	66
Figure 2.6	The cartoon showed the comparison of experiment and molecular simulation.....	68
Figure 2.7	Basic steps on MD simulation study.....	75
Figure 2.8	Periodic boundary conditions. As a particle moves out of the simulation box, an image particle moves in to replace it. In calculating particle interactions within the cutoff range, both real and image neighbors are included.....	76
Figure 2.9	Schematic of the SASA.....	83

LIST OF FIGURES (cont.)

13

		pages
Figure 3.1	Schematic representation of the HIV-1 PR complexed with saquinavir (a) in which molecular structure of Gly48 (b) and Val48 (c) in the G48V and of saquinavir (d) were also displayed.....	86
Figure 3.2	The atom labels of saquinavir.....	87
Figure 3.3	Saquinavir complexed with Mono-25 (a), Mono-25' (b) and Di-pro (c) G48V HIV-1 PR.....	89
Figure 3.4	The complex of G48V HIV-1 PR and saquinavir in solvate box.	90
Figure 3.5	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 400-1000 ps for Mono-25 system.....	93
Figure 3.6	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 400-1000 ps for Mono-25' system.....	94
Figure 3.7	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 400-1000 ps for Di-pro system.....	94
Figure 3.8	Global RMSD values of HIV-1 PR (black line) and saquinavir (red line) with respect to the starting structure over 400-1000 ps for Mono-25 (a), Mono-25' (b) and Di-pro (c) systems.....	95
Figure 3.9	RMSD as a function of simulation time of the two chains, A (black line) and B (gray line), of the G48V HIV-1 PR mutant type complexed with saquinavir in the Mono-25 (a), Mono-25' (b) and Di-pro (c) states.....	96
Figure 3.10	RMSD per residues of the Mono-25 (black), Mono-25' (red) and Di-pro (blue) states where the highly flexible residues were labeled.....	97
Figure 3.11	Definition of five subsites of saquinavir structure that labeled by circle and torsional angle of each subsite was defined by <i>tor</i>	98

LIST OF FIGURES (cont.)

14

	pages	
Figure 3.12	Root-mean-square displacement (RMSD) of each subsite of saquinavir (defined in Figure 3.11) for the three simulated systems, (a) Mono-25, (b) Mono-25' and (c) Di-pro.....	99
Figure 3.13	Changes of torsional angles of four subsites of saquinavir inhibitor, (a) P ₁ subsite; (b) P ₂ subsite; (c) P ₂ ' subsite; (d) P ₃ subsite, (defined in Figure 3.11) for the three simulated systems where solid, dashed and dotted lines in the torsional angle plots represent the Mono-25, Mono-25' and Di-pro systems, respectively. The two conformations corresponding to torsional angles of 90.3° and -75.5° of P ₂ are also given in an inset.....	100
Figure 3.14	Definition of torsional angles <i>tor1</i> - <i>tor4</i> (a) of the P ₂ ' subsite of saquinavir inhibitor and their changes as a function of simulation time; (b) Mono-25; (c) Mono-25'; (d) Di-pro.....	102
Figure 3.15	Detected hydrogen bonds (dash line) in the HIV-1 PR enzyme-saquinavir complex in (a) Mono-25, (b) Mono-25' and (c) Di-pro states, where the average hydrogen bond distances (H---A distance) in Å are numbered.....	103
Figure 3.16	Indirect binding between HIV -1 PR and saquinavir via hydrogen bonds (dash line) through water molecules in the (a) Mono25 and (b) Mono-25' states where the average distance (in Å) are numbered.....	105
Figure 4.1	Schematic representation of wild-type HIV-1 PR complexed with saquinavir (a) and comparison of wild-type (Gly48 and Leu90) and mutant type residues (Val48 and Met90).....	100
Figure 4.2	The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ps for WT-SQV system.....	114

LIST OF FIGURES (cont.)

15

		pages
Figure 4.3	The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ps for G48V-SQV system.....	114
Figure 4.4	The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ps for L90M-SQV system.....	115
Figure 4.5	The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ps for G48V/L90M-SQV system.....	115
Figure 4.6	Global RMSD values of backbone atoms of HIV-1 PR (black line) and all atoms saquinavir (red line) with respect to the starting structure over 1-2 ns for (a) WT-SQV, (b) G84V, (c) L90M and (d) G48V/L90M.....	116
Figure 4.7	According to a conventional classification of the protease subsites, the binding pockets are designated by the inhibitor side chains, P_1 , P_2 , P_3 , P_1' , and P_2' . Torsional angles of each subsites were defined by <i>tor1-tor7</i>	118
Figure 4.8	RMSDs of each SQV subsite with respect to the starting structure versus the simulation time for the wild-type (a), G48V (b), L90M (c) and G48V/L90M (d) mutants, where black, red, green, blue and sky-blue lines represent the RMSDs for the P_3 , P_2 , P_1 , P_2' and P_1' subsites, respectively.....	118
Figure 4.9	Fluctuation of <i>tor1-tor7</i> corresponding to the dihedral angles of the inhibitor side chains, for the P_1 , P_2 , P_3 and P_1' , respectively (see Figure 4.7 for definition).....	121
Figure 4.10	RMSD of Asp25 (black) and Asp25' (red) for WT-SQV (a), G48V-SQV (b), L90M-SQV (c) and G48V/L90M-SQV (d).....	122

LIST OF FIGURES (cont.)

16

		pages
Figure 4.11	The CA-CB-CG-OD1 torsional angle of Asp25 of WT-SQV (a) and G48V/L90M-SQV (b) complexes.....	122
Figure 4.12	RMSD of residues 48 (black), 48' (red), 90 (green) and 90' (blue) for WT-SQV (a), G48V-SQV (b), L90M-SQV (c) and (d) G48V/L90M-SQV	123
Figure 4.13	HIV-1 PR flap structures of the WT-SQV (black), G48V-SQV (red), L90M-SQV (green) and G48V/L90M-SQV (blue). Some selected residues and saquinavir are presented in stick mode. For simplify, hydrogen atoms are not shown.....	124
Figure 4.14	The complex structures representative for WT-SQV, G48V-SQV, L90M-SQV and G48V/L90M-SQV systems with the depicted hydrogen bonds distances $d1-d6$	126
Figure 4.15	The %occupied of hydrogen bonding ($d1-d6$) through the simulations.....	127
Figure 4.16	Figure 4.16 The schematic representation for three-layered ONIOM calculations. The inner layer (A) are SQV, Asp25(25'), Gly48(48') and Leu90(90') as shown in Figure 4.17. The mediate layer (B) is the residues within 5 Å from SQV. The outer layer (C) presents entire enzyme.....	128
Figure 4.17	Schematic structure represents the quantum model for calculations the interaction energy of saquinavir with the key amino acid residues.....	129
Figure 5.1	Schematic representation of wild-type HIV-1 protease complexed with ritonavir (a). In the catalytic residues Asp25/Asp25' and the mutated residues V82/V82' and I84/I84' are shown by stick and ritonavir is shown by the ball and stick.....	133
Figure 5.2	The atom labels of ritonavir.....	134

LIST OF FIGURES (cont.)

17

		pages
Figure 5.3	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 1-2 ns for WT-RTV system.....	138
Figure 5.4	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 1-2 ns for V82F-RTV system.....	139
Figure 5.5	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 1-2 ns for I84V-RTV system.....	139
Figure 5.6	The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 1-2 ns for V82F/I84V-RTV system.....	140
Figure 5.7	Plot of RMSDs versus the time for the wild-type, V82F, I84V and V82F/I84V complexes. The obtained RMSD was computed using the structure at $t = 0$ as a reference.....	140
Figure 5.8	The C α -trace of the wild-type (black), V82F (red), I84V (green) and V82F/I84V (blue) mutants. Asp25, Asp25', Ile50 and Ile50' are displayed in the ball and stick. Dashed lines demonstrate the measured distances of C α -Ile50/C β -Asp25 and of C α -Ile50'/C β -Asp25'.....	141
Figure 5.9	Distributions of the C α -Ile50/C β -Asp25 (a) and of the C α -Ile50'/C β -Asp25' (b) distances (details see text and labels in Figure 5.8) sampled during 1-2 ns.....	142
Figure 5.10	(a) The trajectory of Ile50' torsion defined by the C β -C γ bond (arrow in (b)). (b) Ile50', ritonavir and residues 81-84 of the wild-type HIV-1 PR.....	143
Figure 5.11	Definition of five subsites of ritonavir structure that label by P ₁ , P ₂ , P ₃ , P ₁ ' and P ₂ '.....	144

LIST OF FIGURES (cont.)

18

	pages
<p>Figure 5.12 RMSDs of each ritonavir subsite with respect to the starting structure versus the simulation time for the wild-type (a), V82F (b), I84V (c) and V82F/I84V (d) mutants. P₃, P₂, P₁, P₁' and P₂' subsites are given in black, red, green, blue and sky-blue lines, respectively.....</p>	144
<p>Figure 5.13 Structure comparison of residues 82, 84 and ritonavir in HIV-1 protease: (a) the MD structures of wild-type (blue); V82F (red); I84V (green) and V82F/I84V (black), (b) and (c) the x-ray structures of the V82F (1BV7) and the V82F/I84V mutants (1BWA), respectively. Arrows show different orientation of the phenyl ring of the P₁' and of Phe82 between the single and double mutants.....</p>	145
<p>Figure 5.14 The HF/6-31G potential energy versus RTV conformations defined by the C1-C2-C3-C4 torsion of the P₁' sidechain.....</p>	148
<p>Figure 5.15 Hydrogen bonding (<i>d1-d6</i>) between HIV-1 PR and ritonavir and the percent occurrence of the corresponding hydrogen bonds calculated from the MD trajectories.....</p>	150

NOTATIONS

CHAPTER 1		Section
HIV-1	Human immunodeficiency virus type 1	1.1
AIDS	Acquired immunodeficiency syndrome	1.1
RT	Reverse transcriptase	1.1
PR	Protease	1.1
IN	Integrase	1.1
FDA	The Food and Drug Administration	1.1
WT	Wild-type	1.1
MD	Molecular dynamic	1.1
PCP	<i>Pneumocystis pneumonia</i>	1.2
KS	Kaposi's sarcoma	1.2
PGL	Persistent generalised lymphadenopathy	1.2
MMWR	Morbidity and Mortality Weekly Report	1.2
CDC	The United States Centers for Disease Control and Prevention	1.2
UNAIDS	The Jointed United Nations Programme on HIV/AIDS	1.2
WHO	World Health Organization	1.2
HTLV-III	Human T-cell Leukemia Virus Type 3	1.2
ARV	AIDS related virus	1.2
DNA	Deoxyribonucleic acid	1.3
RNA	Ribonucleic acid	1.3
SIV	Simian immunodeficiency virus	1.3
HIV-2	Human immunodeficiency virus type 2	1.3
CD4	A large glycoprotein molecule on the surface of T lymphocyte	1.3.1
FDC	Follicular dendritic cells	1.3.1
hCyp-18	Cyclophilin A	1.3.3.2
mRNA	Messenger RNA	1.3.3.6
tRNA	Transfer RNA	1.3.3.7
GBV-C	Hepatitis G virus	1.3.3.9

NOTATIONS

HAART	Highly Active Anti-Retroviral Therapy	1.3.4
CA	Capsid protein	1.4.1
MA	Matrix protein	1.4.1
NC	Nucleocapsid	1.4.1
ZDV	Zidovudine	1.6
G48V	Valine-for-glycine substitution at position 48	1.7
L90M	Methionine-for-leucine substitution at residue 90	1.7
K_i	Inhibition constant	1.7
V82F	Phenylalanine-for-valine substitution at position 82	1.8
I84V	Valine-for-isoleucine substitution at position 84	1.8
CHAPTER 2		
WCA	Water Conservation Area	2.1.2
MD	Molecular dynamics	2.3
NVE	Microcanonical ensemble	2.3
NVT	Canonical Ensemble	2.3
NPT	Isobaric-Isothermal Ensemble	2.3
MVT	Grand canonical Ensemble	2.3
BPTI	bovine pancreatic trypsin inhibitor	2.3
PME	Particle Mesh Ewald	2.3.6
FFT	Fast Fourier transform	2.3.6
MM	Molecular mechanics	2.4
PB	Poisson-Boltzmann	2.4
SA	Surface area	2.4
GB	generalized Born	2.4
NMR	Nuclear Magnetic Resonance	2.4.1

NOTATIONS

CHAPTER 3

PDB	Protein data bank	3.1.1
UHBD	University of Houston Brownian Dynamics program	3.1.2
Mono-25	Monoprotonated state at Asp25	3.1.3
Mono-25'	Monoprotonated state at Asp25'	3.1.3
Di-pro	Diprotonated state at Asp25 and Asp25'	3.1.3
SQV	Saquinavir	3.1.5
CPX	Complex	3.1.5
SASA	Solvent-accessible surface area	3.2.1
RMSD	Root mean square deviation	3.2.1
DFT	Density functional theory	3.2.4
B3LYP	Scheme for hybrid Hartree-Fock/density functional theory	3.2.4

CHAPTER 4

WT	Wild-type HIV-1 protease	4.1.1
NMA	Normal mode analysis	4.1.3
ONIOM	Our own N-layered integrated molecular orbital + molecular Mechanics	4.2.5
PM3	Parametrisation 3 of modified neglect of diatomic overlap	4.2.5
UFF	Universal force field	4.2.5

CHAPTER 5

RTV	Ritonavir	5.1.1
-----	-----------	-------

CHAPTER 1

INTRODUCTION

1.1 Research rationale

The human immunodeficiency virus type 1 (HIV-1) is the causative agent in acquired immunodeficiency syndrome (AIDS). The spread of the AIDS has increasingly become one of the biggest social, economic and health challenges in the world. This disease has killed more than 25 million people since it was first recognized in 1981 in U.S.(1). There are three essential enzymes involved in the replication cycle of this virus which are reverse transcriptase (RT), protease (PR) and integrase (IN). They were, then, become important targets for the drug development. Although effective drugs have been developed against HIV-1 PR and HIV-1 RT, however, it was reported that inhibitors at the first target are more potent (2). Therefore, HIV-1 PR is an attractive target for antiviral therapy.

Currently, a large number of HIV-1 PR inhibitors, both pseudopeptidic and non-peptidic in nature, have been designed, many of which are extensively used in anti-AIDS chemotherapy. However, a major limiting factor in the treatment of the HIV infection with protease inhibitor is the emergence of viral strains that show resistance to protease inhibitors (2). It has been estimated that one out of every seven new infections in America is caused by viral strains resistant to at least one type of inhibitor (2). The current medical strategy for delaying the incidence of resistance is to use a combination of potent antiviral agents acting on different targets. However, the PR inhibitors used in combination therapy were designed to act against the wild-type enzyme and are less effective against the mutants. Hence, the new inhibitors which are active also against the mutants responsible for resistance were developed. Recently, one of FDA (The Food and Drug Administration) approved PR inhibitors, darunavir, actives against PR inhibitor-resistant HIV-1 variants (3). Moreover, brexanavir which is new potent HIV-1 PR against wild type (WT) virus and multiple PR inhibitor-resistant isolates is in trial phase (3).

The future success of development those inhibitors depends on understanding the molecular basis of resistance. So far, drug discovery groups have used the

structural characteristics of only WT PR to design inhibitors. Detailed structural information of mutant HIV-1 PR inhibitor complexes has played a key role in structure-based drug design of more potent inhibitors that against both of WT and mutant PR.

To investigate the molecular basic of drug resistance, the molecular dynamics (MD) simulation is a suitable and powerful tool. This method can provide structural and dynamics properties of mutant complexes on the atomic level. Thus, MD simulations of WT and mutant PR complexed with inhibitor were carried out in this thesis.

1.2 Acquired immunodeficiency syndrome (AIDS)

In 1981, doctors in the United States began to notice a series of unusual infections in gay men in Los Angeles, San Francisco, New York, and other big cities. These infections had previously been extremely rare, except amongst people whose immune systems had been seriously weakened in some way. The most life-threatening seemed to be *Pneumocystis pneumonia* (PCP) and Kaposi's sarcoma (KS). But a whole range of other severe protozoa, fungal, bacterial and viral infections and tumors had also been detected amongst gay men in these cities. (1)

In addition, doctors had begun to find persistently swollen lymph glands (persistent generalised lymphadenopathy (PGL)) in otherwise perfectly healthy gay men. The majority of these men stayed healthy for months or years after diagnosis. But some went on quickly to develop the infections mentioned above. (1)

These problems were first reported in the June 5th 1981 edition of Morbidity and Mortality Weekly Report (MMWR), the United States Centers for Disease Control and Prevention (CDC) bulletin for doctors. The report, from doctors in Los Angeles, came about as a result of a sharp increase in requests for a drug used to treat PCP, which was only available through the CDC at that time. Staff spotted a pattern, and the MMWR report subsequently brought the syndrome to the attention of doctors in other cities of the United States. (1)

In 1982, the CDC in the United States acknowledged that there were an epidemic, and formally defined the 'acquired immune deficiency syndrome' (AIDS). The number of people who developed AIDS seemed to be doubling every six months in the United States and AIDS was soon found in every part of the country. (4)

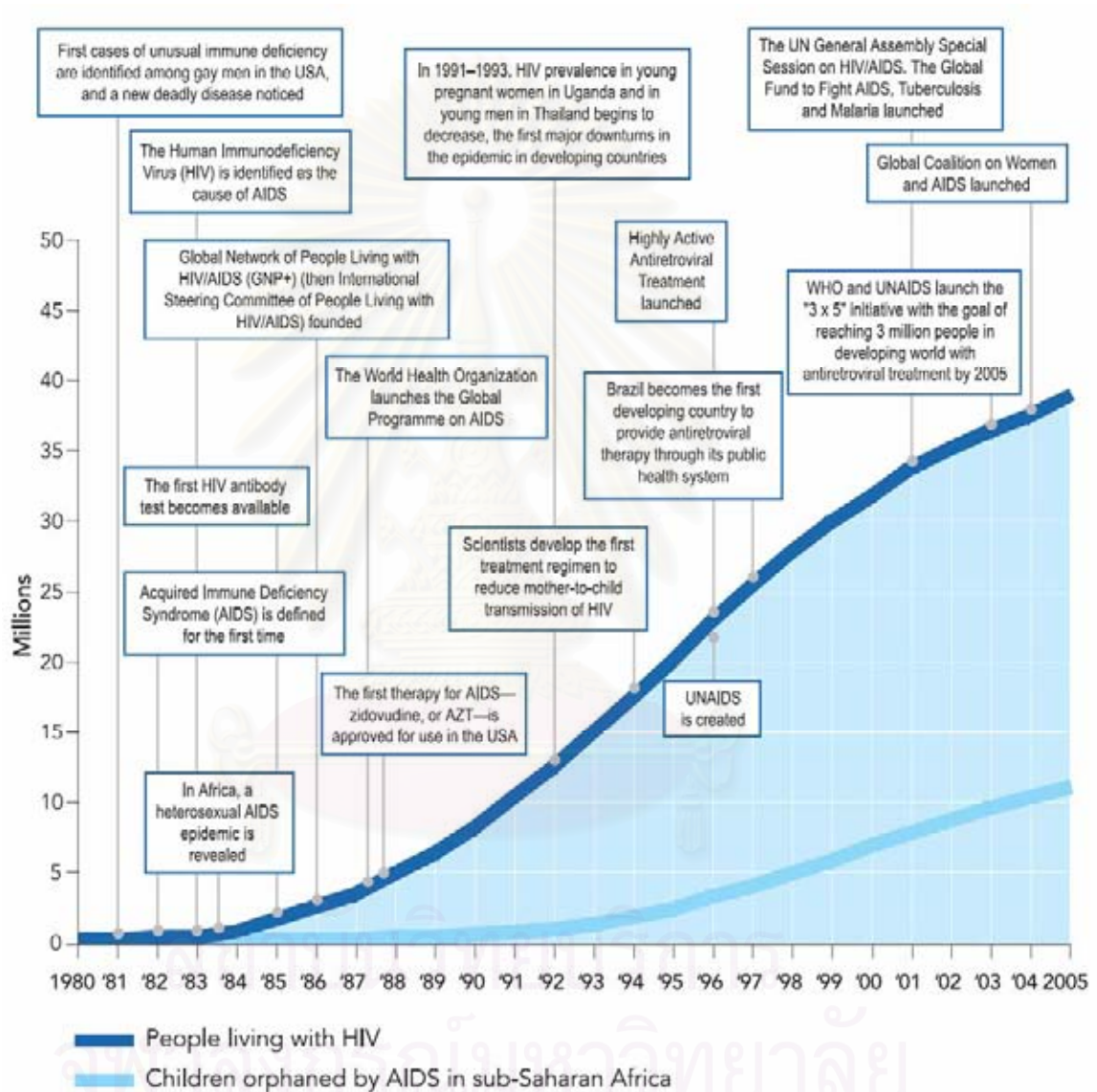


Figure 1.1 The latest statistics on the world epidemic of AIDS & HIV were published by UNAIDS/WHO in May 2006, and refer to the end of 2005. (5)

The spread of the AIDS has increasingly become one of the biggest social, economic and health challenges in the world. It is great concern owing largely to the failure of chemotherapies and a global emergency claiming over 8,000 lives every day.

In fact, 5 people die of AIDS every minute. The number of AIDS deaths continues to increase every year (see Table 1.1).

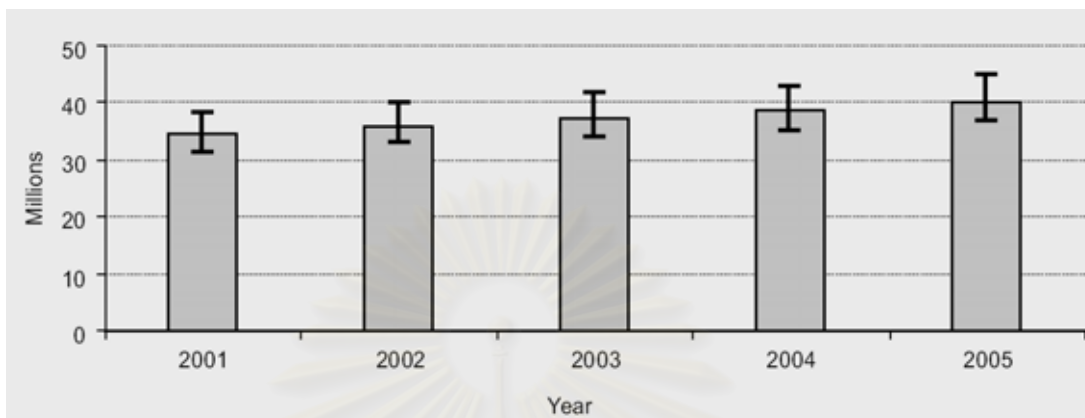


Figure 1.2 Estimated global numbers of people living with HIV, 2001-2005. (6)

From the UNAIDS/WHO report update in the end of 2005, it was shown that AIDS has killed more than 25 million people since it was first recognized in 1981 despite recent, improved access to antiretroviral treatment and care in many regions of the world, the AIDS epidemic claimed ~2.8–3.6 million lives in 2005; more than half a million (570,000) were children. The total number of people living with the human immunodeficiency virus (HIV) is ~36.7–45.3 million. Close to 5 million people were newly infected with the virus in 2005. (5)

Table 1.1 Estimates of the global HIV & AIDS epidemic at the end of 2005 (million). (7)

	Estimate	Range
People living with HIV/AIDS in 2005	38.6	33.4-46.0
- Adults	36.3	31.4-43.4
- Women	17.3	14.8-20.6
- Children	2.3	1.7-3.5
- Newly infected	4.1	3.4-6.2
- Deaths	2.8	2.4-3.3

1.3 Human immunodeficiency virus (HIV)

Beginning with the isolated of a novel retrovirus in 1983 which was later associated with AIDS, a clearer picture of the disease began to emerge. This virus which has been known as HTLV-III (Human T-cell Leukemia Virus Type 3) and ARV (AIDS related virus) is now known as human immunodeficiency virus, HIV. (8)

HIV is a member of the genus lentivirus, part of the family of retroviridae. Lentiviruses have many common morphologies and biological properties. Many species are infected by lentiviruses, which are characteristically responsible for long duration illnesses associated with a long period of incubation. Lentiviruses are transmitted as single-stranded negatively-sensed enveloped RNA viruses. Upon infection of the target-cell, the viral RNA genome is converted to double-stranded DNA by a virally encoded reverse transcriptase which is present in the virus particle. This viral DNA is then integrated into the cellular DNA for replication using cellular machinery. Once the virus enters the cell, two pathways are possible: either the virus becomes latent and the infected cell continues to function or the virus becomes active, replicates and a large number of virus particles are liberated which can infect other cells.(9)

Two distinct types of HIV have been identified, HIV-1 and HIV-2. HIV-1 has been further divided into three virus groups: the predominant M group, which is responsible for most of the epidemic, and N and O groups. HIV-1 is the more virulent and easily transmitted, and is the source of the majority of HIV infections throughout the world. HIV-2 is largely confined to west Africa (10). Both species originated in west and central Africa, jumping from primates to humans in a process known as zoonosis. HIV-1 has evolved from a simian immunodeficiency virus (SIV) found in the chimpanzee subspecies, *Pan troglodyte troglodyte* (11). HIV-2 crossed species from a different strain of SIV, found in sooty mangabey monkeys in Guinea-Bissau (10).

1.3.1 The clinical course of HIV-1 infection

The primary target of HIV seems to be CD4 T-lymphocytes which are part of the machinery of our immune system. Infection with HIV-1 is associated with a progressive loss of CD4 T-cells. This rate of loss can be measured and is used to determine the stage of infection. The loss of CD4 T-cells is linked with an increase in viral load. The clinical course of HIV-infection generally includes three stages: primary infection, clinical latency and AIDS (Figure 1.3). HIV plasma levels during all stages of infection range from just 50 to 11 million virions per ml (12).

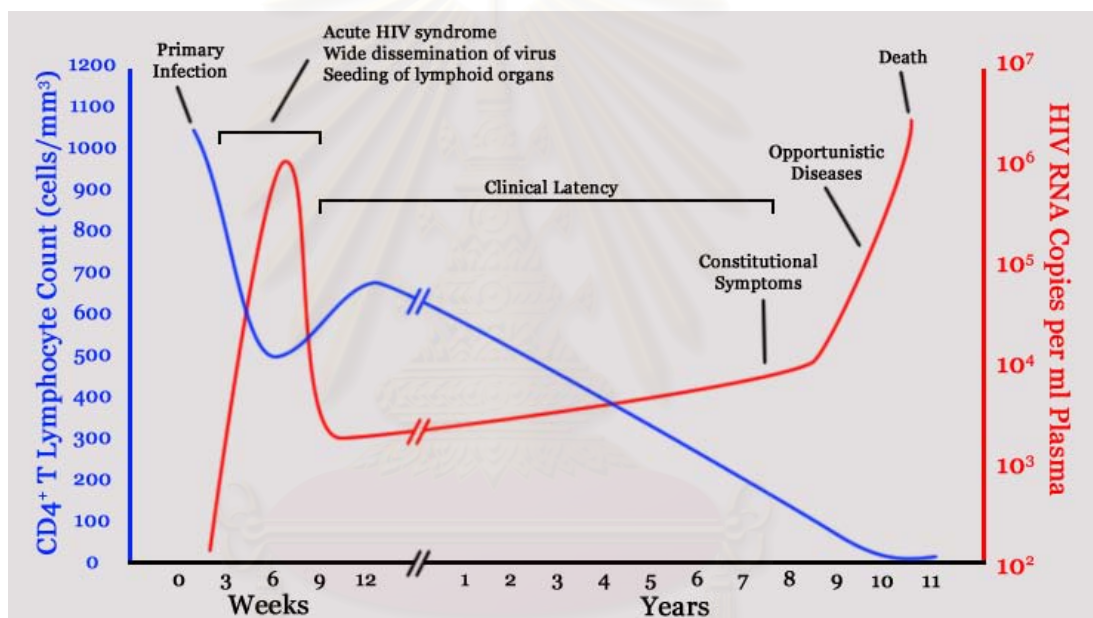


Figure 1.3 Graph showing HIV virus and CD4 levels over the course of an untreated infection (9)

1.3.1.1 Primary infection

Primary or acute infection is a period of rapid viral replication that immediately follows the individual's exposure to HIV. During primary HIV infection, most individuals (80 to 90 %) develop an acute syndrome characterized by flu-like symptoms of fever, malaise, lymphadenopathy, pharyngitis, headache, myalgia, and sometimes a rash (13). Within an average of three weeks after transmission of HIV-1, a broad HIV-1 specific immune response occurs that includes seroconversion. Because of the

nonspecific nature of these illnesses, it is often not recognized as a sign of HIV infection. Even if patients go to their doctors or a hospital, they will often be misdiagnosed as having one of the more common infectious diseases with the same symptoms. Since not all patients develop it, and since the same symptoms can be caused by many other common diseases, it cannot be used as an indicator of HIV infection. However, recognizing the syndrome is important because the patient is much more infectious during this period

1.3.1.2 Clinical latency

As a result of the strong immune defense, the number of viral particles in the blood stream declines and the patient enters clinical latency (Figure 1.3). Clinical latency is variable in length and can vary between two weeks and 20 years. During this phase HIV is active within lymphoid organs where large amounts of virus become trapped in the follicular dendritic cells (FDC) network early in HIV infection. The surrounding tissues that are rich in CD4 T-cells also become infected, and viral particles accumulate both in infected cells and as free virus. Individuals who have entered into this phase are still infectious.

1.3.1.3 The declaration of AIDS

AIDS is the most severe manifestation of infection with HIV. Acute HIV infection progresses over time to clinical latent HIV infection and then to early symptomatic HIV infection and later, to AIDS, which is identified on the basis of certain infections.

1.3.2 Structure of HIV

HIV is different in structure from previously described retroviruses. It is an essential spherical particle with around 120 nm in diameter (120 billionths of a meter; around 60 times smaller than a red blood cell). HIV-1 is composed of two copies of single-stranded RNA enclosed by a conical capsid, which is in turn surrounded by a

plasma membrane that is formed from part of the host-cell membrane. Other enzymes contained within the virion particle includes RT, IN and PR. (9)

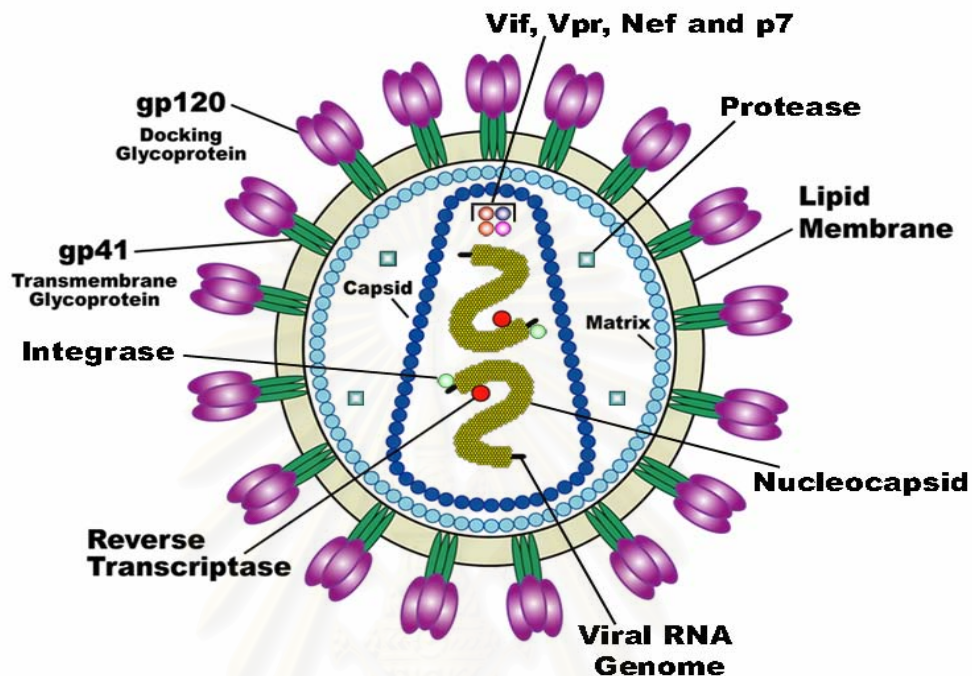


Figure 1.4 Structure of HIV. (9)

Table 1.2 HIV-gene functions. (14)

Genes	Functions
Structural	
<i>gag</i>	matrix, virion maturation and stability, capsid, virus particle maturation and release
<i>pol</i>	protease, reverse transcriptase, integrase
<i>env</i>	external envelope, receptor binding, virion infectivity
Regulatory	
<i>tat</i>	transcription transactivator of gene expression
<i>rev</i>	regulator of protein expression
<i>tev/tmv</i>	undefined
<i>nef</i>	negative factor, virus propagation
<i>vpr</i>	early regulatory protein
Accessory	
<i>vif</i>	cell-free virus transmission, <i>env</i> processing
<i>vpu</i>	virus maturation/release

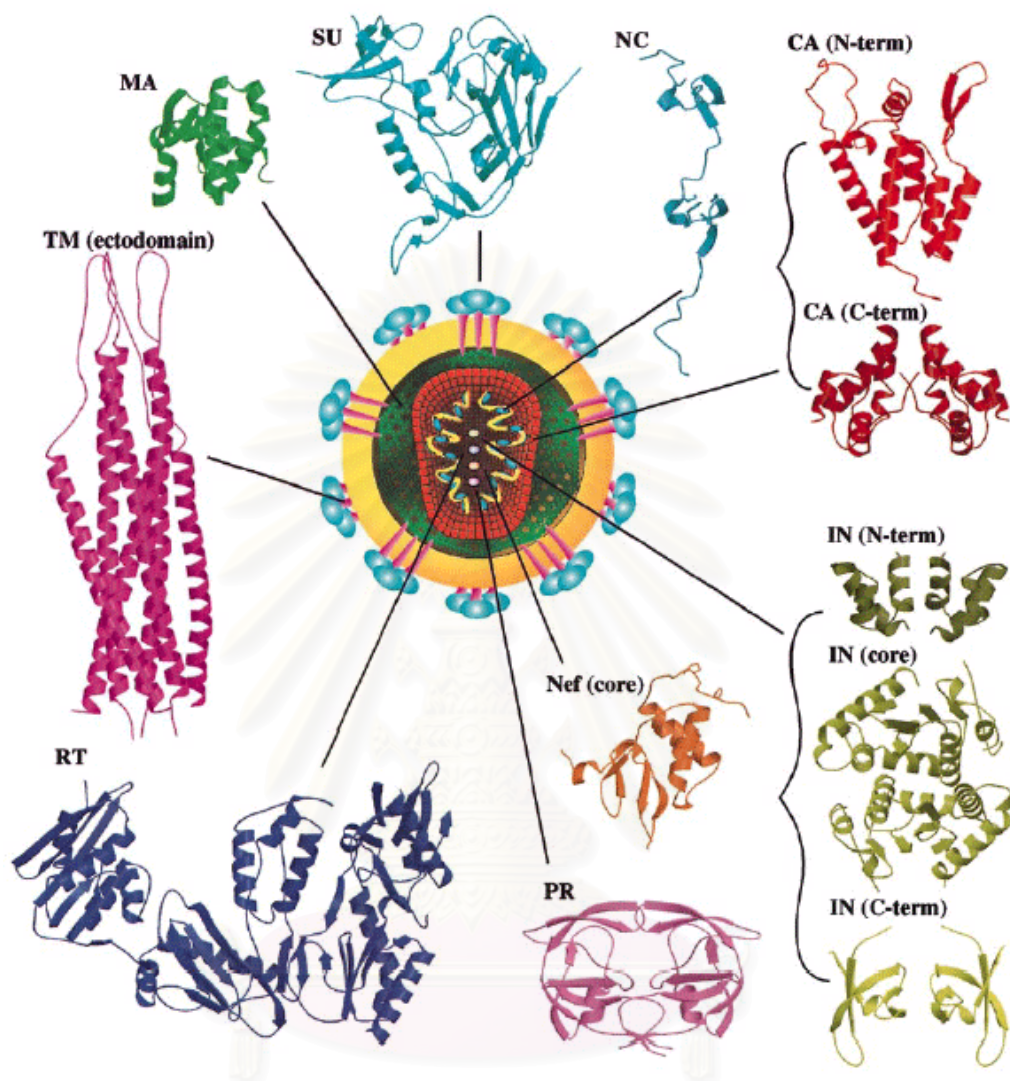


Figure 1.5 Drawing of the mature HIV virion surrounded by ribbon representations of the structurally characterized viral proteins and protein fragments. The protein structures have been displayed to the same scale. (15)

HIV has several major genes coding for structural proteins that are found in all retroviruses, and several nonstructural (accessory) genes that are unique to HIV. The *gag* gene provides the physical infrastructure of the virus; *pol* provides the basic enzymes by which retroviruses reproduce; the *env* gene supplies the proteins essential for viral attachment and entry into a target cell. The accessory proteins *tat*, *rev*, *nef*, *vif*, *vpr*, and *vpu* enhance virus production. Although called accessory proteins, *tat* and *rev* are essential for virus replication. In some strains of HIV, a mutation causes the

production of an alternate accessory protein, *tev*, from the fusion of *tat*, *rev*, and *env* (Table 1.2). The 3-dimensional structures of structural proteins of HIV were displayed in Figure 1.5.

The *gp120* and *gp41* proteins, both encoded by the *env* gene, enable the virus to attach to and fuse with target cells to initiate the infectious cycle. Both, especially *gp120*, have been considered as targets of future treatments or vaccines against HIV.

1.3.3 Replication of HIV (8)

A schematic representation of the replication cycle of HIV appears in Figure 1.6, a myriad of cellular machinery is used to augment HIV's special tool. With over 175,000 articles indexed for HIV and/or AIDS on Medline, it is certainly one of the most thoroughly studied systems today. As such, many details of the biology of HIV will be omitted for the sake of brevity.

1.3.3.1 Virus entry

The entry of HIV into a host cell may be divided into 3 distinct steps: attachment, coreceptor interaction and fusion. Attachment of HIV-1 to the host cell surface is mediated through *gp120* on the virion surface binding to a CD4 antigen on the host cell. Endogenous CD4 is present on the surface of many lymphocytes, which make up a critical part of the body's immune system. This *gp120-CD4* complex interacts with a coreceptor on the cell surface, typically chemokine *CXCR4* or *CCR5*. Transmembrane glycoprotein *gp140* mediates membrane fusion to complete virus entry into the host cell.

1.3.3.2 Uncoating the capsid core

Following fusion, the *p24* encased capsid core is disrupted to dump the contents into the cytoplasm of the host cell. It seems that this is accomplished

with the help of a cytoplasmic peptidyl-prolyl cis-trans isomerase called cyclophilin A (hCyp-18) which had been incorporated into the virion.

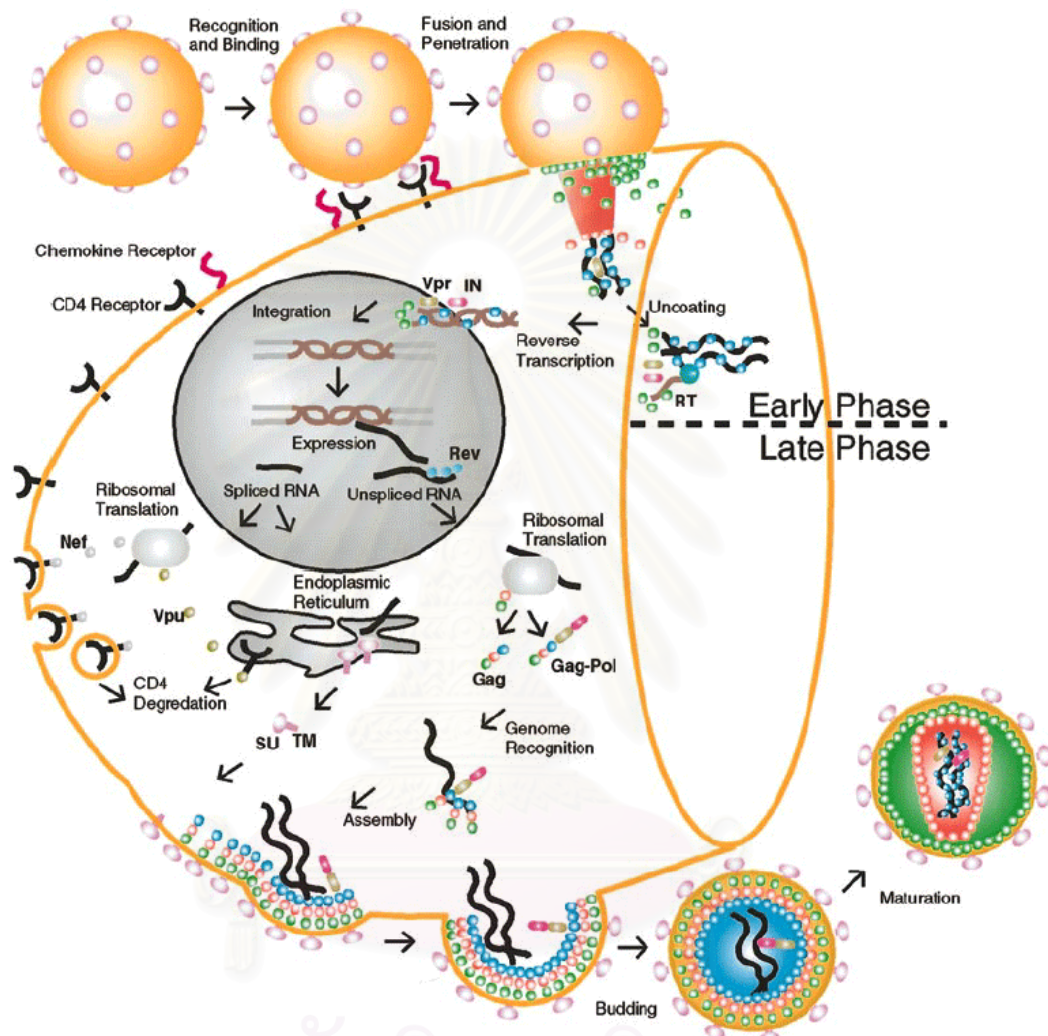


Figure 1.6 Schematic representation of the replication cycle of HIV. (15)

1.3.3.3 Reverse transcription

Successful entry of the contents of the viral capsid core is followed by the reverse transcription of complementary DNA strand from the viral RNA template by the viral enzyme reverse transcriptase (RT) in a complex with other viral proteins. RT then degrades the RNA and produces the double-stranded viral DNA. RT is

highly error-prone since it is unable to catalyze the proof-reading which a normal DNA polymerase performs.

1.3.3.4 Nuclear import

The newly synthesized viral DNA is then imported into the nucleus of the host cell. A short triple-helical region, made from a flap of about 99 bases, synthesized during an interruption of reverse transcription seems to be necessary for this event. Other viral proteins, such as *vpr*, are also thought to be involved but the system is complex.

1.3.3.5 Integration

The property placed viral DNA is processed and transferred to the host genome by the viral enzyme IN as the central agent. Once the viral DNA has been inserted, infection in that cell is for all intents and purpose permanent since finding a way to selectively remove that little patch of DNA from the host genome would seem to be a monumental task.

1.3.3.6 Transcription and translation

Once the viral DNA has been inserted into the host cell's genome, HIV may persist in a latent, proviral state for many years in unstimulated T cell. Activation of the host cells results in transcription of the viral DNA by the host cell machinery into messenger RNA (mRNA). Early genes to be activated express auxiliary proteins *tat*, *nef*, *rev*, and few others, *tat* act as a strong promoter of viral transcription, *nef* acts as a weak negative regulator and *rev* promotes switching to the expression involve a variety of interactions with the cellular proteins. The auxiliary proteins have also been implicated in other role such as the down-regulation and degradation of cell-surface CD4 in infected cell by *vpu* and *nef*, respectively, to promote the release of new viruses. The second phase of transcription produces the unspliced mRNA for the precursor proteins *gag* ($Pr55^{gag}$) and *gag-pol* ($Pr180^{gag}$), which is the results of a translation frame shifting event, in an approximately 20:1 ratio. The unspliced RNA is

also intended to be used as the genome of the next generation of the virus. *Gag* and *gag-pol* are transported out of the cell nucleus by a poorly understood mechanism and anchor to the wall through linkage with myristate their N-termini. The precursor for the envelope glycoproteins *gp120* and *gp140* are treated like cellular membrane proteins: synthesized, processed (glycosylated and cleaved) and transported to the cell surface in the endoplasmic reticulum and the golgi apparatus though some interaction with the *gag* precursor has been implicated.

1.3.3.7 Production of a new virion

Assembly of a new virus particle begins at the cell surface with the clustering of roughly 2,000 *gag* protein, processed envelope proteins *gp120* and *gp41*, two copies of the viral genomic RNA, some viral tRNA and some other components like the *gag* protein mediates the budding process. Some other components like cyclophilin A which will be used after infection of the next cell. It appears that the *gag* protein mediates the budding process. Some of the details involving cellular and viral components have recently been elucidated. Release is assisted by viral *vpu* in an incompletely understood process.

1.3.3.8 Virion maturation

The immature virion is a not-quite spherical blob with an outer membrane derived from the host cell but including the viral coat proteins *gp120* and *gp41*. The inside has roughly alignment of the *gag* protein surrounding the RNA though older work point to a more ordered structure. HIV protease (PR) is required at this stage to cleave the *gag* and *gag-pol* polyproteins into constituent structural (*p17*, *p24*, *p7*, *p6*, *p2*, *p1*) and functional (PR, RT, IN) proteins.

1.3.3.9 External factors

The roles of outside factors have not been outlined here but they certainly should not be discounted. For example, some narcotics have been shown to

act at least as cofactors in AIDS. On the other hand, coinfection with hepatitis G virus (GBV-C) may actually improve chance for survival of AIDS.

1.3.4 Treatment

HIV infection is a chronic infectious disease that can be treated, but not yet cured. There are effective means of preventing complications and delaying, but not preventing, progression to AIDS. At the present time, not all persons infected with HIV have progressed to AIDS, but it is generally believed that the majority will. People with HIV infection need to receive education about the disease and treatment so that they can be active partners in decision making with their health care provider.

A combination of several antiretroviral agents, termed Highly Active Anti-Retroviral Therapy HAART, has been highly effective in reducing the number of HIV particles in the blood stream (as measured by a blood test called the viral load). This can improve T-cell counts. This is not a cure for HIV, and people on HAART with suppressed levels of HIV can still transmit the virus to others through sex or sharing of needles. There is good evidence that if the levels of HIV remain suppressed and the CD4 count remains greater than 200, then the quality and length of life can be significantly improved and prolonged. Improved antiretroviral inhibitors against proteins such as Reverse transcriptase, Integrase and *tat* are being researched and developed. One of the most promising new therapies is a new class of drugs called fusion or entry inhibitors. (16)

As yet, no vaccine has been developed to prevent HIV infection or disease in people who are not yet infected with HIV, but the potential worldwide public health benefits of such a preventive vaccine are vast. Researchers in many countries are seeking to produce such a vaccine, including through the International AIDS vaccine initiative.

1.4 HIV-1 Protease

A protease, PR, is an essential enzyme for the viral replication that cleaves proteins to their component peptides. The HIV-1 Protease hydrolyzes viral polyprotein into functional protein products that are essential for viral assembly and subsequent activity. In 1989, the protease enzyme of HIV-1 was crystallized and its three-dimensional structure was determined by Navia *et al.* from Merck laboratories (17). Subsequently, a more accurate structure was reported by Kent and coworker (18). Since then, PR structure has been extensively characterized in terms of function, substrate specificity and inhibitor binding. This knowledge makes HIV-1 protease is a very important target for drug design strategies to combat AIDS monotherapy.

1.4.1 Functions

HIV protease processes *gag* and *gag-pol* (*p160*) polyprotein, which must be cleaved before nascent viral particles (virions) can mature, products into functional core proteins and viral enzymes (19,20). Cleavage of these polyprotein by protease at different cleavage sites produces three large proteins (*p24*, *p17* and *p7*) that contribute to the structure of the virion and to RNA packaging, and three smaller proteins (*p6*, *p2* and *p1*) of uncertain function as well as the viral enzymes reverse transcriptase, integrase and protease. Although mammalian cells contain aspartyl proteases, none efficiently cleave the *gag* polyprotein. Three of the HIV-cleavage sites are phenylalanine–proline or tyrosine–proline bonds, which are unusual sites of attack for mammalian proteases (21). Proteolytic cleavage of the polyprotein results in morphologic changes in the virion and condensation of the nucleoprotein core. The protease is packaged into virions, and the cleavage events it catalyzes occur simultaneously with or soon after the budding of the virion from the surface of an infected cell. Proviral DNA lacking functional protease produces immature, noninfectious viral particles.

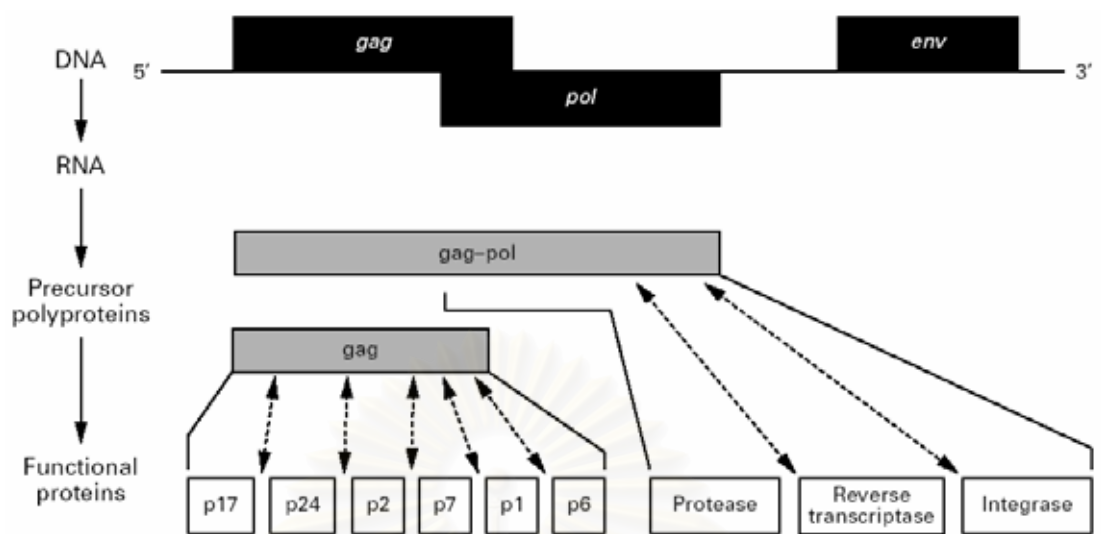


Figure 1.7 The translational products of the HIV *gag-pol* gene and the sites at which the gene product is cleaved by the virus-encoded protease. *p17* denotes capsid protein (CA), *p24* matrix protein (MA), and *p7* nucleocapsid (NC); *p2*, *p1*, and *p6* are small proteins with unknown functions. The arrows denote cleavage events catalyzed by the HIV-specific protease. (21)

1.4.2 Structure

The initial knowledge of the structures of retroviral proteases came from the crystallographic studies described above. The HIV-1 PR was postulated to belong to the family of aspartic acid proteases based on the identification of the Asp-(Ser/Thr)-Gly catalytic triad. Other members of this family, including the endogenous enzymes Pepsin, Cathepsin D and Renin, are single chain proteins of over 300 residues folded into two domains; each of which supplies a catalytic triad of Asp-(Ser/Thr)-Gly. PR is much smaller at one 99 residues in length and possesses only a single Asp-Thr-Gly triad so a homodimeric structure was proposed. (8)

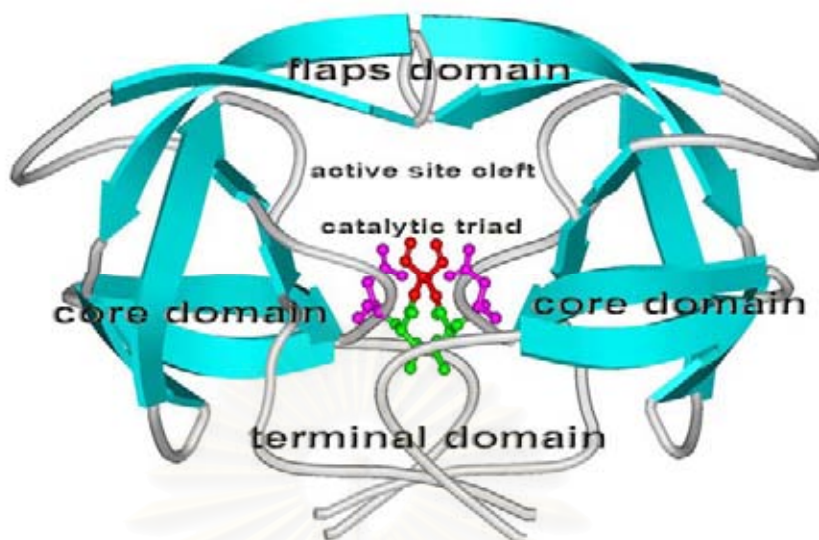


Figure 1.8 Structure of HIV-1 PR.

The HIV-1 PR enzyme functions as a homodimer composed of two identical 99-amino acid chains (22), with each chain containing the characteristic Asp-Thr-Gly active-site sequence at positions 25 to 27 (23). The crystal structure of HIV-1 protease reveals a dimer exhibiting two-fold rotational C_2 symmetry (17). The conserved active-site motifs are located in loops that approach the center of the dimer (Figure 1.8). The two subunits are linked by a fourstranded antiparallel β -sheet involving both the amino and the carboxyl termini of each subunit. Upon binding, both subunits form a long cleft where the catalytically important aspartic acids are located in a coplanar configuration on the floor of the cleft. In addition, the enzyme contains a so-called “flap structure” in each subunit, an antiparallel β -hairpin with a β -turn that extends over the substrate binding site (24,25). The HIV-1 PR structure can be separated into three parts:

1.4.2.1 The terminal domains

Terminal domain or dimerization domain consists of the termini four-stranded beta-sheet; (residues 1-4, and 95-99 of each monomer), the turn encompassing residues 4-9, and the helix (residues 86-94 of each monomer). This domain is quite crucial in dimer formation and stabilization of an active PR.

1.4.2.2 The core domains

These are one core domain from each monomer, composed of primarily four beta-strand structures; and are quite compact. Residues 10-32 and 63-85 from each monomer make up the domain sequence. The conserved Asp25-Thr26-Gly27 catalytic triad is situated at the interface of the core domains from the two monomers. This domain is quite useful in dimer stabilization, as well as the catalytic site stability. The interface between the core and terminal domains is composed primarily of small hydrophobic residues. The helix of the terminal domain packs against several beta-strands of the core domain.

In addition, the dimer is stabilized by non-covalent interactions, hydrophobic packing of side chains and interactions involving the catalytic residues. Aspartyl proteases share a conserved scaffold of hydrogen bonds, termed as "fireman's grip", which involves the residues at the catalytic Asp-Thr-Gly triplets. "Fireman's grip" is indispensable for the stability of the active site (β -1 turn) as well as the dimer.

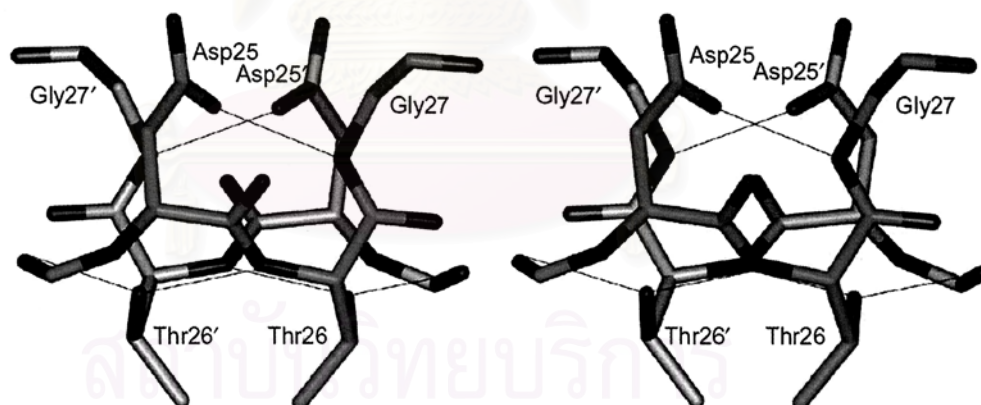


Figure 1.9 The 'fireman's grip', a stereotypical rigid network structure involving the Asp-Thr-Gly signature sequence in the retroviral proteases. Amino acids are identified by three-letter codes. The catalytic aspartic acid residue (Asp) is hydrogen-bonded to the backbone NH group of the glycine (Gly) two amino acids further along in the sequence. In addition, the OH groups of threonine (Thr) are hydrogenbonded to two points on the opposite domain or monomer, the backbone NH group (blue) of the threonine and to the carbonyl oxygen (red) of the residue before the catalytic aspartic acid. (26)

1.4.2.3 The flap domains

This domain includes a mostly solvent exposed loop (residue 33-43) preceding the beta hairpin containing the flaps (residue 44-63). Flexible flaps enclose the active site and provide important ligand binding interactions. Residues 46-56 in both monomers form the flexible part of the flap domain and they contain three characteristic regions: side chains that extend outwards (Met46, Phe53), hydrophobic chains extending inward (Ile47, Ile54), and Glycine rich region (48-52 residues), with Ile50 remaining at the turn. The Glycine rich tips of the flaps are highly flexible. Molecular dynamics studies have shown that the tips are highly mobile in a free enzyme (unliganded), and they maintain the PR in an alternating closed and open conformation. The high flexibility and mobility of the flaps is thought to be necessary for substrate binding and product release. Flaps also play an important role in the enzyme hydrolysis reaction stabilization, whereby both Ile50 at the tip of the flaps and some residues in the substrate form hydrogen bonds via a water molecule in the active site, thus maintaining an active conformation.

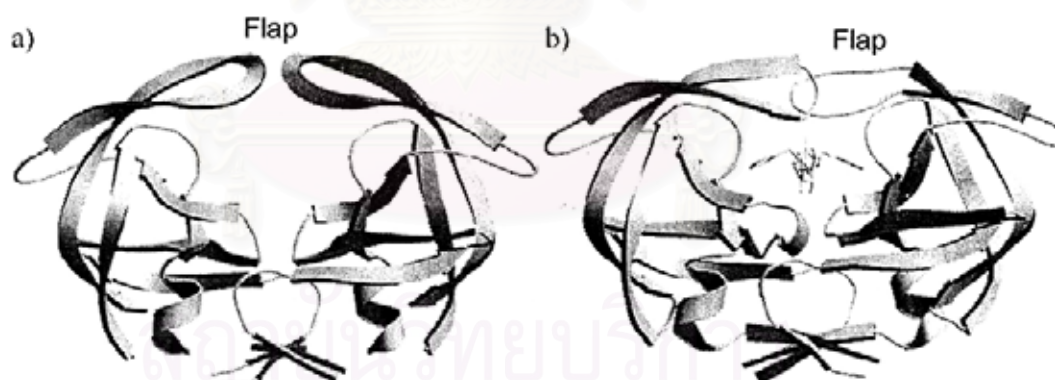


Figure 1.10 Ribbon drawings of (a) apo- and (b) inhibited HIV protease showing the relatively open and closed positions of the flap (top of the images). These images were produced with MolScript and Raster3D on the PDB files 3HVP and 1AJX. (8)

1.4.3 Hydrolysis mechanism of HIV-1 PR

The substrate binds in its extended conformation, in which its interactions with the different amino acid side chains determine the specificity of the

enzyme. Using standard nomenclature (Figure 1.11), the S_1 and S_1' (S_2 and S_2' , etc.) subsites are structurally equivalent. The two S_1 subsites are very hydrophobic, the S_2 subsites are mostly hydrophobic except Asp-29, Asp-29', Asp-30 and Asp-30'. The S_3 subsites are adjacent to S_1 subsites and are also mostly hydrophobic. (27)

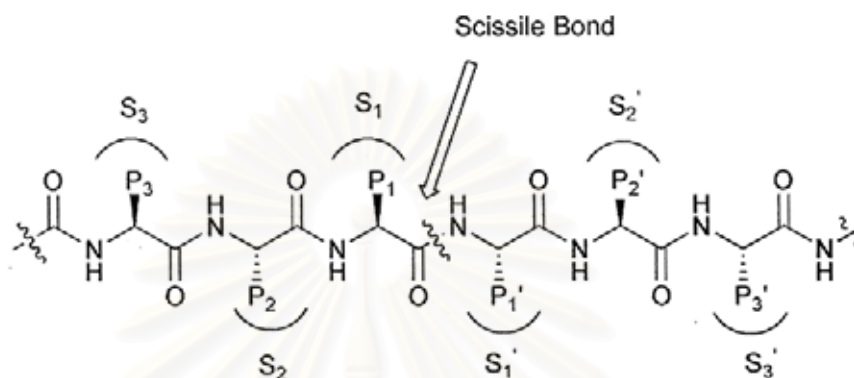


Figure 1.11 Standard nomenclatures $P_1\dots P_n$ - $P_1'\dots P_n'$ is used to designate amino acid residues of peptide substrates. The corresponding binding sites on the protease are referred to as $S_1\dots S_n$ - $S_1'\dots S_n'$ subsites. (27)

The hydrolysis mechanism for HIV protease (Figure 1.12), like all aspartic proteases, is dependent on the catalytic activity of two aspartic acid residues (Asp's), one in each subunit (28,29). They activate a substrate water molecule in the active site that performs an attack on a carbonyl atom in the peptide bond. The water molecule is polarized into a nucleophilic state by one of the aspartic acids and attacks the polarized carbonyl at the scissile bond. During catalysis a tetrahedral intermediate, the transition-state, is formed and broken.

ศึกษาหน่วยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

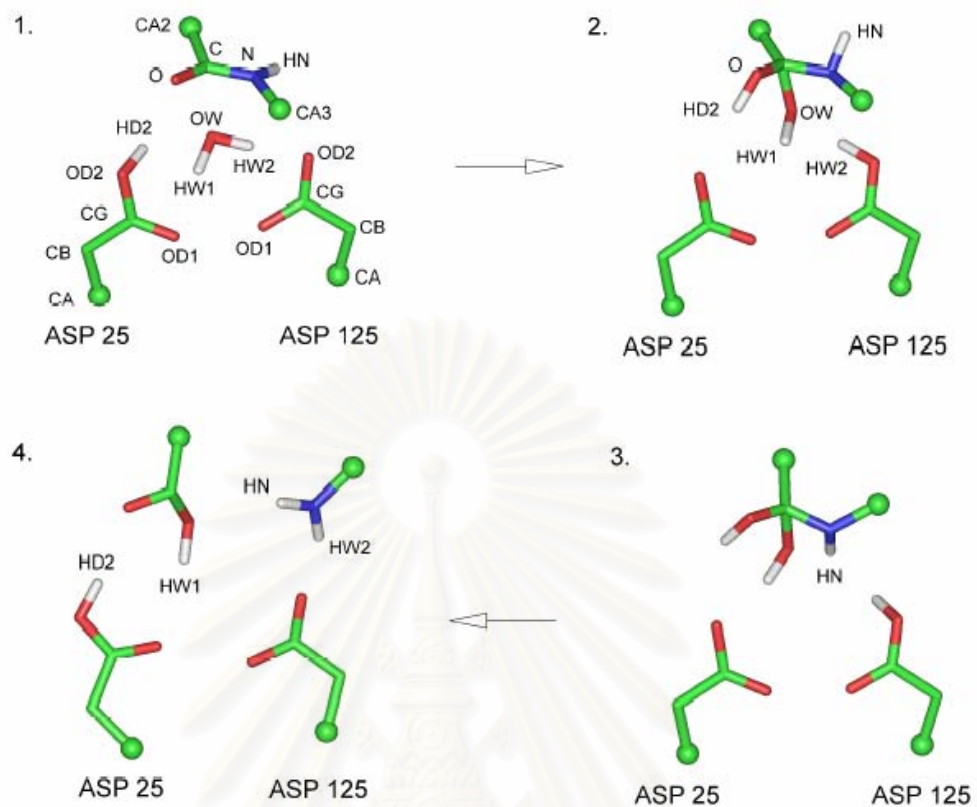


Figure 1.12 Mechanism of HIV-1 PR One of the proposed reaction mechanisms. 1) Enzyme reaction centre: catalytic aspartates (25 and 25'), peptide bond and lytic water molecule. 2) Tetrahedral intermediate following HW2 transfer between water molecule and OD2 of Asp25', nucleophilic attack of the resulting hydroxyl anion (OW-HW1) on the peptide carbon (C), and HD2 transfer between Asp25 and the carbonyl oxygen of the peptide bond. 3) An *anti-gauche* to *gauche* transition of the C-N bond allowing the nitrogen lone pair to accept HW2 proton of Asp25'. 4) C-N bond rupture and hydrogen transfer between one of the hydroxyl groups of the tetrahedral intermediate onto Asp25. Balls represent carbon atoms bonded to other fragments of the protein or the substrate. (30)

HIV-1 protease also contains a structural water molecule in the active site, which forms a bridge between the substrate and the flaps (Figure 1.13). This water molecule donates its hydrogen to a carbonyl atom in the substrate and hydrogen bonds to the amide hydrogen in residues Ile50 and Ile50', thus is stabilizing the closed conformation of the enzyme (Figure 1.13).

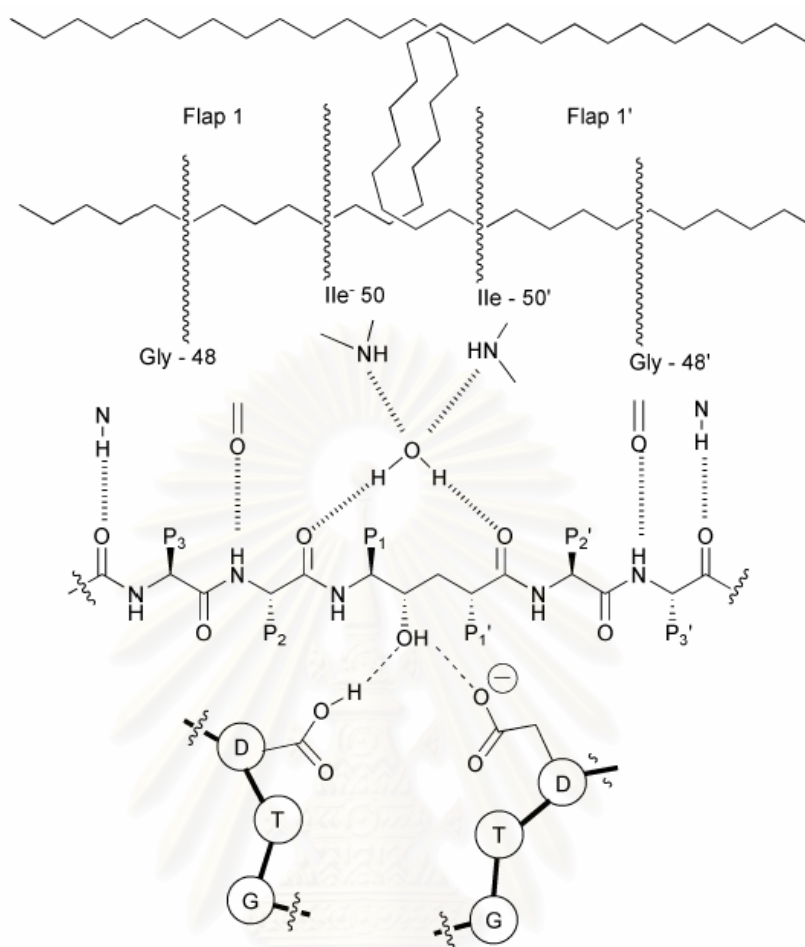


Figure 1.13 schematic representations of two flaps of HIV-1 PR and their hydrogen bonds with the water molecule. The same water molecule makes another two hydrogen bonds with the peptidyl inhibitor or substrate. (31)

1.5 HIV-1 PR inhibitor

The crucial role of the PR in the viral life cycle has made it an important therapeutic target. PR inhibitors block the activity of the enzyme and prevent viral replication, therefore, represent some of the most potent agents available as therapeutic strategies designed to inhibit HIV-1 replication. All the available protease inhibitors were invented after the first three dimensional atomic structure of HIV-1 PR was solved in 1989; this indicates a great application in structure-based drug design. The catalytic mechanisms of protease inhibitors are different, depending on the type of inhibitor (32):

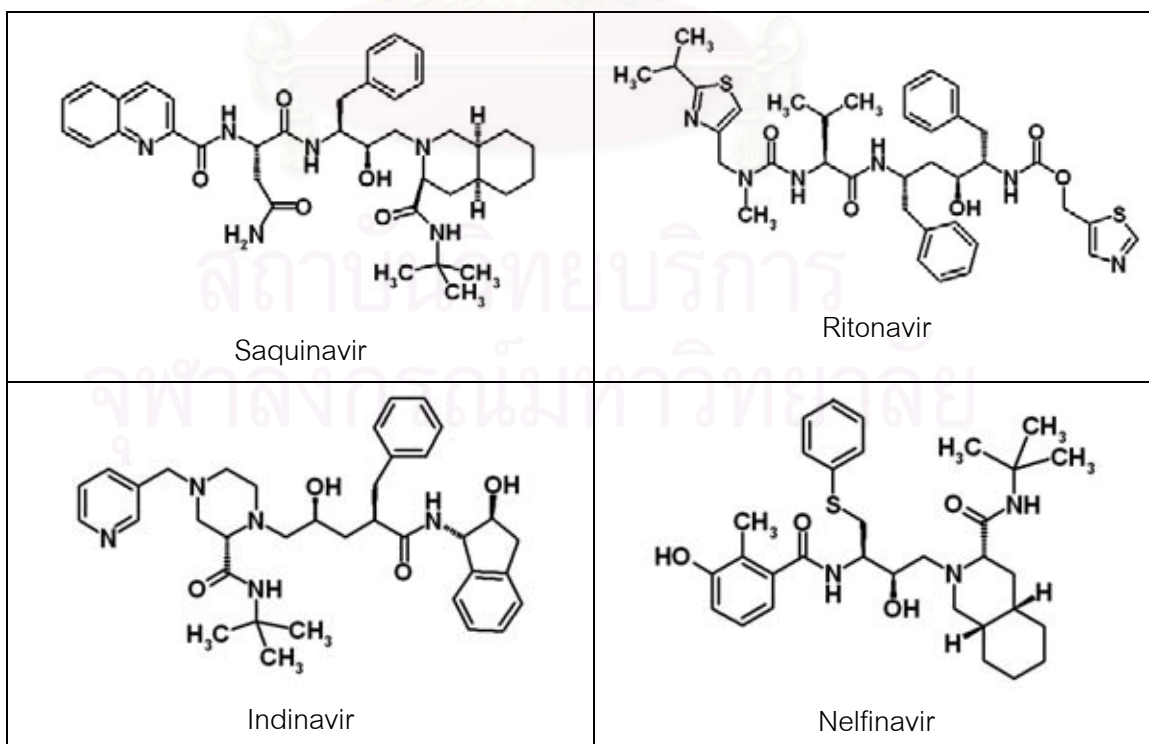
- **Transition-state analogues** work by directly blocking the catalytic activity by non-covalent and reversible binding to the active site.
- **Non-peptide active site inhibitors (28,33)** bind covalently and irreversibly to residues in the active site.
- **Allosteric inhibitors** interact with the enzyme by binding at a site distant from the active site. The binding induces a change in the overall structure of the protease, *e. g.*, locking the structure in a certain position, thereby preventing catalysis.
- **Dimerization inhibitors (34)** that act at the dimerization interface and prevent dimerization.
- **Monoclonal antibodies (35)** bind specifically to non active site epitopes on the protease and prevent catalysis.

The clinical inhibitors (Figure 1.14) are all linear transition-state analogues that bind to the active site. They are asymmetric peptide mimetics with a central secondary hydroxy group that binds to the catalytic aspartic acid residues and mimics the oxygen in the tetrahedral transition-state for amide cleavage (36). The asymmetry of inhibitors may be of importance for resistance development, since the enzyme is a symmetrical dimer and thus has problems in adapting to an asymmetrically inhibitor. The first HIV-1 PIs to be approved by the U.S. Food and Drug Administration (FDA) was saquinavir in 1995. Recently, there are ten FDA approved protease inhibitors in clinical uses:

- Saquinavir (Ro 31-8959) (37)
- Amprenavir (VX-478) (38)
- Indinavir (MK-639) (39)
- Nelfinavir (AG1343) (40)

- Ritonavir (ABT-538) (41)
- Lopinavir (ABT-378) (42)
- Atazanavir (BMS-232632) (43)
- Tipranavir (PNU-140690) (44)
- Fosamprenavir (GW-433908) (45)
- Darunavir (TMC-114) (46)

These drugs are competitive inhibitors binding at the substrate binding site of the PR, *i.e.*, all have been designed to compete with natural substrates for the active site, on the basis of structural and chemical complementarities, as well as thermodynamic stabilities. Efficacy of these drugs is limited due to the natural selection of protease variants that are still catalytically competent but have low affinity of the drugs than wild type enzyme.



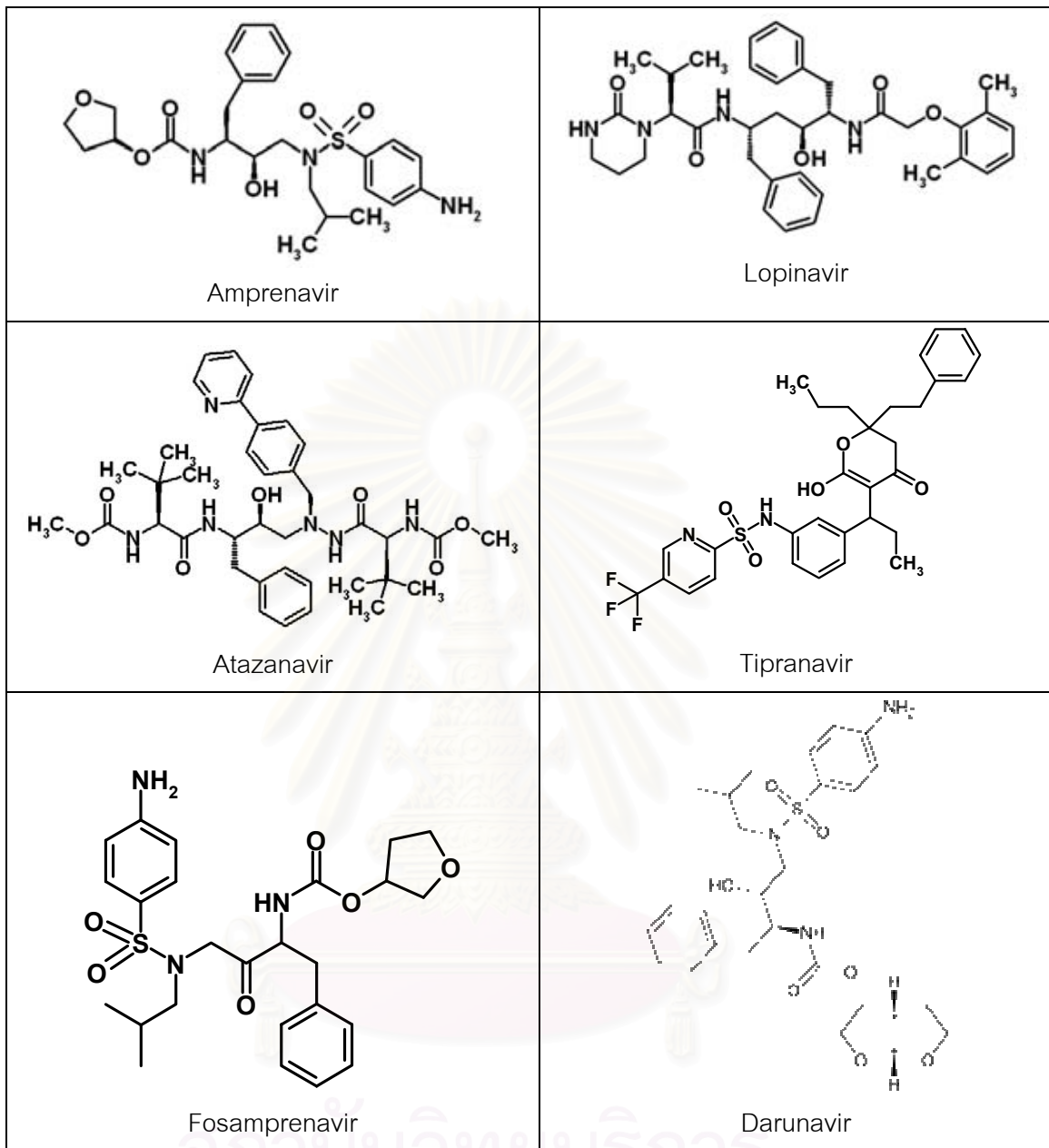


Figure 1.14 Clinical inhibitors of HIV-1 PR.

Table 1.3 Inhibition of HIV-1 PR.

Generic name	Brand name	Approval year	IC_{50} (nM)	IC_{90} (nM)	K_i (nM)
Saquinavir	Invirase®	1995		16	0.15
Ritonavir	Norvir®	1996		137	0.17
Indinavir	Crixivan®	1996		27.2	0.14
Nelfinavir	Viracept®	1997		25.8	0.28
Amprenavir	Agenerase®	1999	80		0.015
Lopinavir	Kaletra®	2000			0.0013
Atazanavir	Reyataz®	2003	2-5	8-12	
Fosamprenavir	Telzir®, Lexiva®	2003		57	
Tipranavir	Aptivus®	2005			8.9
Darunavir	Prezista®	2006			0.016

1.6 HIV-1 drug resistance

HIV-1 drug resistance is a result of mutations occurring within the wild-type viral genome that produce variants capable of efficiently replicating in the presence of antiretroviral agents. The first report of HIV-1 drug resistance was to zidovudine (ZDV) in 1989 (47). Subsequently, drug resistance to all therapeutic antiretroviral agents has been observed. Although highly active antiretroviral therapy for the treatment of HIV-1 infection has produced substantial decreases in morbidity and mortality in recent years, as many as 50% of patients fail therapy within 1 year of initiation. As drug resistant HIV-1 variants are often selected during the course of antiretroviral therapy, drug resistance is considered a major contributor to treatment failure.

1.6.1 Evolution of HIV-1 drug resistance

The evolution of HIV-1 drug resistance within an individual depends on the generation of genetic variation in the virus and on the selection of drug-resistant variants during therapy. HIV-1 genetic variability is a result of the inability of HIV-1 reverse transcriptase (RT) to proofread nucleotide sequences during replication (48). This variability is compounded by the high rate of HIV-1 replication, the accumulation of

proviral variants during the course of HIV-1 infection, and genetic recombination when viruses of different sequence infect the same cell. As a result, innumerable genetically distinct variants (quasispecies) evolve within an individual in the months following primary infection (49).

Development of drug resistance depends on the extent to which virus replication continues during drug therapy, the ease of acquisition of a particular mutation (or set of mutations), and the effect of drug-resistance mutations on drug susceptibility and virus fitness. Some mutations selected during drug therapy confer measurable phenotypic resistance by themselves, whereas other mutations compensate for the diminished replicative activity that can be associated with drug resistance or cause resistance only when present in combination with other mutations.

It is estimated that every possible single point mutation occurs between 10⁴ and 10⁵ times per day in an untreated HIV-1-infected individual and that double mutants also occur commonly (49). Therefore, most drug-resistance mutations are probably present prior to the start of therapy. However, viruses containing sufficient drug-resistance mutations to replicate in the presence of multiple drugs do not appear to exist in previously untreated persons infected with wild type viruses. Indeed, once potent HIV-1 suppression is achieved in previously untreated persons, it usually persists indefinitely if therapy is not interrupted (50).

1.6.2 HIV-1 protease drug resistance

Early development of antiretroviral therapy focused on inhibitors of reverse transcriptase. Both nucleoside and nonnucleoside inhibitors of this enzyme showed significant antiviral activity. However, the clinical benefit of these drugs had been limited due to drug resistance, limited potency, and host cellular factors. Thus, inhibitors targeted against a second essential enzyme of HIV-1 were urgently needed.

In 1988, the protease enzyme of HIV-1 was crystallized and its three-dimensional structure was determined, allowing for the rapid development of protease

inhibitor. Initially, it was hypothesized that HIV-1 protease, unlike reverse transcriptase, would be unable to accommodate mutations leading to drug resistance. This is not the case, and to date more than 45 possible amino acid substitutions in the HIV-1 protease have been observed during treatment with the currently available protease inhibitors. Some mutated residues are located adjacent to the pocket in the protease molecule where inhibitors bind, while others are located outside of the binding pocket. Thus, the mutations in the protease gene are classified as either major or minor:

- Major mutation: In general, major mutations are either (1) selected first in the presence of the drug; or (2) shown at the biochemical or virologic level to lead to an alteration in drug binding or an inhibition of viral activity or viral replication. By themselves, major mutations have an effect on phenotype. In general, these mutations tend to be the major contact residues for drug binding.
- Minor mutation: In general, minor mutations appear later than major mutations, and by them do not have a significant effect on phenotype. In some cases, their effect may be to improve replicative fitness of virus carrying major mutations.

The major amino acid mutations associated with clinical resistance have been mapped. A substitution of phenylalanine for valine at position 82 is the initial mutation associated with resistance to indinavir and ritonavir and mutations in glycine at position 48 and leucine at position 90 results in initial resistance to saquinavir (Table 1.4). In general, initial single amino acid mutations yield only a slight change in drug sensitivity. However, secondary mutations accumulate in the virus and can lead to high-level drug resistance. Unlike the initial mutations, the secondary mutations in virus from patients with resistance to different protease inhibitors overlap. HIV protease can tolerate a substantial amount of mutation; at least one third of its 99 amino acids can deviate from the wild-type sequence without altering function.

Table 1.4 HIV-1 resistances to protease inhibitors (51).

Drug	Position(s) in protease	
	Critical substitutions	Secondary substitutions ^a
Saquinavir	48, 90	10, 36, 63, 71
Ritonavir	82, 84	20, 36, 46, 54, 63, 71, 90
Indinavir	46, 82	10, 20, 24, 32, 54, 63, 71, 84, 90
Nelfinavir	30	46, 63, 71, 88, 90
Amprenavir	50	10, 46, 47

^a The role of secondary substitutions in the development of resistance is not completely understood. Some substitutions are known to be compensatory, enhancing the replicative efficiency of the virus. Other changes exist as naturally occurring polymorphisms in untreated patients.

An obvious and expected mechanism for the development of resistance to HIV PR inhibitors is caused by the changes of specificity-determining residues that can directly interfere with the binding of the inhibitor to the enzyme. The mutations that do not directly change the shape or character of the binding cavity can indirectly influence inhibitor binding via long-range structural perturbations of the active site, or they can change the efficiency of catalysis and the stability of the enzyme.

As shown in Figure 1.15, The International AIDS Society–USA Drug Resistance Mutations Group reviews new data on HIV drug resistance in order to maintain a current list of mutations associated with clinical resistance to HIV. This list includes mutations that may contribute to a reduced virologic response to a drug. The mutations listed have been identified by 1 or more of the following criteria: (1) in vitro passage experiments or validation of contribution to resistance by using sitedirected mutagenesis; (2) susceptibility testing of laboratory or clinical isolates; (3) genetic sequencing of viruses from patients in whom the drug is failing; (4) correlation studies between genotype at baseline and virologic response in patients exposed to the drug.

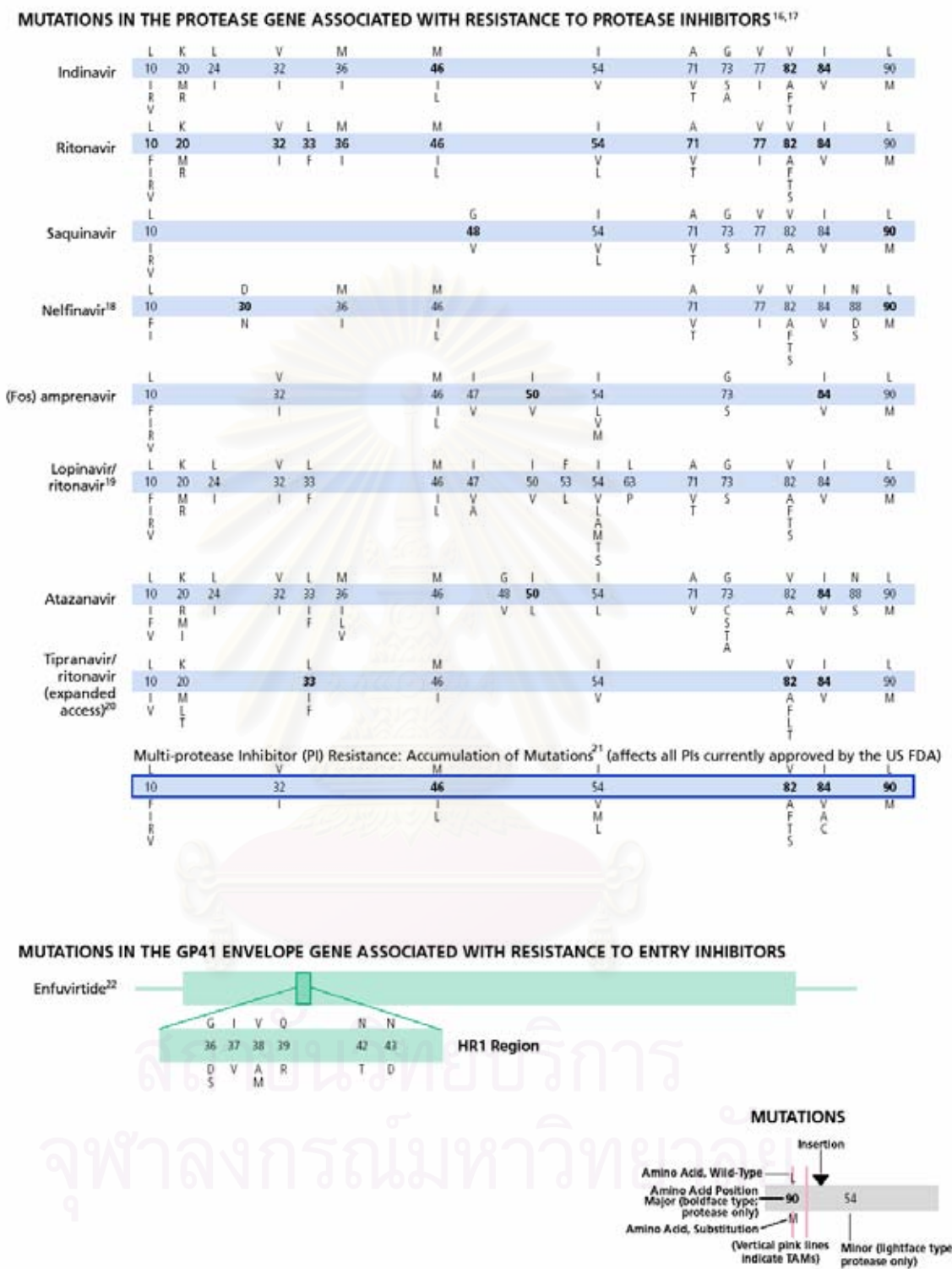


Figure 1.15 Mutations in the protease gene associated with resistance to protease inhibitor. (52)

1.7 Saquinavir resistance

In vivo saquinavir therapy appears to select almost exclusively for mutations at codons 90 and 48 (53-55). Saquinavir-resistant variants emerge in approximately 45% of patients after 1 year of monotherapy with 1,800 mg daily (56). The frequency of genotypic resistance is lower (22%) in patients receiving combination therapy with zidovudine, zalcitabine, and saquinavir (57). In contrast to in vitro-selected virus, where the valine-for-glycine substitution at position 48 (G48V) is the first step to resistance, the methionine-for-leucine substitution at residue 90 (L90M) exchange is the predominant mutation selected in vivo while the G48V (2%) or the double mutant (<2%) is rarely found. In another recent study of in vivo resistance during saquinavir monotherapy no patient was found to harbor a G48V mutant virus. Interestingly, Winters *et al.* (58) observed a higher frequency of the G48V mutation in patients receiving higher saquinavir doses as monotherapy. All patients (six of six) who initially developed G48V also acquired a V82A mutation either during saquinavir treatment or after switching to either indinavir or nelfinavir. An identical mutational pattern was found in another study during saquinavir monotherapy. Some residues represent sites of natural polymorphism of the HIV-1 protease (positions 10, 36, 63, and 71) and appear to be positively correlated to the L90M mutation. Another substitution, G73S, has been recently identified and may play a role in saquinavir resistance in vivo. Isolates from five patients with early saquinavir resistance and those from two patients with induced saquinavir resistance after a switch of therapy to indinavir carried the G73S and the L90M substitutions (59).

In vitro, saquinavir selects for variants with one or both of two amino acid substitutions in the HIV-1 protease gene, a G48V, a L90M, and the double substitution G48V-L90M (54). In most cases, G48V is the first mutation to appear, and continued selection results in highly resistant double-mutant variants. A substitution at either residue results in a 3- to 10-fold decreased susceptibility to the inhibitor, whereas the simultaneous occurrence of both substitutions causes a more severe loss of susceptibility of 100-fold. Amino acid residue 48 is in the flexible flap loop of the

enzyme, while residue 90 is located outside the binding pocket of the enzyme. Substitutions at position 48 could result in less conformational freedom and greater rigidity of the flap. The L90M substitution may induce conformational perturbations in the enzyme altering binding of the inhibitor. Ermolieff *et al.* reported that the inhibition constants K_i of the constructed mutants were significantly higher than those of wild-type virus: 3-fold for L90M, 13.5-fold for G48V, and 419-fold for G48V/L90M (60). Maschera *et al.* determined the affinity of the wild type and the three mutant proteases L90M, G48V, and L90M/G48V (61). The affinity values for saquinavir were 1/20, 1/160, and 1/1,000 of that of the wild type, respectively. The catalytic efficiencies (k_{cat}/K_m) of the mutant proteases G48V and L90M/G48V were markedly reduced, 1/10 to 1/20 of that of the wild-type protease. These findings document the deleterious effects of these mutations on enzyme function and significantly reduced binding of the inhibitor. Of note, the decreased catalytic efficiencies resulted primarily from increased K_m values, as opposed to changes in k_{cat} . This suggests that the enzyme will maintain function in the presence of substrate excess, explaining the viability of the mutant viruses.

Table 1.5 Comparison of K_i values with association and dissociation rate constants for WT and mutant protease with saquinavir. (61)

Enzyme	k_{on} $\times 10^{-7} M^{-1} s^{-1}$	k_{off} s^{-1}	$K_{d(calc)}$ <i>nM</i>	K_i <i>nM</i>
WT	4.2 ± 0.2	0.0014 ± 0.0002^a	ND ^b	0.033 ± 0.005
L90M	4.2 ± 0.2	0.019 ± 0.005	0.5 ± 0.1	0.68 ± 0.04
G48V	2.7 ± 0.1	0.128 ± 0.008	4.7 ± 0.3	5.4 ± 0.6
G48V/ L90M	2.35 ± 0.08	0.54 ± 0.01	22.9 ± 0.9	33 ± 1

^a Calculated from K_i and k_{on} .

^b ND, not determined.

1.8 Ritonavir resistance

Phenotypically resistant virus generated by serial in vitro passages is associated with specific mutations at positions 84, 82, 71, 63, and 46 (62). The I84V substitution appeared to be the major determinant of resistance, resulting in a 10-fold reduction in

sensitivity to ritonavir. Addition of the V82F mutation confers an even greater level of resistance, 10- to 20-fold. The substitutions M46I, L63P, and A71V, when introduced into the protease coding region of wild-type NL4-3, did not result in significant changes in drug sensitivity. Based on replication kinetics experiments, these changes are likely to be compensatory for active-site mutations, restoring the impaired replicative capacity of the combined V82F and I84V mutations. Computer modeling of ritonavir binding to the HIV-1 protease, based on the crystal structure of the related inhibitor A-78791 (24), demonstrated the structural effects of the V82F and I84V substitutions in HIV-1 protease on the interaction with ritonavir at the S1 and S1' subsites (Figure 1.15). Conformational adjustments due to these mutations result in decreases in enzyme-inhibitor interactions which are associated with loss of drug activity.

Molecular dynamics simulation studies suggest that the M46I substitution results in a closed conformation of the flap domain relative to the wild-type enzyme, when a substrate or an inhibitor is bound. The role of the L63P and A71V substitutions is not obvious, since these mutations are away from the immediate vicinity of the active site and are not involved in substrate binding. The catalytic efficiencies (k_{cat}/K_m) for the mutant HIV proteases M46I, V82F, I84V, and M46I/V82F are reduced relative to wild-type enzyme by factors of 2.2, 3.8, 2.6, and 1.6, respectively. Schock *et al.* observed that the mutant with the double substitution M46I/L63P had a greater catalytic efficiency (per mole per second) than that of the wild-type enzyme for every substrate (110 to 360%), strongly suggesting that these changes likely play a compensatory role for other deleterious mutations.

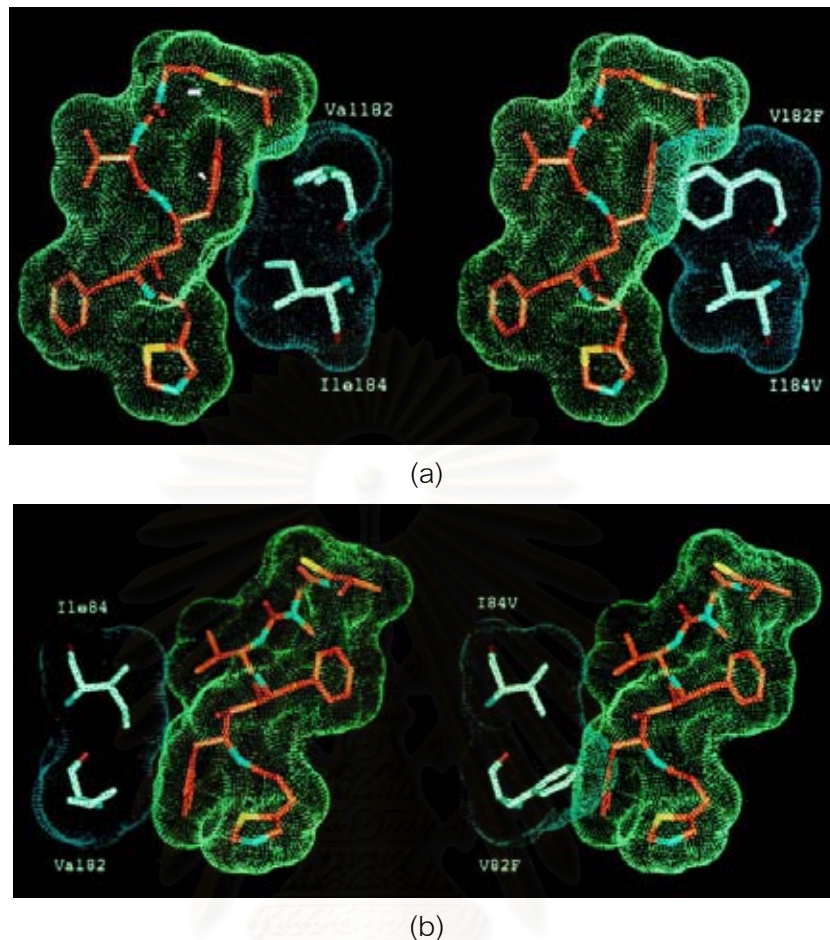


Figure 1.16 Computer-generated model of HIV-1 protease-ritonavir complexes of wild-type HIV-1_{NL4-3} and the V82F/I84V double mutant. (a) Interaction between ritonavir and the protease at the S₁' subsite; (b) the same interactions at the S₁' binding subsite. I84V decreases the interaction with the C_β group of the benzyl side chain of ritonavir, whereas V82F results in a severe spatial overlap with the phenyl ring of the inhibitor at the P₁' site. Note the effects of the double mutant on the van der Waals interactions between the enzyme and the inhibitor. (51)

Phase I/II clinical trials of ritonavir monotherapy demonstrated a rapid decline in plasma HIV RNA to ~1% of baseline levels (63). After 32 weeks of treatment, only patients receiving the highest dose of ritonavir (1,200 mg daily) maintained a mean HIV plasma RNA reduction .0.8 log unit. Genotypic analysis of the HIV protease of viral isolates from patients with a rebound in viremia revealed a stepwise, ordered accumulation of multiple mutations at nine different codons resulting in amino acid

substitutions at residues 20, 33, 36, 46, 54, 71, 82, 84, and 90. An initial mutation at position 82 was consistently observed and appeared to be necessary for the primary loss of antiviral effect. There was no patient with a rise in viral load without variation at residue 82, as either a single substitution or part of a complex mutational pattern. This observation suggests that the primary mechanism for resistance is the selection of preexisting V82 single mutants with incompletely suppressed replication in the presence of ritonavir. The substitution at position 82 is followed most frequently by mutations at positions 54, 71, and 36. In contrast to *in vitro* experiments with ritonavir, *in vivo* I84V emerged late and less frequently. The L90M mutation was also observed less frequently and was always accompanied by at least two additional mutations. Emergence of I54V was not predicted by *in vitro* selection. The coappearance of I54V with V82A/F *in vivo* suggests a possible compensatory role of I54V (64).

1.9 Research objectives

In view of the fact that the experimental findings indicate the rapid emergence of drug resistance to most of the PR inhibitors in use, due to site specific mutations in the enzyme occurring at one or more residues and these mutations are broadly conserve in nature and are a serious obstacle for efficient protease inhibitor salvage therapy with alternative protease inhibitors. This study was focus on the major mutation of saquinavir and ritonavir that observe on the different position (residues 48(48') and 90(90') for saquinavir, 82(82') and 84(84') for ritonavir) around the binding pocket. For this reason the effect from these positions around the binding pocket were observed.

These project pursuers the following goals:

1. The problem of the protonation state of the active site residues, Asp25 and Asp25', was addressed using the MD simulations of G48V complexed with saquinavir in the three possible protonation states, protonate at each of the two Asp and diprotonate.
2. MD simulations of wild-type, single mutants (G48V and L90M) and double mutant (G48V/L90M) complexed with saquinavir at the protonation state obtained from (1) were performed. The binding free energy calculations were also applied to evaluate the relative binding affinity of each HIV-1 PR-inhibitor complex.

3. The same methods in (2) were applied for ritonavir complexed with wild-type and mutants, V82F, I84V and V82F/I84V.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER 2

THEORY

2.1 Potential energy functions (65)

To investigate the relationships between structure, function and dynamics at the atomic level of biological molecules, it is not yet feasible to treat the systems using quantum mechanics calculation. Therefore, the problems become much more tractable when turning to empirical potential energy functions which are much less computationally demanding. However, one of the disadvantage of the use of empirical force field is that no drastic changes in electronic structure are allowed, *i.e.*, no events like bond making or breaking can be modeled.

Recently generation potential energy functions (or force fields) provide a reasonably good compromise between accuracy and computational efficiency. They are often calibrated to experimental results and quantum mechanical calculations of small model compounds. Their ability to reproduce physical properties measurable by experiment is tested; these properties include structural data obtained from x-ray crystallography and NMR, dynamic data obtained from spectroscopy and inelastic neutron scattering, and thermodynamic data. The most commonly used potential energy functions are the AMBER, CHARMM, GROMOS and OPLS/AMBER force fields. The continuing development of force fields remains an intense area of research with implications for both fundamental as well as applied researches in the pharmaceutical industry.

The energy, E , is a function of the atomic positions, R , of all the atoms in the system, these are usually expressed in term of Cartesian coordinates. The value of the energy is calculated as a sum of internal or bonded, terms E_{bonded} , which describe the bonds, angles and bond rotations in a molecule, and a sum of external or nonbonded terms, $E_{non-bonded}$. These terms account for interactions between nonbonded atoms or atoms separated by 3 or more covalent bonds. These are drawn together to form a potential energy function:

$$V(R) = E_{\text{bonded}} + E_{\text{non-bonded}} \quad (2.1)$$

2.1.1 Intramolecular interactions

The E_{bonded} term is a sum of the following three terms

$$E_{\text{bonded}} = E_{\text{bond-stretch}} + E_{\text{angle-bend}} + E_{\text{rotate-along-bond}}, \quad (2.2)$$

which correspond to three types of atom movement, r_{23} , θ_{234} and ϕ_{1234} shown in Figure 2.1.

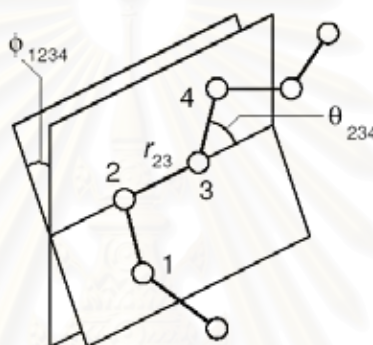


Figure 2.1 Geometry of a simple chain molecule, illustrating the definition of interatomic distance r_{23} , bend angle θ_{234} , and torsion angle ϕ_{1234} . (66)

The E_{bonded} is the energy function for stretching a bond between two atoms, atoms 2 and 3 in Figure 2.1. It is a harmonic potential representing the interaction between atomic pair where atoms are separated by one covalent bond, *i.e.*, 2, 3-pair. This is the approximation to the energy of a bond as a function of displacement from the equilibrium bond length, b_0 . The force constant, K_b , determines the strength of the bond. Both ideal bond lengths and force constants are specific for each pair of bound atoms, *i.e.*, their value only depend on chemical type of atoms-constituents.

$$E_{\text{bond-stretch}} = \sum_{1,2\text{-pairs}} K_b (b - b_0)^2 \quad (2.3)$$

Values of force constant are often evaluated from experimental data such as infrared stretching frequencies or from quantum mechanical calculations.

Values of bond length can be inferred from high resolution crystal structures or microwave spectroscopy data.

The second term in Equation 2.2 is associated with alteration of bond angles θ from equilibrium values θ_0 , which is also represented by a harmonic potential. Values of θ_0 and K_θ depend on chemical type of atoms constituting the angle.

$$E_{\text{bond} - \text{bend}} = \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 \quad (2.4)$$

These two terms describe the deviation from equilibrium geometry; effectively, they are penalty functions and that in a perfectly optimized structure, the sum of them should be close to zero.

The third term in Equation 2.2 represents the torsion angle potential function which models the presence of steric barriers between atoms separated by 3 covalent bonds (1,4-pairs). The motion associated with this term is a rotation, described by a dihedral angle and coefficient of symmetry $n=1, 2, 3$ around the middle bond. This potential is assumed to be periodic and is often expressed as a cosine function.

$$E_{\text{torsion-angle}} = \sum_{1,4\text{-pair}} K_\phi (1 - \cos(n\phi)) \quad (2.5)$$

2.1.2 Non-bonded interactions

The energy term representing the contribution of non-bonded interactions in the potential function has two components which are the Van der Waals interaction energy and the electrostatic interaction energy.

$$V_{\text{non-bonded}}(R) = E_{\text{vdw}} + E_{\text{electrostatic}} \quad (2.6)$$

The van der Waals energy, E_{vdw} , describes the repulsion or attraction between atoms that are not directly bonded. This term can be interpreted as the part of the interaction which is not related to atomic charges. The van der Waals interaction between two atoms arises from a balance between repulsive and attractive forces. The

repulsive force arises at short distances where the electron-electron interaction is strong. The attractive force, also referred to as the dispersion force, arises from fluctuations in the charge distribution in the electron clouds. The fluctuation in the electron distribution on one atom or molecules gives rise to an instantaneous dipole which, in turn, induces a dipole in a second atom or molecule giving rise to an attractive interaction. Each of these two effects is equal to zero at infinite atomic separation R and become significant as the distance decreases. The attractive interaction is longer range than the repulsion but as the distance become short, the repulsive interaction becomes dominant. This gives rise to a minimum in the energy. Positioning of the atoms at the optimal distances stabilizes the system. Both value of energy at the minimum E^* and the optimal separation of atoms r^* (which is roughly equal to the sum of van der Waals radii of the atoms) depend on chemical type of these atoms.

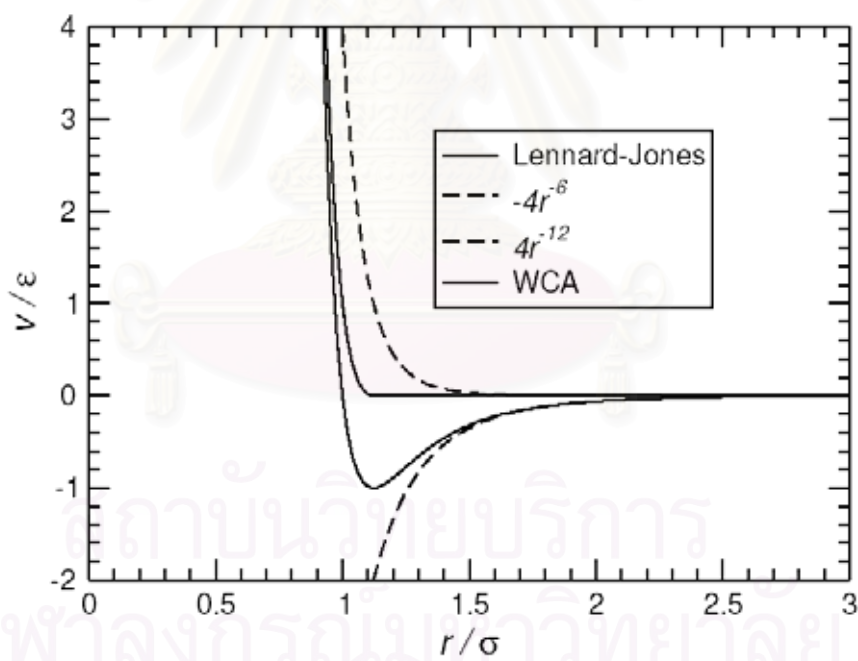


Figure 2.2 Lennard-Jones pair potential showing the r^{-12} and r^{-6} contributions. Also shown is the WCA (Water Conservation Area) shifted repulsive part of the potential. (66)

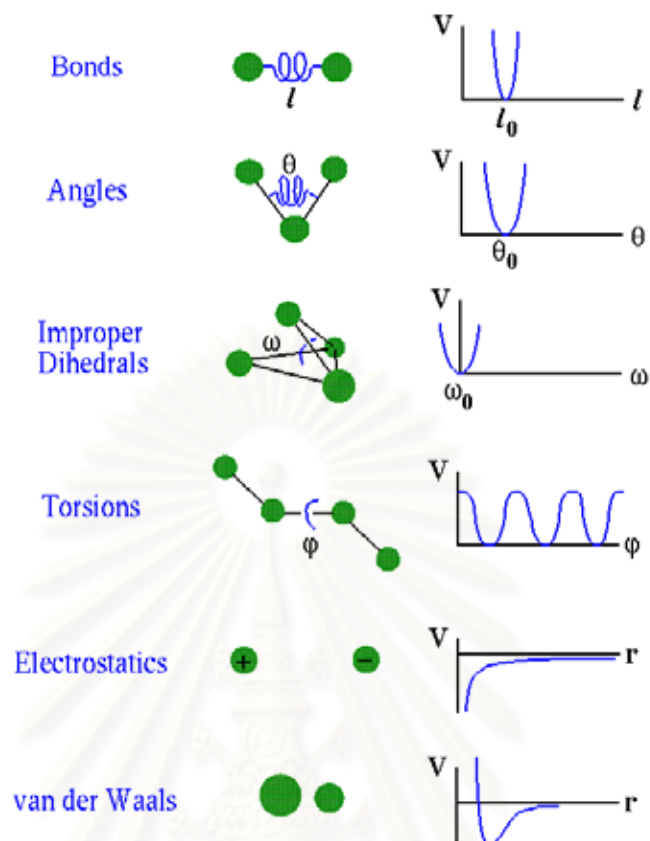


Figure 2.3 Empirical potential energy function and their components.

The van der Waals interaction is most often modeled using the Lennard-Jones 6-12 potential which expresses the interaction energy using the atom-type dependent constants, A and C . Values of A and C may be determined by a variety of methods, like non-bonding distances in crystals and gas-phase scattering measurements. The van der Waals interactions are one of the most important for the stability of the biological macromolecules.

$$E_{vdw} = \sum_{non-bonded} \left(\frac{A_{ik}}{r_{ik}^{12}} - \frac{C_{ik}}{r_{ik}^6} \right) \quad (2.7)$$

The second term of the non-bonded interaction is $E_{electrostatic}$ which is due to internal distribution of the electron. It is created by positive and negative parts of the molecule. The electrostatic interaction between a pair of atoms is represented by

Coulomb potential; D is the effective dielectric function for the medium and r_{ik} is the distance between two atoms having charges q_i and q_k .

$$E_{electrostatic} = \sum_{non - bonded} \frac{q_i q_k}{Dr_{ik}} \quad (2.8)$$

The empirical potential energy function is differentiable with respect to the atomic coordinates; this gives the value and the direction of the force acting on an atom and thus it can be used in a molecular dynamics simulation.

2.2 Energy minimization (67)

An important method for exploring the potential energy surface is to find configurations that are stable points on the surface. This means finding a point in the configuration space where the net force on each atom vanishes, *i.e.*, the derivative or gradient $\nabla V(R) = 0$. By adjusting the atomic coordinates and unit cell parameters (for periodic models, if requested) so as to reduce the model potential energy, stable conformations can be identified. Perhaps more important, the addition of external forces to the model in the form of restraints allows for the development of a wide range of modeling strategies using minimization strategies as the foundation for answering specific questions. For example, the question "How much energy is required for one molecule to adopt the shape of another?" can be answered by forcing specific atoms to overlap atoms of a template structure during an energy minimization.

Derivatives provide information that can be very useful in minimization procedure. There can be more than one minimum for a large molecule. The minima are called local minima. The ideal solution of geometry minimization is the global minimum. Due to numerical limitations, however, it is impossible to exactly reach the global minimum or even the local minimum. In practice, local minimum refers to a point on the potential energy surface where the applied minimization procedure cannot further reduce the function value. Mostly, the magnitude of the first derivative is a rigorous way to characterize convergence. The minimum has converged when the derivatives are

close to zero. The typical tolerance, for example, in AMBER program is in the range of 10^{-5} to 10^{-6} kcal/mol.Å. To reach the minimum the structure must be successively updated by changing the coordinates (take a step) and checking for convergence. Each complete cycle of differentiation and stepping is known as minimization iteration. The efficiency of minimization can be judge by both the number of iterations required to converge and number of function evaluation needed per iteration. Normally, thousands of iterations are required for macromolecules to reach the convergence.

Two first-order minimization methods, which are frequently used in molecular modeling, are *steepest descents* and *conjugate gradient* methods. Both techniques use the first derivative of the potential function. Additionally, the Newton-Raphson method which uses both the first and the second derivatives to locate the minimum, namely the second-order method, is also widely used.

2.2.1 Steepest descent method

In this method, the energy was minimized by repeating minimization along the direction of the force. The first derivative of the potential energy surface was used with respect to the Cartesian coordinates. The method moves in the parallel direction to the net force by considering moving down the steepest slope of the interatomic forces on the potential energy surfaces. By adding an increment to the negative gradient of the potential energy ($-\nabla V(R)$) or forces, the descent is accomplished (see Figure 2.4).

The method is commonly used in the initial step for relaxing the poorly refined structure either resolved from the crystallography or model building since it reasonably converges at the initial step and requires minimal computing time. However, the progress becomes slow when approaching the minimum. This leads to another minimization method known as conjugate gradient to be used.

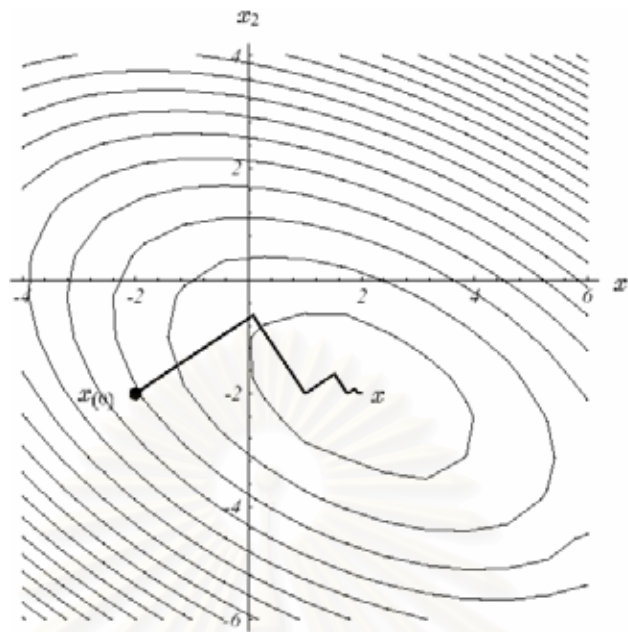


Figure 2.4 The method of steepest descents.

2.2.2 Conjugate gradient method

The conjugate gradient method also uses the first derivatives of the potential energy. But instead of local gradient for downhill as in the steepest descent method, the conjugate gradient technique define the new gradient direction for each iteration by using information from previous gradient directions to determine the optimum direction for the line search. It is considered to select a successive search direction in order to eliminate repeated minimization along the same direction. This made the method more efficient and gives smaller number of iterations to reach the convergence, comparing to the steepest descent method. The conjugate gradient method is displayed in Figure 2.5.

Generally, this method converges in approximately M steps for a quadratic function, where M is the number of degrees of freedom of the function. Note that several terms in the potential energy are quadratic. Nevertheless, the disadvantage is that the line minimizations need to be performed accurately in order to ensure that the conjugate direction is set up correctly and thus time consuming. In addition, the method can be unstable if conformation is so far away from a local minimum.

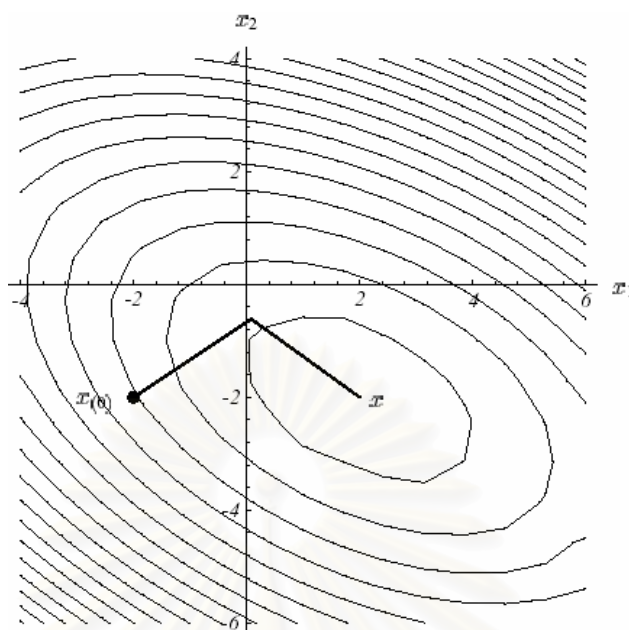


Figure 2.5 The method of conjugate gradient.

2.2.3 Newton-Raphson method

The Newton-Raphson method uses the second-order derivative providing information about curvature of the function, as well as the first derivatives providing the gradient. In addition to the use of the gradient to identify a search direction, the curvature of the function is also applied to predict where the function passes through a minimum along that direction, *i.e.*, to predict where along the gradient the function will change direction. This gives rise to the Newton-Raphson method to converge faster (if the starting geometry is not too far from the minimum) and more accurate with a tolerance up to 10^{-9} kcal/mol.Å in comparing to that of 10^{-6} kcal/mol.Å of the conjugate gradient method. However, for a molecule of N atoms it requires not only the vector of $3N$ first derivatives but also the Hessian matrix of $(3N)^2$ second derivatives, which must be inverted. The need to calculate the Hessian matrix, the iteration makes this algorithm computationally expensive and large storage requirement. In particular, calculating, inverting and storing the second derivative matrix are prohibitive for large systems (usually more than 100 atoms). Furthermore, when a structure is far from the minimum, the minimization can become unstable.

From the above description, it can be seen that different minimization algorithms have different advantage and disadvantages. To optimize the minimization procedure it is usually best to combine several algorithms in the minimization scheme. For instance, the steepest descents method should be used for the first 10-1000 steps of minimization procedure to obtain the configuration close to the local minimum, then the conjugate gradient or Newton-Raphson method (depending on how large the system) is employed to reach the minimum.

2.3 Molecular dynamics simulation

One of the principal tools in the theoretical study of biological molecules is the method of molecular dynamics simulations (MD simulation). This computational method calculates the time dependent behavior of a molecular system. MD simulation has provided detailed information on the fluctuations and conformational changes of proteins and nucleic acids. These methods are now routinely used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes. They are also used in the determination of structures from x-ray crystallography and from NMR experiments.

An experiment is usually made on a macroscopic sample that contains an extremely large number of atoms or molecules sampling an enormous number of conformations, whereas the molecular dynamics simulation generate information at the microscopic level, including atomic positions and velocities. The conversion of this microscopic information to macroscopic observables such as pressure, energy, heat capacities, *etc.*, requires statistical mechanics. Statistical mechanics is fundamental to the study of biological systems by molecular dynamics simulation.

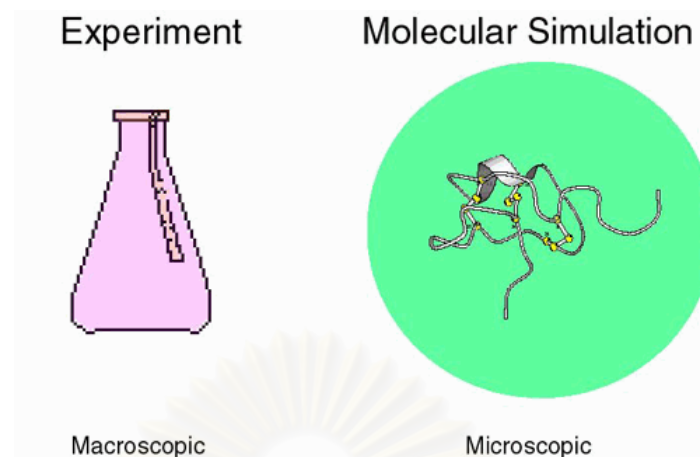


Figure 2.6 The cartoon shown the comparison of experiment and molecular simulation.

In statistical mechanics, averages corresponding to experimental observables are defined in terms of ensemble averages; one justification for this is that there has been good agreement with experiment. An ensemble average is average taken over a large number of replicas of the system considered simultaneously and given by

$$\langle A \rangle_{ensemble} = \iint dp^N dr^N A(p^N, r^N) p(p^N, r^N) \quad (2.9)$$

where $A(p^N, r^N)$ is the observable of interest and it is expressed as a function of the momenta, p , and the positions, r , of the system. The integration is over all possible variables of r and p . There exist different ensembles with different characteristics.

- Microcanonical ensemble (NVE) : The thermodynamic state characterized by a fixed number of atoms, N , a fixed volume, V , and a fixed energy, E . This corresponds to an isolated system.
- Canonical Ensemble (NVT): This is a collection of all systems whose thermodynamic state is characterized by a fixed number of atoms, N , a fixed volume, V , and a fixed temperature, T .
- Isobaric-Isothermal Ensemble (NPT): This ensemble is characterized by a fixed number of atoms, N , a fixed pressure, P , and a fixed temperature, T .

- Grand canonical Ensemble (MVT): The thermodynamic state for this ensemble is characterized by a fixed chemical potential, μ , a fixed volume, V , and a fixed temperature, T .

The molecular dynamics method was first introduced by Alder and Wainwright in the late 1950's (68) to study the interactions of hard spheres. Many important insights concerning the behavior of simple liquids emerged from their studies. The next major advance was in 1964, when Rahman carried out the first simulation using a realistic potential for liquid argon (69). The first molecular dynamics simulation of a realistic system was done by Rahman and Stillinger in their simulation of liquid water in 1974 (70). The first protein simulations appeared in 1977 with the simulation of the bovine pancreatic trypsin inhibitor (BPTI) (71). Today in the literature, one routinely finds molecular dynamics simulations of solvated proteins, protein-DNA complexes as well as lipid systems addressing a variety of issues including the thermodynamics of ligand binding and the folding of small proteins. The number of simulation techniques has greatly expanded; there exist now many specialized techniques for particular problems, including mixed quantum mechanical-classical simulations that are being employed to study enzymatic reactions in the context of the full protein. Molecular dynamics simulation techniques are widely used in experimental procedures such as x-ray crystallography and NMR structure determination.

2.3.1 Basic theory of molecular dynamics

MD simulation technique is based on the classical equation of motion, i.e. Newton's equation (Equation 2.10). Basically, from the force acting on each atom, the acceleration of each atom in the system can then be determined. By integrating the equation of motion, a trajectory describing the position, velocities and the accelerations over time can be obtained. The average values of many properties can then be determined from this trajectory.

$$\vec{F} = m\vec{a} \quad (2.10)$$

Where \vec{F} is the force exerted on particle, m is particle mass and \vec{a} is the acceleration of the particle and equal to the second derivative of the position over time. The force can be expressed as the gradient of the potential energy.

$$\vec{F} = -\nabla V \quad (2.11)$$

where V is the potential energy of the system.

By combining Equation 2.10 and 2.11, the Newton's equation of motion thus relate the derivative of the potential energy to the changes in position as a function of time, Equation 2.12.

$$m \frac{d^2 \vec{r}}{dt^2} = -\frac{dV}{dr} \quad (2.12)$$

Since, the acceleration equals to the derivative of velocity over time and velocity equal to the derivative of position over time, the acceleration can then be related to the position, velocity and time as follow:

$$\begin{aligned} \vec{a} &= \frac{d\vec{v}}{dt} \\ \vec{v} &= \vec{a}t + \vec{v}_0 \\ \vec{v} &= \frac{d\vec{x}}{dt} \\ \vec{x} &= \vec{v}t + \vec{x}_0 \\ \vec{x} &= \vec{a}t^2 + \vec{v}_0 \cdot t + \vec{x}_0 \end{aligned} \quad (2.13)$$

where x is the position of the particle, x_0 and v_0 are initial position and velocity, respectively. To calculate the trajectory, the initial position of atoms, initial distribution of velocities and the acceleration are needed. The acceleration can be determined from the gradient of the potential energy while the initial position can be obtained from experimentally resolved structures. The initial distribution of velocities is resolute from a Maxwell-Boltzmann distribution at a given temperature (Equation 2.14)

$$p(v) = \left(\frac{m}{2\pi k_B T} \right)^{\frac{1}{2}} \exp \left[-\frac{1}{2} \frac{mv^2}{k_B T} \right] \quad (2.14)$$

$$T = \frac{1}{(3N)} \sum \frac{|P|^2}{2m}$$

where N is the number of atoms in the system.

For this purpose we need to be able to carry out the forces F that acting on the atoms and these are usually derived from a potential energy $V(R)$, where $R = (r_1, r_2, \dots, r_N)$ represents the complete set of $3N$ atomic coordinates. The potential energy is a function of the atomic positions ($3N$) of all the atoms in the system. Due to the complicated nature of this function, there is no analytical solution to the equations of motion. They must be solved numerically. Numerous numerical algorithms have been developed for integrating the equations of motion.

Time integration algorithms are required to integrate the equation of motion of the interacting particles and follow their trajectories. They are the engine of a molecular dynamics program, which are based on finite difference methods. The equations are solved step-by-step in discrete time interval that is called *time step*, δt . Knowing the positions and some of their time derivatives at time t , the integration scheme gives the same quantities at later time $t + \delta t$. By iterating the procedure, the time evolution of the system can be followed for long times. There are many integration methods widely used in a molecular dynamics simulation.

2.3.2 Integration algorithms (66)

The integration algorithms approximated by a Taylor series expansion (Equation 2.15) was introduced to solve the equation numerically as for the complicated system (function of atomic position) that dose not have analytical solution.

$$\begin{aligned}
x(t + \delta t) &= x(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \frac{1}{6}b(t)\delta t^3 + \dots \\
v(t + \delta t) &= v(t) + a(t)\delta t + \frac{1}{2}b(t)\delta t^2 + \frac{1}{6}c(t)\delta t^3 + \dots \\
a(t + \delta t) &= a(t) + b(t)\delta t + \frac{1}{2}c(t)\delta t^2 + \dots
\end{aligned} \tag{2.15}$$

where x is the position, v is the velocity (the first derivative with respect to time), a is the acceleration (the second derivative with respect to time).

In molecular dynamics, the most commonly used time integration algorithm is probably the so-called Verlet algorithm. The basic ideal is to two third-order Taylor explanations for the position $x(t)$, one forward $x(t+\delta t)$ and one backward $x(t-\delta t)$ in time. The relation can be written as,

$$\begin{aligned}
x(t + \delta t) &= x(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \dots \\
x(t - \delta t) &= x(t) - v(t)\delta t + \frac{1}{2}a(t)\delta t^2 - \dots
\end{aligned} \tag{2.16}$$

Summing these two equations, one obtains

$$x(t + \delta t) = 2x(t) - x(t - \delta t) + a(t)\delta t^2 \tag{2.17}$$

This is the basic form of the Verlet algorithm. Because we are integrating Newton's equation, acceleration, $a(t)$, is just the force divided by the mass (Equation 2.10). As one can see the truncation error of the algorithm when evolving the system by δt is of the order of four, even if third derivatives do not appear explicitly. This algorithm is at the same time simple to implement, accurate and stable, explaining its large popularity among molecular dynamics simulations.

The advantages of the Verlet algorithm are, *i*) it is straightforward, and *ii*) the storage requirements are modest. The disadvantage is that the algorithm is of moderate precision. A problem with this version of the Verlet algorithm is that velocities are not directly generated. While they are not needed for the time evolution, their

knowledge is sometimes necessary. Moreover, they are required to compute the kinetic energy K , whose estimate is necessary to test the conservation of the total energy $E=K+V$. This is one of the most importance tests to verify that a MD simulation is proceeding correctly. One could compute the velocities from the position using

$$v(t) = \frac{x(t + \delta t) - x(t - \delta t)}{2\delta t} \quad (2.18)$$

To overcome this difficulty, some variants of the Verlet algorithm have been developed. They give rise to exactly the same trajectory, and the stored variables are different in memory and at what times. The leap frog algorithm (Equation 2.19) is one of such variants where velocities are handled somewhat well. In this algorithm, the velocities are first calculated from $t + \frac{1}{2}\delta t$ and then used to calculate the position x at time $t + \delta t$. The velocities at time t can then be approximated by Equation 2.20.

$$\begin{aligned} x(t + \delta t) &= x(t) + v(t + \frac{1}{2}\delta t)\delta t \\ v(t + \frac{1}{2}\delta t) &= v(t - \frac{1}{2}\delta t) + a(t)\delta t \end{aligned} \quad (2.19)$$

The advantage of this algorithm is that the velocities are explicitly calculated, however, the disadvantage is that they are not calculated at the same time as the positions. The velocities at time t can be approximated by the relationship:

$$v(t) = \frac{1}{2} \left[v(t - \frac{1}{2}\delta t) + v(t + \frac{1}{2}\delta t) \right] \quad (2.20)$$

An improved integrating, which was also used in simulations, is the Velocity Verlet algorithm which is designed to further improve on the velocity evaluations. The positions, velocities and accelerations at time $t+\delta t$ are obtained from the same quantities at time t in the following way:

$$\begin{aligned}
 x(t + \delta t) &= x(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) \\
 v(t + \delta t) &= v(t) + \frac{1}{2} \delta t [a(t) + a(t + \delta t)]
 \end{aligned}
 \tag{2.21}$$

The advantage is the best evaluation of velocities but there is also the disadvantage which is computationally more expensive than other integration algorithm (Verlet and Leap-Frog).

2.3.3 The basic steps in MD simulation

The MD method is deterministic in which the state of the system at any time are predictable once positions and velocities of each atom are known. MD simulations are sometime time consuming and computational expensive. Nevertheless, the faster and cheaper of the computer today bring up the calculation to the nanosecond time scale.

The basic step in MD simulation was shown in Figure 2.7. The initial system, in which the coordinated can be normally obtained from x-ray or NMR data or built up using molecular modeling method, is minimized in order to get rid of bad atomic contacts due to the addition of some newly atoms and residues, *i.e.*, hydrogen atoms, amino acid residues as well as water molecules. The initial velocities of the system were assigned to the system using Maxwell-Boltzmann distribution as described above. Prior to the production dynamics, the heating and equilibration dynamics were performed with the aim to decreases the probability that localized fluctuations in the energy such as hot spots in which will persist throughout the simulation. Following the initial thermalization/equilibration, some more period of time for equilibrations were applied until the system is reached the equilibrium by adjusting the temperature and rescaling the velocities. The trajectory of the systems is calculated and used for analysis.

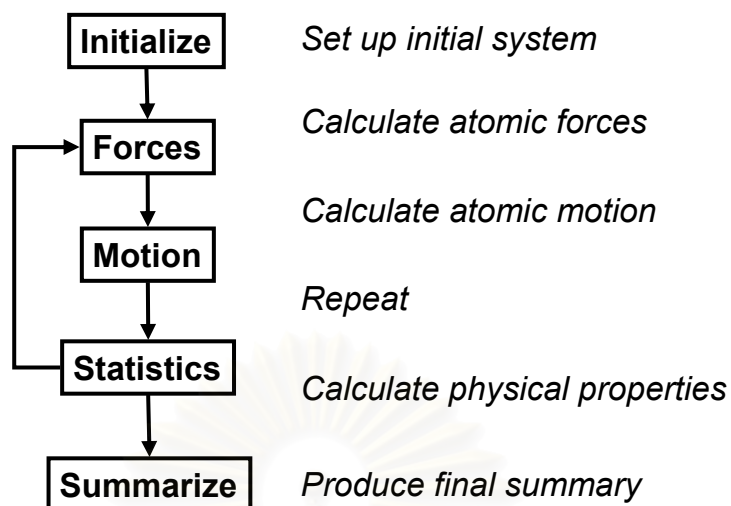


Figure 2.7 Basic steps on MD simulation study.

2.3.4 Periodic boundary condition

The periodic boundary condition was introduced in the MD simulations for reducing the boundary effect at the edge of the simulation box where the molecules which stay close to the edge of the box receive the incomplete interaction (one side interaction). The periodicity was applied to the system of interest in which locating at the central cell. The interactions of the molecule nearby the edge of the box were then complete by involving the interaction from the next box. Furthermore, the number of molecules was kept constant by entering of the image of the leaving cell.

This ideal is represented in Figure 2.8. The original box contains a solute and solvent molecules which is surrounded with identical images of itself. The positions and velocities of corresponding particles in all of the boxes are identical. The common approach is to use a cubic or rectangular parallelepiped box, but other shapes are also possible, e.g., truncated octahedron. By this approach, we obtain what is in effect an infinite sized system.

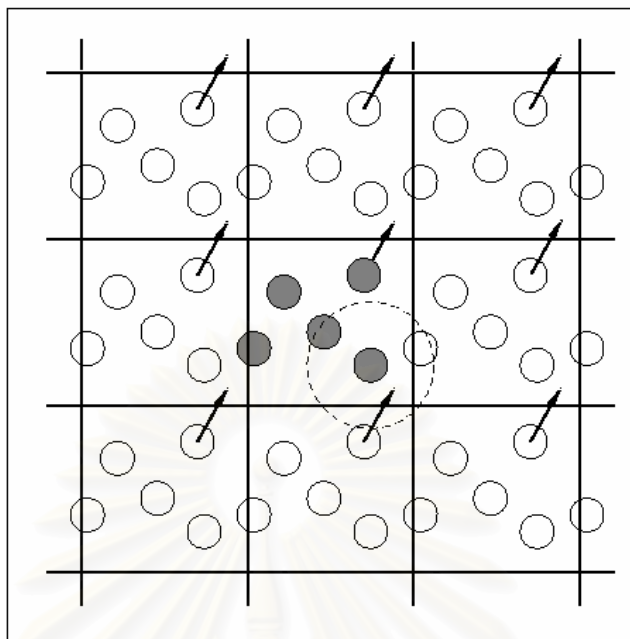


Figure 2.8 Periodic boundary conditions. As a particle moves out of the simulation box, an image particle moves in to replace it. In calculating particle interactions within the cutoff range, both real and image neighbors are included. (66)

2.3.5 Potential cut-off for non-bonded interaction

The calculation of non-bonded interactions is the most time-consuming part in the MD simulation since all pair interactions are calculated for every pair of atoms in the system. In order to reduce the calculation time, the total force acting on a particle from neighboring particles are of the main contribution. The interactions are evaluated between each pair of particles with a distance less than a cut-off radius, r_{co} . This compromises between the correction and efficiency.

2.3.6 Treatment of long range interaction

Presently simulations of bimolecular systems include a very large numbers of atoms ten to hundred of thousands atoms. They are involved over timescales of many nanoseconds. The accurate computation of electrostatic and van der Waals interactions is the most difficult task in computer modeling. The most time consuming part of any molecular dynamics simulation is the calculation of the

electrostatic interactions. These interactions fall off as $1/r$, where r is the separation between charges, and have consequently to be considered as long-range.

The need for efficient treatment of long-range electrostatic interactions in the molecular size has been clearly well-known in the last decade. $E_{electrostatic}$ represents a sum over $N(N-1)/2$ pairs; i.e., it is an $O(N^2)$ calculation necessary to evaluate all non-bonded pair in macromolecular systems. The treatment long-range forces were ignored in the macromolecular simulations, being with the use of 'truncation' or 'cutoff', r_{off} . These methods were developed to limit the computational effort needed by the evaluation of the long-range forces. However, there is problem to select the truncation technique. One can use a straight truncation method, which the electrostatic interactions are zero at r_{off} . The truncation simulations can perform in the old version of AMBER package. Another truncation method, the shifting function $S(r)$, scales the interaction potential to zero at the specific distance

$$S(r) = \begin{cases} (1 - (r/r_{off})^2)^2, & r \leq r_{off} \\ 0, & r > r_{off} \end{cases} \quad (2.22)$$

One can see that the short-range interactions are disturbed. The distortion and overestimation of the short-range interactions are the drawbacks of the shifting function scheme. The switching function is other smoothing scheme, which brings the potential to zero between a $switches_{on}$ and $switches_{off}$ distance. With a suitable $switches_{on}$ the short-range interactions are not distorted, giving continuous force or potential energy. Although the truncation methods can significantly reduce the amount of computational time for evaluating the electrostatic interactions, these methods are inaccurate because of finite cutoff distance which restricts severely the infinite character of the system. This may result an unstable geometry for a long simulation.

In order to improve the treatment of the long-range electrostatic interaction, the Particle Mesh Ewald (PME) method was used to calculate the full

electrostatic energy of a periodic box in a macroscopic lattice of repeating images. The method is based on the Ewald summation method and particle mesh method, which gives the exact result for the electrostatic energy of a periodic system containing an infinite replicated neutral box of charged particles. The method is usually used for the complex molecular system with periodic boundary condition.

Consider the Coulomb's law:

$$V_{r_i} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.23)$$

where $\epsilon = 4\pi\epsilon_0$, the Ewald method split the Coulomb potential which is slowly converged into three exponentially converged contributions as described in Equation (2.24)

$$V_t = V_f + V_r + V_s \quad (2.24)$$

where V_t , V_f , V_r and V_s are potential for total, Fourier Space part, Real Space part and Self interaction, respectively. Each contributor is given as follows:

$$\begin{aligned} V_f &= \frac{1}{2v\epsilon_0} \sum_{k \neq 0} \frac{\exp\left[-k^2 / 4\alpha^2\right]}{k^2} \left| \sum_{i=1}^{i=N} \exp\left[-ik \cdot r_i\right] \right|^2 \\ V_r &= \frac{1}{4\pi\epsilon_0} \sum_{i < j} \frac{q_i q_j \operatorname{erfc}(\alpha r_{ij})}{r_{ij}} \\ \operatorname{erfc}(x) &= \frac{2}{\sqrt{\pi}} \times \int_x^\infty dt \exp[-t^2] \\ V_s &= -\frac{1}{4\pi\epsilon_0} \sum_{i=1}^{i=N} \frac{q_i^2 \alpha}{\sqrt{\pi}} \end{aligned} \quad (2.25)$$

Application of the PME method allows fast Fourier transform (FFT) to be applied, thus, the method is better. The advantages of the PME are the method for the treatment of long-range forces in macromolecular systems. The high accuracy can be

obtained with relatively little increase in computational cost and efficiently implemented into usual MD algorithm such as AMBER package.

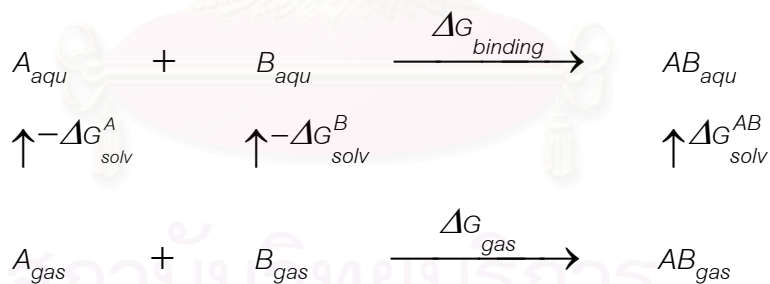
2.4 The molecular mechanics Poisson-Boltzmann surface area (MM/PBSA)

The MM/PBSA approach represents the post processing method to evaluate free energies of binding or to calculate absolute free energies of molecules in solution. The sets of structures are usually collected with molecular dynamics or Monte Carlo methods. However, the collections of structure should be stored in the format of an AMBER trajectory file.

In general, the free energy of the inhibitor binding, $\Delta G_{binding}$, is obtained from the difference between the free energy of the receptor-ligand complex (G_{cpx}), and the unbound receptor (G_{rec}) and ligand (G_{lig}) as follows:

$$\Delta G_{binding} = \Delta G_{cpx} = G_{cpx} - (G_{rec} + G_{lig}) \quad (2.26)$$

In the MM/PBSA methodology, the free energy of $A + B \rightarrow AB$ is calculated using the following thermodynamic cycle:



$$\begin{aligned}
 \Delta G_{binding} &= \Delta G_{gas} - \Delta G_{sol}^A - \Delta G_{sol}^B + \Delta G_{sol}^{AB} \\
 &= \Delta H_{gas} - T\Delta S - \Delta G_{PBSA}^A - \Delta G_{PBSA}^B + \Delta G_{PBSA}^{AB} \\
 &= \Delta H_{gas} - T\Delta S + \Delta\Delta G_{PB} + \Delta\Delta G_{SA} \quad (2.27)
 \end{aligned}$$

$$\Delta G_{gas} = \Delta H_{gas} - T\Delta S \quad (2.28)$$

$$\Delta H_{gas} \approx \Delta E_{gas} = \Delta E_{int ra} + \Delta E_{electrostatic} + \Delta E_{vdW} \quad (2.29)$$

$$\Delta \Delta G_{PB} = \Delta G_{PB}^{AB} - (\Delta G_{PB}^A + \Delta G_{PB}^B) \quad (2.30)$$

$$\Delta \Delta G_{SA} = \Delta G_{SA}^{AB} - (\Delta G_{SA}^A + \Delta G_{SA}^B) \quad (2.31)$$

ΔG_{gas} is the interaction energy between A and B in the gas phase that related to the enthalpic changes in the gas phase and ΔG_{solv}^A , ΔG_{solv}^B and ΔG_{solv}^{AB} are the solvation free energies of A, B and AB, which are estimated using a continuum approach (Poisson-Boltzmann/surface area),

$$\Delta G_{solv}^{AB} = \Delta G_{PBSA}^{AB} = \Delta G_{PB}^{AB} + \Delta G_{SA}^{AB} \quad (2.32)$$

where ΔG_{PB}^{AB} is the polar solvation energy, which is computed in continuum solvent, usually using a finite-difference Poisson-Boltzmann (PB) model and ΔG_{SA}^{AB} is nonpolar solvation term, which can be derived from the solvent-accessible surface area (SA).

For a series of compounds with similar structures and binding modes, the entropy contribution can be omitted if one is only interested in the relative order of binding affinities. However, if one intends to obtain the absolute binding free energy or the compounds are very different, the contribution of entropy cannot be neglected. There are several approaches to estimate the entropy, including the normal-mode analysis, the quasi-harmonic analysis, and the quasi-Gaussian approach. The first two approaches, which are more practical for the biological systems, are much similar in theory except that the atomic fluctuation matrix in quasi-harmonic analysis is not from a normal mode calculation, but is obtained from the snapshots of MD simulations. Estimates of entropy from normal-mode analysis have some limitations, for example, the anharmonic contribution is not taken into account and low frequency modes that lead to large displacements are not treated accurately in the harmonic limit. On the other hand, quasi-harmonic analysis, which is an alternative to harmonic analysis, incorporates some effects due to the anharmonic nature of macromolecules. However, there are some serious problems for the quasi-harmonic analysis if the microstates are not well sampled

with MD simulations. Considering that normal-mode analysis has been successfully applied in estimating the binding entropy for several biological systems. By now, MM/PBSA has been successfully applied to binding free energy calculations for several systems.

In AMBER program, the molecular mechanical energies are determined with the SANDER module and represent the internal energy (bond, angle and dihedral), van der Waals and electrostatic interaction. An infinite cutoff for all interactions is used. The electrostatic contribution to the solvation free energy is calculated with the Poisson-Boltzmann (PB) method, for example, as implemented in Delphi program (72) or by generalized Born (GB) method implemented in SANDER module in AMBER program. The hydrophobic contribution to the solvation free energy is determined is determined with solvent accessible surface area dependent term (73). The surface area is computed with Paul Beroza's molsurf program, which based on analytical ideas primarily developed by Mike Connolly. Estimates of conformational entropies can be made with the nmode module from AMBER.

2.4.1 Poisson-Boltzmann (PB) model (30)

Calculations of electrostatic potentials required for evaluating the ionization and interaction energies are based on the Poisson-Boltzmann model of electrostatic interactions. The foundation of the PB model constitutes the fundamental equation of electrostatics, differential Poisson equation that describes the electrostatic potential $\Phi(r)$ in a medium with a dielectric scalar field $\epsilon(r)$ and with a charge density $\rho(r)$.

$$\nabla \cdot \epsilon(r) \nabla \Phi(r) = -4\pi\rho(r) \quad (2.33)$$

When one deals with a macromolecule (in this case the protein) immersed in an aqueous medium with mobile ions, the charge density $\rho(r)$ can be separated into two components:

$$\rho(r) = \rho_{int}^f + \rho_{ext}^m \quad (2.34)$$

where ρ_{int}^f describes the interior charge distribution of the fixed (f) positions of all the charges in the protein and ρ_{ext}^m is a mobile (m) exterior charge density modeled by a Boltzmann distribution. Equation 2.33 takes the form

$$\nabla \cdot \varepsilon(r) \nabla \Phi(r) = -4\pi \left[\rho_{int}^f(r) + \lambda(r) \sum_i q_i n_i \exp(-q_i \Phi(r) / kT) \right] \quad (2.35)$$

where n_i is the bulk number density of type- i ions, q_i is the ionic charge, k is the Boltzmann constant, and λ equals 1 for ion-accessible regions and 0 elsewhere.

Linearizing the exponential terms in the mobile charge distribution and assuming equal number of positively and negatively charged ions one obtains

$$\nabla \cdot \varepsilon(r) \nabla \Phi(r) = -4\pi \left[\rho_{int}^f(r) + \lambda(r) \sum_i q_i n_i q_i \Phi(r) / kT \right] \quad (2.36)$$

Next, the modified Debye-Hückel parameter is introduced

$$\bar{\kappa}^2 = \frac{4\pi \sum_i n_i q_i^2}{kT} = \frac{8\pi e^2 N_A I}{1000kT} \quad (2.37)$$

where N_A is the Avogadro constant, e the charge of an electron, and I the ionic strength. The ionic strength is represented as

$$I = \frac{1}{2} \sum_i c_i z_i^2 \quad (2.38)$$

where c_i and z_i are the molar concentration and the valency of the mobile ions, respectively.

Introducing $\bar{\kappa}^2$ in Equation 2.37, one obtains the linearized Poisson-Boltzmann equation where the protein is represented as a low dielectric medium containing fixed charges and the solvent is represented as a medium of dielectric

constant of 80 which contains mobile ions that screen the fixed charges according to the Debye-Hückel model:

$$\nabla \cdot \epsilon(r) \nabla \phi(r) = -4\pi\rho_{int}(r) + \lambda(r)\bar{\kappa}^2\phi(r) \quad (2.39)$$

The charges are usually located at nucleic positions determined by x-ray crystallographic methods or high-resolution Nuclear Magnetic Resonance (NMR) spectroscopy.

2.4.2 Solvent accessible surface area (SASA) (74)

The accessible surface is traced out by the probe sphere center as it rolls over the protein. It is a kind of expanded van der Waals surface. If one increases each atom's van der Waals radius by the probe radius, one get so-called expanded atom radii. The union of the expanded atoms is what Tim Richmond (1984) calls the solvent-excluded volume. It is the region enclosed by the accessible surface.

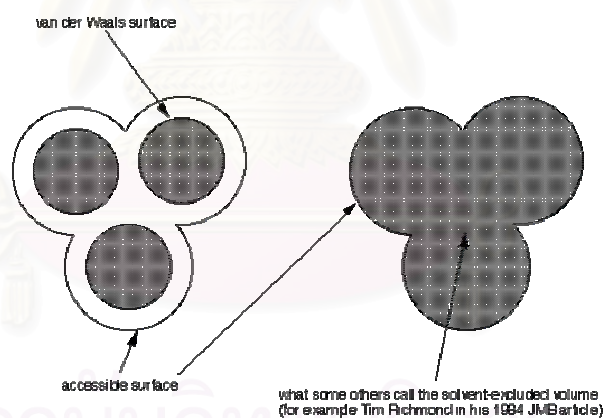


Figure 2.9 Schematic of the SASA

CHAPTER 3

MD SIMULATIONS OF DIFFERENT PROTONATION STATES OF G48V HIV-1 PR COMPLEXED WITH SAQUINAVIR IN AQUEOUS SOLUTION

The ability of protein residues to exchange protons with their environment and the dependence of this process on pH influence the electrostatic properties of enzymes. At extreme values of pH , for instance, proteins unfold and enzymes cease to function. The rate of enzymatic catalysis and also the binding of inhibitors, as has been observed experimentally, depend upon pH . These processes are, therefore, thermodynamically linked to protonation states of the system. Protonation states of ionizable residues determine the overall charge distribution of enzymes. This distribution may play a crucial role in the reaction path, *e.g.*, drawing the substrate near the active site or stabilizing possible transition states. Also the appropriate protonation state of active site residues is often required to activate an enzyme because most enzymes contain residues that promote acid-base catalysis. Moreover, the protonation state of ionizable residues surrounding the active site may also be important since their electrostatic field can stabilize transition states of the reaction path.

The properties of HIV-1 PR are also affected by pH . The activity of this enzyme varies with pH and is peaked at $pH \sim 5 - 6$ (75). Studies of the pH dependence of the kinetics of HIV-1 PR have been used to deduce the titration behavior of the active site aspartyl dyad, Asp25 and Asp25'. These studies yield two distinct pK_a s for the dyad in the free enzyme (76). Competitive inhibitors of HIV-1 PR bind in the active site, displacing water and forming contacts with the carboxyl groups of the dyad. As a consequence, they also can alter the pK_a s of the dyad. Thus, studies of protonation states of dyad of HIV-1 PR are essential for further modeling of its properties, *e.g.*, molecular dynamics for both wild-type and mutant type.

Up to date, many previous studies proposed the protonation state of catalytic dyad of HIV-1 PR as shown in Table 3.1. Different protonation states were found depending on the local environment of the enzyme-inhibitor complex. The single

protonation state at one of two acidic acids has been most commonly observed with the binding of inhibitor. However, the calculations of ionization constants of titratable residues in peptides or proteins proposed that the crucial aspartates in the active site of the enzyme may have different protonation states depending on the type of the bound ligand (30). Unfortunately, the information about protonation state of saquinavir and ritonavir complexes were not available. Thus, molecular dynamics simulations of the G48V HIV-1 PR complexed with saquinavir in explicit aqueous solution were carried out in this study for all three possible protonation states of the two aspartic residues, Asp25 and Asp25'.

Table 3.1 The proposed protonation states of HIV-1 PR from the previous studies.

Year	System of interest	Methods	Proposed results	Ref.
1995	HIV-1 PR complexed with the S- and R-isomers of the U85548E inhibitor	The free energy simulation	Only one catalytic aspartic acid residue is protonated	(77)
1996	HIV-1 protease complexed with pepstatin inhibitor	NMR measurements	At least, monoprotated	(78)
2000	HIV-1 PR complexed with hydroxyethylene isostere inhibitors	The free energy simulation	Only the case with Asp25 OD2 protonation results in the reasonable binding affinity.	(79)
2001	HIV-1 PR complexed with Pepstatin A inhibitors	<i>Ab Initio</i> Molecular Dynamics-Based Assignment	Both Asp groups in the complex are protonated.	(80)
2003	HIV-1 PR complexed with A74704 inhibitor.	Quantum mechanics and molecular mechanical method.	Protonated state of complex is monoprotated state on OD2 of Asp25.	(81)

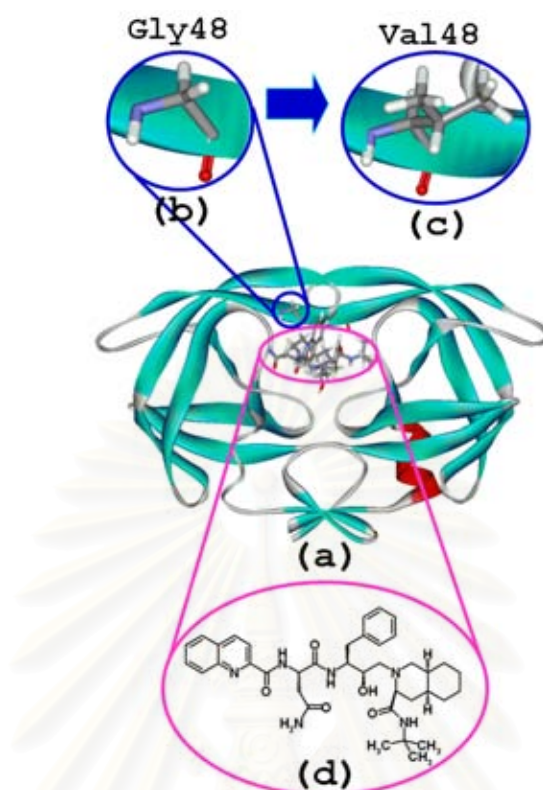


Figure 3.1 Schematic representation of the HIV-1 PR complexed with saquinavir (a) in which molecular structure of Gly48 (b) and Val48 (c) in the G48V and of saquinavir (d) were also displayed.

3.1 Computational methods

3.1.1 Atomic charges of saquinavir

Saquinavir is not standard residues in the AMBER residue libraries; therefore, their force field parameters are not available in the AMBER package. Thus, potential parameters for saquinavir were developed using the following steps. Geometry of the starting structure of saquinavir was taken from the crystal structure of saquinavir bound to wild-type protease (PDB code: 1HXB). Then, optimized at the Hartree-Fock level with 6-31G** basis functions to adjust the bond-length involving hydrogens. Finally, the RESP charge fitting procedure was employed to calculate partial atomic charges of the inhibitor (82). All electronic structure calculations were carried out with the program package Gaussian98 (83). The results were summarized in Table 3.2.

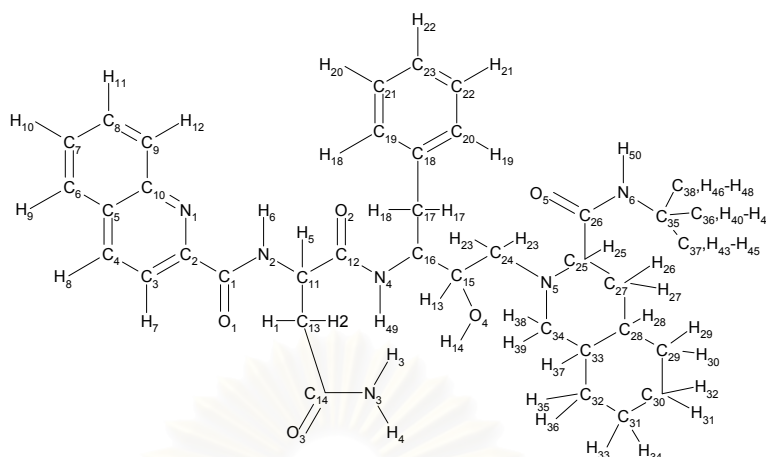


Figure 3.2 The atom labels of saquinavir.

3.1.2 The pK_a calculations of HIV-1 PR

An assumption used for assigning the protonation state of the ionizable residues in the simulations was inspected by the prediction of the pK_a values. The method estimates the pK_a shift by calculating the electrostatic free energy of ionizable residues in the neutral and the charge states in solution (84). The computations were done by solving finite different Poisson-Boltzmann equations implemented in the University of Houston Brownian Dynamics program (UHBD) (85). The protocols describe as follows. Polar hydrogens were added to the x-ray model using the Python based Molecular Viewer (PMV) (86). For generating electrostatic potentials, the model was then placed in a $65 \times 65 \times 65$ dimension with a grid spacing of 2.5 Å. The focusing technique was additionally employed using finer grid spacing of 1.2, 0.75, and 0.25 Å for a cubic dimension of 15, 15, and 20, respectively (84,87). Atomic radii and charges available in UHBD program were originally derived from optimized potentials for liquid simulations and CHARMM22 parameter sets (88). The 1.4 Å probe radius with a resolution of 500 dots/atomsphere was used. The calculations employed a solvent dielectric of 80 with 150 mM ionic strength, and a temperature of 298 K. A dielectric constant of HIV-1 PR was examined by varying to 1, 4, and 20. We found that a protein dielectric constant of 20 produced the best pK_a prediction. A dielectric constant of 1 and 4 yielded unusual pK_a values due to an overestimation of electrostatic potentials. This phenomenon is thoroughly discussed in an early work (84).

Table 3.2 RESP charges of saquinavir molecule obtained from single point HF/6-31G* calculation on the HF/6-31G** optimized geometry.

Atom label	Charge	Atom label	Charge	Atom label	Charge
C9	-0.377	H5	0.116	C35	0.876
C10	0.592	C12	0.970	C36	-0.531
N1	-0.636	O2	-0.686	C37	-0.531
H12	0.201	N4	-1.102	H43-H45	0.121
C8	-0.074	H49	0.409	C38	-0.531
H11	0.144	C16	1.005	H46-H48	0.121
C7	-0.151	C17	-0.443	H50	0.342
H10	0.149	C18	0.193	H25	0.094
C6	-0.226	C19-C20	-0.201	C27	-0.322
H9	0.164	C21-C22	-0.139	H26-H27	0.131
C5	-0.048	C23	-0.156	C28	0.106
C4	-0.016	H18-H19	0.145	H28	0.006
H8	0.149	H20-H21	0.138	C29	-0.170
C3	-0.414	H22	0.132	H29-H30	0.033
H7	0.199	H16-H17	0.113	C30	0.104
C2	0.359	H15	-0.063	H31-H32	-0.008
C1	0.727	C15	0.014	C31	-0.043
O1	-0.573	O4	-0.743	H33-H34	0.009
N2	-0.785	H14	0.460	C32	-0.163
H6	0.358	H13	0.091	H35-H36	0.098
C11	0.184	C24	-0.040	C33	0.034
C13	-0.745	H23-H24	0.077	H37	0.030
C14	1.088	N5	-0.575	C34	0.040
O3	-0.653	C25	0.054	H38-H39	0.037
N3	-1.080	C26	0.682	H1-H2	0.206
H3-H4	0.420	O5	-0.528	N6	-0.880

3.1.3 Preparation of the initial structures

In order to investigate the relative dynamics properties among different protonation states of the mutant HIV-1 PR complexed with the inhibitor (saquinavir), three MD simulations were performed, monoprotionate on Asp25 (Mono-25), monoprotionate on Asp25' (Mono-25') and diprotionate on both aspartic acids (Di-pro). The crystal structure of saquinavir bound to wild-type protease was taken from the Protein Data Bank PDB (1HXB) (89) and used as the reference structure. In the 1HXB crystal structure, two orientations of the drug molecule are presented (each with roughly 40-50% occupancy). The first orientation of saquinavir molecule was chosen. This orientation of saquinavir in complex with HIV-1 PR fits very well to the definition of chains A and B given by Krohn *et.al.*, who published this crystals structure (89). This structure was commonly used in almost all publications.

All missing atoms of the protein were added using the LEaP module in the AMBER 7 software package (90). Hydrogen atoms were then added to the two catalytic aspartic residues in order to generate the Mono-25, Mono-25' and Di-pro states using the LEaP module in AMBER 7 software package. The Mono-25, Mono-25' and Di-pro systems were solvated by TIP3P water molecules (91). The crystallographic waters were also included in the simulations. The sodium and chloride ions were added to neutralize the system.

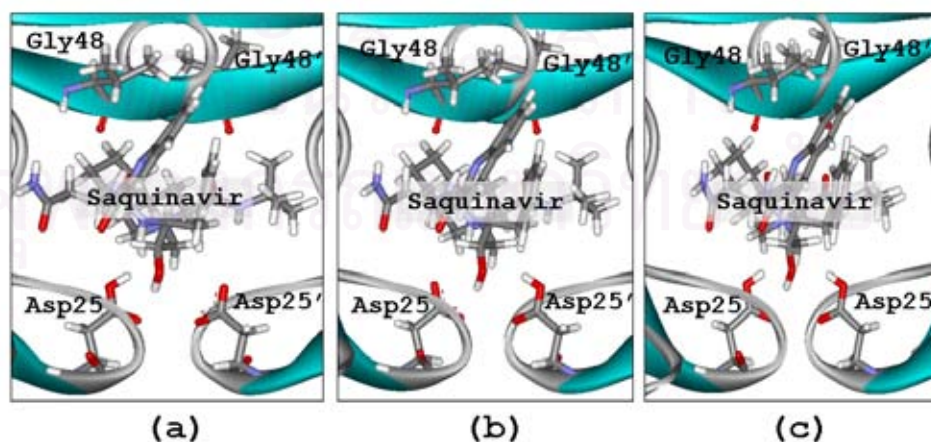


Figure 3.3 Saquinavir complexed with Mono-25 (a), Mono-25' (b) and Di-pro (c) HIV-1 PR.

It should be noted that the x-ray structure of the double mutant, G48V/L90M-SQV complex (1FB7) could be considered as an alternative template (92). However, the x-ray coordinates of the second monomer of the double mutant are not available. Thus, 1HXB is considered to be more appropriate as a template. The mutant protease enzyme was modeled from this structure replacing glycine by valine at residue 48 (G48V) using the AMBER 7.

Table 3.3 Number of water molecules, counter ions and total atoms together with the dimension of simulation boxes of the three simulations, Mono-25, Mono-25' and Di-pro.

system	Number of			Box dimension (Å ³)
	Waters	Counter ions	Total atoms	
Mono-25	9,633	8 Cl ⁻ , 3 Na ⁺	32,150	76 × 78 × 68
Mono-25'	9,627	8 Cl ⁻ , 3 Na ⁺	32,134	73 × 77 × 68
Di-pro	9,627	8 Cl ⁻ , 3 Na ⁺	32,133	76 × 78 × 68

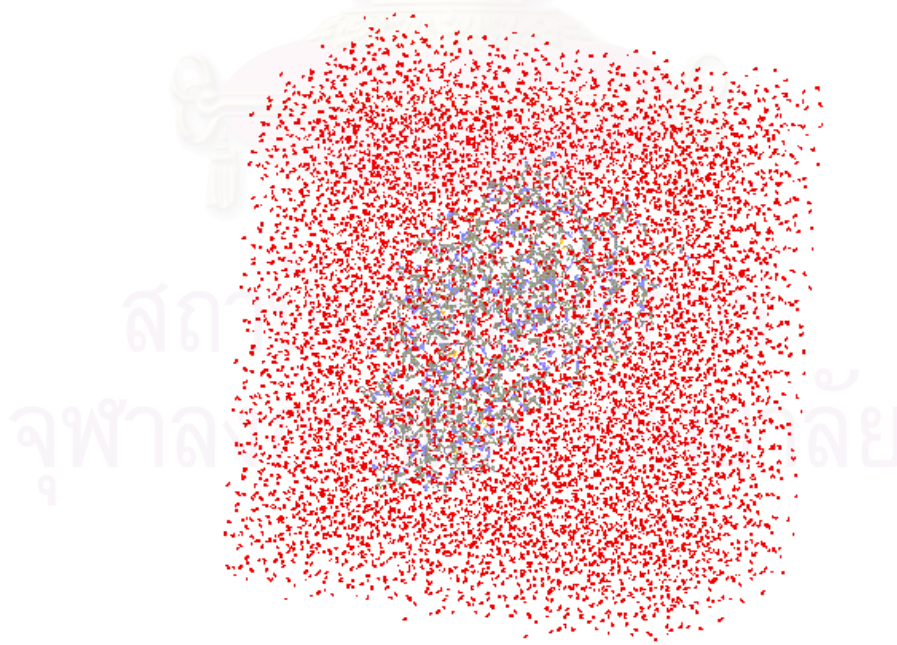


Figure 3.4 The complex of G48V HIV-1 PR and saquinavir in solvate box.

3.1.4 Molecular dynamics simulations

Three MD simulations were carried out using the SANDER module of AMBER 7 (93) with the Cornell force for the mutant, G48V, HIV-1 PR complexed with saquinavir in the above mentioned states. Before starting the MD simulations, two steps of energy minimizations were performed: (i) only water molecules were subjected to energy minimization using 3,000 steps for steepest descent method and then 2,000 steps for conjugate gradient method. (ii) The whole systems energy minimization was employed with 3,000 steps for steepest descent method and then 2,000 steps for conjugate gradient method. It should be noted that position-restrained minimizations of some particular regions were carried out for systems that clashed during minimization because of incidental overlay of atoms. This procedure was repeated until there was no sign of bad contacts.

The MD simulations were performed employing the periodic boundary condition with the NPT ensemble. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and pressure of the systems (94). The SHAKE algorithm (95) was employed to constrain all bonds involving hydrogens. The simulation time step of 1 fs was used. All MD simulations were run with a 12 Å residue-based cutoff for nonbonded interactions and the particle-mesh Ewald method was used for an adequate treatment of long-range electrostatic interactions (96).

The simulation consists of thermalization, equilibration, and production phases. Initially, the temperature of the system was gradually heated from 0 to 298 K during the first 60 ps. Then, the systems were maintained at 298 K until MD reached 400 ps of the simulation. Finally, the production phase started from 400 ps to 1 ns. The convergence of energies, temperature, and global RMSD was used to verify the stability of the systems. The MD trajectories were collected every 0.1 ps. The 600 ps trajectory of the production phase was used to calculate the average structure. All MD simulations were carried for 1 ns. MD trajectories were evaluated in term of RMSD, distances, torsion angles using the CARNAL and Ptraj modules of AMBER 7.

3.1.5 Quantum chemical calculations

Quantum chemical calculations were performed to investigate the interaction energy between catalytic residues (Asp25 and Asp25') and saquinavir in the different protonated states of the complexes using the Gaussian 98 program (83). Initial structures of the complex in the three protonation states were averaged from the whole trajectory, then; the geometry was relaxed using energy minimization from the AMBER 7 program. The selected residues, Asp25 and Asp25', were capped by CH₃NH- and -COCH₃ groups at the C- and the N-terminal, respectively. The geometry of the hydrogen atoms in the cap methyl group were optimized using semi-empirical calculations at the PM3 level. Then, single point calculation with the density functional theory B3LYP was applied to investigate the total energy of the system. The basis sets used are the 6-31G(d,p) for unprotonated residue and the extended 6-31+G(d,p) for protonated residue. The interaction energies between saquinavir and the catalytic dyad residues (ΔE_{cpx}) was calculated following equation:

$$\Delta E_{cpx} = E_{cpx} - (E_{Asp25+Asp25'} + E_{SQV}) \quad (3.1)$$

where E on the right hand side represents total energy of the systems given in the subscript (cpx = complex and SQV = saquinavir) and $E_{Asp25+Asp25'}$ is total energy of the Asp25 and Asp25' dimer. Evaluations were also performed to understand the contribution of pair interactions, $\Delta E_{Asp25-SQV}$ and $\Delta E_{Asp25'-SQV}$ to the ΔE_{cpx} i.e.,

$$\Delta E_{cpx} = \Delta E_{Asp25-SQV} + \Delta E_{Asp25'-SQV} + E_{3bd} \quad (3.2)$$

where the ΔE_{3bd} denotes an error due to three-body correction. In other words, the third particle was excluded from the calculation of the pair interactions which are defined as

$$\left. \begin{aligned} \Delta E_{Asp25-SQV} &= E_{Asp25+SQV} - (E_{Asp25} + E_{SQV}) \\ \Delta E_{Asp25'-SQV} &= E_{Asp25'+SQV} - (E_{Asp25'} + E_{SQV}) \end{aligned} \right\} \quad (3.3)$$

where $E_{Asp25+SQV}$ and $E_{Asp25'+SQV}$ are total energy of the SQV complexed with Asp25 and SQV complexed with Asp25', respectively.

In addition the pair interaction energy between two catalytic residues (Asp25 and Asp25') was calculated separately according to the following equation:

$$\Delta E_{Asp25-Asp25'} = E_{Asp25+Asp25'} - (E_{Asp25} + E_{Asp25'}) \quad (3.4)$$

where $E_{Asp25+Asp25'}$ is total energy of the Asp25 and Asp25' dimer.

3.2 Results and discussion

3.2.1 Reliability of the simulations

The total, potential and kinetic energies as well as temperature over 600 ps simulation time were plotted and shown in Figures 3.5-3.7. The results indicate that our system propagates properly. The mean values for the total energy, kinetic energy and potential energy are summarized in Table 3.4. The average temperatures of the systems during the simulations are also in Table 3.4.

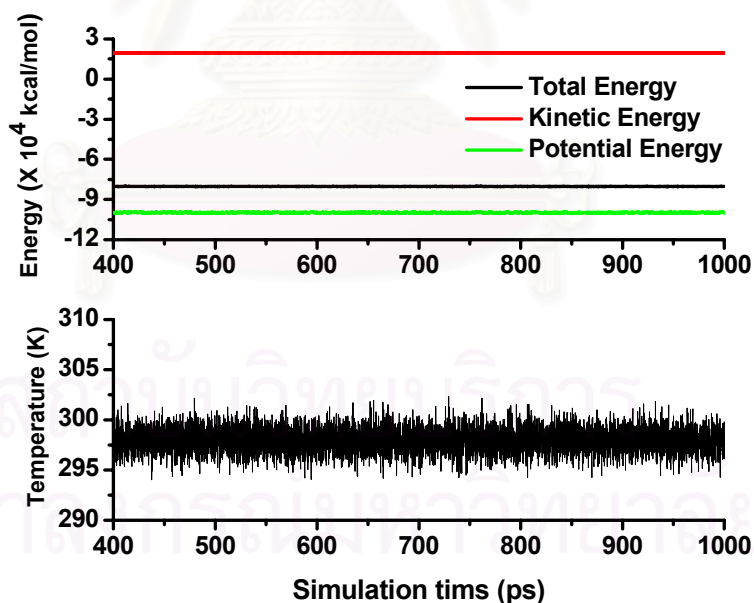


Figure 3.5 The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 400-1,000 ps for Mono-25 system.

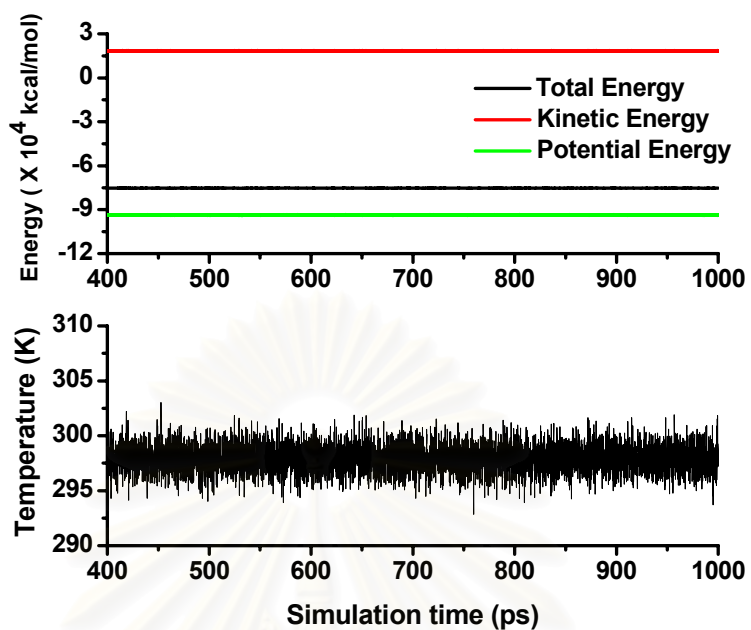


Figure 3.6 The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 400-1,000 ps for Mono-25' system.

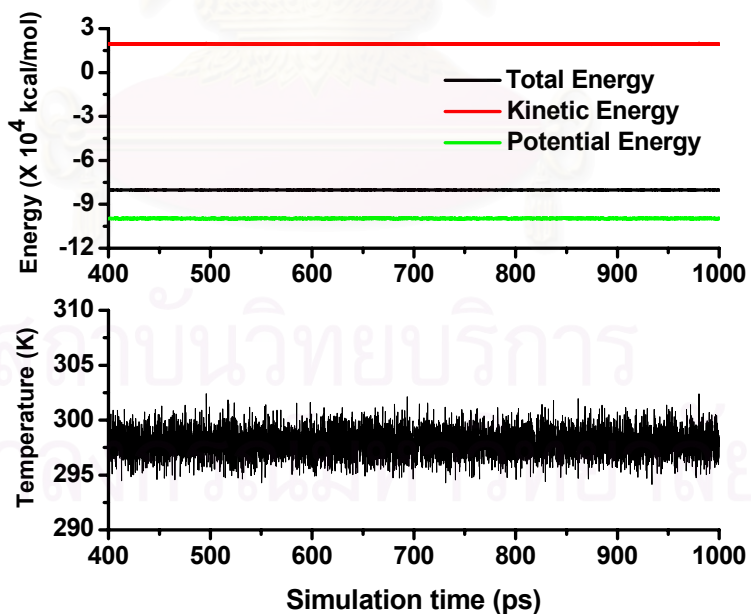


Figure 3.7 The plots of total (back), kinetic (red) and potential (green) energies (a) as well as temperature (b) over 400-1,000 ps for Di-pro system.

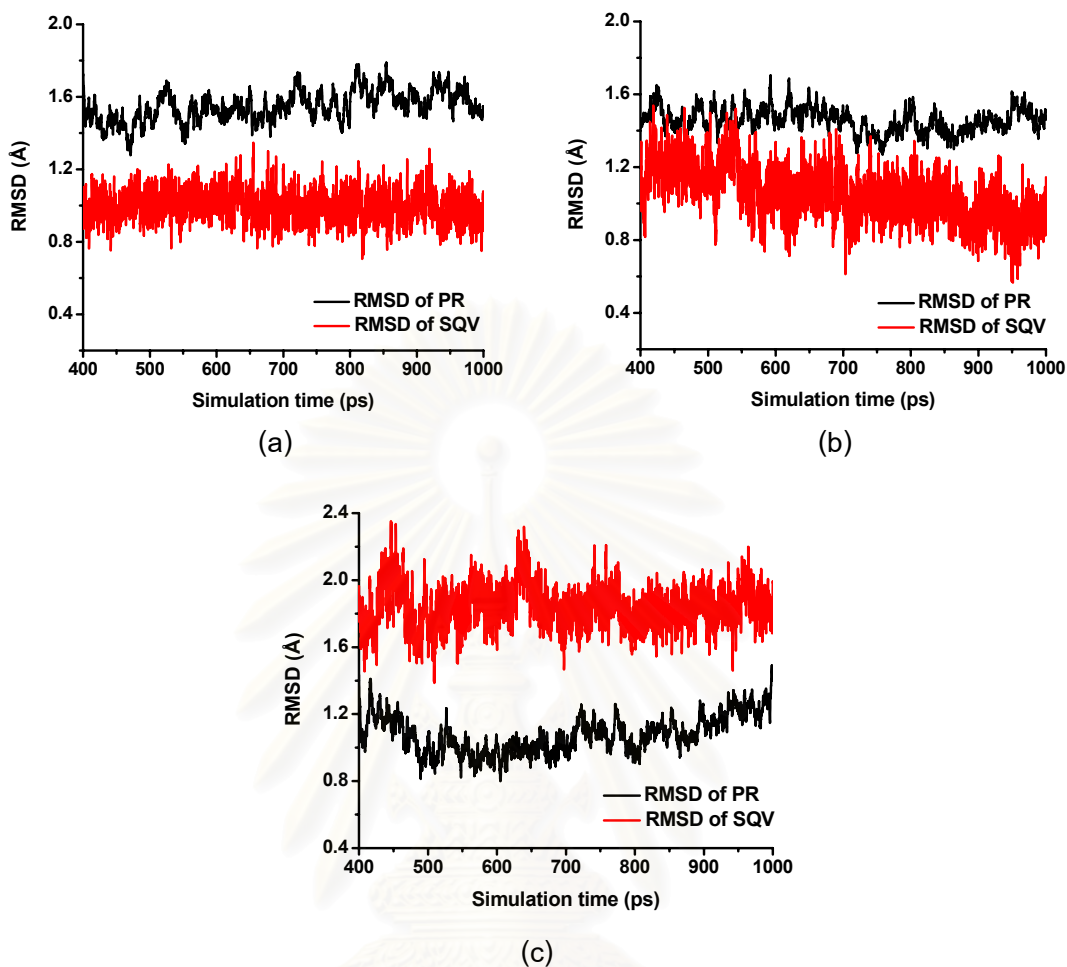


Figure 3.8 Global RMSD values of HIV-1 PR (black line) and saquinavir (red line) with respect to the starting structure over 400-1,000 ps for Mono-25 (a), Mono-25' (b) and Di-pro (c) systems.

The global root mean square deviation (RMSD) values of protease enzyme and saquinavir with respect to the starting structure were plotted against the 400-1,000 ps simulation time as shown in Figure 3.8. The mean RMSD values of PR and saquinavir are summarized in Table 3.4.

Table 3.4 Mean values of total, potential and kinetic energy as well as temperature

Values	Mono-25	Mono-25'	Di-pro
Total Energy (kcal/mol)	-8.02×10^4	-7.52×10^4	-8.02×10^4
Potential Energy (kcal/mol)	-9.97×10^4	-9.36×10^4	-9.96×10^4
Kinetic Energy (kcal/mol)	1.95×10^4	1.84×10^4	1.95×10^4
Temperature (K)	297.98	297.97	297.96
RMSD of PR (Å)	1.55 ± 0.08	1.46 ± 0.07	1.07 ± 0.11
RMSD of SQV (Å)	0.99 ± 0.08	1.05 ± 0.15	1.84 ± 0.13

3.2.2 Flexibility of enzyme

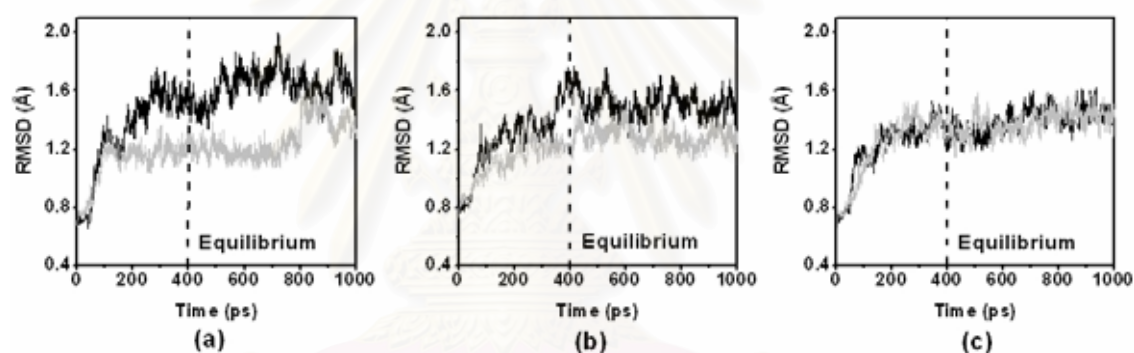


Figure 3.9 RMSD as a function of simulation time of the two chains, A (black line) and B (gray line), of the G48V HIV-1 PR mutant type complexed with saquinavir in the Mono-25 (a), Mono-25' (b) and Di-pro (c) states.

The RMSD of all atoms with respect to the starting geometry as a function of simulation time was evaluated separately for each chain of the HIV-1 PR. The results were shown in Figure 3.9. The plots describe the structure relaxations when the molecule was dissolved in the solution. The RMSD values of all plots increase rapidly for the first 100 ps indicating a change in enzyme geometry relative to the starting x-ray structure. Although the value of almost all plots remain constant after 200 ps, however, the RMSDs for chain A of Mono-25 (Figure 3.9a) and both chains of Mono-25' still increase slowly (Figure 3.9b). This data suggest that monoprotinated states, Mono-25 and Mono-25', require longer time than the Di-pro state to approach equilibrium. To

enhance the reliability of the data, all systems were considered to be in equilibrium after 400 ps. Cooperative effects due to different protonation states and asystematic geometry of the saquinavir inhibitor lead to higher RMSD on chain A (Figures 3.9a and 3.9b) when compared with chain B of the two monoprotonated states. This difference was not detected for the Di-pro state (Figure 3.9c) where both catalytic residues, Asp25 and Asp25' were protonated.

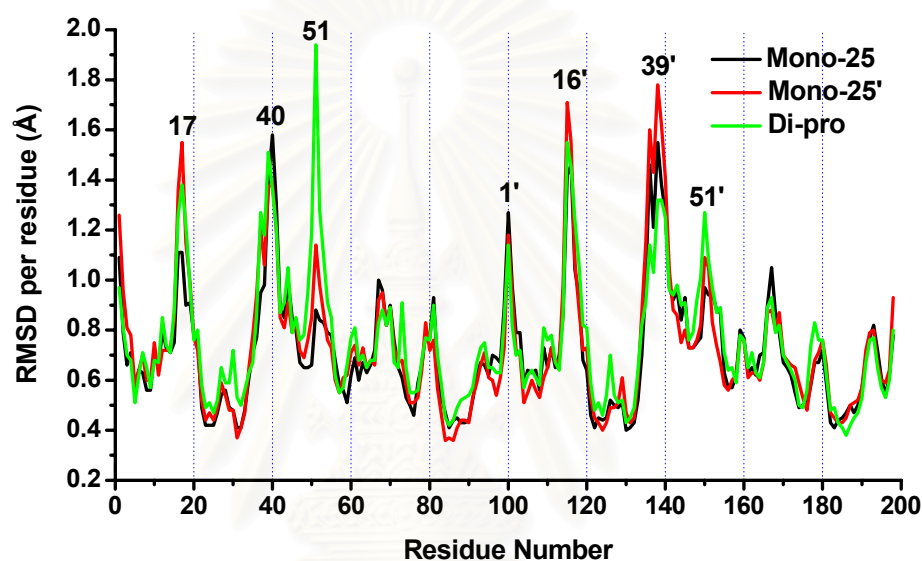


Figure 3.10 RMSD per residues of the Mono-25 (black), Mono-25' (red) and Di-pro (green) states where the highly flexible residues were labeled.

To ascertain for more details on the flexibility of each residue in the enzyme structure, the RMSD for each residue were calculated and displayed in Figure 3.10. From the plot, one can see that RMSD values of most residues are less than 1.0 Å, however, there are three highly flexible regions with RMSD more than 1.0 Å. Those regions are terminal (residues 1-5), cheek (cheek turn: residues 11-22) and flap (residues 43-58) regions that generally known as the flexible regions of HIV-1 PR. To be clearly seen, the highly flexible residues were labeled in Figure 3.10. For the three states, the highly flexible residues are found in the same regions that located on the solvent accessible surface of the enzyme. The significantly difference were observed on residues 17, 51 and 37'-39'. In the catalytic binding pocket (residues 25-30), which have

low RMSD values, slightly difference between monoprotonated and diprotonated state were detected. Note that the RMSDs of these residues for Di-pro state are higher than that of Mono-25 and Mono-25' states.

3.2.3 Flexibility of the saquinavir inhibitor

The structure of saquinavir is known to consist of five subsites which are P_1 , P_2 , P_3 , P_1' and P_2' subsites as shown in Figure 3.11.

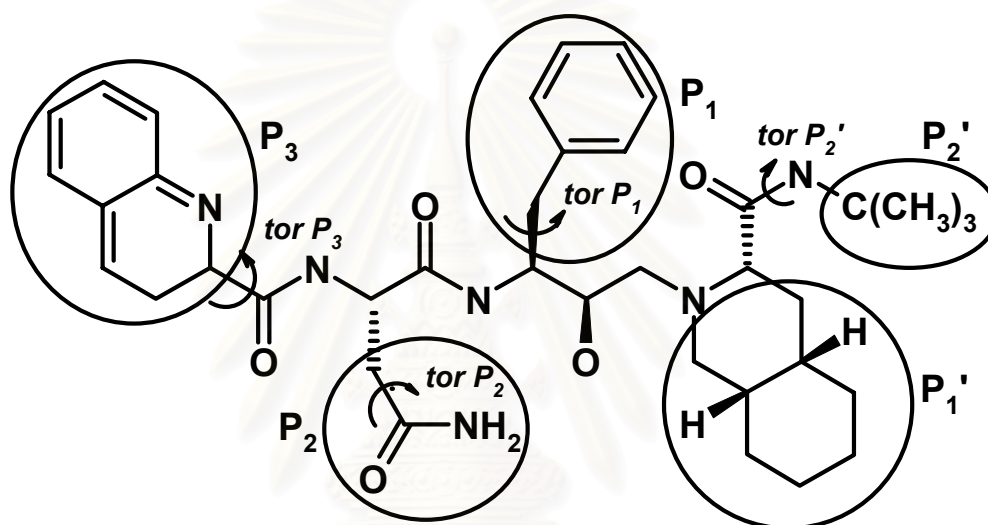


Figure 3.11 Definition of five subsites of saquinavir structure that labeled by circle and torsional angle of each subsite was defined by *tor*.

The root-mean-square displacements of each subsite with respect to the starting coordination, the x-ray structure, were plotted in Figure 3.12 and their distributions (labeled by *tor* in Figure 3.11) were plotted in Figure 3.13. Evaluations were carried out for each subsite of the three protonation states. Note that due to the rigid nature of the planar P_1' subsite, therefore, its torsional plot was not taken into consideration. The RMSD plots through the simulation time indicate clearly that conformations of all subsites in the three protonated states are almost unchanged, except for the P_2 subsite of Mono-25 and P_2' subsite in all states.

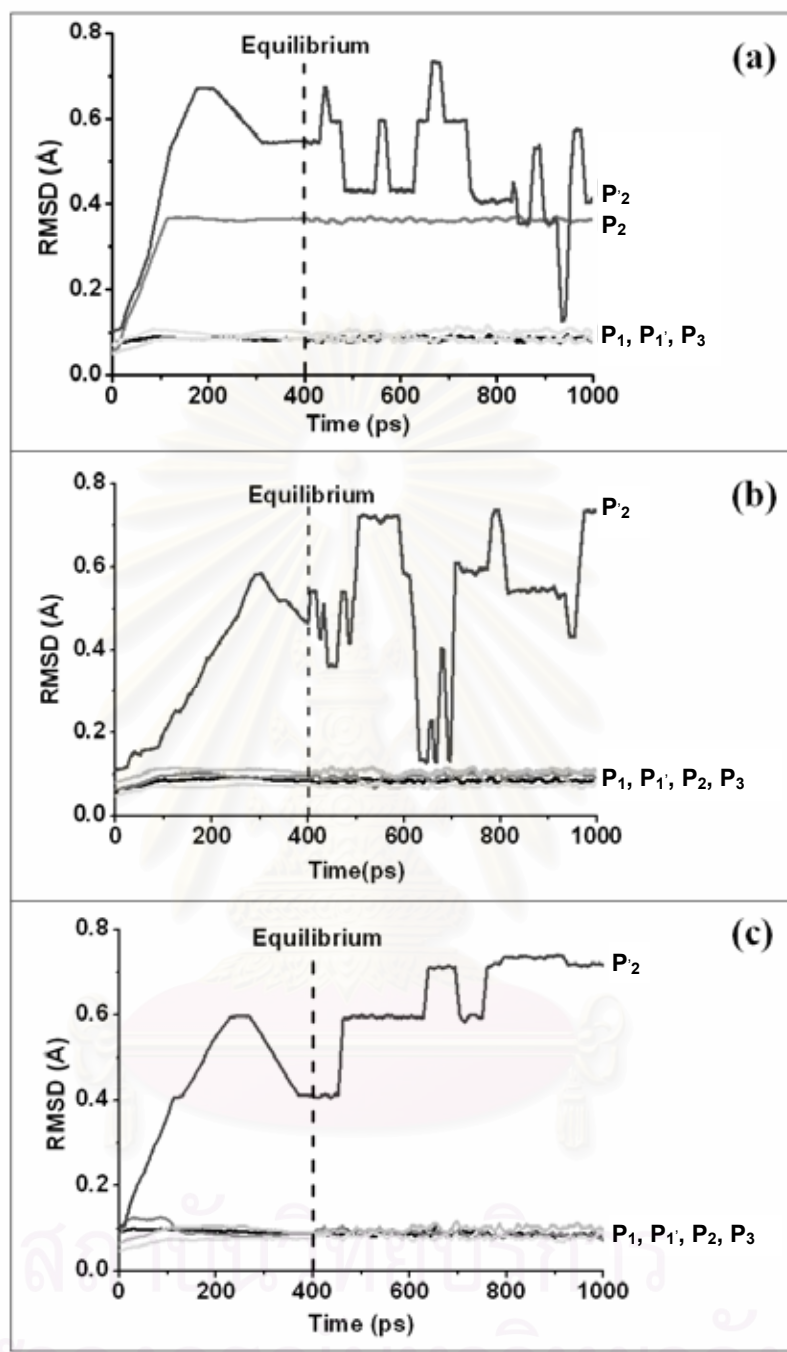


Figure 3.12 Root-mean-square displacements (RMSD) of each subsite of saquinavir (defined in Figure 3.11) for the three simulated systems, (a) Mono-25, (b) Mono-25' and (c) Di-pro.

As a result of these observations, attention was focused on P_2 and P_2' subsites. The RMSD plot for P_2 subsite for the Mono-25 state (Figure 3.12a) is significantly higher than the other two states and the plot for P_2' for the three states are

highly flexible when compared with the other subsites. The first event indicates a change of P_2 subsite of Mono-25 from one (RMSD ~ 0.05 Å) to another (RMSD ~ 0.36 Å) conformations (more details are discussed later). Furthermore, the high flexibility of P_2' subsite for all protonation states (Figures 3.12a - 3.12c) seems to indicate the detection of three preferential conformations for this subsite of saquinavir, *e.g.*, at RMSDs *ca.* 0.4, 0.6 and 0.7 Å for Di-pro system *etc.* This assumption was confirmed by the plot shown in Figure 3.13.

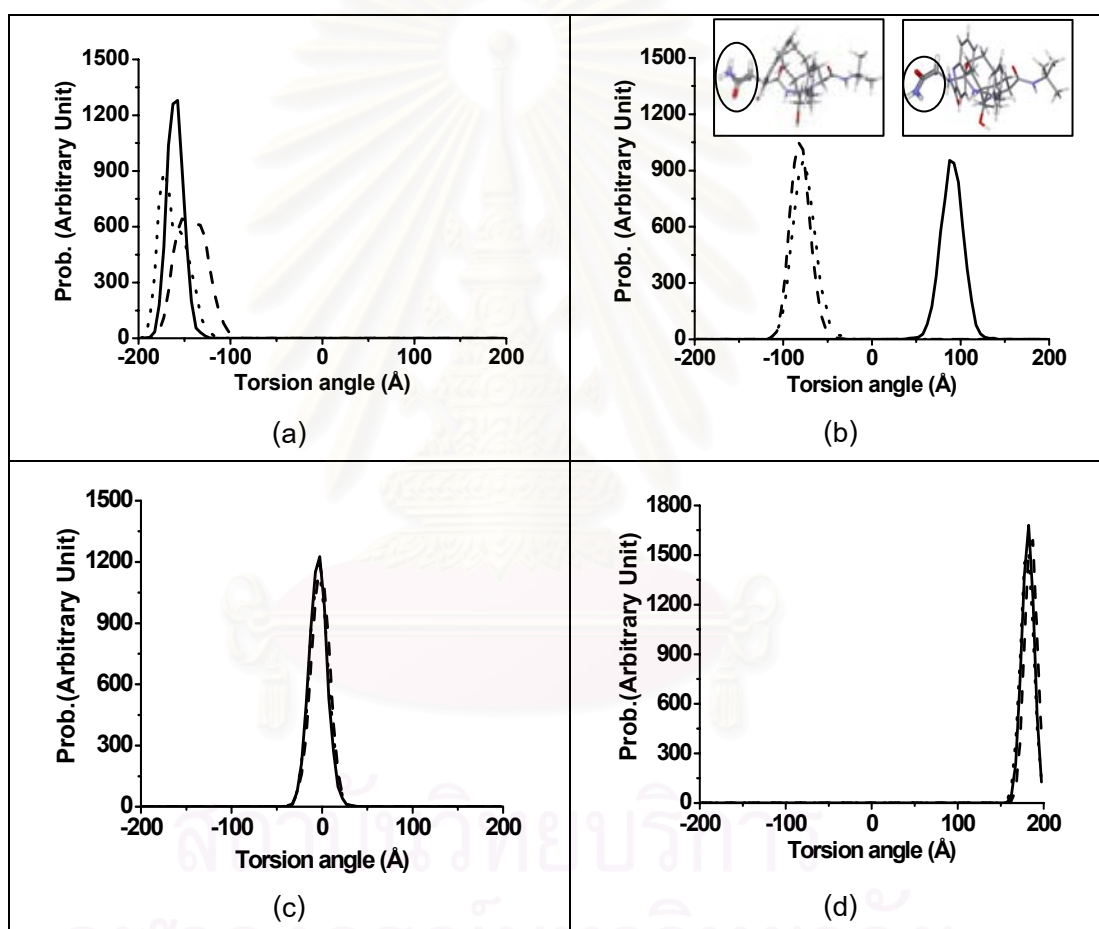


Figure 3.13 Changes of torsional angles of four subsites of saquinavir inhibitor, (a) P_1 subsite; (b) P_2 subsite; (c) P_2' subsite; (d) P_3 subsite, (defined in Figure 3.11) for the three simulated systems where solid, dashed and dotted lines in the torsional angle plots represent the Mono-25, Mono-25' and Di-pro systems, respectively. The two conformations corresponding to torsional angles of 90.3° and -75.5° of P_2 are also given in an inset.

Detailed information on the conformational flexibility of each subsite of saquinavir can be seen in terms of distributions of the torsional angles (Figure 3.13) after equilibrium, defined in Figure 3.13. All plots show pronounced and sharp peaks indicating slight flexibility of each subsite in a narrow range. For P_2' (Figure 3.13c) and P_3 (Figure 3.13d) subsites, their conformations are not sensitive to the change of protonation states, leading to the most probable torsional angles of 0° and 180° , respectively. Figure 3.13b, the torsional angle of P_2 subsite of the Mono-25 of $\sim 90^\circ$ (solid line, Figure 3.13b) differs substantially from that of $\sim 75^\circ$ for the other two states. The difference between the positions of the two peaks, $\sim 166^\circ$, suggests that the orientation of the P_2 subsite of saquinavir in the Mono-25 state complex is almost opposite to that of the other states. This statement is supported by the snapshot shown in an inset of Figure 3.13b. Note that single preferential conformation after equilibrium of P_2 subsite for the Mono-25 state took place at the torsion angle of $\sim 90^\circ$ (solid line, Figure 3.13b) is consistent with the RMSD = 0.36 Å shown in Figure 3.14a. Furthermore, the sharp peak at the torsional angles of $\sim 75^\circ$ for the Mono-25' and Di-pro (Figure 3.13b) states corresponds to the conformation represented by the RMSDs of ~ 0.08 Å (Figures 3.12b-3.12c).

It is interesting to note that, the observed conformation of P_2 subsite of saquinavir in the Mono-25 state is similar to the x-ray structure found in the double mutant, G48V/L90M, HIV-1 PR complex (92). It was also proposed by the x-ray data that rotation of the P_2 subsite of saquinavir in the G48V/L90M mutation depends only on the position of Gly48 which lies much closer to the inhibitor than the Leu90. However, x-ray structure for the present system (single G48V mutation of HIV-1 PR complexed with saquinavir) is not available.

In addition, the three peaks of P_1 subsite for the three states are slightly different (Figure 3.13a). A sharper and higher peak at the torsional angle of -157° (Mono-25, solid line) than the other at -172° (Di-pro, dot line) and -142° (Mono-25', dash line) indicate a higher rigidity of P_1 subsite in the Mono-25 than the other two states.

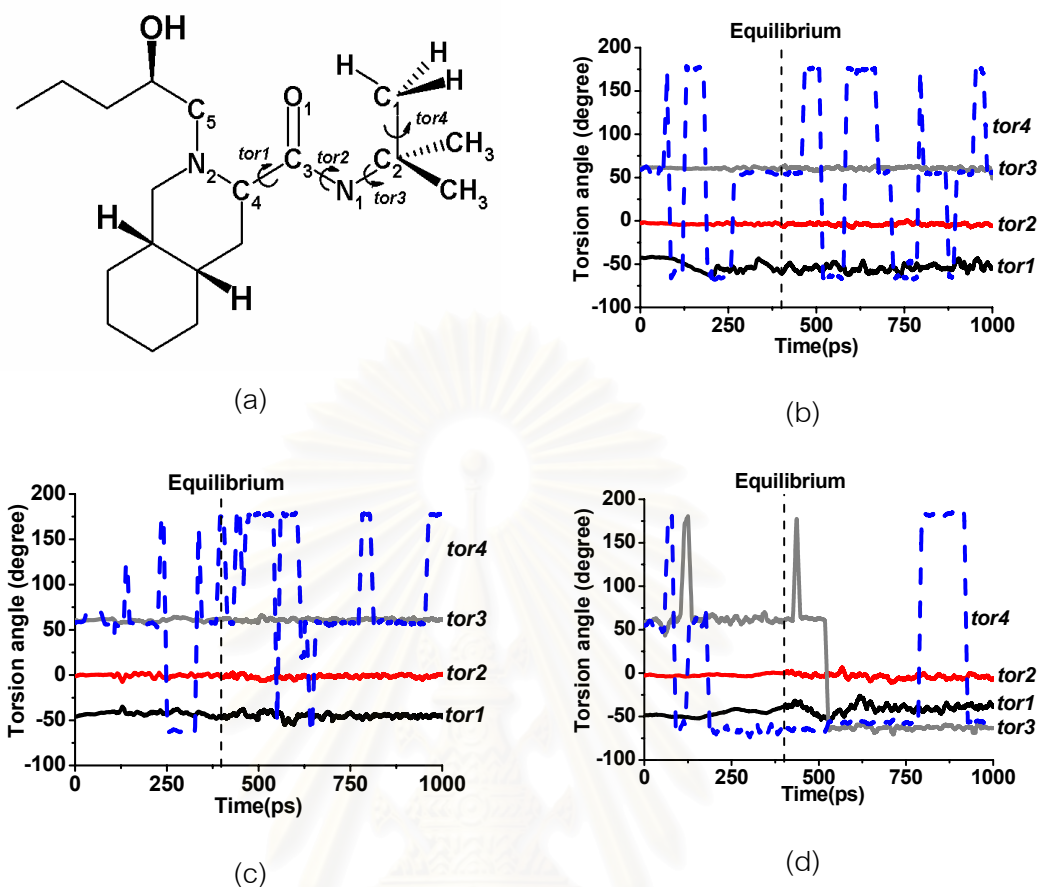


Figure 3.14 Definition of torsional angles *tor1* - *tor4* (a) of the P_2' subsite of saquinavir inhibitor and their changes as a function of simulation time; (b) Mono-25; (c) Mono-25'; (d) Di-pro.

To seek more information on the high flexibility of P_2' subsite, the four possible torsional angles, *tor1* ($N_2-C_4-C_3-O_1$), *tor2* ($O_1-C_3-N_1-C_2$), *tor3* ($C_3-N_1-C_2-C_1$) and *tor4* ($N_1-C_2-C_1-H$) were defined (Figure 3.14a) and evaluated. The results were plotted in Figures 3.14b-3.14d. The *tor1* and *tor2* for all protonated states are highly stable showing preferential orientations at approximately -50° and 0° , respectively. These results suggested that the $-C_3=O_1$ functional group of P_2' subsite (see Figure 3.14a) was tilted by 50° from the $C_3-C_4-N_2$ plan while C_2 was detected to locate almost in the $O_1-C_3-N_1$ plane (*tor2* = 0°). Fluctuations were observed for *tor3* of Di-pro and *tor4* for all states. Particular, *tor4* shows three preferential conformations at -60° , 60° and 180° . This fact can be understood by a free rotation of the $-H_3$ and $-(CH_3)$ functional groups. An

appearance of *tor4* at the three torsional angles for all states indicates a staggered conformation among the three hydrogen atoms of the $-CH_3$. The same conclusion can also be made for *tor3* of the diprotonated state, Di-pro, in which the three $-CH_3$ groups are observed to rotate freely with the three preferential states in the staggered conformation.

3.2.4 Enzyme – saquinavir interaction

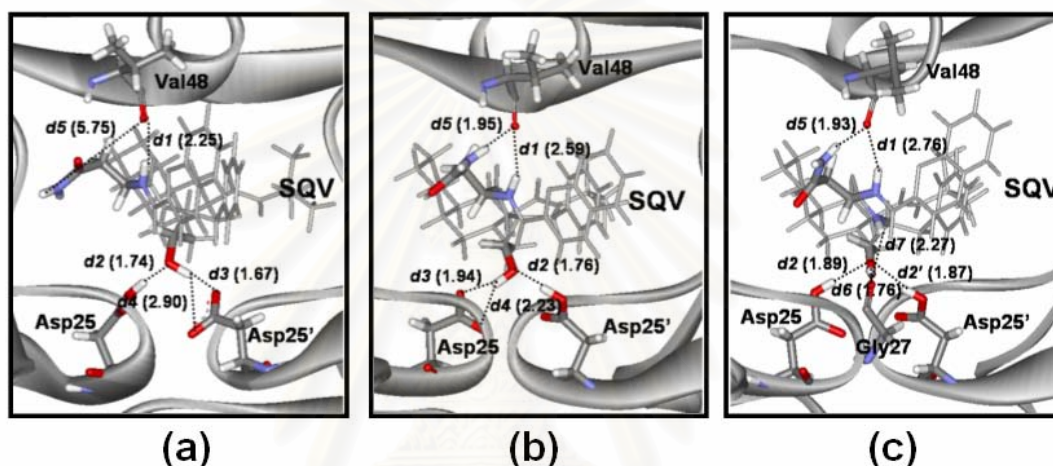


Figure 3.15 Detected hydrogen bonds (dash line) in the HIV-1 PR enzyme-saquinavir complex in (a) Mono-25, (b) Mono-25' and (c) Di-pro states, where the average hydrogen bond distances (H---A distance) in Å are numbered.

It is well known that inhibitor or substrate was held in the active site of enzyme via the intermolecular forces, such as electrostatic, hydrophobic, dispersion forces, *etc.* For the investigated system where the hydrophilic OH group of saquinavir binds directly to the catalytic residues, the catalytic region contains residues 25–30: Asp25-Thr26-Gly27-Ala28-Asp29-Asp30 and some water molecules. It appears that the electrostatic forces in the catalytic pocket play a key role of the complex. Therefore, attention was focused on the hydrogen bonding between inhibitor and enzyme. In addition, interaction between saquinavir and the catalytic dyad was also investigated.

Hydrogen bonding between HIV-1 PR and saquinavir was determined based on the CARNAL module in the AMBER 7 using the following criteria, (i) the distance between proton donor (D) and acceptor (A) atoms was less than or equal to 3.5 Å and (ii) the D-H...A angle was greater than or equal to 120°. Analysis was carried out over the entire trajectory after equilibration. The detected hydrogen bonds with the H...A distances were shown in Figures 3.15a-3.15c. It is interesting to note that hydrogen bonds between saquinavir and HIV-1 PR were formed only in the two regions that are flap and active site region. The number of hydrogen bonds were found 4, 5 and 6 for the Mono-25, Mono25' and Di-pro states, respectively. In the flap region, hydrogen bonds were formed only with the Val48 of the HIV-1 PR. In the Mono-25 state (Figure 3.15a), the deformation of $d5$ bond (5.75 Å) was observed. This observation is due to the rotation of the NH_2 functional group of P_2 subsite of saquinavir (inset of Figure 3.13b). In addition, the remaining hydrogen bond ($d1$) at the average distance of 2.25 Å is slightly stronger than those of 2.59 Å and 2.76 Å for the Mono-25' and Di-pro states, respectively. For the active site region of the two monoprotonated states, the OH functional group of saquinavir was found to form three hydrogen bonds ($d2$ - $d4$) with the two catalytic residues in which, one hydrogen bond with the protonated and the other two with the unprotonated carbonyl group. An exception was found for the Di-pro state where, the unprotonated group is not available, therefore, the other bond was formed with Gly27 ($d6$ and $d7$ in Figure 3.15c).

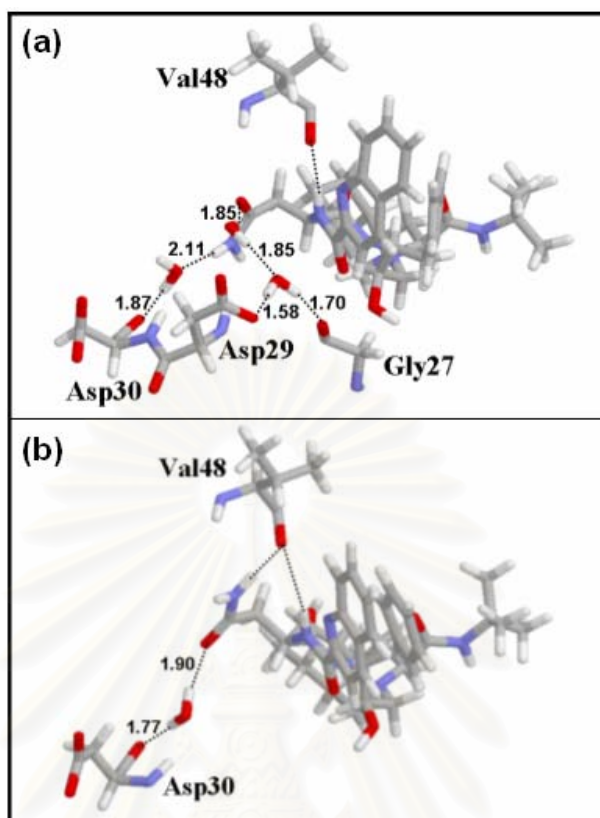


Figure 3.16 Indirect binding between HIV -1 PR and saquinavir via hydrogen bonds (dash line) through water molecules in the (a) Mono-25 and (b) Mono-25' states where the average distance (in Å) are numbered.

Furthermore, indirect binding between saquinavir and the three residues, Gly27, Asp29 and Asp30, of the HIV-1 PR via the hydrogen bonds through water molecules was detected and displayed in Figure 3.18. In the Mono-25 state (Figure 3.16a), the conformation of P₂ subsite was found to rotate and indirectly bind to the three residues of enzyme. The oxygen atom and hydrogen atom of P₂ subsite form hydrogen bonds with Asp29 and Gly27 via two water molecules and with Asp30 via one water molecule (Figure 3.16a), respectively. Even though those hydrogen bonds were not detected for Mono-25'. However, the nitrogen atom of the NH₂ group of P₂ subsite for Mono-25' was found to form a hydrogen bond with Asp30 via one water molecule (Figure 3.16b), instead.

Table 3.5 Interaction between saquinavir (SQV) and the catalytic residues, Asp25 and Asp25', of HIV-1 PR for the three protonated states yielded from *ab initio* calculations (see text for more details).

Interaction energy (kcal/mol)	Mono-25	Mono-25'	Di-pro
ΔE_{cpx}	-28.4	-28.0	-11.5
$\Delta E_{Asp25-SQV}$	-8.0	-18.0	-6.2
$\Delta E_{Asp25'-SQV}$	-19.2	-8.0	-7.9
$\Delta E_{Asp25-Asp25'}$	-8.8	-2.6	+0.8
ΔE_{3bd}	-1.2	-2.1	+2.6

To evaluate the interaction between saquinavir and the catalytic residues, quantitatively, DFT calculations were performed. The interaction energy between saquinavir and catalytic dyad (ΔE_{cpx}) and pair interactions as well as error due to three body effects (ΔE_{3bd}) are summarized in Table 3.5 (see method of calculation for more details).

Table 3.6 Average RMSDs as well as their deviations for the ASP25 and Asp25' of the Mono-25, Mono-25' and Di-Pro systems evaluated after equilibration, 400-1,000 ps.

Res. No.	Mono-25		Mono-25'		Di-pro	
	RMSD	Deviation	RMSD	Deviation	RMSD	Deviation
Asp25	0.56	0.10	0.38	0.08	0.18	0.05
Asp25'	0.37	0.06	0.47	0.12	0.35	0.16

The average RMSDs of the ASP25 and ASP25' residues for the three investigated systems with the deviations of less than 0.16 Å were summarized in Table 3.6. This indicates reliability of the DFT energy calculations in which the MD snapshots can be represented by the average structure.

In terms of complexation energies, the obtained order of the ΔE_{cpx} is Mono-25 > Mono-25' > Di-pro, *i.e.*, interaction energy of -28.4 kcal/mol for the Mono-25 state is significantly lower than those of -28.0 and -11.5 kcal/mol for the Mono-25' and

Di-pro states, respectively. This evidence is in good agreement with that previously reported for saquinavir complexed with wild type HIV-1 PR in which only Asp25 monoprotinated state is in the reasonable binding affinity (79,97,98). In addition, quantum mechanics and molecular mechanics calculations (99) for the HIV-1 PR – A74704 inhibitor complex leads also to the conclusion that the protonated state of the active site of the complex is monoprotinated state on Asp25.

In Table 3.5, the total interaction energy can be separated into three terms of pair interaction and additional three-body correction terms. As expected, the main contribution to interaction energy between SQV and catalytic residues is due to the electrostatic interactions between saquinavir and non-protonated state aspartic acids which amount to -19.2 and -18.0 kcal/mol for the Mono-25 and Mono-25', respectively. In addition, interactions with the two protonated residues are -8.0 kcal/mol for both Mono-25 and Mono-25'. Summation of the two terms of the pair interaction, $\Delta E_{SQV-Asp25}$ and $\Delta E_{SQV-Asp25'}$, yields the total interaction of -27.2 kcal/mol for Mono-25 and -26.0 kcal/mol for Mono-25'. These two values are almost consistent with the total interaction between saquinavir and the two aspartic residues calculated from ΔE_{cpx} . The difference is due to many-body correction term, Asp25' and Asp25 were excluded in the calculation of the $\Delta E_{SQV-Asp25}$ and $\Delta E_{SQV-Asp25'}$, respectively. Taking into account the data summarizing above, no significant difference was found among the interaction between saquinavir and the catalytic residues of the HIV-1 PR of the two monoprotinated states.

Interest is focused on the Asp25-Asp25' interaction, a significantly lower Asp25-Asp25' interaction energy of -8.8 kcal/mol for the Mono-25 than that of -2.6 kcal/mol for the Mono-25' suggests us to make a clear conclusion that protonation at the Asp25 leads to a better arrangement in the catalytic dyad. Interestingly, such rearrangement does not make any difference on the $\Delta E_{SQV-Asp25}$ or $\Delta E_{SQV-Asp25'}$ interactions (Table 3.5) in comparison between the two monoprotinated residues. This finding also agrees well with the hydrogen bond formation between saquinavir and catalytic residues of the Mono-25 and Mono-25' where three hydrogen bonds were

observed and the hydrogen bond distance for the two systems are almost equivalent, d_2 , d_3 and d_4 in Figure 3.15.

Furthermore, rearrangement of the catalytic dyad was supposed to be the source of conformation change of the P_2 subsite of saquinavir in the Mono-25 state (Figure 3.13) and of the vanishing of one hydrogen bond between Val48 and saquinavir in the flap region (Figure 3.15a). The above data suggest us to conclude that interaction in the catalytic site should be used as criteria to enhance capability in drug designing and drug screening instead of using the total inhibitor/enzyme interaction. In addition, interaction between the inhibitor and the catalytic region of the enzyme is supposed to relate directly to the activity of the inhibitor in the catalytic process.

3.3 Conclusion

We have reported results of 1 ns MD simulations of three protonation states (Mono-25, Mono-25' and Di-pro) for G48V mutant HIV-1 PR complexed with the saquinavir inhibitor. Detailed analyses of structure and dynamics characters among the three protonation states of G48V mutant were given in terms of the flexibility, hydrogen bonding and conformational changes as well as binding energy between inhibitor and enzyme. The results indicate that the three complexes display a distinct dynamical behavior. A major structure change via complexation was found at P_2 subsite of saquinavir which affected the hydrogen bonding for enzyme–inhibitor interaction and water molecule. The interaction data yielded from quantum chemical calculations supports the previous conclusion which state that the protonated state of the active site of the mutant-type HIV-1 PR complexed with saquinavir is a monoprotinated state on Asp25 (Mono-25). The energy data indicates also that protonation at the Asp25 leads to a better arrangement in the catalytic dyad. In addition, conformation of the P_2 subsite in the Mono-25 is consistent with that obtained from the x-ray structure of G48V/L90M mutant HIV-1 PR complexed with saquinavir.

CHAPTER 4

MD SIMULATIONS OF WILD-TYPE AND MUTANTS HIV-1 PR COMPLEXED WITH SAQUINAVIR IN AQUEOUS SOLUTION

Saquinavir resistances were observed to take place at Gly48 and Leu90 during *in vivo* and *in vitro* passage with saquinavir, respectively. Residues Gly48 locates in the flap region of the HIV-1 PR while Leu90 is outside the binding pocket of the enzyme. Saquinavir selects for variants with one or both of two amino acid substitutions in the HIV-1 protease gene, a valine-for-glycine substitution at position 48 (G48V), a methionine-for-leucine substitution at residue 90 (L90M), and the double substitution G48V-L90M (51). In most cases, G48V is the first mutation to appear, and continued selection results in highly resistant double-mutant variants. The kinetic study data on L90M, G48V and G48V/L90M mutants indicate the following order of resistance. The corresponding inhibition constants (K_i) for saquinavir are WT, L90M, G48V and G48V/L90M, respectively.

In this study, we employed a computational approach to access information regarding molecular structure and dynamics of the G48V, L90M and G48V/L90M HIV-1 PR conferring to saquinavir resistance. The four MD simulations were carried out for the wild-type, G48V, L90M and G48V/L90M HIV-1 PR complexed with saquinavir in explicit aqueous solution. The comparison of MD results provides insightful details of how the mutants are associated with saquinavir resistance. The study provided fundamental principles on the molecular mechanism of inhibitor binding and resistance, which will be useful for designing an anti-HIV inhibitor to combat AIDS. Moreover, the MM/PBSA methods have been performed to evaluate the binding free energy of all protein-ligand complexes and compared with the experimental data, K_i and IC_{50} values.

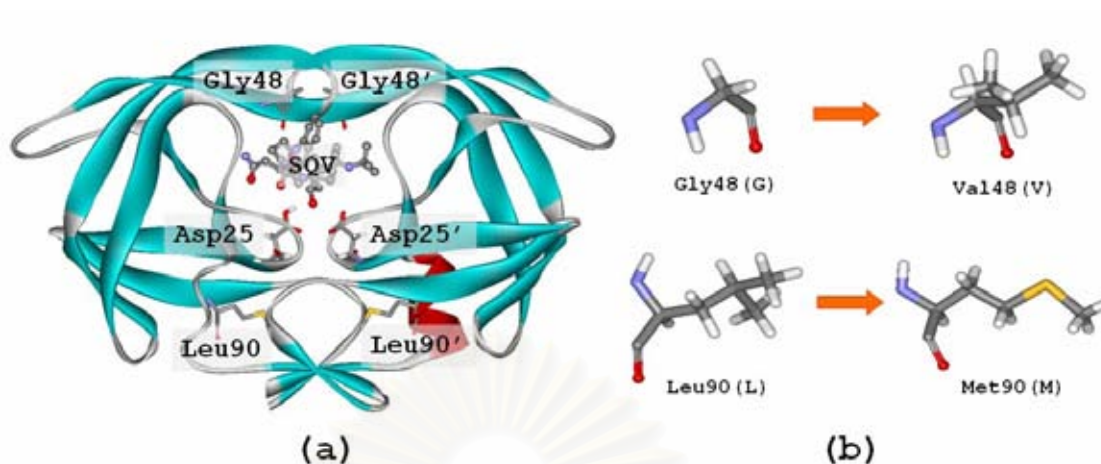


Figure 4.1 Schematic representation of wild-type HIV-1 PR complexed with saquinavir (a) and comparison of wild-type (Gly48 and Leu90) and mutant types residues (Val48 and Met90).

4.1 Computational methods

4.1.1 Preparation of the initial structures

The MD simulations of four different complexes are performed in this chapter. Those complexes are between saquinavir and a wild-type HIV-1 PR (WT-SQV), two of single mutants, G48V-SQV and L90M-SQV, and a double mutant (G48V/L90M-SQV). The x-ray structure of the wild-type HIV-1 protease-saquinavir complex (PDB code: 1HXB) was chosen from the PDB and used as an initial structure for all complexes. Three mutant protease complexes were modeled from that crystal structure, 1HXB, replaced Gly48 by Val (G48V) and Leu90 by Met (L90M) for both chains using the LEaP module in AMBER 7 software package (90).

Regarding the MD results among the four protonation state of HIV-1 PR complexed with saquinavir given in chapter 3, the complex was observed to prefer the monoproteination state at Asp25, *i.e.*, only the catalytic Asp25 residue was protonated. All initial structures of this chapter were produced to satisfy this situation. In this study, hydrogen atom was added to Asp25 in order to build the monoprotinated state at Asp25 by AMBER 7 simulation package. For the ionization states of other ionizable

residues (excepted Asp25 and Asp25') computed by the pK_a calculation (calculation details were in chapter 3).

From the initial structure, hydrogen atoms were added using LEaP module. The crystallographic water molecules were maintained throughout the preparation process. The complexes were immersed in a 10 Å radius sphere of TIP3P water model (91) to allow the protein-ligand complexes in an aqueous solution. In addition, the sodium and chloride ions were added to neutralize the systems. The number of water molecules, counter ions, total atoms and dimension of simulation boxes were given in Table 4.1.

Table 4.1 Number of water molecules, counter ions and total atoms together with the dimension of simulation boxes of the four simulations, WT, G48V, L90M and G48V/L90M.

System	Number of			Box dimension (Å ³)
	Waters	Counter ions	Total atoms	
WT	9387	8 Cl ⁻ , 3 Na ⁺	31394	67 × 72 × 82
G48V	9384	8 Cl ⁻ , 3 Na ⁺	31403	67 × 72 × 82
L90M	9387	8 Cl ⁻ , 3 Na ⁺	31390	67 × 72 × 82
G48V/L90M	9384	8 Cl ⁻ , 3 Na ⁺	31399	67 × 72 × 82

4.1.2 Molecular dynamics simulations

Energy-minimization and MD simulations were performed using SANDER module of AMBER 7 with the Cornell force field (100). First, energy minimization was applied for water molecules using 3,000 steps of the steepest descents and then 2,000 steps of the conjugate gradients. Subsequently, the whole systems were subjected to energy minimization with 3,000 steps of the steepest descents and 2,000 steps of the conjugate gradients.

The MD simulations were performed employing the periodic boundary condition with the NPT ensemble. A Berendsen coupling time of 0.2 ps was used to

maintain the temperature and pressure of the systems (94). The SHAKE algorithm (95) was employed to constrain all bonds involving hydrogens. The simulation time step of 1 fs was used. All MD simulations were run with a 12 Å residue-based cutoff for nonbonded interactions and the particle-mesh Ewald (PME) method was used for an adequate treatment of long-range electrostatic interactions (96).

The MD simulation consists of thermalization and equilibration. During 0-60 ps, the temperature of the system increases from 0 to 298 K using NVT condition. Then the temperature was maintained at 298 K during equilibration employing at NPT conditions. The reference pressure and temperature were set to 1 atm and 298 K, respectively. The MD trajectories were collected every 0.1 ps. The convergence of energies, temperature and pressure of the systems, and the atomic root-mean square deviation of the enzyme and the inhibitor (RMSD) were used to verify the stability of the systems. The series of snapshot between 400 and 1,000 ps of the equilibrium phase were used for free energy calculations and structure evaluation.

4.1.3 Calculation of the binding free energy

In an evaluation of the binding free energy (ΔG_{bind}) of HIV-1 PR and saquinavir, MM/PBSA method was employed the complexes, using MM/PBSA module in AMBER 7 program, which interfaces the DelPhi4 program (72). The MM/PBSA approach calculates ΔG_{bind} on the basis of a thermodynamic cycle.

The enthalpic change in the gas phase upon binding (E^{MM}) was computed with no cutoff. The electrostatic free energy of solvation was computed using the Poisson-Boltzmann solver implemented in the DelPhi 4 program (72). The grid spacing was set to 1/3 Å with the boundary conditions of Debye-Hückel potentials. Atomic charges were taken from the Cornell force field. The water and protein dielectric constants were set to 80.0 and 4.0, respectively. The solvent-accessible surface area (SASA) was calculated using a 1.4 Å probe radius. Atomic radii were taken from the PARSE parameter set (73). The nonpolar free energy of solvation was obtained by $(0.00542 \times \text{SASA}) + 0.92$ kcal/mol (73).

In the study, the entropy contribution ($T\Delta S$) was not included. Considering the entropic effects, structural reorganization and solvent entropy are important contribution in enzyme-ligand bindings. However, a number of studies published by Kollman's and Warshel's groups showed that related contribution due to this effect, the same protein complexed with different inhibitors, is rather small (101,102). Some example are the thermodynamic study of four FDA approved drugs (indinavir, nelfinavir, saquinavir and ritonavir) where the experimental $T\Delta S$ values of the V82F/I84V are ~ 0.7 kcal/mol less favorable than that of the wild-type (103). Therefore, explanation of the entropies term to understand the ΔG_{bind} changes, which is the main goal of our work, should be a reasonable assumption. In addition to the above information, normal mode analysis (NMA) is a useful tool to estimate entropic changes of the solute molecule but the calculation is considered problematic and time-consuming. In addition, this approach does not take the solvent entropy into consideration. With those reasons, NMA would not greatly improve the correlation between the experimental K_i and the binding free energy. Thus the calculation of the solute $T\Delta S$ term has been omitted in the study.

4.2 Results and discussion

4.2.1 Reliability of the simulations

The total, potential and kinetic energies as well as temperature over 400-1,000 ps simulation time were plotted and shown in Figures 4.2-4.5. The results demonstrated a well-behaved MD simulation for all systems. The mean values for the total energy, kinetic energy and potential energy were summarized in Table 4.2. The average temperatures of the system for all systems were also shown in Table 4.2.

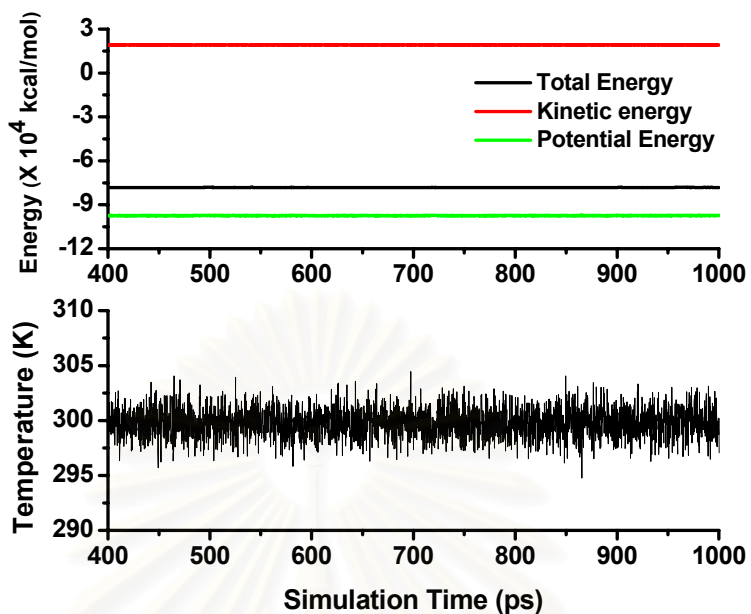


Figure 4.2 The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ps for WT-SQV system.

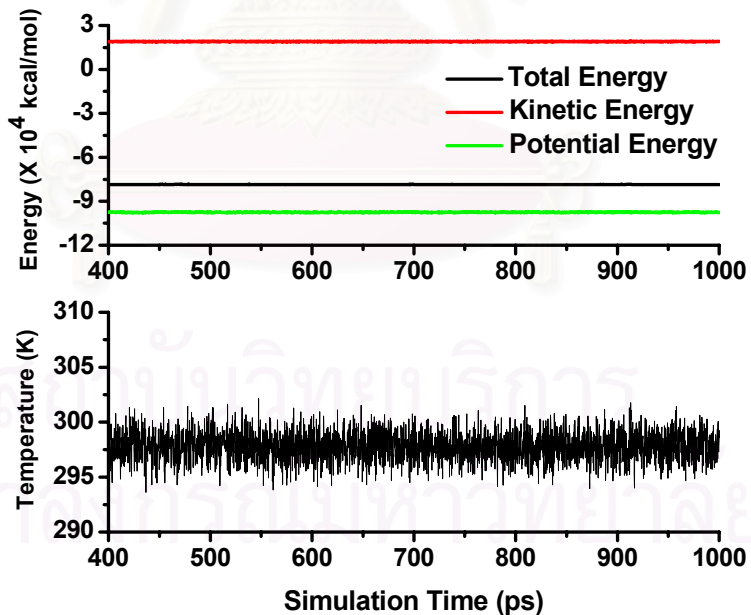


Figure 4.3 The plots of total (back), kinetic (red) and potential (green) energy (top) as well as temperature (button) over 400-1,000 ps for G48V-SQV system.

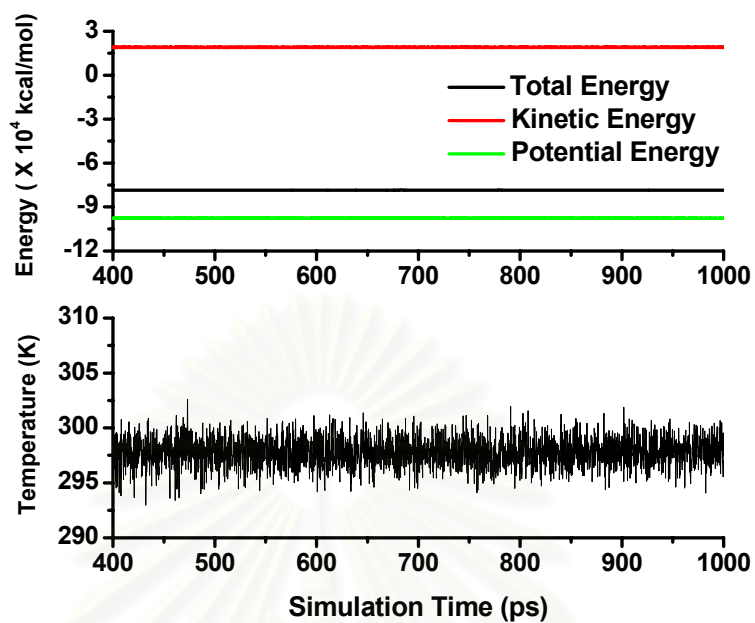


Figure 4.4 The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ns for L90M-SQV system.

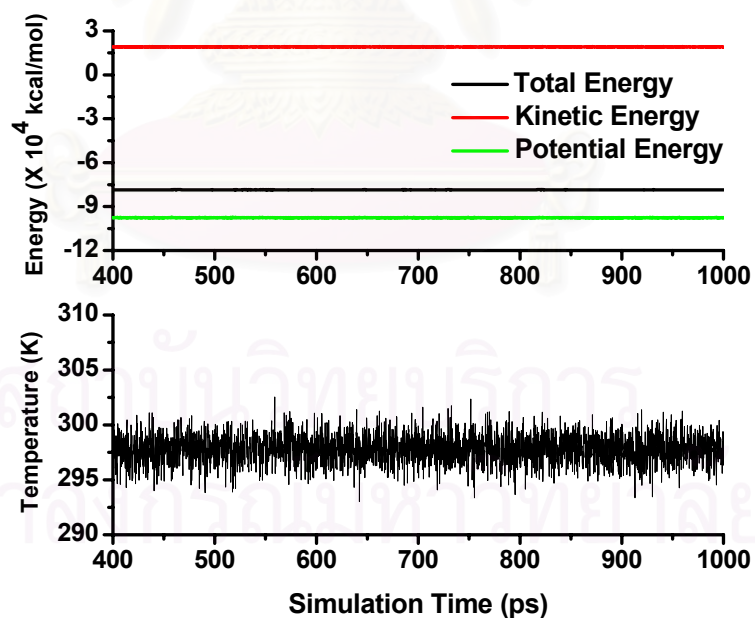


Figure 4.5 The plots of total (back), kinetic (red) and potential (green) energies (top) as well as temperature (button) over 400-1,000 ps for G48V/L90M-SQV system.

The global RMSD values of backbone atoms of protease enzyme and all atoms of saquinavir with respect to the starting structure were plotted against the 400-1,000 ps simulation time as shown in Figure 4.6. The mean RMSD values of PR and saquinavir over 400-1,000 ps are summarized in Table 4.2

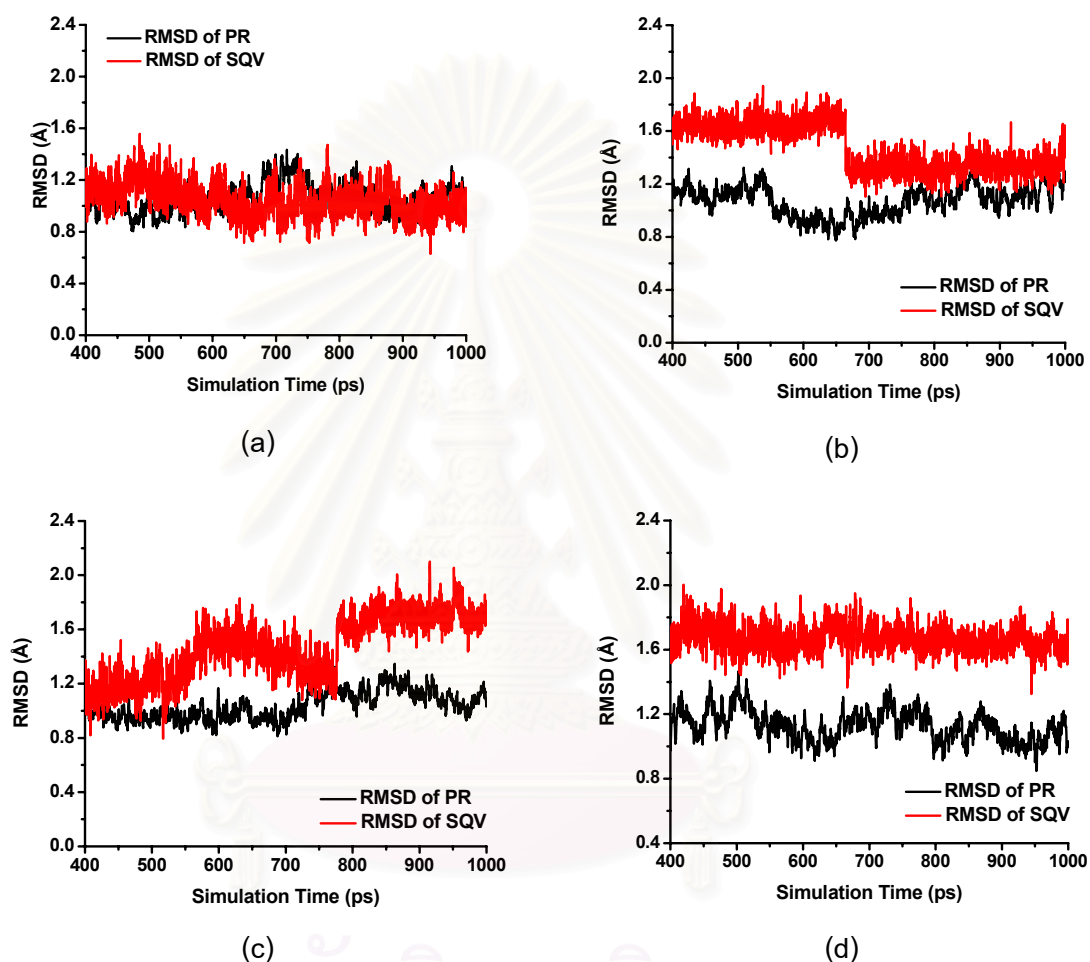


Figure 4.6 Global RMSD values of backbone atoms of HIV-1 PR (black line) and all atoms saquinavir (red line) with respect to the starting structure over 400-1000 ps for (a) WT-SQV, (b) G84V, (c) L90M and (d) G48V/L90M.

Table 4.2 Mean values of total, potential and kinetic energies as well as temperatures for the four simulated systems.

Values	WT-SQV	G48V-SQV	L90M-SQV	G48V/L90M-SQV
Total Energy	-7.82×10^4	-7.83×10^4	-7.58×10^4	-7.93×10^4
Potential Energy	-9.74×10^4	-9.75×10^4	-9.81×10^4	-9.86×10^4
Kinetic Energy	1.92×10^4	1.93×10^4	1.96×10^4	1.96×10^4
Temperature	299.75	298.75	299.85	298.35
RMSD of PR	1.12 ± 0.10	1.23 ± 0.16	1.15 ± 0.08	1.23 ± 0.09
RMSD of SQV	1.82 ± 0.70	1.79 ± 0.55	1.60 ± 0.40	1.74 ± 0.13

The energies and RMSD plots (Figures 4.2-4.6) demonstrated a well-behaved MD simulation for all systems. After 400 ps, the RMSD fluctuations of backbone of HIV-1 PR are less than 1.3 Å. This structural fluctuation is not uncommon in the typical MD simulation of protein, indicating the reliable equilibrium of the system in this chapter.

4.2.2 Flexibility and conformational changes of SQV subsites

In Schechter & Berger's nomenclature, residues to the left of the isostere are referred to as the P_1 , P_2 , P_3 , *etc.* subsite, and those to the right are given the designation P'_1 , P'_2 , P'_3 , *etc.* Similarly, HIV-1 PR binding pockets are designated as S_1 , S_2 , S_3 and S'_1 , S'_2 , S'_3). For PR-SQV complexes, the definitions of each subsite of saquinavir were displayed in Figure 4.7. In this study, flexibility and conformational changes of the inhibitor side chains P_1 , P'_1 , P_2 , P'_2 , and P_3 were investigated in terms of torsion angle fluctuations that labeled by *tor1-tor7* in Figure 4.7 and RMSD value of each subsite.

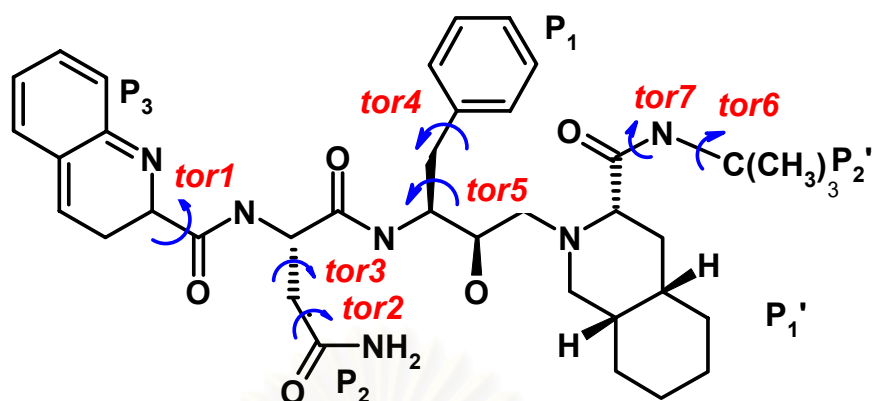


Figure 4.7 According to a conventional classification of the protease subsites, the binding pockets are designated by the inhibitor side chains, P₁, P₂, P₃, P₁' and P₂'. Torsional angles of the subsites were defined by *tor1-tor7*.

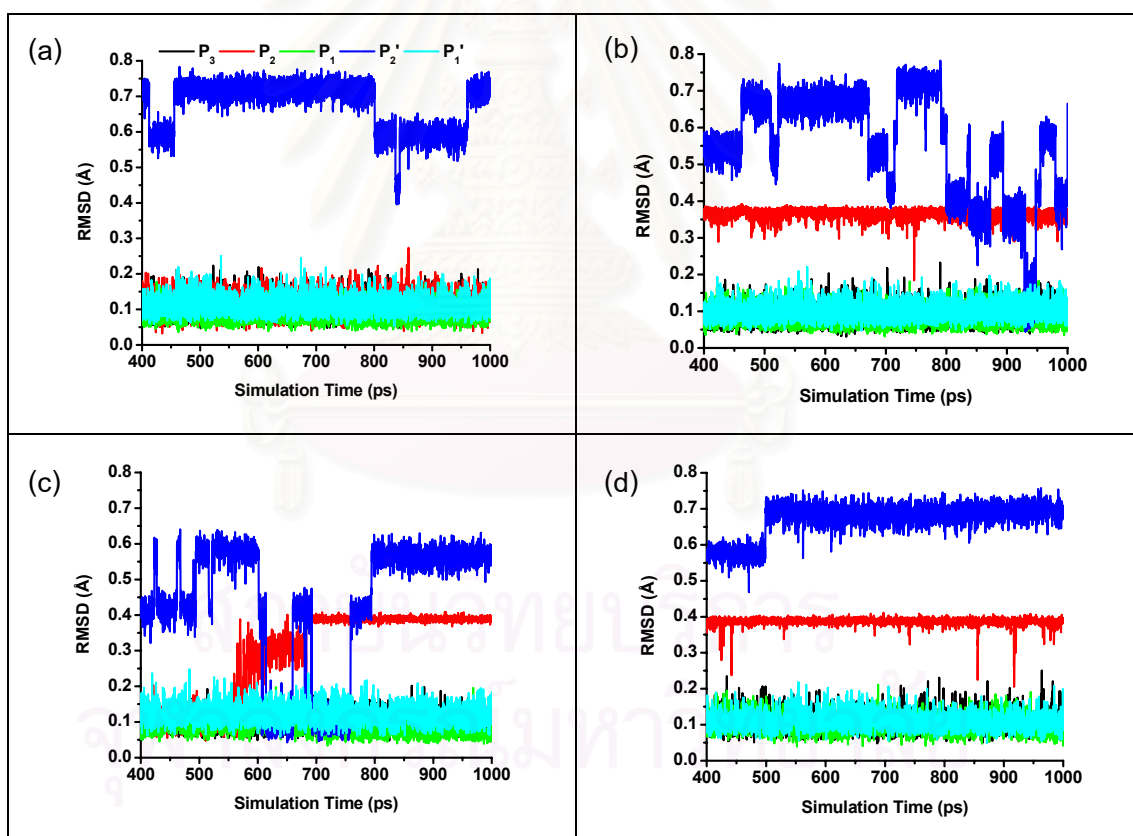
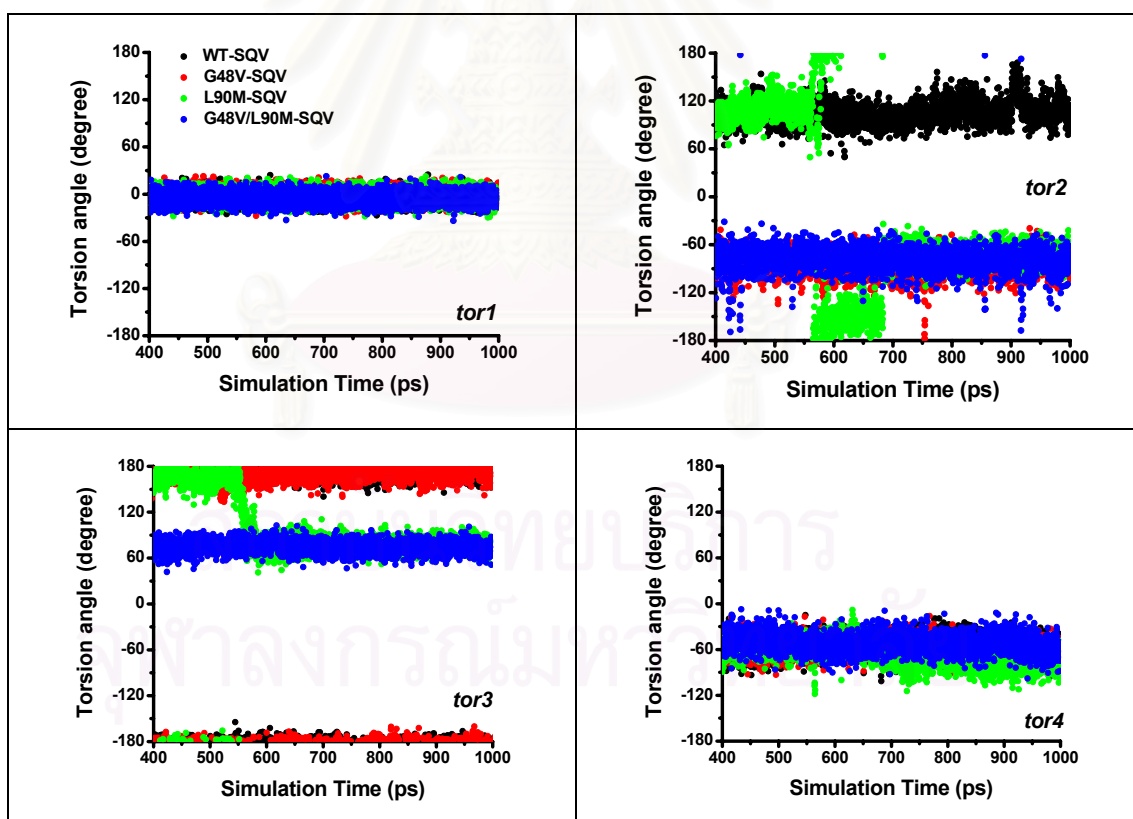


Figure 4.8. RMSDs of each SQV subsite with respect to the starting structure versus the simulation time for the wild-type (a), G48V (b), L90M (c) and G48V/L90M (d) mutants, where black, red, green, blue and sky-blue lines represent the RMSDs for the P₃, P₂, P₁, P₂' and P₁' subsites, respectively.

The RMSD plots show that conformation of the P_3 (black), P_1 (green), and the P_1' (sky-blue) subsites for all simulations do not change significantly during the simulations. This indicates that the inhibitor sidechains of the bound SQV in the binding pocket of the enzyme are considerably rigid. Meanwhile, P_2' (blue) sidechains for all systems show larger degree of atomic displacements. These are due to the rotation of the three methyl groups of the isopropyl chain. Thus, the flexibility of these subsites is relatively higher upon binding. It appears that the most significant difference between the wild-type and the mutants was found at the P_2 subsite of SQV. The RMSD changes in the single and double mutants are higher than that in the wild-type. This indicates that there is a greater degree of conformational change of the P_2 subsite taking place on the G48V, L90M and G48V/L90M complex only.



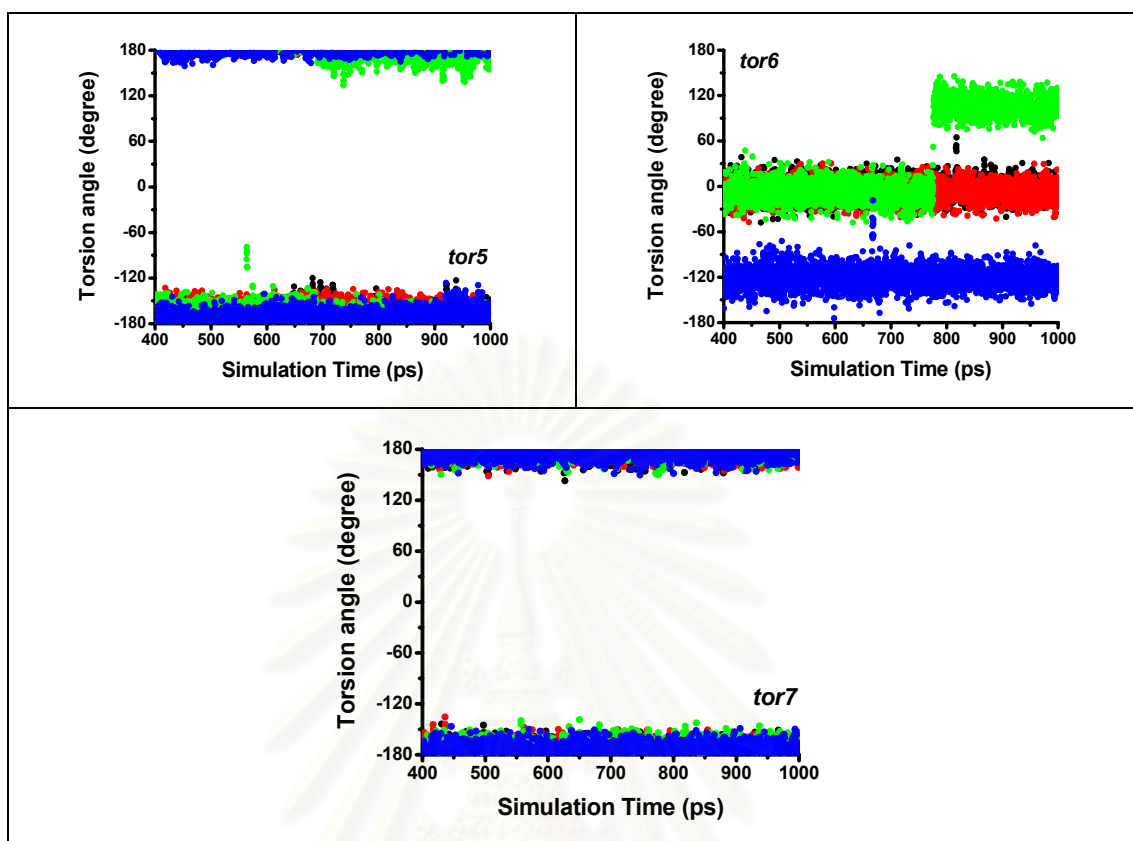


Figure 4.9 Fluctuation of the *tor1-tor7*, corresponding to the dihedral angles of the inhibitor side chains, for the P_1 , P_2 , P_3 and P_1' , respectively (see Figure 4.7 for definition).

From Figure 4.9, an oscillation of all dihedral angles throughout the simulations suggested that in the WT-SQV complex, all SQV sidechains undergo a narrow range of dynamic fluctuation. In other words, the inhibitor side chains were inflexible and retained their starting conformation during the course of the MD trajectory. In the case of the mutants, all torsion angles except for *tor2*, *tor3* and *tor6* adopted the values similar to that of the WT-SQV (Figure 4.9). This indicated that the conformations of these subsites are unchanged. However, a remarkable shift in *tor2*, *tor3* and *tor6* imply a rotation of the P_2 and P_2' subsites. For *tor6* of P_2' subsite, the L90M-SQV system shows two preferential conformations at 0° and 120° , while the wild-type, G48V and G48V/L90M show only one preferential conformations at 0° , 0° and -120° , respectively. This fact can be understood by a free rotation of the $-C(CH_3)_3$ functional groups. Thus, the three preferential conformations (120° , 0° and -120°) are identical due to molecular symmetry.

The results of *tor2* and *tor3* suggested a substantial rotation of the P₂ of SQV in the all mutants (Figure 4.9). The different of torsional angle (*tor2*) was observed during MD simulation, ~90° for WT-SQV and ~-90° for G48V-SQV and G48V/L90M-SQV. In case of L90M-SQV, *tor2* changes from 90° to -90° as shown in Figure 4.9. For *tor3*, the sidechains of WT-SQV and G48V-SQV located at the same torsion angle (180°), whereas, those of the G48V/L90M-SQV and L90M-SQV are 80°. This rotation of P₂ subsite involves a decrease of the strength of the hydrogen bond between the mutated residue and the P₂ subsite (described in the next topic). The overall change in terms of torsion angles is supposed to explain a decrease of saquinavir sensitivity.

4.2.3 Flexibility and conformational change of HIV-1 PR

4.2.3.1 Catalytic Site of HIV-1 PR

The catalytic cavity of the HIV-1 PR located at the dimer interface and formed by Asp25 and Asp25' is shown in Figure 4.1. To investigate flexibility and conformational changes of the active site, the RMSD value of the Asp25 and Asp25' were calculated and displayed in Figure 4.10. From the plots, it indicates that RMSD of Asp25 is larger than that of Asp25' for all systems. This is due to hydrogen atom that added to generate Mono-25 protonation state for the simulated system (see 4.1.4 for details). Moreover, the RMSD of Asp25 in WT-SQV complex fluctuates in widely range and shows highest RMSD when compared with other three complexes. However, such high flexibility dose not effect the binding between hydroxyl group of SQV and active site of the HIV-1 PR because the high RMSD is due to the rotation of the CB-CG bond of the Asp25 (Figure 4.11) that dose not disturb the essential hydrogen bonding (more details in next topic).

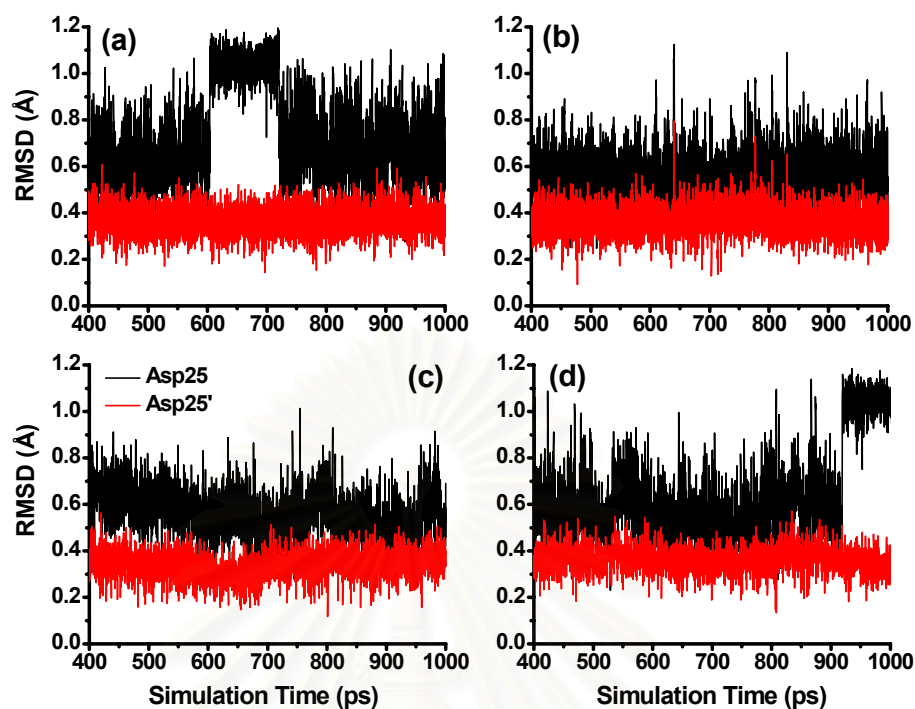


Figure 4.10 RMSD of Asp25 (black) and Asp25' (red) for WT-SQV (a), G48V-SQV (b), L90M-SQV (c) and G48V/L90M-SQV (d).

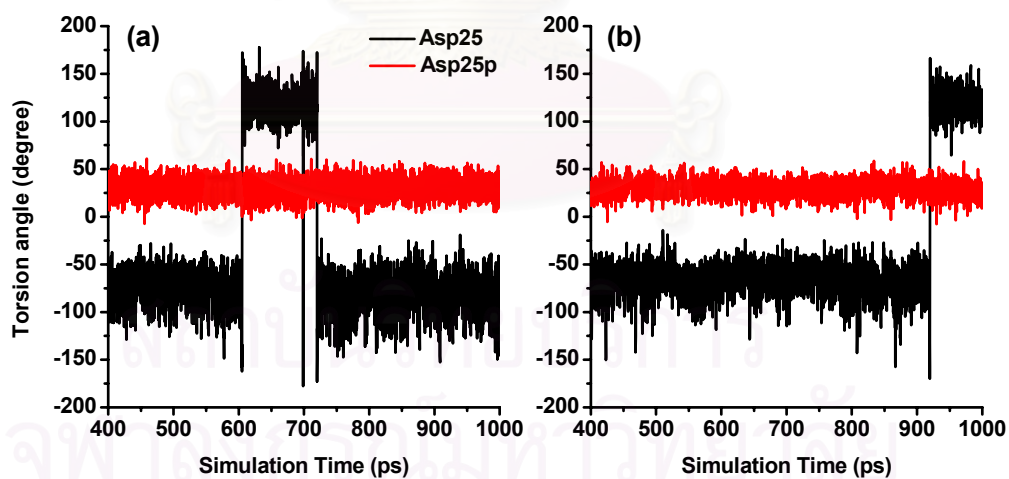


Figure 4.11 The CA-CB-CG-OD1 torsional angles of Asp25 of WT-SQV (a) and G48V/L90M-SQV (b) complexes.

4.2.3.2 The mutate residues of HIV-1 PR

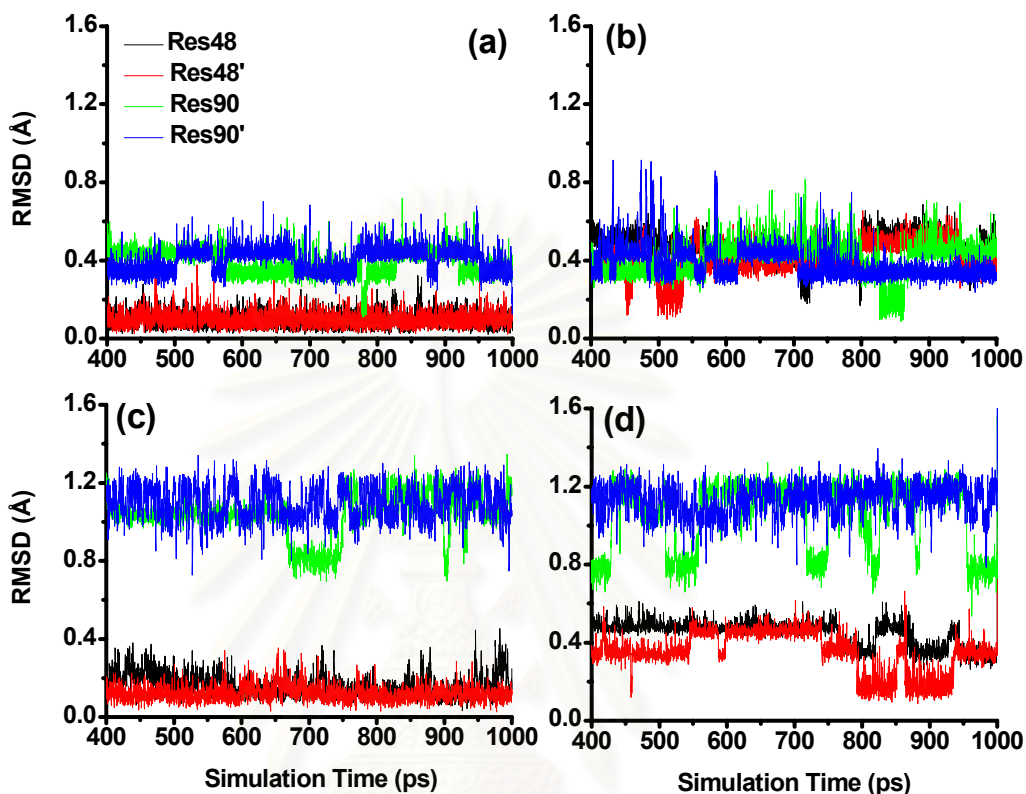


Figure 4.12 RMSD of residues 48 (black), 48' (red), 90 (green) and 90' (blue) for WT-SQV (a), G48V-SQV (b), L90M-SQV (c) and (d) G48V/L90M-SQV.

In Figure 4.12, the conformational change and flexibility of mutate residues which are Res48, Res48', Res90 and Res90' were evaluated in term of RMSD plots with respect to the initial structure. The apparent different between the wild-type and the mutants HIV-1 PR-SQV complexes was observed for all mutate residues. From the plots, RMSDs of all four residues thought the simulation suggested that all four residues in the WT-SQV complex undergo a narrow range of dynamic fluctuation (Figure 4.12a). The observation fir the WT-SQV system is completely in contrast with that of the G48V/L90M-SQV complex as shown in Figure 4.12d. For single mutants (G48V (Figure 4.12b) and L90M (Figure 4.12c), the RMSDs of mutate residues Val48 (Val48') for G48V and Met90 (Met90') for L90M are higher than that in the WT-SQV complex. It indicates that the structure of mutate residues of mutant complexes differ from that of the WT-SQV

complex. Surprisingly, all four residues show higher flexibility and larger RMSD values for G48V/L90M-SQV complex (Figure 4.12d). This behavior can directly affect the binding between drug and catalytic pocket of HIV-1 PR. Moreover, the high RMSD values of double mutant complex support its resistant fold and explain a decrease of saquinavir's sensitivity in the HIV-1 PR inhibition.

4.2.3.3 Flap region of HIV-1 PR

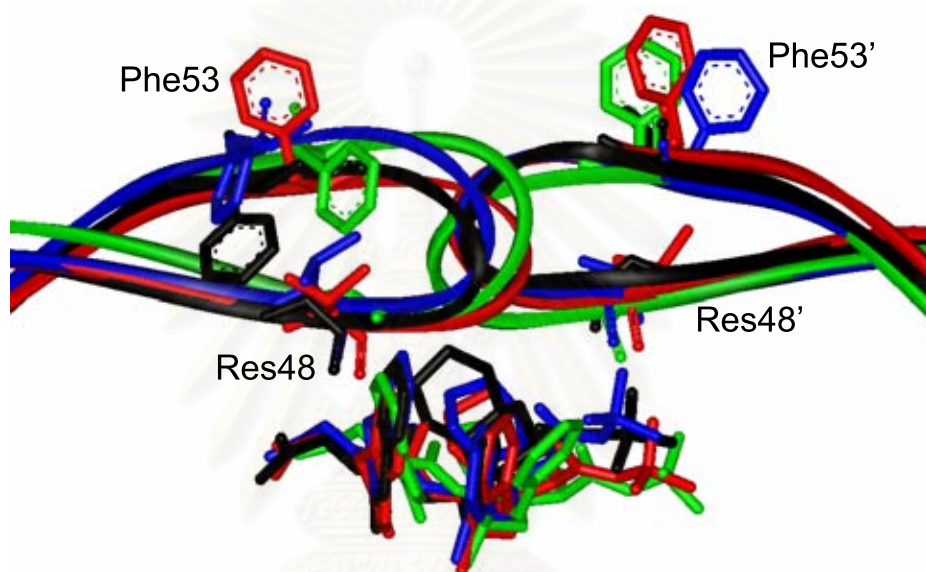


Figure 4.13 HIV-1 PR flap structures of the WT-SQV (black), G48V-SQV (red), L90M-SQV (green) and G48V/L90M-SQV (blue), some selected residues and saquinavir are presented in stick mode. For simplicity, hydrogen atoms are not shown.

Significant changes of atomic positions were observed at the protein flaps. Moreover, the flap structure of chain A is different from that of chain B. This is caused by the interactions of the enzyme to the asymmetric inhibitor. Particularly, the perturbation adopts heavily on the flap conformation of chain A rather than chain B due to a bulky group P_3 sidechains. As shown in Figure 4.13, position 48 was in close contact with the Phe53 side chain of the enzyme and the P_2 and P_3 groups of SQV. Since the substituted valine of the mutant cannot be entirely accommodated in the hydrophobic pocket due to steric conflict of the dimethyl groups with the Phe53 and Phe53' side chains, the flap, therefore, shifted toward the symmetric axis of the enzyme.

This rearrangement additionally destabilizes hydrogen bonding between the backbone CO of residue 48 of the enzyme and the P₂ subsite of the inhibitor. In addition to the flap movement, the side chain of hydrophobic Phe53 became solvent-exposed to avoid steric clashes with Val48 and Met46, whereas an orientation of Phe53' in flap B of the mutant was not significantly changed. The role of the HIV-1 PR mutation at the flexible flap has been considerably debated about whether it would facilitate the binding reaction or reduce stability of the inhibitor, or both (92,104). The crystal structure of the double mutant G48V/L90M complexed with SQV revealed side-chain rearrangement of the P₂ subsite and the Phe53 of the enzyme similar to our study (92). Particularly, the missing of hydrogen bonding between the P₂ subsite and the backbone C=O of residue 48 was also found in the x-ray structure of the double mutant. The crystal structure of G48H complexed with peptidic inhibitor U-89360E reported a decrease of flap mobility to stabilize the ligand (92).

4.2.4 Hydrogen bonding in binding site

The observation of hydrogen bonding between HIV-1 PR and SQV based on the donor-acceptor distance criteria in CARNAL module of AMBER 7 (distance < 3.5 Å and angle < 60 degrees) along the simulation are demonstrated in Figure 4.14 and 4.15. In the WT-SQV complex, SQV was held in the cavity of the HIV-1 PR by the three hydrogen bonds of ~100% occupation with Gly48 in flap region and two hydrogen bonds with catalytic site (Asp25 and Asp25'). The MD results show also that the conformational difference in SQV subsites between the wild-type and the mutant complexes was located at the P₂ side chain. Among residues surrounding the P₂ subsite, position 48 is critical to conformational change of the inhibitor subsite. For the single mutation at G48V which is located at the flap region, the rotation of P₂ subsite of saquinavir (see Figure 4.14) give significantly result compared to wild-type. One is the disappearance of hydrogen bonding formed between NH-group of P₂ subsite and the C=O peptide bond of the residue 48 backbone as indicated by distance *d*₆. The plot clearly shown that the G48V/L90M double mutation cause the weaker hydrogen bonding formed between the saquinavir and the catalytic dyad residue (Asp25 and Asp25')

indicated by distance $d1-d3$ and hydrogen bonding to the flap region (residue 48) can not be observed throughout the simulations as indicated by distance $d4-d6$. In case of L90M, the observed hydrogen bonds are similar to WT-SQV system, however, the occupations of hydrogen bonds with Gly48 are lower.

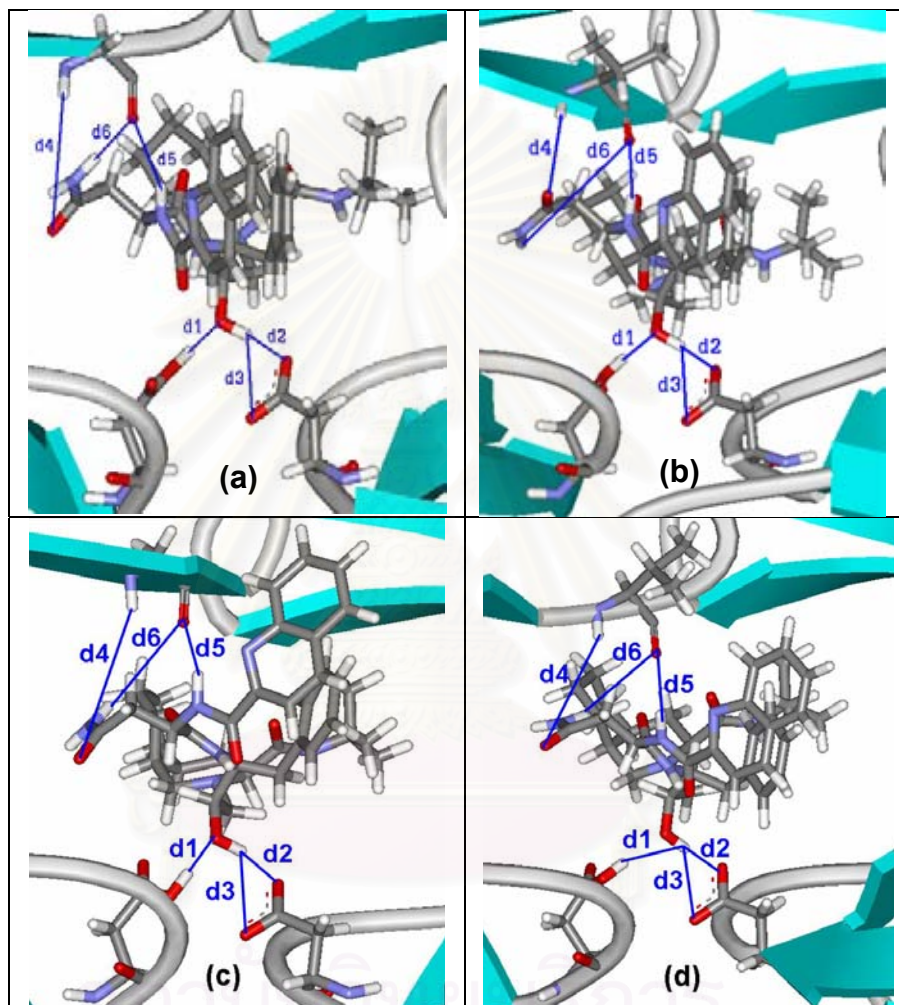


Figure 4.14 The complex structures representative for WT-SQV, G48V-SQV, L90M-SQV and G48V/L90M-SQV systems with the depicted hydrogen bonds distances $d1-d6$.

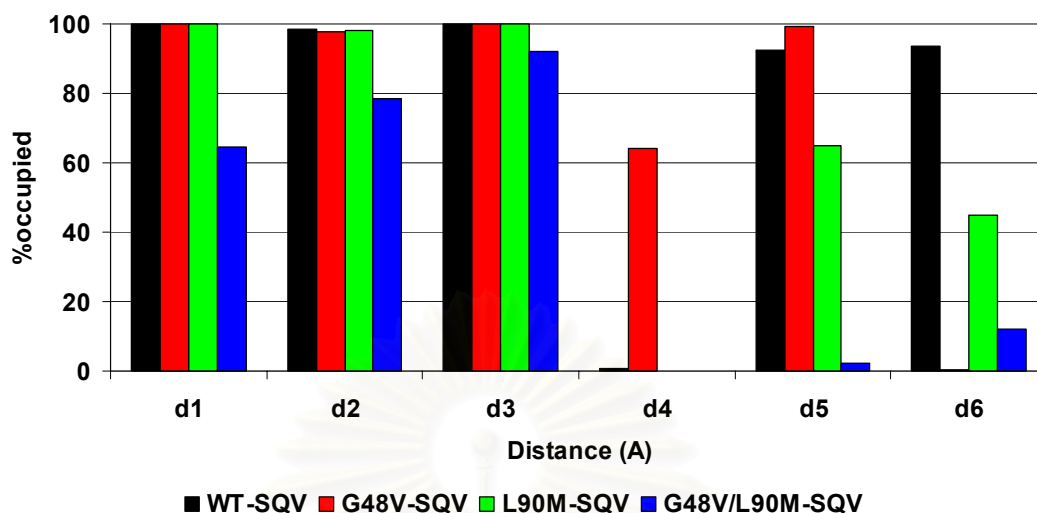


Figure 4.15 The %occupied of hydrogen bonding ($d1-d6$) through the simulations.

The lost of ligand-enzyme hydrogen bonds agree well with the calculated binding free energy in which (i) the G48V/L90M leads significantly to decrease the ΔG_{bind} from -95.10 kcal/mol of WT to -84.15 kcal/mol of double mutation, and (ii) The ΔG_{bind} for WT (-95.10 kcal/mol) and single mutants (-94.80 kcal/mol for G48V and -95.25 kcal/mol for L90M) are almost the same, supporting the above hydrogen bond data where both WT and single mutants demonstrate the same number of hydrogen bond, only slightly difference in terms of %occupation.

4.2.5 Specific interaction energy at the catalytic site of the HIV-1 PR

In order to seek for more detailed and precise information on the ligand-enzyme binding, interaction between saquinavir and the catalytic as well as mutated residues of the HIV-1 PR were calculated. Here, the three-layer ONIOM approach (ONIOM3) was performed. The average structures from 1-ns equilibrated MD trajectories were used as the starting geometry of the system interest.

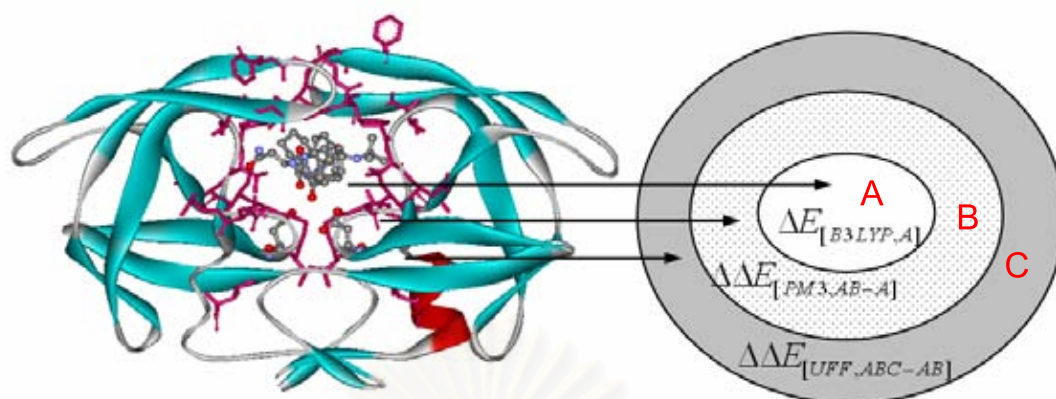


Figure 4.16 The schematic representation for three-layered ONIOM calculations. The inner layer (A) are SQV, Asp25(25'), Gly48(48') and Leu90(90') as shown in Figure 4.17. The mediate layer (B) is the residues within 5 Å from SQV. The outer layer (C) presents entire enzyme.

In Figure 4.16, the systems of three possible protonation state of HIV-1 PR/SQV complexes were divided in to three parts represented by inner layer (A), intermediate layer (AB) and *real* layer (ABC). The inner layer (Figure 4.17) representing the direct binding strength of SQV to catalytic dyad (Asp25/Asp25'), flap region (Gly48/Gly48') and Leu90 /Leu90' was treated at the high level of quantum chemical calculation using density functional theory (B3LYP/6-31G(d,p)). The surrounding amino acid residues interacting to SQV within the distance of 5 Å were included to the intermediate layer representing the contributed interactions such as the Van de Walls interaction from the environment, and treated with the lower level of calculation using semiempirical method (PM3). These amino acid residues added to the intermediate layer were taken from both chain A and B which were totally 36 amino acid residues (Figure 4.16 shown as in magenta stick model: AB). The real layer or the entire enzyme system (Figure 4.16 shown as the ribbon model: ABC) was able to carried out by the molecular mechanic method using universal force field (UFF). All calculations based on ONIOM approach were carried out using GAUSSIAN98 program package.

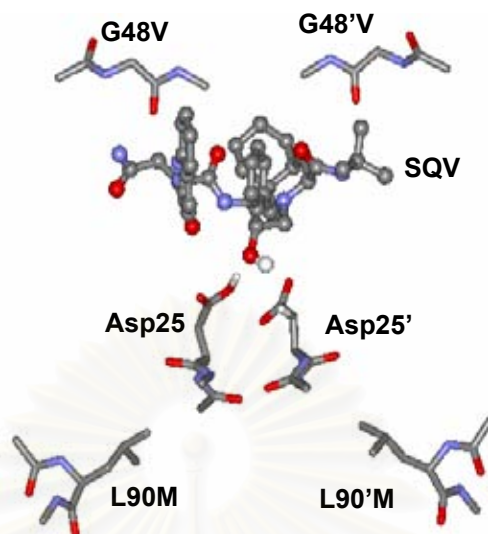


Figure 4.17 Schematic structure represents the quantum model for calculations the interaction energy of saquinavir with the key amino acid residues.

Here, the interaction energy (ΔE) of SQV to the HIV-1 PR was defined by:

$$\Delta E = E_{\text{complex}} - E_{\text{PR}} - E_{\text{SQV}} \quad (4.1)$$

where E_{complex} , E_{PR} and E_{SQV} represents total energy of the HIV-1 PR/SQV complex, HIV-1 PR and SQV, respectively. Hence, the total energy obtained from ONIOM3 calculations (E^{ONIOM3}) herein can be expressed as:

$$E^{\text{oniom3}}[ABC] = E_{[\text{UFF},ABC]} + E_{[\text{PM3},AB]} + E_{[\text{B3LYP},A]} - E_{[\text{UFF},AB]} - E_{[\text{PM3},A]} \quad (4.2)$$

Therefore, the total interaction energy ($\Delta E_{\text{total}}^{\text{oniom3}}[ABC]$) between SQV and the HIV-1 PR using ONIOM3 method can be written by the expression of independent energy components as following:

$$\Delta E_{\text{total}}^{\text{oniom3}}[ABC] = \Delta E_{[\text{UFF},ABC]} - \Delta E_{[\text{UFF},AB]} + \Delta E_{[\text{PM3},AB]} - \Delta E_{[\text{PM3},A]} + \Delta E_{[\text{B3LYP},A]} \quad (4.3)$$

$$\Delta E_{total}^{oniom3} [ABC] = \Delta E_{[B3LYP,A]} + \Delta\Delta E_{[PM3,AB-A]} + \Delta\Delta E_{[UFF,ABC-AB]} \quad (4.4)$$

where $\Delta E_{[B3LYP,A]}$ is the interaction energy in the region A, particularly for the direct interaction between SQV with the catalytic dyad residues, Asp25 and ASP25', evaluated at the B3LYP/6-31G(d,p) level. $\Delta\Delta E_{[PM3,AB-A]}$ is the interaction energy between the region A and B evaluated at the PM3 level, and $\Delta\Delta E_{[UFF,ABC-AB]}$ is the interaction between regions AB and C evaluated at the molecular mechanics (UFF), all in the HIV-1 PR/SQV complex.

In Table 4.3, the ONIOM3 results show the lowest ΔE_{total} of SQV binding to PR taken place on the G48V mutant. This is by 2.52 kcal/mol slightly lower than that of WT but, by 25.87 kcal/mol much lower than that of G48V/L90M mutant. In case of G48V/L90M, this result clearly reveals the large decreasing in the binding affinity of SQV to PR which is well corresponding to the 419-fold of resistance over WT. For the G48V, an order does not relatively correlate to the biological relevant if using the direct comparison of the ΔE_{total} of WT and G48V mutant. However, this results can be explained by decomposition of the ΔE_{total} into three interaction terms ($\Delta E_{[B3LYP,A]}$, $\Delta\Delta E_{[PM3, AB-A]}$ and $\Delta\Delta E_{[UFF,ABC-AB]}$) that contributed from 3 different layers of protein complex (see also Figure 4.16). By focusing on the specific interaction at catalytic site, the result of $\Delta E_{[B3LYP,A]}$ shows that SQV still binds tighter to Asp25/Asp25', Leu90/Leu90' at catalytic site and Gly48/Gly48' at flap region in WT than that in G48V mutant which is still in a good agreement with the previous results from AMBER force field and DFT. $\Delta\Delta E_{[PM3, AB-A]}$ and $\Delta\Delta E_{[UFF,ABC-AB]}$ represents the interaction contributed from the protein environment in 5 Å and the rest of entire protein. In case of comparison between G48V and WT, the interaction contributed from these protein effects brings the ΔE_{total} of G48V to be lower than WT. This is not in case for G48V/L90M although, there is a larger stabilization energy come from $\Delta\Delta E_{[UFF,ABC-AB]}$ of about -22.55 kcal/mol compared to -16.38 kcal/mol and -17.74 kcal/mol of WT and G48V, respectively. This due to the much lower of the direct interact at $\Delta E_{[B3LYP,A]}$ compared to WT and G48V mutant in which the differences, G48V/L90M-WT and G48V-WT, are about 25.04 kcal/mol and 22.43 kcal/mol, respectively. Note that $\Delta\Delta E_{[PM3,AB-A]}$ and $\Delta\Delta E_{[UFF,ABC-AB]}$ represent the interaction

contributed from the protein environment in 5 Å and the rest of entire protein (see also Figure 4.17), respectively. A conclusion is that the interaction contributed from the entire protein brings the ΔE_{total} of G48V to be lower than WT. This is not true for the G48V/L90M where is a larger stabilization energy comes from $\Delta\Delta E_{[UFF,ABC-AB]}$ of about -22.55 kcal/mol compared to -16.38 kcal/mol and -17.74 kcal/mol of WT and G48V, respectively. The implication from those results of stabilization energy contributed from the catalytic site and protein effects between WT to the mutant system is that *i*) the disruption of the direct interaction between drug and pocket, e.g., SQV--G48 and SQV--D25 by mutated residue in the direct contact (G48→V) play the major effect to causes the significant change of the drug binding affinity, *ii*) as in the mutated residue such as L90→M, repositioning of the protein structure can be not only due to the close contact but also through the entire protein.

Table 4.3 The interaction energy calculated by ONIOM3 and the component of decomposition.

Systems	Interaction energy (kcal/mol)				
	WT	G48V	$(\Delta\Delta E)^a$	G48V/L90M	$(\Delta\Delta E)^b$
ΔE_{total}	-84.52	-87.04	(-2.52)	-58.65	(25.87)
$\Delta E_{[B3LYP,A]}^c$	-41.08	-38.47	(2.61)	-16.04	(25.04)
$\Delta\Delta E_{[PM3, AB-A]}^d$	-27.05	-30.82	(-3.77)	-20.05	(7.00)
$\Delta\Delta E_{[UFF,ABC-AB]}^e$	-16.38	-17.74	(-1.37)	-22.55	(-6.17)

$$^a \Delta\Delta E = \Delta E_{G48V \text{ mutant}} - \Delta E_{wild-type}$$

$$^b \Delta\Delta E = \Delta E_{G48V/L90M \text{ mutant}} - \Delta E_{wild-type}$$

^c $\Delta E_{[B3LYP,A]}$ accounts for interaction of SQV and Asp25/25', Gly48/48' and Leu90/90'

^d $\Delta\Delta E_{[PM3, AB-A]}$ represents interactions of SQV and the 5 Å surrounding residues, excluding Asp25/25', Gly48/48' and Leu90/90'

^e $\Delta E_{[UFF,ABC-AB]}$ represents interactions of SQV and the remaining residues

4.3 Conclusion

Molecular dynamics simulations of WT, G48V, L90M and G48V/L90M HIV-1 PR complexed with SQV were carried out to investigate the molecular basis of drug resistance. The MD results combined with quantum chemical calculations extend the capability of molecular modeling methods to study some biological systems. Overall tertiary structure of the WT and the mutant PR were not significantly altered. Particularly, structure and interactions at the central active site are almost unchanged. However, conformational differences between the WT and the G48V mutant were on the protein flap, which slightly change was observed for the G48V/L90M. The conformational change of the P₂ subsite decreases the strength of hydrogen bonding of the backbone CO of the residue 48 to SQV. The change in binding free energies for G48V/L90M was comparable to the experimental K_i data. These observations provide useful information for designing potent HIV-1 PR inhibitors.

CHAPTER 5

MD SIMULATIONS OF WILD-TYPE AND MUTANTS HIV-1 PR COMPLEXED WITH RITONAVIR IN AQUEOUS SOLUTION

In this chapter, structural and dynamics properties of the primary HIV-1 PR resistance against ritonavir were investigated using molecular dynamic simulations. The simulations were carried out for the four systems; WT, V82F, I84V and V82F/I84V HIV-1 protease complexed with ritonavir in explicit aqueous solution. In addition, the free energy of binding was performed using molecular mechanics Poisson-Boltzmann Surface Area (MM/PBSA) approach. The MM/PBSA method offers an efficient computation in calculating the binding free energy of biomolecular systems. It has been extensively used to explore many receptor-drug complexes.

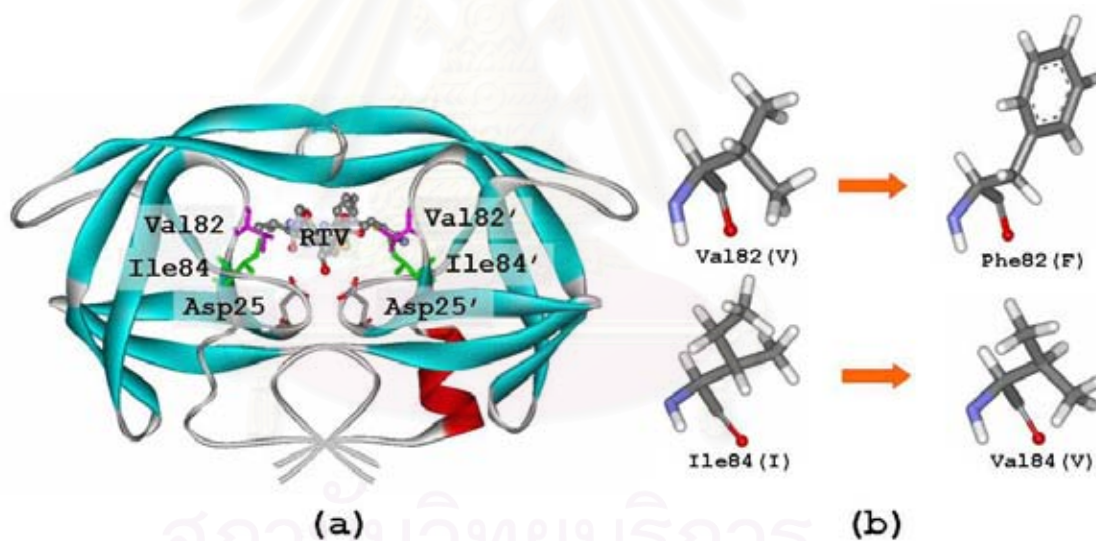


Figure 5.1 Schematic representation of wild-type HIV-1 protease complexed with ritonavir molecule (a). In the catalytic residues Asp25/Asp25' and the mutated residues Val82/Val82' and Ile84/Ile84' are shown by stick and ritonavir is shown by the ball and stick.

5.1 Computational methods

5.1.1 Atomic charges of ritonavir

Ritonavir is not standard residues in the AMBER residue libraries; therefore, their force field parameters are not available in the AMBER package. Thus, potential parameters for ritonavir were developed using the following steps. Geometry of the starting structure of ritonavir was taken from the crystal structure of saquinavir bound to wild-type protease (PDB code: 1HXW). Then, it was optimized at the Hartree-Fock level with 6-31G** basis functions to adjust the bond-length involving hydrogens. Finally, the RESP fitting procedure was employed to calculate partial atomic charges of the inhibitor. All electronic structure calculations were carried out with the program package Gaussian98 (83). Partial charge generation and assignment of the force field were performed using the Antechamber suite (105). The results were summarized in Table 5.1.

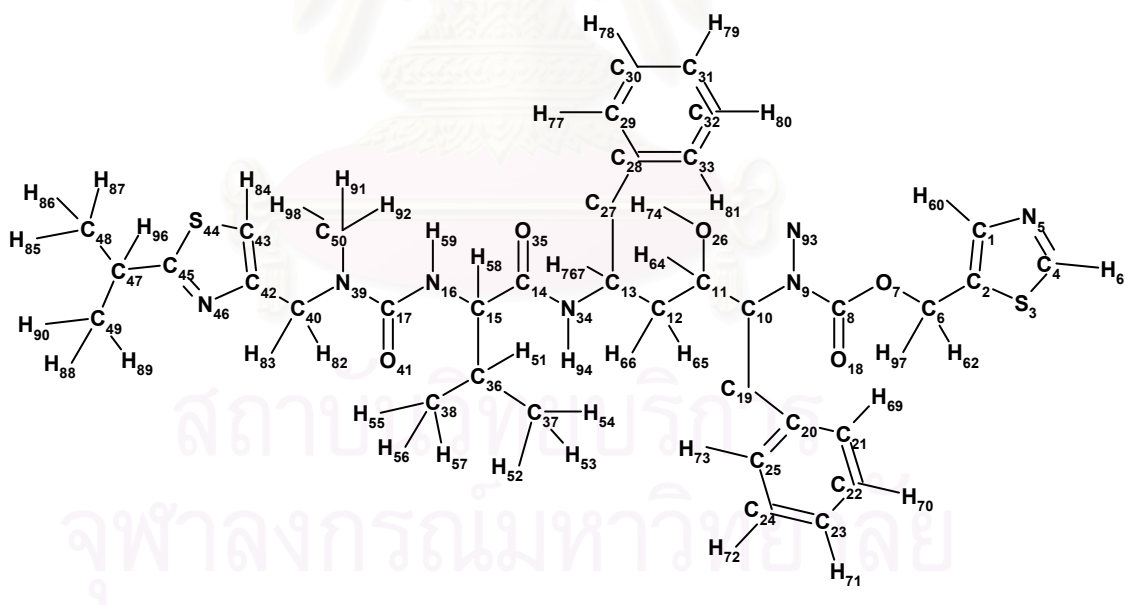


Figure 5.2 The atom labels of ritonavir.

Table 5.1 RESP charges of ritonavir molecule obtained from single point HF/6-31G* calculation on the HF/6-31G** optimized geometry.

Atom label	Charge	Atom label	Charge	Atom label	Charge
C48-C49	-0.345	H58	0.042	C21	-0.256
H85-H87	0.068	C14	0.571	C22	-0.083
C47	0.614	O35	-0.557	C23	-0.201
H88-H90	0.068	N34	-0.590	C24	-0.083
H96	-0.075	H94	0.277	C25	-0.256
C45	0.039	C13	0.875	H73	0.148
N46	-0.409	C27	-0.711	H72	0.129
S44	-0.028	C28	0.325	H71	0.140
C43	-0.235	C29	-0.306	H70	0.129
H84	0.217	C30, C32	-0.077	H69	0.148
C42	0.225	C31	-0.185	H68, H95	-0.034
C40	-0.266	C33	-0.306	H63	0.056
H82-H83	0.146	H77, H81	0.171	N9	-0.652
N39	-0.191	H78, H80	0.129	H93	0.289
C50	-0.424	H79	0.132	C8	1.056
H91-H92	0.161	H75-H76	0.176	O18	-0.640
H98	0.161	H67	-0.074	O7	-0.431
C17	0.852	C12	-0.437	C6	-0.059
O41	-0.626	H65-H66	0.025	H62, H97	0.139
N16	-0.768	C11	0.770	C2	-0.203
H59	0.369	O26	-0.794	C1	0.258
C15	0.040	H74	0.459	H60	0.122
C36	0.393	H64	-0.142	N5	-0.569
C37-C38	-0.401	C10	0.018	C4	0.193
H52-H57	0.097	C19	0.031	H61	0.133
H51	-0.025	C20	0.189	S3	0.006

5.1.2 Preparation of the initial structures

The structures of the HIV-1 protease-ritonavir (RTV) in this chapter include the wild-type, two of single mutants (V82F and I84V) and a double mutant (V82F/I84V). The 1.8 Å resolution crystallographic structure of the WT HIV-1 PR-RTV complex (PDB entry 1HXW) was used as the initial structure (41). All missing atoms and hydrogens of the enzyme were added using the Leap module in AMBER 7 package (90). Because the x-ray structure of the mutants is unknown, therefore the V82F, I84V and V82F/I84V structures were constructed using the WT model as the template. According to the pK_a calculation, the ionization states of all ionizable residues computed by the UHBD (University of Houston Brownian Dynamics) program (85) were assigned to their common charge, *i.e.*, Glu⁻¹, Asp⁻¹, Lys⁻¹ and Arg⁺¹, except for the catalytic Asp25 (97,98). The single protonation at Asp25 residue has experimentally and theoretically been proposed for a number of the HIV-1 protease-inhibitor complexes, particularly inhibitor containing the hydroxyl-ethylene isostere (79,99). In this study, one of the two aspartyl residues, Asp25, of the enzyme was assigned as a neutral state. Thus, the four simulated systems were in the monoprotection state. Each complex was immersed in a 10 Å radius of TIP3P water model (91). The crystallographic water molecules were kept in the systems. Counter ions were added to neutralize the system using LEaP module.

Table 5.2 Number of water molecules, counter ions and total atoms together with the dimension of simulation boxes of the four simulations, WT, V82F, I84V and V82F/I84V.

system	Number of			Box dimension (Å ³)
	Waters	Counter ions	Total atoms	
WT	9027	8 Cl ⁻ , 3 Na ⁺	30327	83 × 72 × 64
V82F	9027	8 Cl ⁻ , 3 Na ⁺	30333	83 × 72 × 64
I84V	9027	8 Cl ⁻ , 3 Na ⁺	30319	83 × 72 × 64
V82F/I84V	9027	8 Cl ⁻ , 3 Na ⁺	30327	83 × 72 × 64

5.1.3 Molecular dynamics simulations

Energy-minimization and MD simulations were performed using SANDER module of AMBER 7 (93) with the Cornell force field (100). First, energy minimization was applied for water molecules using 500 steps of the steepest descents and then 500 steps of the conjugate gradients. Subsequently, the whole systems were subjected to energy minimization with 1000 steps of the steepest descents and 1500 steps of the conjugate gradients.

The MD simulations were performed employing the periodic boundary condition with the NPT ensemble. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and pressure of the systems (94). The SHAKE algorithm (95) was employed to constrain all bonds involving hydrogens. The simulation time step of 2 fs was used. All MD simulations were run with a 12 Å residue-based cutoff for nonbonded interactions and the particle-mesh Ewald method was used for an adequate treatment of long-range electrostatic interactions.

The MD simulation consists of thermalization and equilibration. During 0-60 ps, the temperature of the system increased from 0 to 298 K using NVT condition. Then the temperature was maintained at 298 K during equilibration employing at NPT conditions. The reference pressure and temperature were set to 1 atm and 298 K, respectively. The MD trajectories were collected every 0.2 ps. The convergence of energies, temperature and pressure of the systems, and the atomic root-mean square deviation of the enzyme and the inhibitor (RMSD) were used to verify the stability of the systems. The series of snapshot between 1 and 2 ns of the equilibrium phase were used for the free energy calculations and structure evaluation.

5.1.4 Calculations binding free energy

The procedures of calculation binding free energy were performed in the same way of saquinavir systems. (see the details in computational method in chapter 4).

5.2 Results and discussion

5.2.1 Reliability of the simulations

The 2 ns MD trajectories of four systems consisting of the WT, V82F, I84V and V82F/I84V HIV-1 PR complexed with ritonavir were generated. The convergence of energies and temperature of the systems during a course of the simulations indicate well-behave of the systems as display in Figure 5.3-5.6. The atomic RMSD of the protein and inhibitor structure were shown in Figure 5.7. The mean values of those parameters were summarized in Table 5.3.

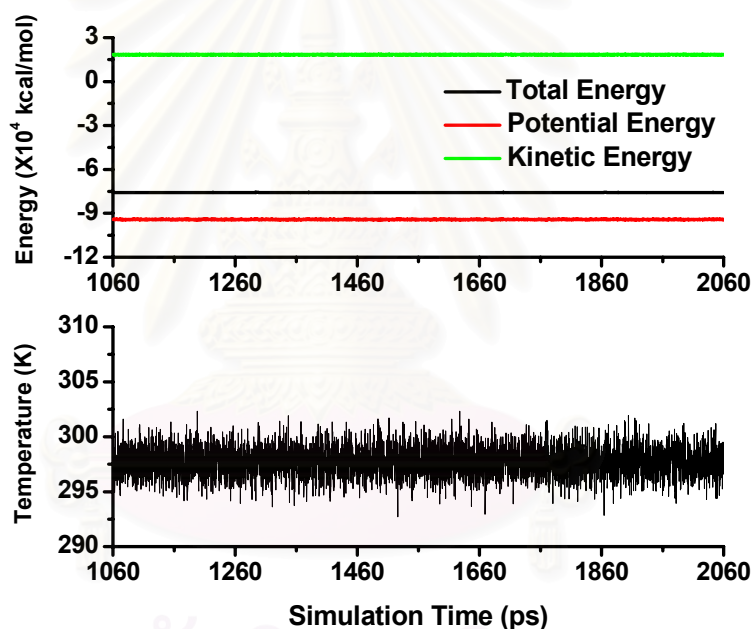


Figure 5.3 The plots of total (black), kinetic (green) and potential (red) energies as well as temperature over 1-2 ns for the WT-RTV system.

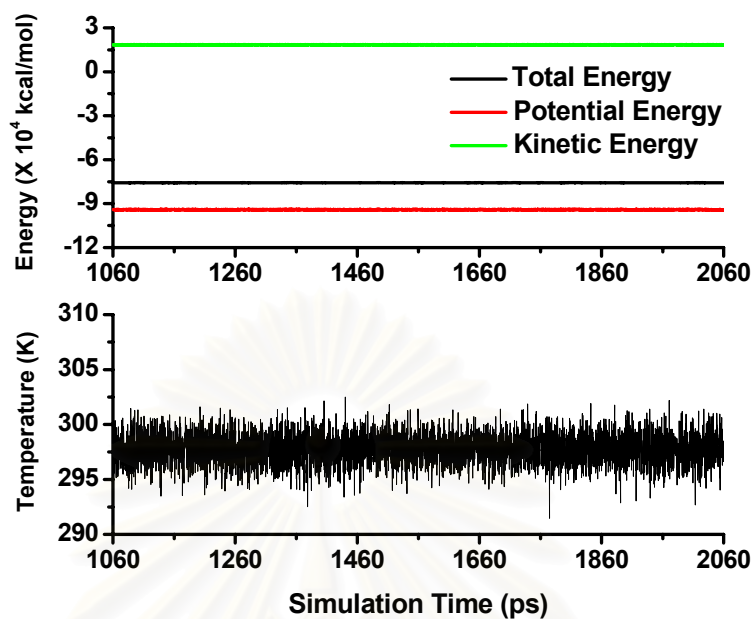


Figure 5.4 The plots of total (black), kinetic (green) and potential (red) energies as well as temperature over 1-2 ns for the V82F-RTV system.

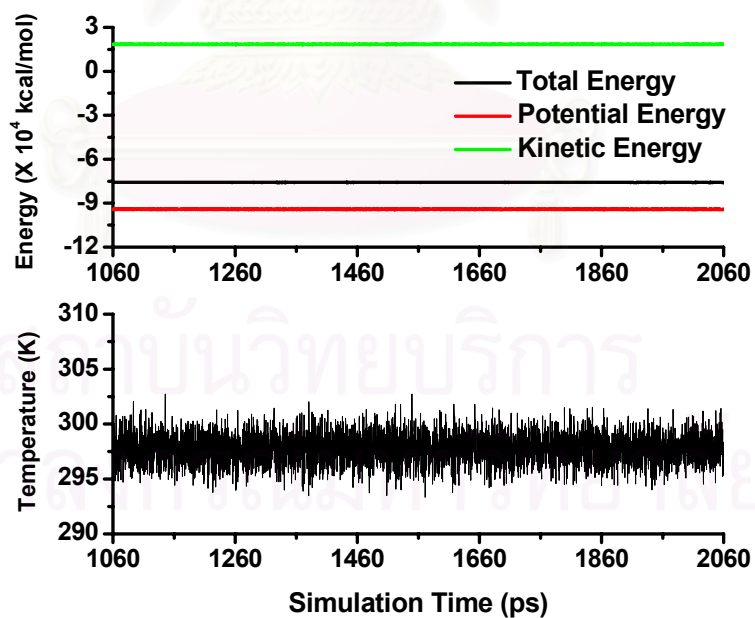


Figure 5.5 The plots of total (black), kinetic (green) and potential (red) energies as well as temperature over 1-2 ns for the I84V-RTV system.

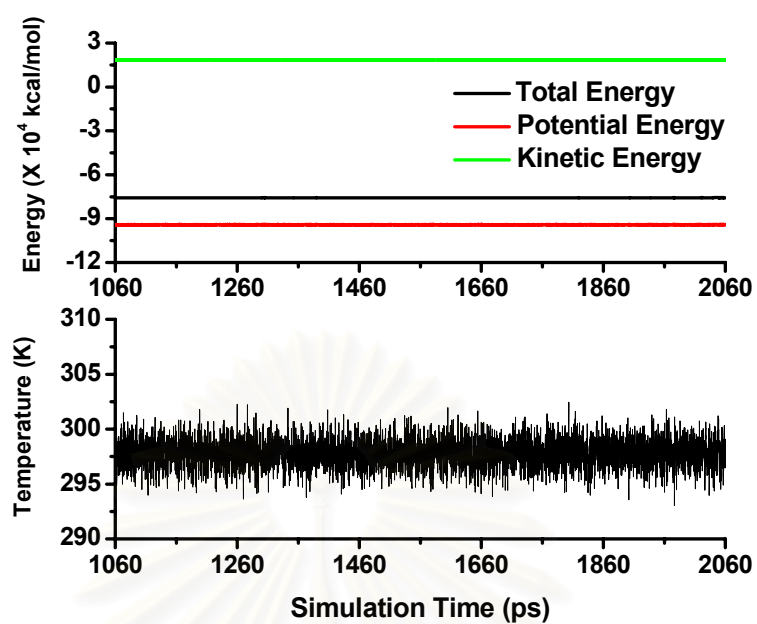


Figure 5.6 The plots of total (back), kinetic (green) and potential (red) energies as well as temperature over 1-2 ns for the V82F/I84V-RTV system.

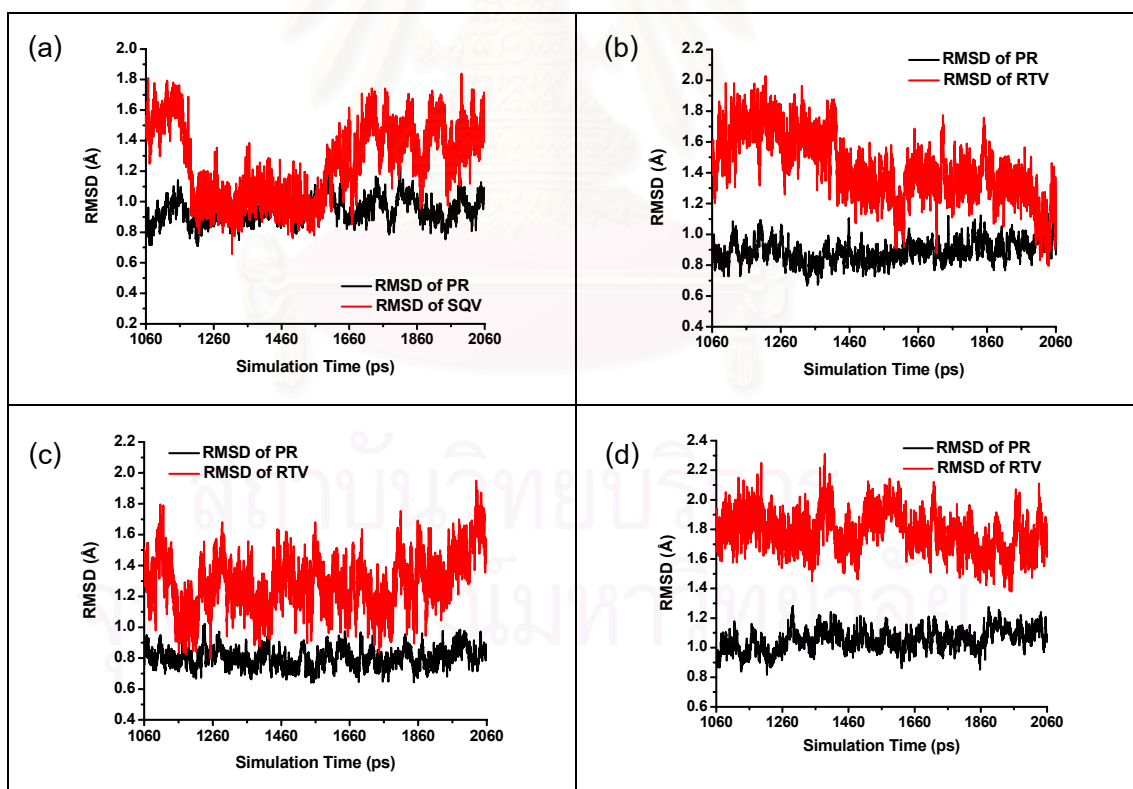


Figure 5.7 Plot of RMSDs versus simulation time for the (a) WT, (b) V82F, (c) I84V and (d) V82F/I84V complexed with ritonavir. The obtained RMSDs were computed using the structure at $t = 0$ as a reference.

Table 5.3 Mean values of total, potential and kinetic energies as well as temperature

Values	WT-RTV	V82F-RTV	I84V-RTV	V82F/I84V-RTV
Total Energy	-7.57×10^4	-7.57×10^4	-7.58×10^4	-7.57×10^4
Potential Energy	-9.42×10^4	-9.41×10^4	-9.42×10^4	-9.42×10^4
Kinetic Energy	1.84×10^4	1.84×10^4	1.84×10^4	1.84×10^4
Temperature	297.75	297.75	297.74	297.78
RMSD of SQV	1.25 ± 0.24	1.43 ± 0.21	1.28 ± 0.18	1.78 ± 0.13
RMSD of PR	0.95 ± 0.08	0.89 ± 0.07	0.80 ± 0.06	1.05 ± 0.07

5.2.2 Structural features of the HIV-1 protease

Structures of the WT and all three mutants taken from a snapshot of the MD trajectories are almost identical (Figure 5.8). From RMSD comparison, the global tertiary structures of all four systems are very similar. The structure of the catalytic triad residues of all three mutants is essentially identical to that of the WT with the RMSD < 0.5 Å (data not shown). However, a large RMSD difference can be seen on the HIV flaps of chain B (residues 49'-54') in the V82F/I84V mutant.

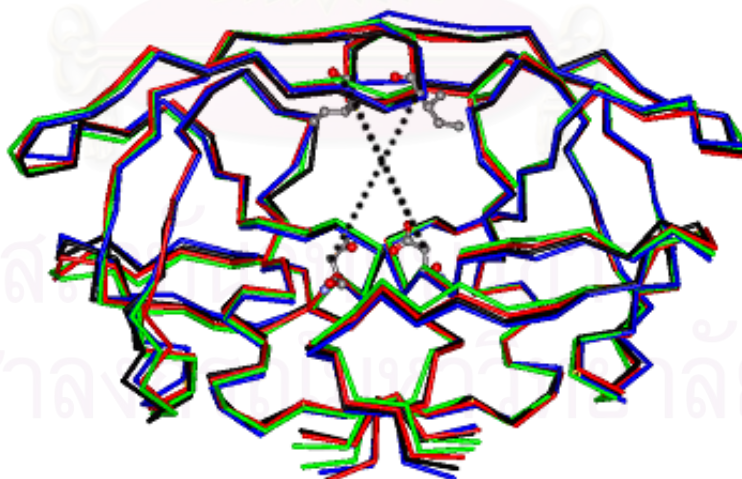


Figure 5.8 The C α -trace of the WT (black), V82F (red), I84V (green) and V82F/I84V (blue) mutants. Asp25, Asp25', Ile50 and Ile50' are displayed in ball and stick. Dashed lines demonstrate the measured distances of C α -Ile50/C β -Asp25 and C α -Ile50'/C β -Asp25'.

The degree of the flap change is simply demonstrated by a measure of the interresidue distances of $C\alpha$ -Ile50/ $C\beta$ -Asp25 (chain A) and of $C\alpha$ -Ile50'/ $C\beta$ -Asp25' (chain B). The distance separation was calculated from the MD trajectory and, then, statistically analyzed as population frequency. From Figure 5.8, distance distribution peaks of the WT and the mutants are mostly overlying for chain A while this is not true for chain B (Figure 5.9). A slight shift of the interresidue distance of chain B was observed for the single mutants but a greater change was found for the double mutant with in a range of 1.0-1.5 Å. This implies that the flap structure may have changed due to the mutation. This has been known that the inhibitor is not symmetric. It is likely that the flap tip of the double mutant has the greater degree of opening than that of the others. This may relate to a loss of interactions of the drug in the binding pocket of HIV-1 V82F/I84V protease as compared with that of the WT and the V82F and I84V mutants.

A number of MD studies published previously has reported large conformational changes of the flap (106). The MD simulations of the apo wild-type and the apo V82F/I84V mutant of HIV-1 protease observed more frequent and more rapid curling of the flap tip in the mutant than in the wild-type. Moreover, the mutant's flap also opened farther than the wild-type's, implying more flexibility of the mutant. The flexibility of the flap is highly depending on the availability and type of ligands.

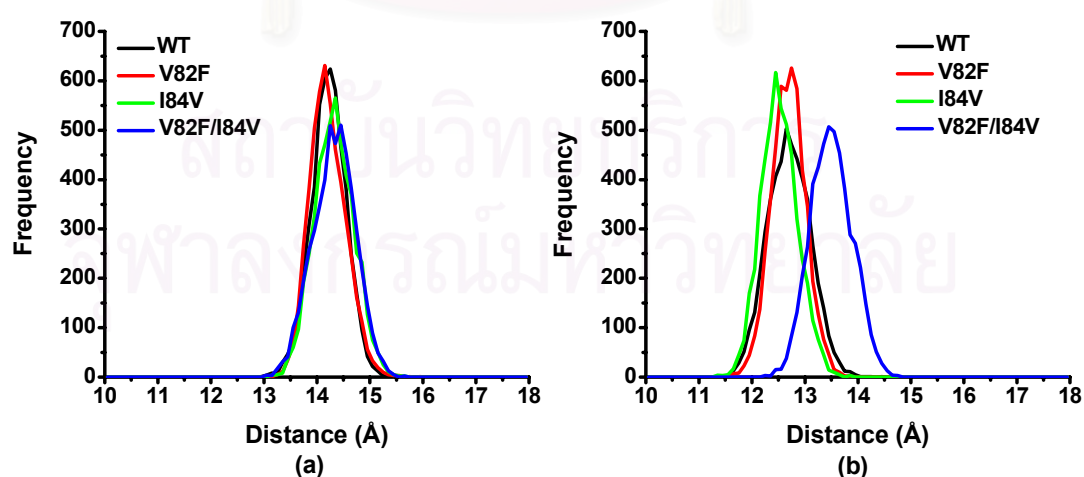


Figure 5.9 Distributions of the $C\alpha$ -Ile50/ $C\beta$ -Asp25 (a) and of the $C\alpha$ -Ile50'/ $C\beta$ -Asp25' (b) distances (details see text and labels in Figure 5.8) sampled during 1-2 ns.

It appears that the Ile50' sidechain of chain B is in close contact with P₁', P₂ subsites of RTV and residues 81 to 84 of chain A (Figure 5.10b). One should note that the interactions in this binding pocket are formed by hydrophobic packing of sidechains. Thus, shape and size of the substituted non-polar sidechains of V82F and I84V are responsible for the binding with drug molecule. Detailed attention was focused on residues 50', 82 and 84 as well as the inhibitor subsites. For Ile50', the most probable of the χ_1 torsion is in a range 50°-90° for the WT and V82F. This suggested no substantial change of Ile50' sidechain. On the other hand, the χ_1 (Ile50') in the I84V and V82F/I84V mutants shows two distinct values with a difference of ~120°, implying two preferable conformations of Ile50'. The two phenyl rings of Phe82 and Phe82' in the V82F/I84V mutant display similar torsional angle of sidechain while they are significantly different for the V82F mutant. This discrepant sidechain orientation of Phe82 implies difference interactions between the V82F and the V82F/I84V mutants and the inhibitor.

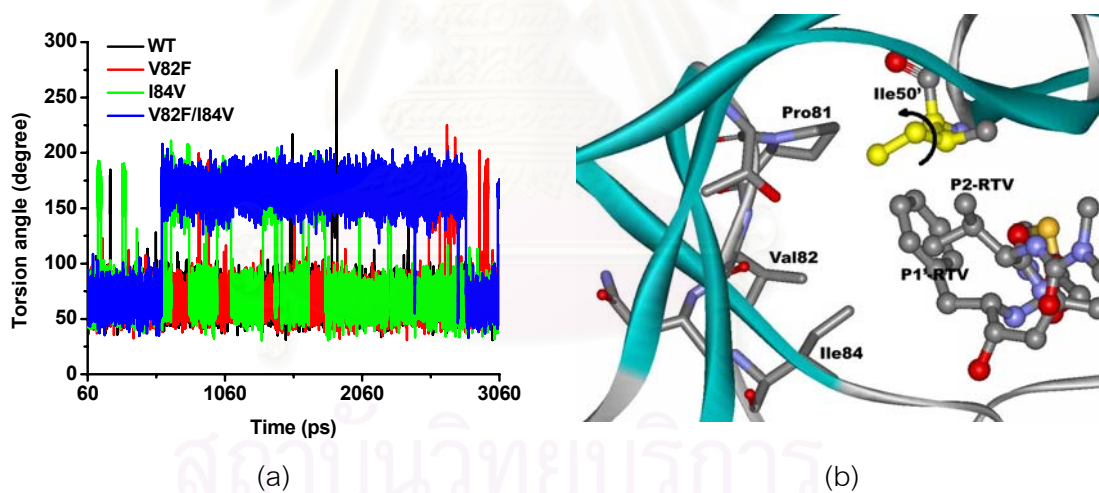


Figure 5.10 (a) The trajectory of Ile50' torsion defined by the C β -C γ bond (arrow in (b)). (b) Ile50', ritonavir and residues 81-84 of the wild-type HIV-1 PR.

5.2.3 Conformational flexibility of ritonavir

From the simulations, a fluctuation of RTV structure compared to the x-ray structure exhibits the RMSD of 1.25 ± 0.24 , 1.43 ± 0.21 , 1.28 ± 0.18 and 1.78 ± 0.13

Å for the WT, V82F, I84V and V82F/I84V, respectively. The RMSD fluctuations of RTV subsite P₁, P₂, P₃, P₁' and P₂' were defined in Figure 5.11 and shown in Figure 5.12.

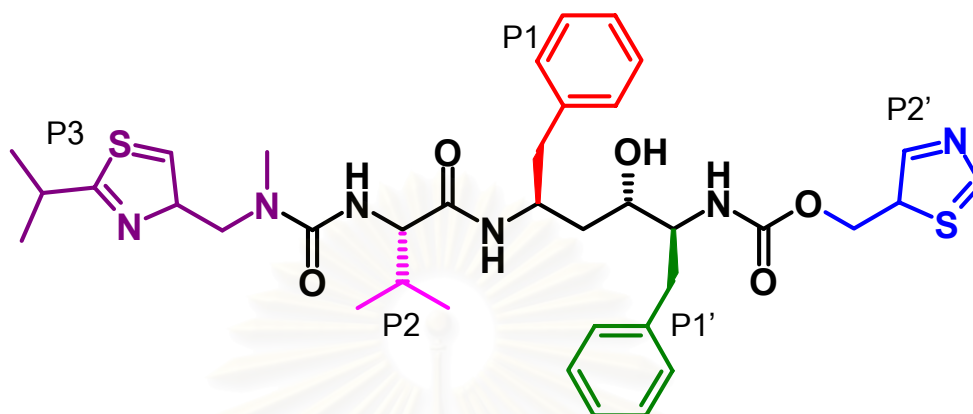


Figure 5.11 Definition of five subsites of ritonavir structure that label by P₁, P₂, P₃, P₁' and P₂'.

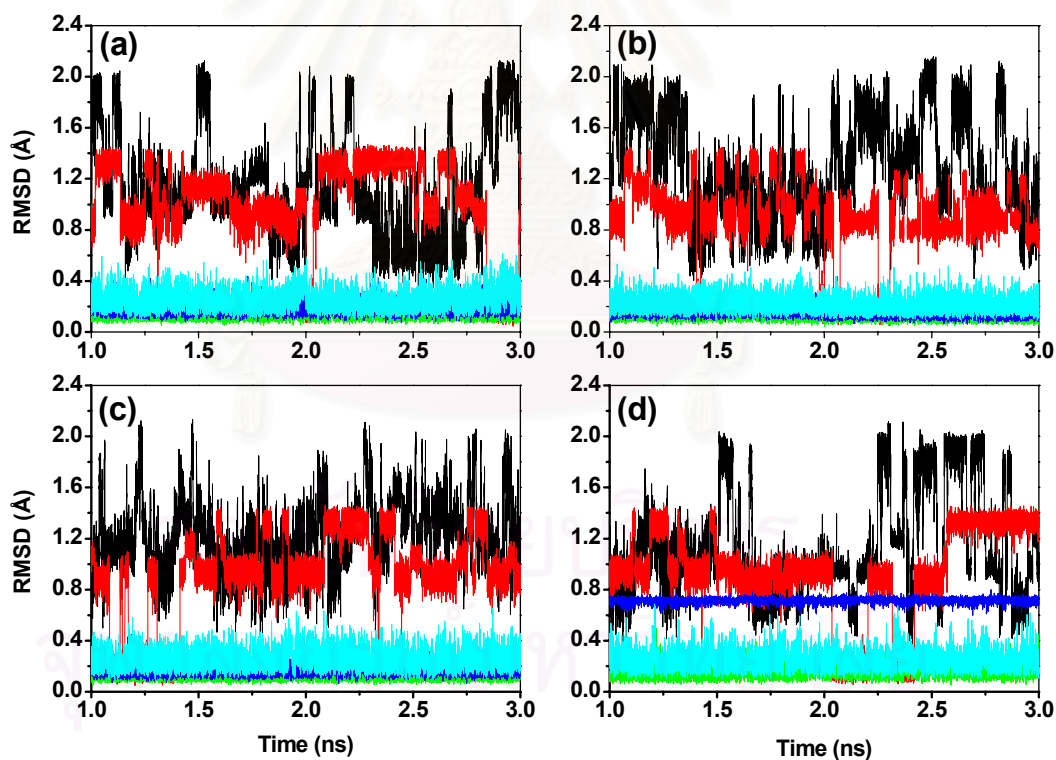


Figure 5.12 RMSDs of each ritonavir subsite with respect to the starting structure versus the simulation time for the wild-type (a), V82F (b), I84V (c) and V82F/I84V (d) mutants. P₃, P₂, P₁, P₁' and P₂' subsites are given in black, red, green, blue and sky-blue lines, respectively.

The RMSD plots show that the phenyl ring of P_1 (green) and P_1' (blue), and the thiazol ring of P_2' (sky-blue) do not change their conformation significantly during the simulations, except the P_1' of the V82F/I84V system. This indicated that the inhibitor sidechains of the bound RTV in the binding pocket of the enzyme are considerably rigid. Meanwhile, P_3 (black) and P_2 (red) sidechains show larger degree of atomic displacements. These are due to the rotation of the two methyl groups of the isopropyl chain. Thus, the flexibility of these subsites is relatively higher upon binding. It appears that the most significant difference between the WT and the mutants was found at the P_1' subsite of RTV. The RMSD change in the double mutant is higher than that in the WT and the two single mutants. This indicates that a greater degree of conformational change of the P_1' phenyl ring takes place on the V82F/I84V complex only.

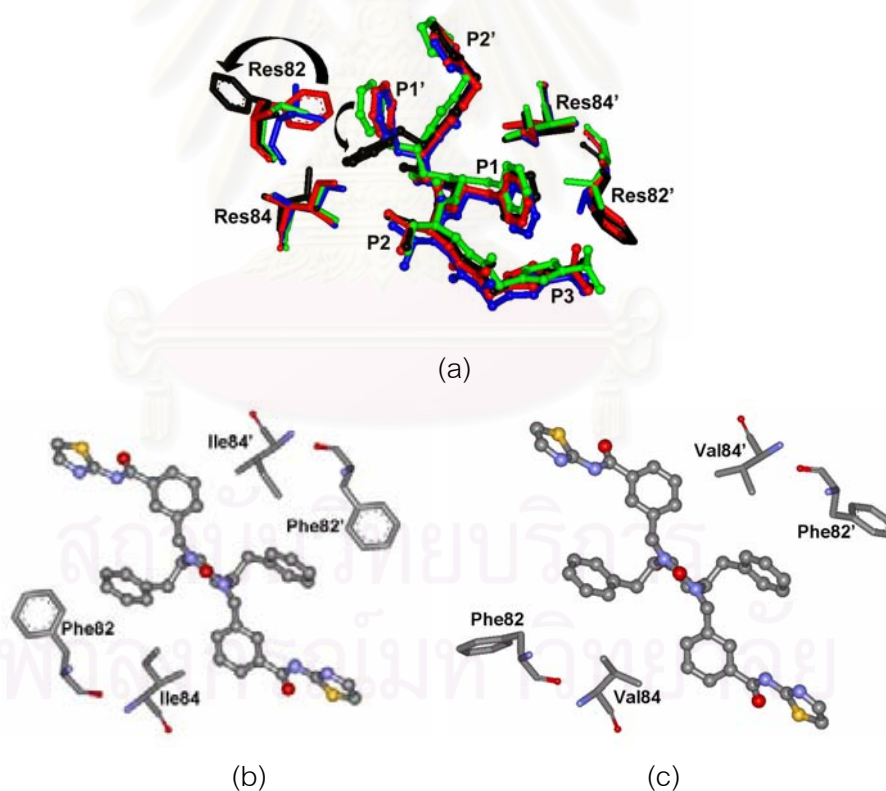


Figure 5.13 Structure comparison of residues 82, 84 and ritonavir in HIV-1 protease: (a) the MD structures of wild-type (blue); V82F (red); I84V (green) and V82F/I84V (black), (b) and (c) the x-ray structures of the V82F (1BV7) and the V82F/I84V mutants (1BWA),

respectively. Arrows show different orientation of the phenyl ring of the P_1' and of Phe82 between the single and double mutants.

Conformational changes of the P_1' phenyl ring in the double mutant can be illustrated in Figure 5.12a. The P_1' ring in the WT and the two single mutants remains largely similar to each other while that in the double mutant is orient differently. Additionally, the ring orientation of Phe82 of the V82F is apparently different from that of the V82F/I84V mutant. Nevertheless, an arrangement of Phe82 sidechain was found to be similar to the crystal structures of the single and double mutations of HIV-1 protease (Figure 5.13b and 5.13c).

It should be noted that the ring rearrangement of the P_1' subsite in the V82F/I84V complex has consequently affected to the binding pocket S_1' of the protein and the conformational energy of the inhibitor. In the V82F/I84V complex, the sidechain rearrangements of Phe82 coupled with the smaller sidechain of Val84 create a larger gap of the S_1' pocket (Figure 5.13). The volumes of the binding cavity calculated using CAST program (107) were shown in Table 5.4. The order of the cavity volumes for all four complexes is V82F/I84V > I84V > Wild-type > V82F. Apparently, the mutation in V82F/I84V has the largest impact on the volume change in the binding cavity of the protein. This may decrease direct contact between RTV to the residues in the binding pocket S_1' . On the contrary, the inhibitor-binding site of the single V82F mutant becomes more restricted as the size of the cavity is shrinking due to the bulkiness of the group. This possibly causes the P_1' ring unchanged in the V82F complex.

A question may arise whether the residue 82' in another subunit of the double mutant and the P_1 ring of the inhibitor undergo similar conformational changes. The fact is that RTV is an asymmetric inhibitor, thus the distinct binding mode in the second monomer is not uncommon. From the simulations, the substituted residues of the protein and the P_1' change of the inhibitor did not significantly alter the conformation of Phe82' and the P_1 ring. In conclusion, the conformation perturbation of residue 82 took place only on one subunit of the HIV-1 protease.

Table 5.4 Calculated volumes in the binding cavity of the protease complexes.

Complexes	Cavity Volume (\AA^3)	ΔV^b (\AA^3)
Wild-type	1257 \pm 80	-
V82F	1241 \pm 84	-16
I84V	1317 \pm 59	60
V82F/I84V	1472 \pm 55	215

^aBecause the cavity volume was obtained from the MD trajectory, the average value and standard deviation are reported. ^bChanges of the cavity volume (V) calculated from $V(\text{wild-type})-V(\text{mutant})$.

The conformational energy of RTV was investigated by calculating the intramolecular potential energy with various conformations of the P_1' subsite. The calculations start with the structure of RTV taken from the MD snapshot and then the C1-C2-C3-C4 torsion defined for the conformation of the P_1' phenyl ring (Figure 5.13) was subjected to change from 0° to 360° . The single-point energy calculations to the whole RTV structure were performed on each generated conformation at HF/6-31G level. A plot of the calculated potential energy versus RTV conformations at the scanned torsional angles (Figure 5.14) indicates three minimum potential energies with the torsion of 70° , 170° and 260° . The results are in good agreement with the P_1' torsions determined from the MD simulations (Figure 5.13). The lowest total energy corresponds to the RTV conformer in the V82F/I84V complex while the minimum at 170° is close to the conformer found in the wild-type and the two single mutants. It should be noted that in the simulations the starting RTV conformer for all systems was $\sim 170^\circ$ corresponding to the second energy minimum. The rearrangement of the P_1' ring in the double mutant can overcome the energy barrier of ~ 3 kcal/mol to reach an energetically more favorable state, although, the lowest energy of the conformation is about 0.5 kcal/mol slightly lower than the second minimum. This conformation change might relax the strain around the P_1' subsite due to the steric conflict of the V82F/I84V mutant.

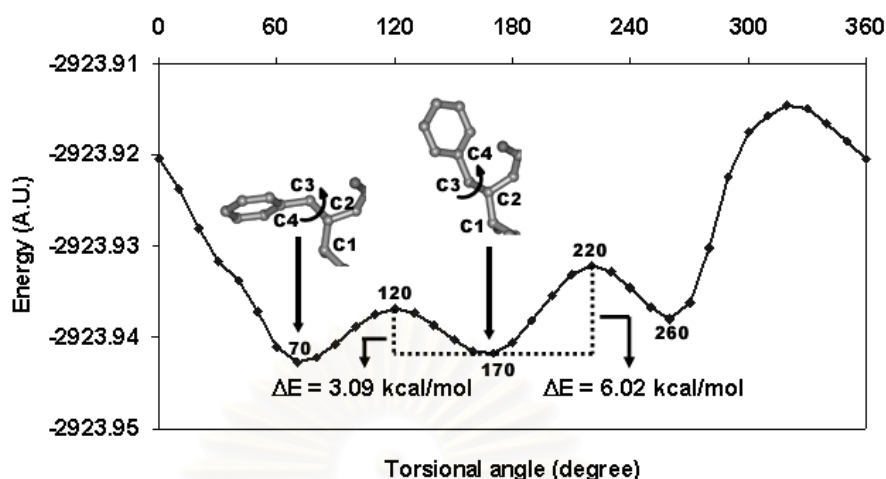


Figure 5.14 The HF/6-31G potential energy versus RTV conformations defined by the C1-C2-C3-C4 torsion of the P₁' sidechain.

5.2.4 Ritonavir-HIV-1 protease binding energy

Table 5.5 shows the estimated binding free energy (ΔG_{bind}) averaged over the trajectory for each complex. Our calculations show a good correlation with experimental results. A net loss of ΔG_{bind} in the mutant complexes with respect to the wild-type corresponds to a decrease in the inhibition potency. The calculated difference in free energy of binding between the WT and each mutant is -3.65 kcal/mol for V82F, -10.81 kcal/mol for I84V and -13.18 kcal/mol for V82F/I84V. This suggested that the binding affinity of RTV is strongest for the native protein and becomes less favorable in the mutants. Particularly, biochemical data indicated that RTV has about 700 times lower inhibition constant (K_i) value for the double mutant. All complexes are correctly rank-ordered, although the magnitudes of the ΔG_{bind} are overestimated.

Detailed analysis suggested that there is a decrease in both van der Waals and electrostatic contributions in the mutants upon binding. It appears for all complexes that van der Waals interactions are the largest contribution to the binding free energy (~70%). This explains an essence of the substituted hydrophobic residues in the primary resistance mutation. In addition, it is in accordance with the fact that the binding pocket of the HIV-1 protease is considerably hydrophobic. However, the magnitude of ΔE_{elec} (Table 5.5) should also be taken into consideration. The differences

in the electrostatic components between the WT and each mutant show an energy loss of -1.10 kcal/mol for V82F, -4.41 kcal/mol for I84V and -8.19 kcal/mol for V82F/I84V, suggesting unfavorable interactions arose from the mutations. In terms of the magnitude, the loss of the non-bonded energy changes correlates well with the volume change of the interior cavity (ΔV in Table 5.4) induced by the mutation.

Table 5.5 Averaged energy contributes (kcal/mol) to the binding free energy for wild-type and mutant HIV-1 protease-ritonavir complexes.

Energy	Wild-type	V82F	I84V	V82F/I84V
ΔE_{vdW}^b	-75.83 ± 3.81^a	-76.28 ± 4.91	-69.81 ± 4.44	-68.89 ± 4.00
ΔE_{elec}^c	-52.59 ± 4.76	-49.35 ± 5.09	-47.93 ± 4.67	-44.07 ± 5.26
$\Delta G_{elec,solv}^d$	23.63 ± 1.67	24.52 ± 1.96	23.40 ± 1.63	20.89 ± 1.05
$\Delta G_{nonpolar,solv}^e$	-6.29 ± 0.18	-6.31 ± 0.28	-5.93 ± 0.36	-5.82 ± 0.41
ΔG_{bind}^f	-111.08 ± 4.89	-107.43 ± 6.11	-100.27 ± 4.21	-97.90 ± 2.40
$K_i (nM)^g$	0.17	0.14	1.90	119

^aStandard error of the mean. ^bVan der Waal energy. ^cElectrostatic energy. ^delectrostatic contribution to solvation. ^eNonpolar contribution to solvation. ^fBinding free energy in the absence of entropic contribution. ^gThe experimental inhibition constant.

5.2.5 Effect of mutation to the binding in the catalytic domain

Residues in the binding domain of HIV-1 PR such as the catalytic triad and the flap region allow several important hydrogen bonds between enzyme and the mainchain RTV to be formed. The hydrogen bonding was determined using the following criteria, (i) the distance between donor and acceptor heavy atoms ≤ 3.5 Å, (ii) the angle of donor-H-acceptor $\geq 120^\circ$. From the histogram in Figure 9, six hydrogen bonds of the WT complex exist over 90% occurrence frequency whereas the single and the double mutants have five and four hydrogen bonds, respectively. Hydrogen bonding between the catalytic residues, Asp25 and Asp25', and the OH of RTV denoted by *d2*, *d3* and *d4* is apparently strong. Particularly, *d2* and *d4* fluctuate in a range of 2.6-2.8 Å with the angle fluctuation of 164-172° for all four MD trajectories. Moreover, in terms of the

observed quantities, the hydrogen bonds *d1* and *d5* have the same magnitude. The difference is depicted by *d6*, which is a hydrogen bond between Asp29 and the P₃ subsite. This hydrogen bond is considerably weak but it appears only in the WT complex.

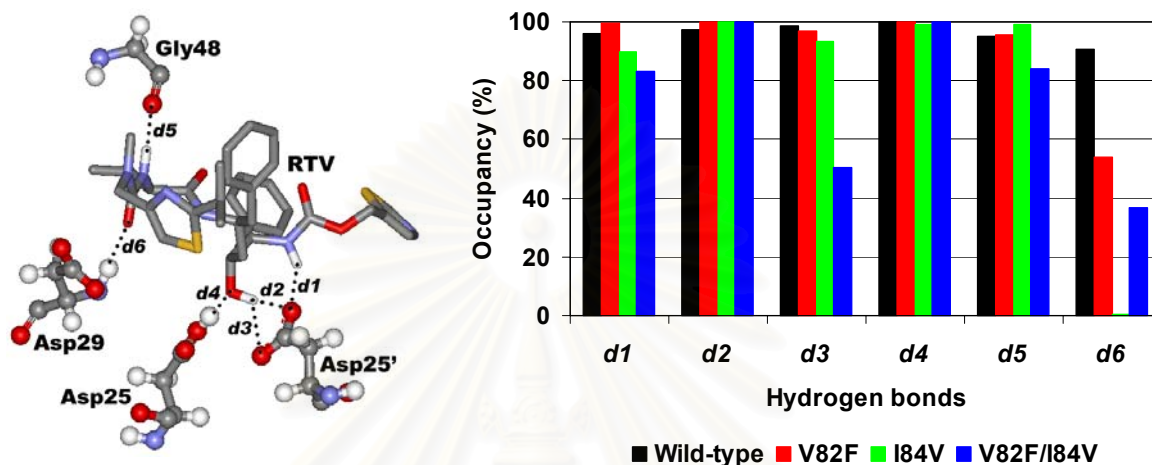


Figure 5.15 Hydrogen bonding (*d1*-*d6*) between HIV-1 PR and ritonavir and the percent occurrence of the corresponding hydrogen bonds calculated from the MD trajectories.

5.3 Conclusion

We have studied structure, dynamics and free energy of binding of wild-type and primary resistance mutants, V82F, I84V and V82F/I84V, HIV-1 protease complexed with ritonavir. Overall structures of the enzyme and the hydrogen bonding of the catalytic residues to ritonavir are similar in all four complexes. Based on the MD data, the tip of the HIV-1 protease flap of the double mutant has the greater degree of opening with respect to that of the wild-type and the two single mutants. The greatest conformational changes in the catalytic binding domain also took place on the double mutant complex. The substituted residues of the resistance mutants require hydrophobic property to maintain the interactions in the binding pocket. The computed free energy of binding using solvent continuum model is in good agreement with the experimental inhibition constant. Those changes induced by the mutations cost free energy penalties for binding of the drug and are responsible for the loss of the drug susceptibility.

CHAPTER 6

CONCLUSION

In this study, we performed MD simulations of three protonation states for G48V mutants HIV-1 PR complexed with SQV to investigate the optimal protonation state of catalytic residues for these complexes. The results lead to a conclusion that the protonated state of the active site of the mutant-type HIV-1 PR complexed with saquinavir is a monoprotinated state on Asp25 (Mono-25). This state was, then, used for the mutation studies.

The MD simulations of wild-type, single and double major mutations HIV-1 PR complexed with saquinavir and ritonavir were preformed. The results indicate that the double mutants of saquinavir and ritonavir show significantly effect on the structural and energy of binding when compared with single mutants. For saquinavir, the major structure changes were observed at the P₂ subsite of SQV and a residue on the flap domain (Phe53) of PR. The changes of P₂ subsite affects the hydrogen bonding between enzyme and inhibitor as well as that with water molecules around P₂ subsite. In case of ritonavir, the P₁' subsite and sidechain of Phe82 of double mutant display the largest conformational changes. In addition, the complexation energy also shows that the complex stability of double mutants are lower than that of wild-type and single mutants.

The above data provides information on drug resistance at molecular level. This information can be used in the structure-based drug designing and drug screening of more potent PR inhibitor that against both of wild-type and mutant PR. To increase reliability of the results, the *ab initio* MD simulation is an interesting and powerful method especially in providing bond breaking and bond formations. The question of protonation state will be clearly understood. However, this method is an extremely time consuming.

REFERENCES

- [1] Available from: <http://www.aidsmap.com/en/docs/898100B5-4DCA-4CC8-B291-08F5471B11F7.asp>.
- [2] Mahalingam, B., Louis, J.M., Hung, J., Harrison, R.W., and Weber, I.T. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. *Proteins*. 43 (2001): 455-464.
- [3] Available from: <http://www.niaid.nih.gov/Daids/dtpdb/FDADRUG.asp>.
- [4] Available from: <http://www.aidsmap.com/en/docs/5D622B02-6D10-4205-9953-AB6C7F6A6BEC.asp>.
- [5] UNAIDS. 2006 Report on the global AIDS epidemic. May 2006.
- [6] UNAIDS/WHO. AIDS epidemic update: Special Report on HIV Prevention. December 2005.
- [7] Available from: <http://www.avert.org/worldstats.htm>.
- [8] Schaal, W. Computational Studies of HIV-1 Protease Inhibitor. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy* 263 (2002): 88 pp.
- [9] Available from: <http://fixedreference.org/2006-Wikipedia-CD-Selection/wp/h/HIV.htm>.
- [10] Reeves, J.D., and Doms, R.W. Human immunodeficiency virus type 2. *J. Gen. Virol.* 83 (2002): 1253-1265.
- [11] Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cimmins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., Sharp, P.M., and Hahn, B.H. Origin of HIV-1 in the chimpanzee *Pan troglodytes* troglodytes. *Nature* 397 (1999): 436-441.
- [12] Piatak, M.J.R., Saag, M.S., Yang, L.C., Clark, S.J., Kappes, J.C., Luk, K.C., Hahn, B.H., Shaw, G.M., and Lifson, J.D. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science* 259 (1993): 1749-1754.

- [13] Kahn, J.O., and Walker, B.D. Acute Human Immunodeficiency Virus Type 1 Infection. N. Engl. J Med. 339 (1998): 33-39.
- [14] Beaulieu, P.L., and Proudfoot, J. HIV and AIDS Therapeutics. 1-6.
- [15] Turner, B.G., and Summers, M.F. Structural Biology of HIV. J. Mol. Biol. 285 (1999): 1-32.
- [16] Available from: <http://www.planetarios.com/Aids-Hiv.htm>.
- [17] Navia, M.A., Fitzgerald, P.M.D., McKeever, B.M., Leu, C.-T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L., and Springer, J.P. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature 337 (1989): 615-620.
- [18] Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J., and Kent, S.B. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science 245 (1989): 616-621.
- [19] Oliveira, T.d., Engelbrecht, S., Rensburg, E.J.v., Gordon, M., Bishop, K., Megede, J.z., Barnett, S.W., and Cassol, S. Variability at Human Immunodeficiency Virus Type 1 Subtype C Protease Cleavage Sites: an Indication of Viral Fitness? J. Virol. 77 (2003): 9422-9430.
- [20] Kawamura, M., Shimano, R., Inubushi, R., Amano, K., Ogasawara, T., Akari, H., and Adachi, A. Cleavage of Gag precursor is required for early replication phase of HIV-1. FEBS letters 415 (1997): 227-230.
- [21] Fiexner, C. HIV-Protease inhibitors. N. Engl. J Med. 338 (1998): 1281-1292.
- [22] Debouck, C., Gorniak, J.G., Strickler, J.E., Meek, T.D., Metcalf, B.W., and Rosenberg, M. Human immunodeficiency virus protease expressed in *Escherichia coli* exhibits autoprocessing and specific maturation of the gag precursor. Proc. Natl. Acad. Sci. USA. 84 (1987): 8903-8906.
- [23] Toh, H., Ono, M., Saigo, K., and Miyata, T. Retroviral protease-like sequence in the yeast transposon Ty1. Nature 315 (1985): 691-692.

- [24] Hosur, M., Bhat, T.N., Kempf, D.J., Baldwin, E.T., Liu, B., Gulnik, S., Wideburg, N.E., Norbeck, D.W., Appelt, K., and Erickson, J.W. Influence of stereochemistry on activity and binding modes for C2 symmetry-based-inhibitors of HIV-1 protease. J. Am. Chem. Soc. 116 (1994): 847-855.
- [25] Ho, D., Toyoshima, T., Mo, H., Kempf, D.J., Norbeck, D., Chen, C.M., Wideburg, N.E., Burt, S.K., Erickson, J.W., and Singh, M.K. Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C2-symmetric protease inhibitor. J. Virol. 68 (1994): 2016-2020.
- [26] Dunn, B.M., Goodenow, M.M., Gustchina, A., and Wlodawer, A. Retroviral proteases. Genome Biology 3 (2002): 3006.3001-3006.3007.
- [27] Schechter, I., and Berger, A. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27 (1967): 157-162.
- [28] Moelling, K., Schulze, T., Knoop, M., Kay, J., Jupp, R., Nicolaou, G., and Pearl, L. In vitro inhibition of HIV-1 proteinase by cerulenin. FEBS. Letters 261 (1990): 373-377.
- [29] Rodriguez, E., Angeles, T., and Meek, T. Use of nitrogen-15 kinetic isotope effects to elucidate details of the chemical mechanism of human immunodeficiency virus 1 protease. Biochemistry 32 (1993): 12380-12385.
- [30] Trylska, J. Computational Modelling of Protonation Equilibria and Reaction Mechanism of HIV-1 Protease. (2001).
- [31] Brik, A., and Wong, C.-H. HIV-1 protease: mechanism and drug discovery. Org. Biomol. Chem. 1 (2003): 5-14.
- [32] Lindgren, M. Exploring Inhibitor of HIV-1 Protease. Interaction Studies with Applications for Drug Discovery. Acta Universitatis Upsaliensis. Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1038 (2004): 47 pp.

- [33] Salto, R., Babe, L., Li, J., Rose, J., Yu, Z., Burlingame, A., Voss, J.D., Sui, Z., Montellano, P.O.D., and Craik, C. In vitro characterization of nonpeptide irreversible inhibitors of HIV proteases. J. Biol. Chem. 269 (1994): 10691-10698.
- [34] Boggetto, N., and Reboud-Ravaux, M. Dimerization inhibitors of HIV-1 protease. Biol. Chem. 383 (2002): 1321-1324.
- [35] Rezacova, P., Brynda, J., Fabry, M., Horejsi, M., Stouracova, R., Lescar, J., Chitarra, V., Riottot, M.M., Sedlacek, J., and Bentley, G.A. Inhibition of HIV Protease by Monoclonal Antibodies. J. Mol. Recognit. 15 (2002): 272-276.
- [36] Wlodawer, A., and Vondrasek, J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. Annu Rev Biophys Biomol Struct. 27 (1998): 249-284.
- [37] Roberts, N., Martin, J., Kinchington, D., Broadhurst, A., Craig, J., Duncan, I., Galpin, S., Handa, B., Kay, J., Krohn, A., Lambert, R., Merrett, J., Mills, J., Parkes, K., Redshaw, S., Ritchie, A., Taylor, D., Thomas, G., and Machin, P. Rational design of peptide-based HIV proteinase inhibitors. Science 248 (1990): 358-361.
- [38] Kim, E., Baker, C., Dwyer, M., Murcko, M., Rao, B., and Tung, R. Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. J. Am. Chem. Soc. 117 (1995): 1181-1182.
- [39] Dorsey, B., Levin, R., McDaniel, S., Vacca, J., Guare, J., Darke, P., Zugay, J., Emini, E., Schleif, W., and Quintero, J. L-735,524: the design of a potent and orally bioavailable HIV protease inhibitor. J. Med. Chem. 37 (1994): 3443-3451.
- [40] Kaldor, S., Kalish, V., Davies, J., Shetty, B., Fritz, J., Appelt, K., Burgess, J., Campanale, K., Chirgadze, N., and Clawson, D. Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. J. Med. Chem. 40 (1997): 3979-3985.

- [41] Kempf, D.J., Sham, H.L., Marsh, K.C., Flentge, C.A., Betebenner, D., Green, B.E., McDonald, E., Vasavanonda, S., Saldivar, A., Wideburg, N.E., Kati, W.M., Lisa Ruiz, Zhao, C., Fino, L., Patterson, J., Molla, A., Plattner, J.J., and Norbeck, D.W. Discovery of Ritonavir, a Potent Inhibitor of HIV Protease with High Oral Bioavailability and Clinical Efficacy. J. Med. Chem. 41 (1998): 602-617.
- [42] Sham, H., Kempf, D., Molla, A., Marsh, K., Kumar, G., and Chen, C. ABT-378, a Highly Potent Inhibitor of the Human Immunodeficiency Virus Protease. Antimicrob. Agents Chemother. 42 (1998): 3218-3224.
- [43] Robinson, B., Riccardi, K., Gong, Y., Guo, Q., Stock, D., Blair, W., Terry, B., Deminie, C., Djang, F., Colonno, R., and Lin, P. BMS-232632, a Highly Potent Human Immunodeficiency Virus Protease Inhibitor That Can Be Used in Combination with Other Available Antiretroviral Agents. Antimicrob. Agents Chemother. 44 (2000): 2093-2099.
- [44] Turner, S., Strohbach, J., and Tommasi, R. Tipranavir (PNU-140690): a potent orally bioavailable nonpeptidic HIV protease inhibitor of the 5,6-dihydro-4-hydroxy-2-pyrone sulfonamide class. J. Med. Chem. 41 (1998): 3467-3476.
- [45] Furfine, E., Baker, C., Hale, M., Reynolds, D., Salisbury, J., Searle, A., Studenberg, S., Todd, D., Tung, R., and Spaltenstein, A. Preclinical pharmacology and pharmacokinetics of GW433908, a water-soluble prodrug of the human immunodeficiency virus protease inhibitor amprenavir. Antimicrob. Agents Chemother. 48 (2004): 791-798.
- [46] Surleraux, D., Tahri, A., Verschueren, W., Pille, G., Kock, H.d., and Jonckers, T. Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor. J. Med. Chem. 48 (2005): 1813-1822.
- [47] Larder, B. Mechanism of HIV-1 drug resistance. AIDS 15 (2001): S27-S34.
- [48] Mansky, L., and Temin, H.M. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J. Virol. 69 (1996): 5087-5094.

- [49] Coffin, J. HIV population in vivo: implications for genetic variation, pathogenesis, and therapy. Science 267 (1995): 483-489.
- [50] Phillips, A., Miller, V., Sabin, C., Lepri, A.C., Klauke, S., Bickel, M., Doerr, H., Hill, A., and Staszewski, S. Durability of HIV-1 viral suppression over 3.3 years with multi-drug antiretroviral therapy in previously drug-naive individuals. AIDS 15 (2001): 2379-2384.
- [51] Boden, D., and Markowitz, M. Resistance to Human Immunodeficiency Virus Type 1 Protease Inhibitors. Antimicrob. Agents Chemother. 42 (1998): 2775-2783.
- [52] Johnson, V.A., Brun-Vézinet, F., Clotet, B., Conway, B., Kuritzkes, D.R., Pillay, D., Schapiro, J., Telenti, A., and Richman, D. Update of the Drug Resistance Mutations in HIV-1: 2005. Special Contribution - Drug Resistance Mutations in HIV-1 13 (2005): 51-57.
- [53] Vella, S., Galluzzo, C., Giannini, G., Pirillo, M.F., Duncan, I., Jacobsen, H., Andreoni, M., Sarmati, L., and Ercoli, L. Saquinavir/zidovudine combination in patients with advanced HIV infection and no prior antiretroviral therapy: CD41 lymphocyte/plasma RNA changes, and emergence of HIV strains with reduced phenotypic sensitivity. Antiviral Res. 29 (1996): 91-93.
- [54] Jacobsen, H., Yasargil, K., Winslow, D., Craig, J., Kroehn, A., Duncan, I., and Mous., J. Characterization of human immunodeficiency virus type 1 mutant with decreased sensitivity to proteinase inhibitor Ro31-8959. Virology 206 (1995): 527-534.
- [55] Jacobsen, H., Hangi, M., Ott, M., Duncan, I.B., Owen, S., Andreoni, M., Vella, S., and Mous., J. In vivo resistance to a human immunodeficiency virus type 1 proteinase inhibitor: mutations, kinetics, frequencies. J. Infect. Dis. 173 (1996): 1379-1387.

- [56] Craig, I., Duncan, I.B., Roberts, N.A., and Whittaker, L. Ro31-8959, an inhibitor of HIV proteinase, appears relatively refractory to the generation of virus mutants with reduced sensitivity. In Abstracts of the 9th International Conference on AIDS, Berlin, Germany. abstr. PO-A26-0694 (1993): 234.
- [57] Collier, A., Coombs, R., Schoenfeld, D.A., Bassett, R.L., Timpone, M.S.J., Baruch, A., Jones, M., Facey, K., Whitacre, C., McAuliffe, V.J., Friedman, H.M., Merigan, T.C., Reichmann, R.C., Hooper, C., and Corey, L. Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. N. Engl. J. Med. 334 (1996): 1011-1017.
- [58] Winters, M., Schapiro, J.M., Lawrence, J., and Merigan, T.C. Genotypic and phenotypic analysis of the protease gene in HIV-1 infected patients that failed long-term saquinavir therapy and switched to other protease inhibitors. In Abstracts of the International Workshop on HIV Drug Resistance, Treatment Strategies and Eradication, St. Petersburg, Fla. abstr. 17 (1997): 11.
- [59] Dulioust, A., Paulous, S., Guillemot, L., Boue, F., Galanaud, P., Clavel, F., and 1997. Selection of saquinavir resistant mutants by indinavir following a switch from saquinavir. In Abstracts of the International Workshop on HIV Drug Resistance, Treatment Strategies and Eradication, St. Petersburg, Fla. abstr. 16 (1997): 11.
- [60] Ermolieff, J., Lin, X., and Tang, J. Kinetic Properties of Saquinavir-Resistant Mutants of Human Immunodeficiency Virus Type 1 Protease and Their Implications in Drug Resistance in Vivo. Biochemistry 36 (1997): 12364-12370.
- [61] Maschera, B., Darby, G., Palu, G., Wright, L.L., Tisdale, M., Myers, R., Blair, E.D., and Furfine, E.S. Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. J. Biol. Chem. 271 (1996): 33231-33235.

- [62] Markowitz, M., Mo, H., Kempf, D.J., Norbeck, D.W., Bhat, T.N., Erickson, J.W., and Ho., D.D. Selection and analysis of human immunodeficiency virus type 1 variants with increased resistance to ABT-538, a novel protease inhibitor. J. Virol. 69 (1995): 701-706.
- [63] Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., P. Deutsch, Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., Saag, M.S., and Shaw., G.S. Viral dynamics in human immunodeficiency virus type 1 infection. Nature 373 (1995): 117-122.
- [64] Schmit, J., Ruiz, L., Clotet, B., Raventos, A., Tor, J., Leonard, J., Desmyter, J., Clerq, E.D., and Vandamme., A.M. Resistance-related mutations in the HIV-1 protease gene of patients treated for 1 year with the protease inhibitor ritonavir (ABT-538). AIDS 10 (1996): 995-999.
- [65] Jorgensen, W.L., and Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. PNAS 102 (2005): 6665-6670.
- [66] Allen, M.P. Introduction to Molecular Dynamics Simulation. NIC Series 23 (2004): 1-28.
- [67] Shewchuk, J.R. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain Edition 5/4. August 4, 1994.
- [68] Alder, B.J., and Wainwright, T.E. Phase Transition for a Hard Sphere System. J. Chem. Phys. 27 (1957): 1208-1209.
- [69] Rahman, A. Correlations in the motion of atoms in liquid argon. Phys. Rev. 136 (1964): 405-411.
- [70] Stillinger, F.H., and Rahman, A. Improved Simulation of Liquid Water by Molecular Dynamics, F. H. Stillinger and A. Rahman, J. Chem. Phys. 60, 1545-1557 (1974). J. Chem. Phys. 60 (1974): 1545-1557.
- [71] McCammon, J., Gelin, B., and Karplus, M. Dynamics of folded proteins. Nature 267 (1977): 585-590.

- [72] Rocchia, W., Alexov, E., and Honig, B. Extending the Applicability of the Nonlinear Poisson-Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions. J. Phys. Chem. B. 105 (2001): 6507-6514.
- [73] Sitkoff, D., Sharp, K.A., and Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. J. Phys. Chem. 98 (1994): 1978-1988.
- [74] Available from: <http://www.netsci.org/Science/Compchem/feature14.html>.
- [75] Hyland, L.J., Tomaszek, T.A., and Meek., T.D. Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. Biochemistry 30 (1991): 8454–8463.
- [76] Ido, E., Han, H., Kezdy, F.J., and Tang., J. Kinetic studies of Human Immunodeficiency Virus type 1 protease and its active-site hydrogen bond mutant A28S. J. Biol. Chem. 266 (1991): 24349–24366.
- [77] Chen, X., and Tropsha, A. Relative binding free energies of peptide inhibitors of HIV-1 protease: the influence of the active site protonation state. J. Med. Chem. 38 (1995): 42-48.
- [78] Smith, R., Brereton, I., Chai, R., and Kent, S. Ionization states of the catalytic residues in HIV-1 protease. Nat. Struct. Biol. 3 (1996): 946-950.
- [79] Won, Y. Binding Free Energy Simulations of the HIV-1 Protease and Hydroxyethylene Isostere Inhibitors. Bull. Korean Chem. Soc. 21 (2000): 1207-1212.
- [80] Piana, S., Sebastiani, D., Carloni, P., and Parrinello, M. Ab Initio Molecular Dynamics-Based Assignment of the Protonation State of Pepstatin A/HIV-1 Protease Cleavage Site. J. Am. Chem. Soc. 123 (2001): 8730-8737.
- [81] Nam, K.Y., Chang, B.H., Han, C.K., Ahn, S.G., and No, K.T. Investigation of the protonated state of HIV-1 protease active site. Bull. Korean Chem. Soc. 24 (2003): 817–823.

- [82] Aruksakunwong, O., Wittayanarakul, K., Sompornpisut, P., Sanghiran, V., Parasuk, V., and Hannongbua, S. Structural and dynamical properties of different protonated states of mutant HIV-1 protease complexed with the saquinavir inhibitor studied by molecular dynamics simulations. JMGM. 25 (2006): 324–332.
- [83] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G., Montgomery, J.A., Stratmann, R.E.J., Burant, J.C., Dapprich, S., Millam, J.M., Daniels, A.D., Kudin, K.N., Strain, M.C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G.A., Ayala, P.Y., Cui, Q., Morokuma, K., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Cioslowski, J., Ortiz, J.V., Baboul, A.G., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, N., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Andres, J.L., Gonzalez, C., Head-Gordon, M., Replogle, E.S., and Pople, J.A. 2002. Gaussian 98. version A.11 ed, Gaussian, Inc. Pittsburgh, PA. Gaussian, Inc. Pittsburgh, PA.
- [84] Antosiewicz, J., Mccammon, J.A., and Gilson., M.K. Prediction of pH-dependent properties of proteins. J. Mol. Biol. 238 (1994): 415–436.
- [85] Davis, M., Madura, J.D., Luty, B.A., and McCammon, J.A. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics Program. Comput. Phys. Comm. 62 (1991): 187-197.
- [86] Sanner, M.F., Duncan, B.S., Carrillo, C.J., and Olson., A.J. Integrating Computation and Visualization for Biomolecular Analysis: An example using Python and AVS. Proc. Pacific Symposium in Biocomputing 99 (1998): 401-412.
- [87] Yang, A., Gunner, M.R., Sampogna, R., Sharp, K., and Honig., B. On the calculation of pK(a)s in proteins. Proteins. 15 (1993): 252-265.

- [88] Brooks, B., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J. Comput. Chem. 4 (1983): 187-217.
- [89] Krohn, A., Redshaw, S., Ritchie, J.C., Graves, B.J., and Hatada, M.H. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. J. Med. Chem. (1991): 3340-3342.
- [90] Case, D.A., Pearlman, J.W., Caldwell, T.E., Cheatham, J., III., Wang, W.S., Ross, C.L., Simmerling, T.A., Darden, K.M., Merz, R.V., Stanton, A.L., Cheng, J.J., Vincent, M., Crowley, V., Tsui, H., Gohlke, R.J., Radmer, Y., Duan, J., Pitera, I., Massova, G.L., Seibel, U.C., Singh, P.K., and Kollman, P.A. 2002. AMBER. version 7.0 ed, University of California, San Francisco, CA. University of California, San Francisco, CA.
- [91] Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79 (1983): 926-935.
- [92] Hong, L., Zhang, X.C., Hartsuck, J.A., and Tang, J. Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance. Protein Sci. 9 (2000): 1898–1904.
- [93] Case, D., Pearlman, J., Caldwell, T., Cheatham, J., Wang, W., Ross, C., Simmerling, T., Darden, K., Merz, R., Stanton, A., Cheng, J., Vincent, M., Crowley, V., Tsui, H., Gohlke, R., Radmer, Y., Duan, J., Pitera, I., Massova, G., Seibel, U., Singh, P., and Kollman, P. 2002. AMBER. version 7.0 ed, University of California, San Francisco, CA. University of California, San Francisco, CA.
- [94] Berendsen, H.J.C., Postma, J.P.M., Gunsteren, W.F.v., DiNola, A., and Haak, J. R. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81 (1984): 3684-3690.

- [95] Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23 (1977): 327-341.
- [96] York, D.M., Darden, T.A., and Pedersen, L.G. The effect of longrange electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. J. Chem. Phys. 99 (1993): 8345-8348.
- [97] Wittayanarakul, K., Aruksakunwong, O., Sompornpisut, P., Sanghiran-Lee, V., Parasuk, V., Pinitglang, S., and Hannongbua, S. Structure, Dynamics and Solvation of HIV-1 Protease/Saquinavir Complex in Aqueous Solution and Their Contributions to Drug Resistance: Molecular Dynamic Simulations. J. Chem. Inf. Model. 45 (2005): 300-308.
- [98] Wittayanarakul, K., Aruksakunwong, O., Saen-oon, S., Chantratita, W., Parasuk, V., Sompornpisut, P., and Hannongbua, S. Insights into Saquinavir Resistance in the G48V HIV-1 Protease: Quantum Calculations and Molecular Dynamic Simulations. Biophys. J. 88 (2005): 867-879.
- [99] Nam, K.-Y., Chang, B.H., Han, C.K., Ahn, S.K., and No, K.T. Investigation of the Protonated State of HIV-1 Protease Active Site. Bull. Korean Chem. Soc. 24 (2003): 817-823.
- [100] Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. A second generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J. Am. Chem. Soc. 117 (1995): 5179-5197.
- [101] Villà, J., Trajbl, M., Glennon, T.M., Sham, Y.Y., Chu, Z.T., and Warshel, A. How important are entropic contributions to enzyme catalysis? PNAS. 97 (2000): 11899-11904.
- [102] Kuhn, B., and Kollman, P.A. Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models. J. Med. Chem. 43 (2000): 3786 - 3791.

- [103] Todd, M.J., Luque, I., Vela'zquez-Campoy, A., and Freire, E. Thermodynamic Basis of Resistance to HIV-1 Protease Inhibition: Calorimetric Analysis of the V82F/I84V Active Site Resistant Mutant. Biochemistry. 39 (2000): 11876-11883.
- [104] Maschera, B., Darby, G., Palu, G., Wright, L.L., Tisdale, M., Myers, R., Blair, E.D., and Furfine, E.S. Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. J. Biol. Chem. 271 (1996): 33231-33235.
- [105] Wang, J., Wang, W., Kollman, P.A., and Case, D.A. Antechamber: an accessory software package for molecular mechanical calculations. J. Am. Chem. Soc. 222 (2001): U403-U403.
- [106] Perryman, A.L., Lin, J.-H., and Mccammon, J.A. 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 Science. 13 (2004): 1108-1123.
- [107] Liang, J., Edelsbrunner, H., and Woodward, C. Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. Protein Science. 7 (1998): 1774-1897.



APPENDICES

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1. Outputs from this thesis

- 1.1 Wittayanarakul, K., **Aruksakunwong, O.**, Saen-oon, S., Chantratita, W., Parasuk, V., Sompornpisut, P., and Hannongbua, S. Insights into saquinavir resistance in the G48V HIV-1 protease: quantum calculations and molecular dynamic simulations. J. Bio. Phys. 88 (2005): 867-879.
- 1.2 **Aruksakunwong, O.**, Wittayanarakul, K., Sompornpisut, P., Lee, V. S., Parasuk, V., and Hannongbua, S. Structure and dynamic properties of different protonated states of mutant HIV-1 protease complexed with the saquinavir Inhibitor studied by molecular dynamic simulations. JMGM. 25 (2006): 324–332.
- 1.3 **Aruksakunwong, O.**, Wolschann, P., Hannongbua, S., and Sompornpisut, P. Molecular dynamic and free energy studies of primary resistance mutations in HIV-1 protease-ritonavir complexes. J. Chem. Inf. Model. (2006) accepted.

2 Related outputs

- 2.1 **Aruksakunwong, O.**, Promsri, S., Wittayanarakul, K., Nimmanpipug, P., Lee, V. S., Wijitkosoom, A., Sompornpisut, P., and Hannongbua S. Currently developed HIV-1 protease inhibitors. Current Computer-Aided Drug Design (2006) accepted.
- 2.2 Saen-oon. S., **Aruksakunwong, O.**, Wittayanarakul, K., Sompornpisut, P. and Hannongbua, S. Insight into analysis of interactions of saquinavir with HIV-1 protease in comparison between wild-type, G48V and G48V/L90M mutant: based on QM and QM/MM calculations. (2006) submitted.

3 Other contributions

- 3.1 Lee, V. S., Wittayanarakul, K., Remsungnen, T., Parasuk, V., Sompornpisut, P., Chantratita, W., Sangma, J., Vannarat, S., Srichaikul, P., Hannongbua, S., Saparpakorn, P., Treesuwan, W., **Aruksakulwong, O.**, Pasomsub, E., Promsri, S., Chuakheaw, D., and Hannongbua, S. Structure and dynamics of SARS coronavirus proteinase: the primary key to the designing and screening for anti-SARS drugs. ScienceAsia 29 (2003): 181-188.
- 3.2 Wittayanarakul, K., **Aruksakunwong, O.**, Sompornpisut, P., Sanghiran Lee, V., Parasuk, V., Pinitglang, S., and Hannongbua, S. Structure, dynamics and solvation of HIV-1 protease/saquinavir complex in aqueous solution and their contributions to drug resistance: molecular dynamic simulations. J. Chem. Inf. Model. 45 (2005): 300-308.
- 3.3 **Aruksakunwong, O.**, Hannongbua, S., and Wolschann, P. Hydrogen bonding in molecular recognition by HIV-1 protease. J. Mol. Structure. 790 (2006): 174-182.
- 3.4 Lee, S. V., Nimmanpipug, P., **Aruksakunwong, O.**, Promsri, S., Sompornpisut, P., and Hannongbua, S. Structure analysis of lead fullerene-based inhibitor bound to human immunodeficiency virus type 1 protease in solutions from molecular dynamics simulations. J. Bio. Phys. (2006) revised.
- 3.5 Lee, V. S., Wittayanarakul, K., Remsungnen, T., Parasuk, V., Sompornpisut, P., Chantratita, W., Sangma, J., Vannarat, S., Srichaikul, P., Hannongbua, S., Saparpakorn, P., Treesuwan, W., **Aruksakulwong, O.**, Pasomsub, E., Promsri, S., Chuakheaw, D., and Hannongbua, S. Dimer of SARS coronavirus main protease as studied by molecular dynamics simulations”, J. Mol. Structure: THEOCHEM. (2006) revised.

- 3.6 Aruksakunwong, O., Malaisree, M., Decha, P., Sompornpisut, P., Parasuk, V., Pianwanit, S., and Hannongbua, S. On the lower susceptibility of oseltamivir to influenza neuraminidase subtype N1 than those in N2 and N9” J. Bio. Phys. (2006) revised.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Insights into Saquinavir Resistance in the G48V HIV-1 Protease: Quantum Calculations and Molecular Dynamic Simulations

Kitiyaporn Wittayanarakul,* Ornjira Aruksakunwong,* Suwipa Saen-oon,* Wasun Chantratita,[†] Vudhichai Parasuk,* Pornthep Sompornpisut,* and Supot Hannongbua*

*Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand; and [†]Virology and Molecular Microbiology, Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok 10400, Thailand

ABSTRACT The spread of acquired immune deficiency syndrome has increasingly become a great concern owing largely to the failure of chemotherapies. The G48V is considered the key signature residue mutation of HIV-1 protease developing with saquinavir therapy. Molecular dynamics simulations of the wild-type and the G48V HIV-1 protease complexed with saquinavir were carried out to explore structure and interactions of the drug resistance. The molecular dynamics results combined with the quantum-based and molecular mechanics Poisson-Boltzmann surface area calculations indicated a monoproteination took place on D25, one of the triad active site residues. The inhibitor binding of the triad residues and its interaction energy in the mutant were similar to those in the wild-type. The overall structure of both complexes is almost identical. However, the steric conflict of the substituted valine results in the conformational change of the P2 subsite and the disruption of hydrogen bonding between the $-NH$ of the P2 subsite and the backbone $-CO$ of the mutated residue. The magnitude of interaction energy changes was comparable to the experimental K_i data. The designing for a new drug should consider a reduction of steric repulsion on P2 to enhance the activity toward this mutant strain.

INTRODUCTION

The spread of acquired immune deficiency syndrome (AIDS) has constantly threatened the world because the disease leads to a significant loss of morbidity and mortality. Unfortunately, chemotherapy for the disease has, in many cases, failed to achieve complete viral suppression (Deeks, 2003). This relies on the fact that the human immunodeficiency virus (HIV) develops resistance to antiretroviral drugs by genetic mutation. The development of novel drug targets and HIV vaccine is promising but the results of those studies remain far from the clinical stage. Understanding the mutations that confer resistance to available drugs is thus an urgent issue in HIV chemotherapy.

The HIV type-1 protease (HIV-1 PR) is an important target for AIDS chemotherapy. This viral protein cleaves the *gag* and *pol* nonfunctional polypeptide into functional proteins essential for maturation of infectious HIV particles (Debouck et al., 1987). The protein is a homodimer. Each protein monomer consists of 99 amino acids (Meek et al., 1989). From x-ray data (Fig. 1 B), the substrate/inhibitor binding site is located at the dimer interface (Hong et al., 1996, 2000; Jaskolski et al., 1991; Krohn et al., 1991; Swain et al., 1990; Vondrasek and Wlodawer, 2002). As a member of the aspartyl protease family, HIV-1 PR is composed of the conserved sequences, so-called the binding triads:

D25-T26-G27 and D25'-T26'-G27', of which D25 and D25' are known as the active site residues. These two ionizable residues play a major role in the catalytic reaction.

Because of the therapeutically important enzyme, structural and functional studies have been carried out to gain understanding of molecular mechanisms of the proteolytic cleavage process (Hyland et al., 1991; Northrop, 2001; Okimoto et al., 1999; Scott and Schiffer, 2000; Smith et al., 1996). The size and the availability of high-resolution x-ray structures of HIV-1 PR are amenable for molecular dynamics (MD) technique to investigate the relationship of structure, dynamics, and function of the enzyme (Collins et al., 1995; Harte et al., 1992, 1990; Levy and Caffisch, 2003; Piana et al., 2002; Scott and Schiffer, 2000) as well as to serve as a test system for developing computational methodology (Piana et al., 2004, 2001; Wang and Kollman, 2000; York et al., 1993a). The MD approach has provided insightful information on the enzyme-substrate interactions and binding conformations, the protonation states of the active site residues, the role of the flexible flap and the binding waters, and drug resistance. Characterizations of structural intermediates have been made useful for rational drug design (Randolph and DeGoey, 2004; Roberts et al., 1990; Rodriguez-Barrios and Gago, 2004).

Saquinavir (SQV, Fig. 1 A), a peptidomimetic protease inhibitor, is clinically used to treat infected HIV patients. The inhibitor containing a nonhydrolyzable hydroxyethylene isostere was designed based on the transition state structure in the enzyme-substrate complex. Combination of PR and reverse transcriptase inhibitors appears to be a highly effective treatment against HIV (Boucher, 1996). The PR inhibitor blocks the maturation step of the HIV life cycle, which is the

Submitted May 18, 2004, and accepted for publication October 26, 2004.

Kitiyaporn Wittayanarakul and Ornjira Aruksakunwong contributed equally to this work.

Address reprint requests to Dr. Pornthep Sompornpisut, Dept. of Chemistry, Faculty of Science, Chulalongkorn University, 254 Phayathai Rd., Pratumwan, Bangkok 10330, Thailand. Tel.: 662-218-7604; Fax: 662-218-7598; Email: pornthep.s@chula.ac.th.

© 2005 by the Biophysical Society

0006-3495/05/02/867/13 \$2.00

doi: 10.1529/biophysj.104.046110

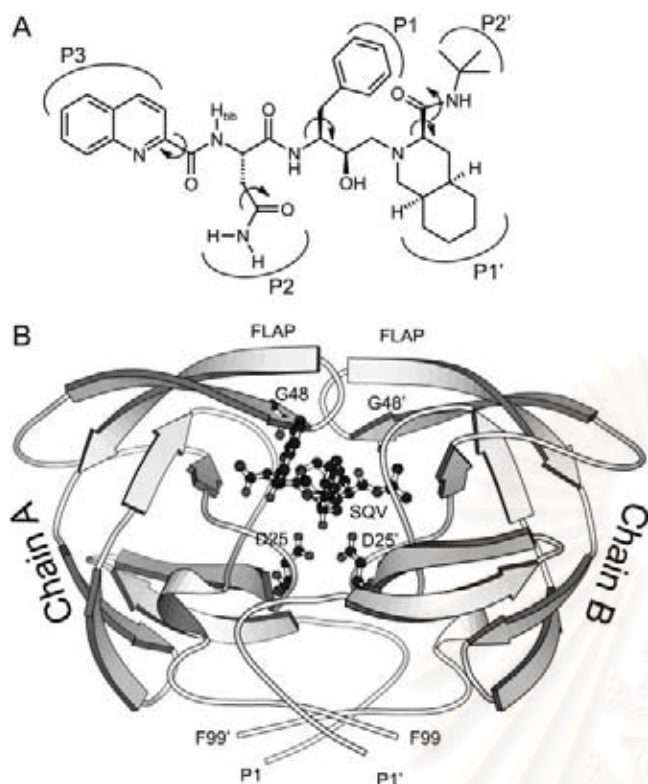


FIGURE 1 Schematic representation of saquinavir (A) and the wild-type HIV-1 protease-saquinavir complex (B). According to a conventional classification of the protease subsites, the binding pockets are designated by the inhibitor side chains P1, P2, P3, P1', and P2'.

crucial stage in the formation of new viral particles. Nevertheless, the current cure with SQV has introduced several resistant variants of HIV-1 PR, some of which can dramatically reduce drug susceptibility (Vondrasek and Wlodawer, 2002). G48V and L90M are considered as the primary mutations commonly occurring *in vivo* or *in vitro* (Eberle et al., 1995; Vaillancourt et al., 1999). These “signature” residue mutations can be associated with a dramatic decrease in drug susceptibility. According to K_i values, the G48V, L90M, and G48V/L90M mutants decrease saquinavir sensitivity by 13.5-, 3-, and 419-fold with respect to that of the wild-type (wt) protease (Ermolieff et al., 1997).

Among common mutations associated with antiretroviral drug resistance, G48V is a unique mutation characteristically generated by SQV. In a view of substituted-type residue, glycine was replaced by a bulkier side-chain residue. The steric conflict of the mutant should involve in a destabilization of the complex. Although several x-ray structures of the HIV-1 PR provided valuable information on the inhibitor binding, this is not the case for the primary resistance to SQV. The crystal structure of the G48V complex is not yet available. With the aid of the available x-ray data, the molecular modeling techniques offer an opportunity to investigate the structural basis of the mutant enzyme (Prabu-Jeyabalan et al., 2003; Swain et al., 1990).

The missing hydrogens in the structural data have led to studies of the ionization state of the active site residues D25/D25' (Smith et al., 1996; Wang et al., 1996; Wlodawer and Vondrasek, 1998; Yamazaki et al., 1994). This subject is important for drug design in a way to optimize the interactions of the inhibitor with the enzyme. Different protonation models were found depending upon the local environment of the enzyme-inhibitor complex. The single protonation at one of the two acidic residues has been most commonly observed with the binding of the hydroxyl-ethylene-based inhibitors (Baldwin et al., 1995; Chen and Tropsha, 1995; Hyland et al., 1991; Smith et al., 1996; Wang et al., 1996; Wlodawer and Vondrasek, 1998). From NMR experiments, the neutral D25/D25' side chain (diprotonation) was determined in the presence of inhibitor diol groups (Yamazaki et al., 1994), whereas the dianionic form (unprotonation) was observed in the free enzyme (Smith et al., 1996; Wang and Kollman, 2000).

In this study, we employed a computational approach to access information regarding molecular structure and dynamics of the G48V HIV-1 protease conferring to saquinavir resistance. The MD simulations were carried out for the wt and the G48V HIV-1 protease complexed with saquinavir in explicit aqueous solution. The study of the protonation state of the HIV-1 PR complexed with SQV has been carried out before exploring the structure and dynamic data of the signature resistance. The density functional theory (DFT), ONIOM, and molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) methods have been performed to identify the protonation model of the active site residues. The MM/PBSA approach offers an efficient computation for calculating the binding free energy of biomolecules (Kollman et al., 2000; Srinivasan et al., 1998; Wang and Kollman, 2000). The method has been extensively used to study protein-ligand complexes. The quantum-based approach, DFT and ONIOM, has been useful in providing accurate energy information of the interested region. In particular, the hybrid quantum mechanical/molecular mechanical (QM/MM) method, ONIOM (our own *N*-layered integrated molecular orbital and molecular mechanics), developed by Morokuma, has been extended from small molecules to biological applications (Friesner and Beachy, 1998; Morokuma, 2002; Prabhakar et al., 2004; Torrent et al., 2002). Its efficiency has been improved over the years. Simply, the concept of the ONIOM approach is partitioning a large molecular system into onion skin-like layers, and applying the quantum mechanics and molecular mechanics methods to the defined different parts (Morokuma, 2002). In the partitioned system, the high-level quantum computations engage the essential part of the central activity, whereas the lower-level energy calculations take into account the contribution of the remaining region. The comparison of MD results of the two systems provides insightful details of how the G48V mutant is associated with saquinavir resistance. The study provided fundamental principles on the molecular mechanism of

inhibitor binding and resistance, which will be useful for designing an anti-HIV inhibitor to combat AIDS.

METHODS

Before starting the MD simulations, the problem of the protonated state of the active site residues was addressed, since such information cannot be directly obtained from the x-ray data. Our approach consists of molecular orbital energy calculations and solvent continuum free-energy calculations. However, structural data of the complexes, particularly the structure of the G48V-SQV, are not available, and the quantum-based computation for the whole enzyme-inhibitor complex is not feasible. A strategy is developed. First, MD simulations of the four protonation states for the wt and for the G48V complexes were performed to obtain the protonation models, which were subsequently subjected to calculate interaction and binding energies. Details of the methodology are described as follows.

Initial structure

The x-ray structure of the wt HIV-1 PR complexed with Ro 31-8959, saquinavir, (Protein Data Bank code 1HXB; 2.3 Å resolution) was used as a starting model. All missing hydrogens of the protein were added using the LEaP module in the AMBER 7 software package (Case et al., 2002). The protonation state of the ionizable residues, the C- and the N-termini, except for D25/25', was assigned based on the predicted pKa values at pH 7. The pKas of ionizable residues were calculated based on the Poisson-Boltzmann free-energy calculations (see the pKa prediction). The results concluded that all Lys, Arg, Glu, Asp, and the terminal groups are charged, whereas His was in the neutral form. Protonation states of D25/25' were explicitly assigned into four different ionizable states, including unprotonation, monoprotation (each site of D25 and D25'), and diprotonation (protonated at both aspartyl residues). For the wt study, the simulated systems were labeled as wt-unpro, wt-mono25, wt-mono25', and wt-dipro, respectively.

The starting structure and force-field parameters for the inhibitor were obtained as follows. Hydrogens were added to the x-ray coordinates of SQV (1HXB) by taking into account the hybridization of the covalent bonds. Geometric optimization was subsequently performed at the Hartree-Fock level with 6-31G** basis functions to adjust the bond-length involving hydrogens. Then, the RESP fitting procedure was employed to calculate partial atomic charges of the inhibitor (Cornell et al., 1993). Force-field parameters of the inhibitor were assigned based on the atom types of the Cornell et al. (1995) force-field model. Gaussian 98 (Frisch et al., 2002) was used to optimize the molecular structure, generate electrostatic potentials, and calculate ab initio energies. Partial charge generation and assignment of the force field were performed using the Antechamber suite (Wang et al., 2001).

The preparation of the initial structure for the simulation of the G48V mutant-SQV complex was similar to that of the wt complex. The comparative model of the mutant was constructed based on 1HXB because the three-dimensional structure of the G48V-SQV complex is not available. It should be noted that the x-ray structure of the double mutant, G48V/L90M-SQV complex (1FB7) could be considered as an alternative template. However, the x-ray coordinates of the second monomer of the double mutant are not available. Thus, 1HXB is considered to be more appropriate as a template. The simulated systems of the mutant complex consist of four protonated states, which were defined similar to those of the wt complex, i.e., unprotonation (mt-unpro), monoprotations (mt-mono25 and mt-mono25'), and diprotonation (mt-dipro).

The next step was to incorporate the solvent and counterions into the models previously prepared. The crystallographic waters were also included in the simulations. Each model was solvated with the TIP3P waters (Jorgensen et al., 1983) and neutralized by the counterions using the LEaP module. The total number of the TIP3P waters in the periodic box for all systems was in a range of 9100–9900 molecules.

Molecular dynamics simulations

Energy minimization and MD simulations were carried out using the SANDER module of AMBER 7 (Case et al., 2002) with the Cornell force field (Cornell et al., 1995). The whole systems were subjected to energy minimization within a range of 200–5000 steepest descent steps to avoid bad contacts. It should be noted that position-restrained minimizations of some particular regions were carried out for systems that clashed during minimization because of incidental overlay of atoms. This procedure was repeated until there was no sign of bad contacts. The resulting protein structure was compared with the before-minimized structure. Root mean-square displacement (RMSD) for nonhydrogen atoms of the compared protein structures showed that there were no RMSDs exceeding 0.1 Å in all systems.

The MD simulation was performed employing the periodic boundary condition with the NPT ensemble. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and pressure of the systems (Berendsen et al., 1984). The SHAKE algorithm (Ryckaert et al., 1977) was employed to constrain all bonds involving hydrogens. The simulation time step of 2 fs was used. All MD simulations were run with a 12 Å residue-based cutoff for nonbonded interactions and the particle-mesh Ewald method was used for an adequate treatment of long-range electrostatic interactions (York et al., 1993a).

The simulation consists of thermalization, equilibration, and production phases. Initially, the temperature of the system was gradually heated from 0 to 298 K during the first 60 ps. Then, the systems were maintained at 298 K until MD reached 400 ps of the simulation. Finally, the production phase started from 400 ps to 1 ns of the simulation. The convergence of energies, temperature, pressure, and global RMSD was used to verify the stability of the systems. The MD trajectory was collected every 0.1 ps. The 600 ps trajectory of the production phase was used to calculate the average structure. All MD simulations were carried for 1 ns. Analysis of all MD trajectories i.e., RMSD, distances, torsion angles, etc. was carried out using the CARNAL and PtraJ modules of AMBER 7. The geometry and stereochemistry of the protein structure were validated using PROCHECK (Laskowski et al., 1996). In summary, a total of eight systems for the MD simulations were carried out.

Graphic visualization and presentation of protein structures were done using RasMol, Swiss-Pdb Viewer (Guex and Peitsch, 1997), WebLab Viewer (Accelrys, San Diego, CA), and MolScript (Kraulis, 1991).

The pKa prediction

An assumption used for assigning the protonation state of the ionizable residues in the simulations was inspected by the prediction of the pKa values. The method estimates the pKa shift by calculating the electrostatic free energy of ionizable residues in the neutral and the charge states in solution (Antosiewicz et al., 1994). The computations were done by solving finite difference Poisson-Boltzmann equations implemented in the University of Houston Brownian Dynamics program (Davis et al., 1991). The protocols describe as follows. Polar hydrogens were added to the x-ray model using Insight II (Accelrys, San Diego, CA). For generating electrostatic potentials, the model was then placed in a 65 × 65 × 65 dimension with a grid spacing of 2.5 Å. The focusing technique was additionally employed using finer grid spacing of 1.2, 0.75, and 0.25 Å for a cubic dimension of 15, 15, and 20, respectively (Antosiewicz et al., 1994; Yang et al., 1993). Atomic radii and charges available in the University of Houston Brownian Dynamics program were originally derived from optimized potentials for liquid simulations and CHARMm 22 parameter sets (Brooks et al., 1983; Jorgensen and Tirado-Rives, 1988). The 1.4 Å probe radius with a resolution of 500 dots/atom-sphere was used. The calculations employed a solvent dielectric of 80 with 150 mM ionic strength, and a temperature of 298 K. A dielectric constant of HIV-1 PR was examined by varying to 1, 4, and 20. We found that a protein dielectric constant of 20 produced the best pKa prediction. A dielectric constant of 1 and 4 yielded unusual pKa values due to an overestimation of electrostatic potentials. This phenomenon is thoroughly discussed in an early work (Antosiewicz et al., 1994).

Protonation state of the HIV-1 PR

In an evaluation of the protonation state of D25/D25', we employed three different approaches: density functional theory, ONIOM, and MM/PBSA methods. The DFT and ONIOM methods provide the interaction energy of the complex, whereas the MM/PBSA calculates the binding free energy ($\Delta G_{\text{binding}}$). It should be noted that properties of the system should be analyzed using the structure ensemble from the MD trajectory. However, quantum chemical methods are too expensive to calculate such enormous structural data. Alternatively, the statistically averaged structure obtained from the 600 ps production phase of the MD trajectory was chosen as the studied model.

The DFT method

In each protonation model, the cluster consists of the two triad residues, D25-T26-G27 and D25'-T26'-G27', and SQV (see Fig. 3A). Atoms that are not within the selected part were removed. To reduce possible terminal-charge effect, both ends of the triad fragments were capped by $\text{CH}_3\text{NH}-$ and $-\text{COCH}_3$ groups. Geometric optimization was performed for the added atoms. The energy for the model was computed using a single-point calculation method with mixed basis sets. B3LYP/6-31 + G** was defined explicitly on the carboxylate oxygens of D25 and D25', and B3LYP/6-31G** was assigned on all the remaining atoms. The quantum calculations were carried out using the program Gaussian 98 (Frisch et al., 2002).

A general form of the interaction energy calculation of the enzyme-inhibitor complex (ΔE_{EI}) is the subtraction of the energy of the complex (E_{EI}) from that of the free enzyme (E_{E}) and that of the free inhibitor (E_{I}).

$$\Delta E_{\text{EI}} = E_{\text{EI}} - E_{\text{E}} - E_{\text{I}}. \quad (1)$$

In this case, the model system contains a cluster of the two triads and SQV. Thus, interaction energy of the complex ($\Delta E_{\text{cluster}}$) was estimated by

$$\Delta E_{\text{cluster}} = E_{\text{cluster}} - E_{\text{triadA}} - E_{\text{triadB}} - E_{\text{SQV}}, \quad (2)$$

where E_{cluster} is the total energy of the cluster, and E_{triadA} , E_{triadB} , and E_{SQV} are the total energy of the isolated triads of chain A and B and the unbound SQV, respectively.

The ONIOM method

To account for an effect of the protein environment, the QM/MM ONIOM method was used. Here, the three-layers approach (ONIOM3) was performed to reduce the boundary effect at the QM/MM junction but maintain considerably accurate energy information. The method is described as follows. The model of the HIV-1 PR-SQV complex was divided into three regions: A, B, and C (see the Appendix). The ONIOM layers were represented by inner (A), intermediate (A + B), and real (A + B + C). The inner layer, the "hot spot" region, consisting of D25, D25', and SQV, was treated at the high-level of quantum chemical calculations using density functional theory (B3LYP/6-31G**). The intermediate layer contains a total of 36 residues including D25-D30, I47-F53, P80-I84, and L90. These residues are located within a 5 Å distance from SQV. This intermediate layer was treated with the semiempirical method using PM3. Lastly, the real layer includes the entire enzyme. The molecular mechanic method using universal force field (UFF) was applied to this layer. All calculations based on the ONIOM approach were carried out using the program Gaussian 98 (Frisch et al., 2002).

Hence, the total energy obtained from the ONIOM3 calculations (E^{ONIOM3}) herein can be defined by (see the Appendix):

$$E^{\text{ONIOM3}}[ABC] = E_{[\text{UFF,ABC}]} + E_{[\text{PM3,AB}]} + E_{[\text{B3LYP,A}]} - E_{[\text{UFF,AB}]} - E_{[\text{PM3,A}]}. \quad (3)$$

For the overlap regions, the subtraction energy terms are introduced to substitute the low-level energy calculations with the high-level one.

Therefore, the total interaction energy ($\Delta E_{\text{total}}^{\text{ONIOM3}}[ABC]$) between SQV and the enzyme using the ONIOM3 method can be expressed as independent energy components from each layer as follows:

$$\Delta E_{\text{total}}^{\text{ONIOM3}}[ABC] = \Delta E_{[\text{UFF,ABC}]} + \Delta E_{[\text{PM3,AB}]} + \Delta E_{[\text{B3LYP,A}]} - \Delta E_{[\text{UFF,AB}]} - \Delta E_{[\text{PM3,A}]} \quad (4)$$

$$\Delta E_{\text{total}}^{\text{ONIOM3}}[ABC] = \Delta E_{[\text{B3LYP,A}]} + \Delta \Delta E_{[\text{PM3,AB-A}]} + \Delta \Delta E_{[\text{UFF,ABC-AB}]} \quad (5)$$

$$\Delta \Delta E_{[\text{PM3,AB-A}]} = \Delta E_{[\text{PM3,AB}]} - \Delta E_{[\text{PM3,A}]} \quad (6)$$

$$\Delta \Delta E_{[\text{UFF,ABC-AB}]} = \Delta E_{[\text{UFF,ABC}]} - \Delta E_{[\text{UFF,AB}]}, \quad (7)$$

where $\Delta E_{[\text{B3LYP,A}]}$ is the interaction energy in the region A evaluated at the B3LYP/6-31G(d,p) level. $\Delta \Delta E_{[\text{PM3,AB-A}]}$ is the interaction energy contributed from the region B evaluated at the PM3 level, and $\Delta \Delta E_{[\text{UFF,ABC-AB}]}$ is the interaction energy contributed from the region C evaluated at the UFF molecular mechanics.

The MM/PBSA method

In general, the free energy of the inhibitor binding, $\Delta G_{\text{binding}}$, is obtained from the difference between the free energy of the receptor-ligand complex (G_{cpx}), and the unbound receptor (G_{rec}) and ligand (G_{lig}) as follows:

$$\Delta G_{\text{binding}} = \Delta G_{\text{cpx}} = G_{\text{cpx}} - (G_{\text{rec}} + G_{\text{lig}}). \quad (8)$$

The MM/PBSA approach calculates $\Delta G_{\text{binding}}$ on the basis of a thermodynamic cycle. Therefore, Eq. 8 can be approximated as

$$\Delta G_{\text{binding}} = \Delta E^{\text{MM}} - T\Delta S + \Delta G_{\text{sol}}, \quad (9)$$

where ΔE^{MM} is related to the enthalpic changes in the gas phase upon binding and obtained from molecular mechanics van der Waals and electrostatic energies, $T\Delta S$ involves the entropy effect, and ΔG_{sol} is the free energy of solvation. The ΔG_{sol} is composed of the electrostatic and nonpolar contributions (Srinivasan et al., 1998), and therefore can be expressed as

$$\Delta G_{\text{sol}} = \Delta G^{\text{PB}} + \Delta G^{\text{SA}}, \quad (10)$$

where ΔG^{PB} is calculated using a continuum solvent model with Poisson-Boltzmann solution (Gilson et al., 1987), and ΔG^{SA} is estimated from the solvent-accessible surface area (SASA) (Sitkoff et al., 1994).

The $\Delta G_{\text{binding}}$ was obtained using the MM/PBSA module in the program AMBER 7, which interfaces the program DelPhi 4 (Rocchia et al., 2001). To calculate electrostatic free energy of solvation, the grid resolution of 0.5 Å with the boundary conditions of Debye-Huckel potentials was employed. Atomic charges were taken from the Cornell force field (Cornell et al., 1995). The water and protein dielectric was set to 80 and 4, respectively. The SASA was calculated using a 1.4 Å probe radius. Atomic radii were taken from the PARSE parameter set (Sitkoff et al., 1994). The nonpolar free energy of solvation was obtained by $0.00542 \times \text{SASA} + 0.92 \text{ kcal mol}^{-1}$ (Sitkoff et al., 1994).

In the study, the contribution of the entropy ($T\Delta S$) was not included. An estimation of the entropy effect from normal mode analysis requires the high computation demands. The effect should be very small because all system models are very similar. In addition, we considered the relative values of the binding free energy.

RESULTS

Hydrogen bonding in the binding site

One of the most important HIV-1 PR-SQV interactions is the formation of the hydrogen bond at the active site. This

observation is estimated from the x-ray data by the close proximity between the terminal side chain of the two aspartyl residues and the hydroxyethylene isostere moiety ($-\text{OH}$) of the inhibitors. Due to the lack of hydrogen position in the structure, the pattern of this typical hydrogen bond is investigated here. Therefore, it is necessary to monitor the active site structure of the MD results and determine the most preferential interactions.

The MD snapshot of the dipro, mono25, mono25', and unpro systems for the wt and the G48V complexes is illustrated in Fig. 2, A–H. The D25/D25' side chains and the $-\text{OH}$ of SQV of all protonated states, except for the unprotonation, occupies the positions suitable for the formation of the hydrogen bond. The distal separation between the active site residues and the hydroxyethylene isostere of SQV maintains similarity to the x-ray structure. The structure of the wt-unpro and the mt-unpro provide the worst scenario of the complex (Fig. 2, D and H). A majority of MD data shows that the active site residues adopted to a conformation completely different from the x-ray data. Hence, the wt-unpro and mt-unpro systems are not an appropriate state for the formation of the HIV-1 PR-SQV complex.

In comparing the four protonated states, the binding pattern of the $-\text{OH}$ of SQV to the D25/D25' was remarkably different. Nevertheless, when comparing the wt with the mutant structure of the same protonated state, the binding pattern at the active site was similar. The results allow drawing a simple scheme of the hydrogen bond pattern at the active site, which is illustrated in Fig. 2, I–L.

Protonation state of the wild-type complex

From the pKa prediction, the D25/D25' in the free wt HIV-1 PR exhibited different protonation states depending on the examined pHs of 4, 5, 6, and 7 because the assay of the HIV-1 PR is measured between pH 5 and 6, and in the basic solution the enzyme precipitates (Wang et al., 1996). However, the dianionic form was the most apparent state between pH 6 and 7. The protonation state at pH 4 and 5 remains inconclusive. In this study, these results were not fully understood. It is possibly associated with the used parameters (atomic charges and radii, dielectric constant, etc.) and the studied model. This awaits further investigation.

Table 1 shows the resulting energy calculated from DFT, ONIOM, and MM/PBSA approaches. Here, the term “the stabilization energy” is used for a simpler and more straightforward definition. The negative sign of the stabilization energy suggests that the protonation model of wt-dipro, wt-mono25, and wt-mono25' systems are energetically favorable. From the full quantum DFT treatment, the $\Delta E_{\text{cluster}}$ wt-mono25 compared to that of wt-dipro and wt-mono25' was lower by 7.62 and 17.22 kcal mol⁻¹, respectively. The lowest ΔE_{cpx} from the ONIOM3 method also took place on the protonation model of the wt-mono25. The energy difference was 14.13 and 16.63 kcal mol⁻¹ more stable than that of the wt-dipro and of the wt-mono25', respectively. From the MM/PBSA method, the stabilization energy of the wt-mono25 system was slightly lower in energy than the other two states. Apparently, all the three approaches showed that the monoprotonation at D25 was the most energetically favorable state.

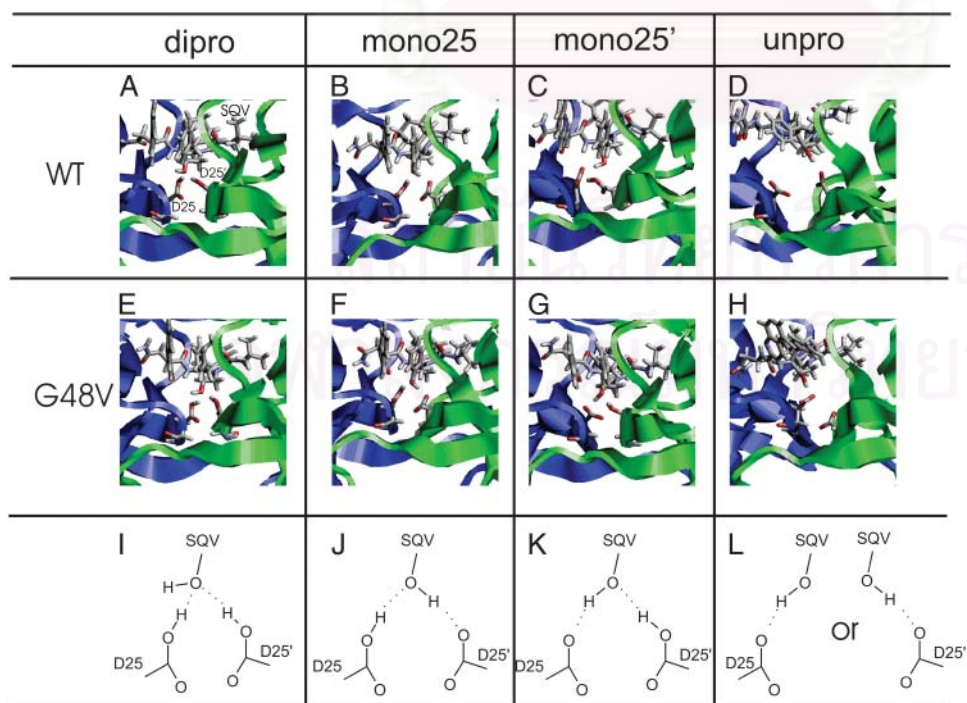


FIGURE 2 MD snapshots showing interactions in the binding pocket of the wt and G48V for various protonation states: diprotonation (dipro), monoprotonation of D25 (mono25), monoprotonation of D25' (mono25'), and unprotonation (unpro). Saquinavir (SQV) and the active site residues D25 and D25' of the enzyme are shown in stick mode. Chain A and B of the enzyme are colored as blue and green.

TABLE 1 The stabilization energy (kcal mol⁻¹)

System	DFT	ONIOM3	MM/PBSA
	$\Delta E_{\text{cluster}}$	ΔE_{cpx}	ΔG_{cpx}
wt-dipro	-59.32	-64.22	-91.62
wt-mono25	-66.94	-78.35	-94.44
wt-mono25'	-49.72	-61.72	-92.91

The stabilization energy shown in Table 1 was decomposed into individual energy components (Table 2) for analyzing quantitatively essential interactions of the different protonation models. From the DFT method, an order of the $\Delta E_{\text{triadA}+\text{SQV}}$ ranging from the lowest value was wt-dipro < wt-mono25 < wt-mono25'. The sum of $\Delta E_{\text{triadA}+\text{SQV}}$ and $\Delta E_{\text{triadB}+\text{SQV}}$ was used to account for interactions of SQV to the triad residues. Still the wt-dipro was the first preferential protonation model with the lowest energy of -34.19 kcal mol⁻¹. Considering the sum of $\Delta E_{\text{triadA}+\text{SQV}}$ and $\Delta E_{\text{triadB}+\text{SQV}}$ of wt-mono25, the energy difference was 7.51 kcal mol⁻¹ greater than that of wt-dipro. On the other hand, its $\Delta E_{\text{triadA,B}}$ value accounting for the triad-triad interactions has significantly gained by 13.16 kcal mol⁻¹ over the wt-dipro system. These data suggested the interactions between the triads were essential to the complex stability.

The ONIOM3 results showed the lowest $\Delta E_{[\text{B3LYP,A}]}$ took place on wt-mono25. It should be noted that $\Delta E_{[\text{B3LYP,A}]}$ has considered the interactions of SQV to D25 and to D25', and between D25/D25'. This is still the case when accounting for interactions of the 5 Å neighboring residues to SQV ($\Delta E_{[\text{PM3,AB-A}]}$). The $\Delta E_{[\text{UFF,ABC-AB}]}$ of the three states are almost equivalent. There was no significant change of the stabilization energy with an addition of $\Delta E_{[\text{UFF,ABC-AB}]}$. Thus, the effect of the protein environment at the long distance-range interactions is negligible.

In the MM/PBSA method, there was no substantial difference of ΔE^{MM} for all three systems, whereas the ΔG_{sol} of the wt-mono25 revealed the lowest values. This observation suggests the solvation free-energy values dominate the contribution of the stabilization energy. The molecular mechanics energy term cannot discriminate the protonation model.

Interaction energy at the catalytic site: the wild-type versus G48V

We have illustrated previously that the quantum-based approach gave promising results to the characterization of the complex model. Here, the approach has been further exploited by calculating the enzyme-inhibitor interaction energy of the G48V complex. The interaction energy of the triads/SQV in the wt complex, $\Delta E_{\text{cluster}}$, previously obtained using the DFT method (Table 1) was used for a comparison.

The results illustrated in Fig. 3 suggested that in the G48V-SQV complex, the lowest $\Delta E_{\text{cluster}}$ also took place at the monoprotonated D25 model. The interaction energy for both the wt and the G48V in the mono25 was ~8 and 17 kcal·mol⁻¹ lower than those from the dipro and the mono25', respectively. At the same protonation state, the $\Delta E_{\text{cluster}}$ of the G48V was no essentially different from that of the wt. This conclusion followed from the fact that the thermal fluctuation at room temperature of ~0.6 kcal·mol⁻¹, equivalent to 1 *kT*, where *T* is 300 K and *k* is the Boltzmann constant. Thus the interactions of the two triad residues and SQV remain unchanged by the G48V mutation.

Those results discussed previously lead us to conclude that the monoprotonation at D25 is the most energetically favorable state. Therefore, further comparison and discussion for both the wt and the mutant complex were focused only on the results of monoprotonation D25.

The energies and RMSD plots (Fig. 4) demonstrated a well-behaved MD simulation for both the wt and G48V-SQV systems. After 400 ps, the RMSD fluctuates 1.27–1.78 Å for the wt and 1.29–1.80 Å for the mutant. The fluctuation is <0.5 Å over the entire production phase. This structural fluctuation is not uncommon in the typical MD simulation of protein, indicating the reliable equilibration of the system in this study. In Table 3, the low RMSD values calculated from 100 snapshot structures taken from the production phase suggested the structures in each set were similar to each other. This allows useful information to be extracted from the MD trajectories. Analysis of some statistic quantities such as structure and dynamics was performed from the trajectories of wt-mono25 and mt-mono25 systems.

TABLE 2 Decomposition of the stabilization energy (kcal mol⁻¹)

System	DFT			ONIOM3			MM/PBSA	
	$\Delta E_{\text{triadA}+\text{SQV}}$	$\Delta E_{\text{triadB}+\text{SQV}}$	$\Delta E_{\text{triadA,B}}$	$\Delta E_{[\text{B3LYP,A}]}$	$\Delta E_{[\text{PM3,AB-A}]}$	$\Delta E_{[\text{UFF,ABC-AB}]}$	ΔE^{MM}	ΔG_{sol}
wt-dipro	-23.81	-10.38	-27.64	-11.45	-36.66	-16.11	-109.42	17.81
wt-mono25	-12.27	-14.41	-40.80	-33.63	-28.34	-16.38	-109.83	15.39
wt-mono25'	-6.33	-6.88	-41.53	-18.33	-27.48	-15.91	-109.88	16.98

$\Delta E_{\text{triadA}+\text{SQV}}$ or $\Delta E_{\text{triadB}+\text{SQV}}$ account for interactions of the triad residues of chain A or B to SQV, and $\Delta E_{\text{triadA,B}}$ accounts for interactions between the triad residues of chain A and those of chain B.

$\Delta E_{[\text{B3LYP,A}]}$ accounts for interactions of D25, D25', and SQV.

$\Delta E_{[\text{PM3, AB-A}]}$ represents interactions of SQV and the 5 Å surrounding residues, excluding D25 and D25', and $\Delta E_{[\text{UFF,ABC-AB}]}$ represents interactions of SQV and the remaining residues.

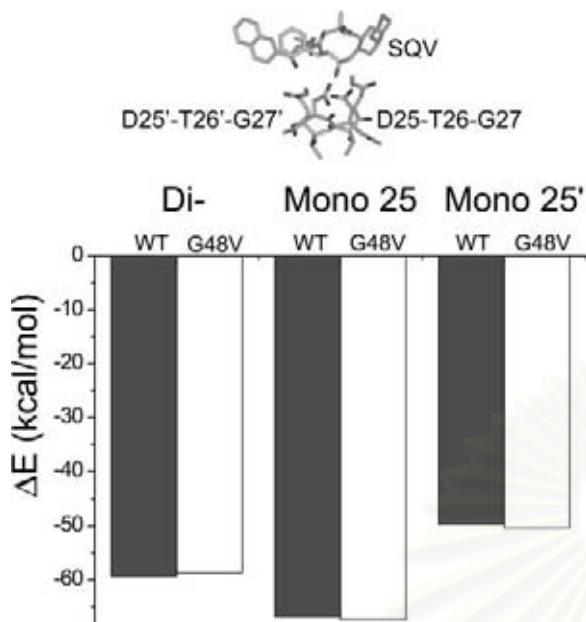


FIGURE 3 ΔE_{cpx} between the triad residues and saquinavir of the wt and the G48V complexes for the dipro, mono25, and mono25' systems.

Structural similarity between the wild-type and the mutant

The global backbone RMSD of the x-ray versus the average MD structure of the wt of 1.04 Å indicates that overall main-chain structures of the complex in the crystal state and in solution are similar. In addition, comparison of the two average MD structures between the wt and the G48V mutant resulted in a backbone RMSD of 0.70 Å. This suggests that the tertiary structure of the G48V mutant was insignificantly different from the native enzyme.

Detailed analysis of RMSD per residue is illustrated in Fig. 5. One can see that most regions of the enzyme, except for the flexible loop of the first subunit, exhibited a small difference in backbone conformation of the wt enzyme with respect to the mutant. Particularly, RMSDs of E21–D30 residues, covering the triad sequence of the enzyme, were in a range from 0.11 to 0.36 Å. This indicates no essential structure alteration of the main-chain hydroxyethylene isostere of SQV and the surrounding residues. Thus, the contacts between the hydroxyl of SQV and the active site residues were independent of the amino acid substitution at residue 48. This result supported the evidence discussed previously that interaction energies of the two triad residues and saquinavir were not significantly changed by mutation of G48V.

Structural difference between the wild-type and the mutant

In Fig. 5, the apparent difference between the wt and the mutant HIV-1 PR-SQV complexes was observed around the

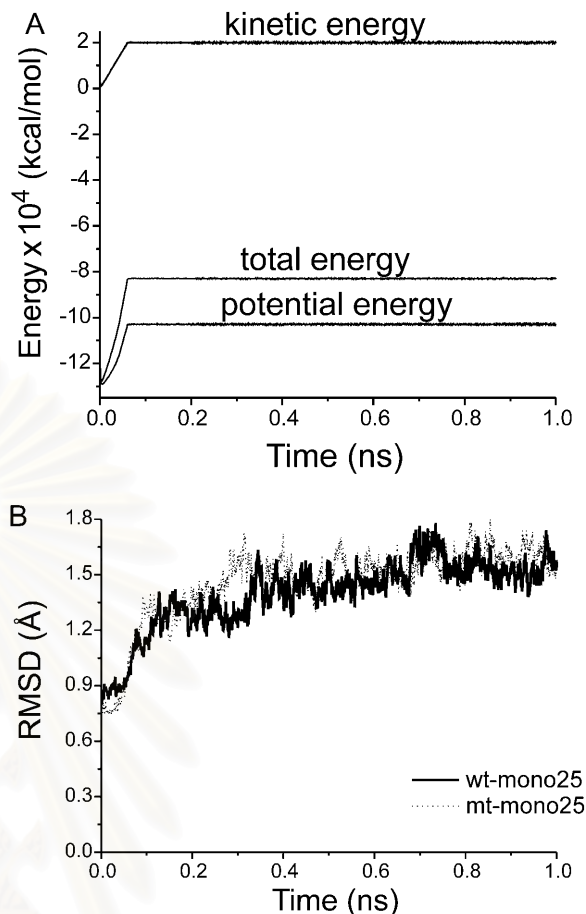


FIGURE 4 Plots of the energies (A) and RMSDs (B) versus the simulation time for the wt and G48V-SQV complex. The obtained RMSD was computed using the structure at $t = 0$ as a reference.

protein flap (residues 46–55). In this region, overall RMSD values of chain A were relatively greater than those of chain B (Fig. 6). The structure of the flap B of the G48V is similar to that of the wt enzyme, whereas the β -hairpin structure of the flap A of the mutant does not superimpose well. Surprisingly, the tip of flap A of the G48V-SQV complex shifted slightly toward the twofold axis of symmetry of HIV-1 PR. However, distance separations between the tips, residues 49–51 in the mutant, were not significantly different from that in the wt complex.

Flexibility and conformational changes of SQV subsites

Flexibility and conformational changes of the inhibitor side chains P1, P1', P2, P2', and P3 were investigated in terms of torsion angle fluctuations of χ_{P1} , $\chi_{P1'}$, χ_{P2} , $\chi_{P2'}$, and χ_{P3} , respectively. From Fig. 7 A, an oscillation of all dihedral angles throughout the simulations, which was no greater than $\pm 10^\circ$, suggested that in the wt complex, all SQV side chains undergo a narrow range of dynamic fluctuation. In other

TABLE 3 Mean global RMSD values calculated from a set of 100 snapshot structures of the 600 ps production phase

		<RMSD>(Å)			
		Dipro	Mono25	Mono25'	Unpro
wt	Backbone	0.90 ± 0.13	0.86 ± 0.12	0.87 ± 0.11	0.83 ± 0.11
	Heavy atoms	1.56 ± 0.22	1.45 ± 0.17	1.48 ± 0.18	1.39 ± 0.15
G48V	Backbone	0.83 ± 0.11	0.86 ± 0.13	0.85 ± 0.12	0.86 ± 0.12
	Heavy atoms	1.39 ± 0.18	1.52 ± 0.22	1.51 ± 0.21	1.47 ± 0.19

words, the inhibitor side chains were inflexible and retained their starting conformation during the course of the MD trajectory. In the case of the G48V mutant, all torsion angles except for χ_{P2} adopted the values similar to that of the wt (Fig. 7 B). This indicated that the conformations of these subsites are unchanged. However, a remarkable shift in χ_{P2} implied a rotation of the P2 subsite starting from $\sim -70^\circ$ to its equilibrium value $\sim 90^\circ$. This suggested a substantial rotation of the P2 of SQV in the G48V (Fig. 8). In addition, the $\pm 40^\circ$ fluctuation of χ_{P2} in the mutant larger than that of the wt indicated that the P2 side chain loses its rigidity. The rearrangement of the P2 was related to the event of the flap motion as describe previously. Moreover, it involves a decrease of the strength of the hydrogen bond between the mutated residue and the P2 subsite (described in the next topic). The overall change in terms of torsion angles is supposed to explain a decrease of saquinavir sensitivity.

Decrease of hydrogen bond strength

The MD results show that the conformational difference in SQV subsites between the wt and the G48V complexes was located at the P2 side chain. Among residues surrounding the P2 subsite, position 48 is critical to conformational change of the inhibitor subsite. Direct contacts from the backbone oxygen of the mutated residue (O(48)) to the side-chain amide group (HN_{P2}) of P2, and to the backbone amide proton (HN_{bb}) of the inhibitor are illustrated by Fig. 8.

The O(48)-HN_{bb}(SQV) distance was, on average, 1.95 Å for the wt and extends to 2.25 Å in the G48V complex (Fig. 9

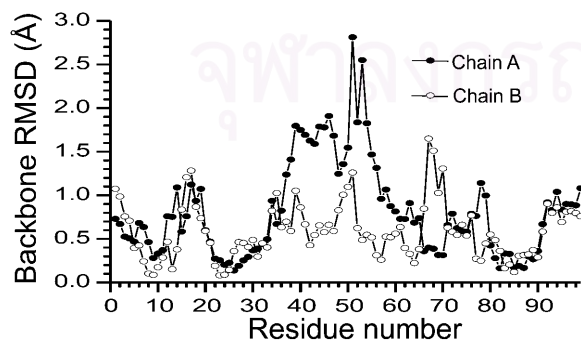


FIGURE 5 Backbone RMSD between the wild-type and the G48V structure.

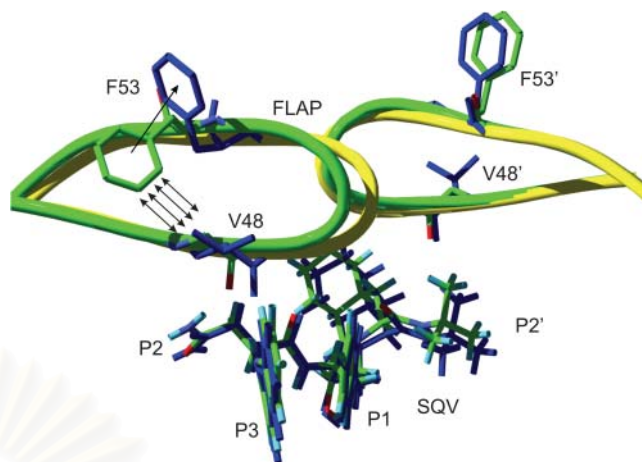


FIGURE 6 HIV-1 PR flap structures of the wt (green) and the G48V (yellow and blue). Some selected residues and saquinavir are presented in stick mode. Hydrogen atoms of the enzyme are not shown for simplification.

A). An increase of the distance is an evidence of decreasing the hydrogen bond strength.

A loss of the hydrogen bond interactions of residue 48 to SQV in the mutant complex was confirmed by a very large distance between O(48) and HN_{P2}(SQV) (Fig. 9 B). The wt complex shows the close proximity between O(48) and HN_{P2}(SQV). On the other hand, it is clear that the O(48)-HN_{P2}(SQV) distance in the mutant was not in a range of hydrogen bonding. The rotation of the P2 side chain as described previously is the cause of the disruption of the hydrogen bond of O(48)-HN_{P2}(SQV).

Analysis of these MD results indicates that the mutation at position 48 decreases the capability of inhibitor binding. A reduction of the interactions was estimated by calculating the ab initio energy. The $\Delta E_{G48-SQV}$ decreased ~ 3.5 kcal·mol⁻¹ with respect to the $\Delta E_{V48-SQV}$. The magnitude of changes in the interaction energy of V48-SQV supports the experimental K_i data that explain a small decrease (13.5-fold) of saquinavir sensitivity.

DISCUSSION

The MD simulations were carried out to compare structure and dynamics of the wt and the G48V HIV-1 PR-saquinavir complex. In addition, the MD simulations for the four protonation systems were carried out to obtain structural models before an evaluation of the ionization form of the active site residues. The study of protonation state of the HIV-PR-SQV complex was achieved with extensive energy calculations using the DFT, ONIOM, and MM/PBSA methods. Based on our data, both quantum chemical and molecular dynamics free-energy calculations confirm that the protonation model of wt-mono25 is the most energetically favorable case.

The results were in agreement with NMR studies of the hydroxyethylene isostere inhibitors, pepstatin and KNI-272,

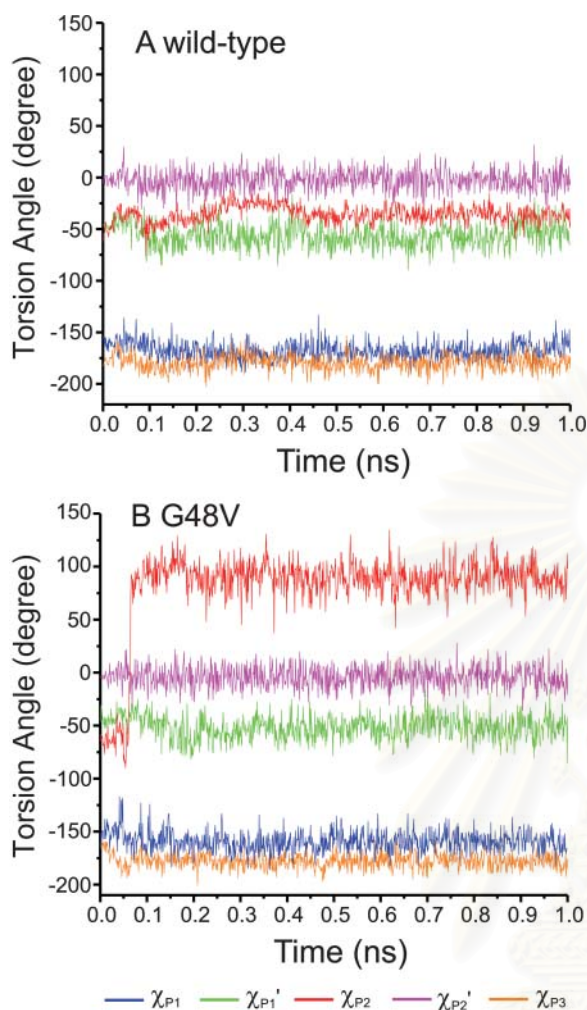


FIGURE 7 Fluctuation of χ_{P1} , $\chi_{P1'}$, χ_{P2} , $\chi_{P2'}$, and χ_{P3} corresponding to the dihedral angles of the inhibitor side chains P1, P1', P2, P2', and P3, respectively.

complexed with the enzyme (Smith et al., 1996; Wang et al., 1996). In this study, the protonation takes place only on the D25 side chain. Although the experimental data are not available for the HIV-1 PR-SQV complex, the structure of

KNI-272 is almost identical to that of saquinavir. They share the most common features of drug specificity, including the capability of binding between the central hydroxyethylene isostere of the protease inhibitor and the catalytic residues of the enzyme.

Recently, the protonation state of the HIV-1 PR-SQV complex was studied using quantum and free-energy perturbation methods (Lepsik et al., 2004; Nam et al., 2003). The interaction energy of SQV and the active-site residues were obtained based on the model taken from the x-ray structure. The protonation model proposed from those studies was also monoprotonated D25.

It is worth nothing that the quantum-based method is a promising tool for the study of receptor-ligand complexes. The determination of the protonation state of the HIV-PR active site residues is achievable on the basis of the QM results. In particular, the ONIOM method has extended the limitation of system size by the pure QM method. As demonstrated by the ONIOM results, the enzyme-inhibitor interactions of the catalytic region and the 5 Å surrounding residues are important for stabilizing the complex. The effect of the long-distance range on the interaction energy was not dominant at the MM level. Nevertheless, the method should be used with some proper care. The effect of the boundary resulting from incompatibility between molecular orbital wave functions in the QM part and the MM region may drive unrealistic energy values (Morokuma, 2002).

An effect of solvent environment, which remains unobvious in quantum approach, has been fulfilled with the calculations of binding free energy. The binding free energy of wt-mono25 obtained from MM/PBSA was ~ 2 kcal mol⁻¹ lower than that of the other states (Table 1). The difference was contributed from ΔG_{sol} rather than ΔE^{MM} . Not surprisingly, ΔE^{MM} of the three protonation model were almost equivalent (Table 2). Among all three systems, the only difference, that is the ionizable groups of D25 and D25', cannot be well described by the empirical force-field energy. In addition, the influence of the solvent to the protonation site was trivial on the basis of the radial distribution function plot (Wittayanarakul et al., 2005). The radial distribution function

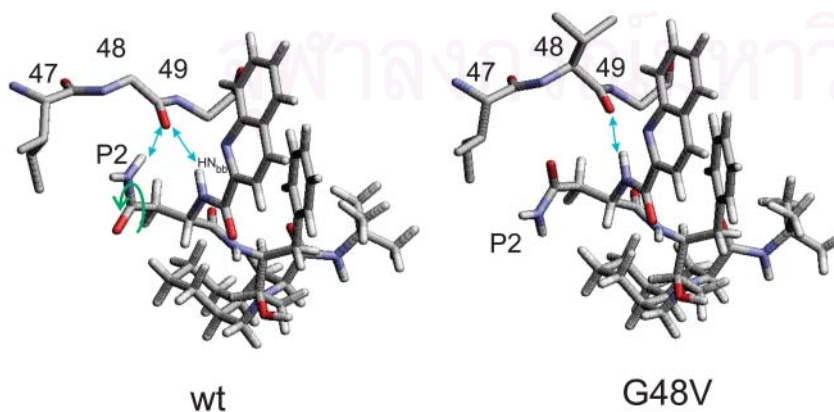


FIGURE 8 Conformational change of SQV at the P2 side chain. The rotation of the P2 subsite and the hydrogen bonds are illustrated.

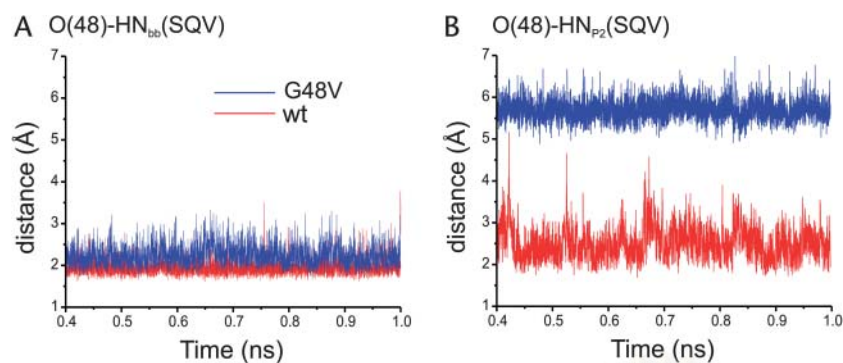


FIGURE 9 Distance trajectories of hydrogen bonding involving the CO backbone of residue 48.

of the hydroxyethylene oxygen of SQV showed an exclusion of water molecules from the protonation site. In this case, the energy information obtained from DFT and ONIOM were reliable. We anticipate that a method for the correction of the MM energy term by the QM treatment in MM/PBSA approach will be valuable for biomolecular research.

The binding pattern at the active site of the wt complex was similar that of the G48V as shown by the very low RMSD of residues E21–D30. Importantly, the interaction energy of the triad residues to SQV was insignificantly different between the wt and the G48V complexes. The results indicated the signature residue mutation developed in the primary resistance does not influence the interactions at the active site.

Significant changes were located at the protein flaps. Moreover, the flap structure of chain *A* was different from that of chain *B*. This is caused by different interactions of the enzyme to the asymmetric inhibitor. Particularly, the perturbation adopts heavily on the flap conformation of chain *A* rather than chain *B*. The slide of flap *A* in the G48V mutant seems to overcome a potential steric conflict caused by the substituted valine. As shown in Fig. 6, position 48 was in close contact with the F53 side chain of the enzyme and the P2 and P3 groups of SQV. Since the substituted valine of the mutant cannot be entirely accommodated in the hydrophobic pocket due to steric conflict of the dimethyl groups with the F53 and P3 side chains, the flap, therefore, shifted toward the symmetric axis of the enzyme. This rearrangement additionally destabilizes hydrogen bonding between the backbone CO of residue 48 of the enzyme and the P2 subsite of the inhibitor. In addition to the flap movement, the side chain of hydrophobic F53 became solvent-exposed to avoid steric clashes with V48 and Met-46, whereas an orientation of F53' in flap *B* of the mutant was not changed dramatically.

The role of the HIV-1 PR mutation at the flexible flap has been considerably debated about whether it would facilitate the binding reaction or reduce stability of the inhibitor, or both (Ermolieff et al., 1997; Hong et al., 1997; Maschera et al., 1996). The crystal structure of the double mutant G48V/L90M complexed with SQV revealed side-chain rearrangement of the P2 subsite and the F53 of the enzyme

similar to this study (Hong et al., 2000). Particularly, the missing of hydrogen bonding between the P2 subsite and the backbone C=O of residue 48 was also found in the x-ray structure of the double mutant. The crystal structure of G48H complexed with peptidic inhibitor U-89360E reported a decrease of flap mobility to stabilize the ligand (Hong et al., 1997).

Apparently the conformational change of the P2 subsite was an influence of steric conflict of the mutation at position 48. Importantly, it reduced saquinavir susceptibility to the mutant by interrupting the hydrogen bond interactions. Thus, a new drug with reduced steric repulsion on P2 could be designed to enhance the activity toward this mutant strain.

The x-ray structural data provided relevant information, insight into the molecular mechanism of HIV resistance to the protease inhibitor. Dynamic details of the full-atomic representation of the complex were required for further investigation. Therefore, MD simulation offers a good opportunity to fulfill basic information that could be useful in understanding the drug resistance mechanism and helpful in designing an anti-HIV inhibitor.

CONCLUSIONS

Molecular dynamics simulations of the wt and the G48V HIV-1 protease complexed with SQV were carried out to investigate the molecular basis of drug resistance. The MD

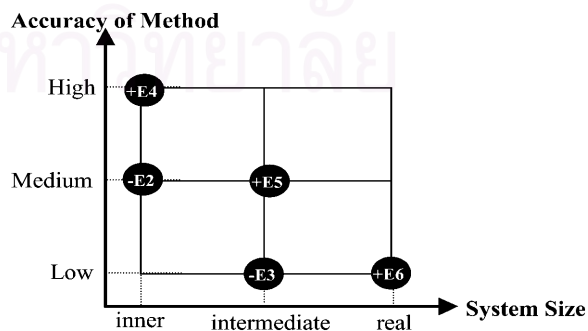


FIGURE 10 Schematic representation of the three-layer ONIOM extrapolation scheme.

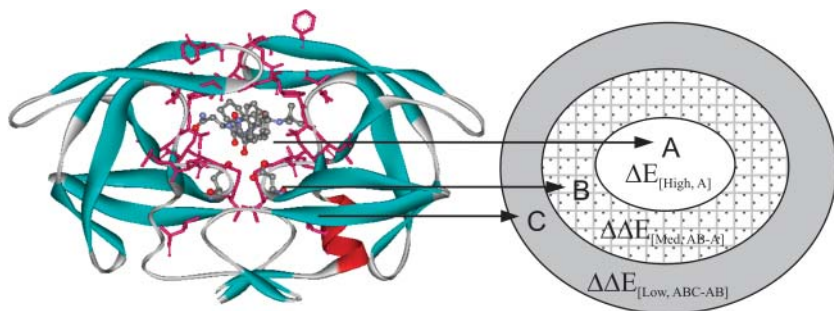


FIGURE 11 Schematic representation for the structure of HIV-1PR-SQV with the three partitioned layers.

results combined with quantum chemical calculations extend the capability of molecular modeling methods to study some biological systems, of which structural information is limited. This study showed that both complexes form the monoprotation on D25. Overall tertiary structure of the wt and the mutant protease was not significantly altered. Particularly, structure and interactions at the central active site remain unchanged. However, conformational differences between the wt and the G48V mutant were on the protein flap. The conformational change of the P2 subsite decreased the strength of hydrogen bonding of the backbone CO of the residue 48 to SQV. The change in interaction energies was comparable to the experimental K_i data. These observations provide useful information for designing potent HIV-1 PR inhibitors.

APPENDIX: ONIOM CALCULATIONS

The key interactions centered on SQV and the catalytic residues were treated at a high level of calculations, whereas the environmental effect of the entire protein was calculated at a lower level of calculations. In this study, three-layered ONIOM (ONIOM3) was employed (Morokuma, 2002).

On the basis of the ONIOM3 method shown in Fig. 10, the total energy of the system can be obtained from five independent calculations as

$$E^{\text{ONIOM3}} = E[\text{Low, real}] + E[\text{Med, intermediate}] + E[\text{High, inner}] - E[\text{Low, intermediate}] - E[\text{Med, inner}] \quad (11)$$

or

$$E^{\text{ONIOM3}} = E6 + E5 + E4 - E3 - E2, \quad (12)$$

where *real* denotes the entire system, of which the energy ($E6$) is calculated at the low level. For the *intermediate* layer, the energy is computed at both the *medium* ($E5$) and *low* ($E3$) level. For the *inner* layer, the energy is obtained at both *high* ($E4$) and *medium* ($E2$) level.

As shown by Fig. 11, the structure of the HIV-1 PR-SQV complex was partitioned into three parts represented by *inner* layer (A), *intermediate* layer (A + B), and *real* layer (A + B + C). The atoms of the enzyme and the inhibitor defined to each layers were described in the Methods section.

To obtain the interaction energy of the HIV-1 PR/SQV complex (ΔE_{cpx}), one can be expressed as

$$\Delta E_{\text{cpx}} = E_{\text{cpx}} - E_{\text{PR}} - E_{\text{SQV}}, \quad (13)$$

where E_{cpx} , E_{PR} , and E_{SQV} are the total energy of the HIV-1 PR/SQV complex, HIV-1 PR and SQV, respectively.

Financial support by the Thailand Research Fund and the generous supply of computer time by the Austrian-Thai Center for Computer-Assisted Chemical Education and Research (Bangkok, Thailand) are gratefully acknowledged.

REFERENCES

- Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1994. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* 238:415–436.
- Baldwin, E. T., T. N. Bhat, S. Gulnik, B. Liu, I. A. Topol, Y. Kiso, T. Mimoto, H. Mitsuya, and J. W. Erickson. 1995. Structure of HIV-1 protease with KNI-272, a tight-binding transition-state analog containing allophenylnorstatine. *Structure.* 3:581–590.
- Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
- Boucher, C. 1996. Rational approaches to resistance: using saquinavir. *AIDS.* 10 (Suppl. 1):S15–S19.
- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* 4:187–217.
- Case, D. A., J. C. D. Pearlman, T. Cheatham III, J. Wang, W. Ross, C. Simmerling, T. Darden, T. Merz, R. Stanton, A. Cheng, J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. Radmer, Y. Duan, J. Pitera, I. Massova, G. Seibel, U. C. Singh, P. Weiner, and P. A. Kollman. 2002. AMBER 7. University of California, San Francisco, CA.
- Chen, X., and A. Tropsha. 1995. Relative binding free energies of peptide inhibitors of HIV-1 protease: the influence of the active site protonation state. *J. Med. Chem.* 38:42–48.
- Collins, J. R., S. K. Burt, and J. W. Erickson. 1995. Flap Opening in HIV-1 protease simulated by activated molecular-dynamics. *Nat. Struct. Biol.* 2:334–338.
- Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* 117:5179–5197.
- Cornell, W. D., P. Cieplak, C. I. Bayly, and P. A. Kollman. 1993. Application of RESP charges to calculate conformational energies, hydrogen-bond energies, and free-energies of solvation. *J. Am. Chem. Soc.* 115:9620–9631.
- Davis, M. E., J. D. Madura, B. A. Luty, and J. A. McCammon. 1991. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics Program. *Comput. Phys. Comm.* 62:187–197.
- Debouck, C., J. G. Gorniak, J. E. Strickler, T. D. Meek, B. W. Metcalf, and M. Rosenberg. 1987. Human immunodeficiency virus protease expressed in *Escherichia coli* exhibits autoprocessing and specific maturation of the gag precursor. *Proc. Natl. Acad. Sci. USA.* 84:8903–8906.

- Deeks, S. G. 2003. Treatment of antiretroviral-drug-resistant HIV-1 infection. *Lancet*. 362:2002–2011.
- Eberle, J., B. Bechowsky, D. Rose, U. Hauser, K. von der Helm, L. Gurtler, and H. Nitschko. 1995. Resistance of HIV type 1 to proteinase inhibitor Ro 31–8959. *AIDS Res. Hum. Retro.* 11:671–676.
- Ermolieff, J., X. Lin, and J. Tang. 1997. Kinetic properties of saquinavir-resistant mutants of human immunodeficiency virus type 1 protease and their implications in drug resistance in vivo. *Biochemistry*. 36:12364–12370.
- Friesner, R. A., and M. D. Beachy. 1998. Quantum mechanical calculations on biological systems. *Curr. Opin. Struct. Biol.* 8:257–262.
- Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, J. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, and others. 2002. Gaussian 98. Gaussian, Inc. Pittsburgh, PA.
- Gilson, M. K., K. Sharp, and B. Honig. 1987. Calculating electrostatic interactions in bio-molecules: method and error assessment. *J. Comput. Chem.* 9:327–335.
- Gueux, N., and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 18:2714–2723.
- Harte, W. E., Jr., S. Swaminathan, and D. L. Beveridge. 1992. Molecular dynamics of HIV-1 protease. *Proteins*. 13:175–194.
- Harte, W. E., Jr., S. Swaminathan, M. M. Mansuri, J. C. Martin, I. E. Rosenberg, and D. L. Beveridge. 1990. Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proc. Natl. Acad. Sci. USA*. 87:8864–8868.
- Hong, L., A. Treharne, J. A. Hartsuck, S. Foundling, and J. Tang. 1996. Crystal structures of complexes of a peptidic inhibitor with wild-type and two mutant HIV-1 proteases. *Biochemistry*. 35:10627–10633.
- Hong, L., X. C. Zhang, J. A. Hartsuck, and J. Tang. 2000. Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance. *Protein Sci.* 9:1898–1904.
- Hong, L., X. J. Zhang, S. Foundling, J. A. Hartsuck, and J. Tang. 1997. Structure of a G48H mutant of HIV-1 protease explains how glycine-48 replacements produce mutants resistant to inhibitor drugs. *FEBS Lett.* 420:11–16.
- Hyland, L. J., T. A. Tomaszek Jr., and T. D. Meek. 1991. Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry*. 30:8454–8463.
- Jaskolski, M., A. G. Tomasselli, T. K. Sawyer, D. G. Staples, R. L. Heinrikson, J. Schneider, S. B. Kent, and A. Wlodawer. 1991. Structure at 2.5-Å resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor. *Biochemistry*. 30:1600–1609.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
- Jorgensen, W. L., and J. Tirado-Rives. 1988. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.
- Kollman, P. A., I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham 3rd. 2000. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33:889–897.
- Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.
- Krohn, A., S. Redshaw, J. C. Ritchie, B. J. Graves, and M. H. Hatada. 1991. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. *J. Med. Chem.* 34:3340–3342.
- Laskowski, R. A., J. A. Rullmann, M. W. MacArthur, R. Kaptein, and J. M. Thornton. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*. 8:477–486.
- Lepsik, M., Z. Kriz, and Z. Havlas. 2004. Efficiency of a second-generation HIV-1 protease inhibitor studied by molecular dynamics and absolute binding free energy calculations. *Proteins*. 57:279–293.
- Levy, Y., and A. Caffisch. 2003. Flexibility of monomeric and dimeric HIV-1 protease. *J. Phys. Chem. B*. 107:3068–3079.
- Maschera, B., G. Darby, G. Palu, L. L. Wright, M. Tisdale, R. Myers, E. D. Blair, and E. S. Furfine. 1996. Human immunodeficiency virus. Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. *J. Biol. Chem.* 271:33231–33235.
- Meek, T. D., B. D. Dayton, B. W. Metcalf, G. B. Dreyer, J. E. Strickler, J. G. Gorniak, M. Rosenberg, M. L. Moore, V. W. Magaard, and C. Debouck. 1989. Human immunodeficiency virus 1 protease expressed in *Escherichia coli* behaves as a dimeric aspartic protease. *Proc. Natl. Acad. Sci. USA*. 86:1841–1845.
- Morokuma, K. 2002. New challenges in quantum chemistry: Quests for accurate calculations for large molecular systems. *Philos. Transact. A Math. Phys. Eng. Sci.* 360:1149–1164.
- Nam, K. Y., B. H. Chang, C. K. Han, S. G. Ahn, and K. T. No. 2003. Investigation of the protonated state of HIV-1 protease active site. *Bull. Kor. Chem. Soc.* 24:817–823.
- Northrop, D. B. 2001. Follow the protons: a low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Acc. Chem. Res.* 34:790–797.
- Okimoto, N., T. Tsukui, M. Hata, T. Hoshino, and M. Tsuda. 1999. Hydrolysis mechanism of the phenylalanine-proline peptide bond specific to HIV-1 protease: Investigation by the ab initio molecular orbital method. *J. Am. Chem. Soc.* 121:7349–7354.
- Piana, S., D. Bucher, P. Carloni, and U. Rothlisberger. 2004. Reaction mechanism of HIV-1 protease by hybrid CarParrinello/classical MD simulations. *J. Phys. Chem. B*. 108:11139–11149.
- Piana, S., P. Carloni, and U. Rothlisberger. 2002. Drug resistance in HIV-1 protease: flexibility-assisted mechanism of compensatory mutations. *Protein Sci.* 11:2393–2402.
- Piana, S., D. Sebastiani, P. Carloni, and M. Parrinello. 2001. Ab initio molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site. *J. Am. Chem. Soc.* 123:8730–8737.
- Prabhakar, R., D. G. Musaev, I. V. Khavrutskii, and K. Morokuma. 2004. Does the active site of mammalian glutathione peroxidase (GPx) contain water molecules? An ONIOM study. *J. Phys. Chem. B*. 108:12643–12645.
- Prabu-Jeyabalan, M., E. A. Nalivaika, N. M. King, and C. A. Schiffer. 2003. Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy. *J. Virol.* 77:1306–1315.
- Randolph, J. T., and D. A. DeGoey. 2004. Peptidomimetic inhibitors of HIV protease. *Curr. Top. Med. Chem.* 4:1079–1095.
- Roberts, N. A., J. A. Martin, D. Kinchington, A. V. Broadhurst, J. C. Craig, I. B. Duncan, S. A. Galpin, B. K. Handa, J. Kay, A. Krohn, and others. 1990. Rational design of peptide-based HIV proteinase inhibitors. *Science*. 248:358–361.
- Rocchia, W., E. Alexov, and B. Honig. 2001. Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. *J. Phys. Chem. B*. 105:6507–6514.
- Rodriguez-Barrios, F., and F. Gago. 2004. HIV protease inhibition: Limited recent progress and advances in understanding current pitfalls. *Curr. Top. Med. Chem.* 4:991–1007.
- Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341.

- Scott, W. R., and C. A. Schiffer. 2000. Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance. *Struct. Fold. Des.* 8:1259–1265.
- Sitkoff, D., K. A. Sharp, and B. Honig. 1994. Accurate calculation of hydration free-energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988.
- Smith, R., I. M. Brereton, R. Y. Chai, and S. B. H. Kent. 1996. Ionization states of the catalytic residues in HIV-1 protease. *Nat. Struct. Biol.* 3: 946–950.
- Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* 120:9401–9409.
- Swain, A. L., M. M. Miller, J. Green, D. H. Rich, J. Schneider, S. B. Kent, and A. Wlodawer. 1990. X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proc. Natl. Acad. Sci. USA.* 87:8805–8809.
- Torrent, M., T. Vreven, D. G. Musaev, K. Morokuma, O. Farkas, and H. B. Schlegel. 2002. Effects of the protein environment on the structure and energetics of active sites of metalloenzymes. ONIOM study of methane monooxygenase and ribonucleotide reductase. *J. Am. Chem. Soc.* 124: 192–193.
- Vaillancourt, M., D. Irlbeck, T. Smith, R. W. Coombs, and R. Swanstrom. 1999. The HIV type 1 protease inhibitor saquinavir can select for multiple mutations that confer increasing resistance. *AIDS Res. Hum. Retroviruses.* 15:355–363.
- Vondrasek, J., and A. Wlodawer. 2002. HIVdb: a database of the structures of human immunodeficiency virus protease. *Proteins.* 49:429–431.
- Wang, J. M., W. Wang, and P. A. Kollman. 2001. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* 222:U403–U403 (Abstr).
- Wang, W., and P. A. Kollman. 2000. Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. *J. Mol. Biol.* 303:567–582.
- Wang, Y. X., D. I. Freedberg, T. Yamazaki, P. T. Wingfield, S. J. Stahl, J. D. Kaufman, Y. Kiso, and D. A. Torchia. 1996. 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.
- Wittayanarakul, K., O. Aruksakunwong, P. Sompornpisut, V. Sanghiran-Lee, V. Parasuk, S. Pinitglang, and S. Hannongbua. 2005. Structure, dynamics and solvation of HIV-1 protease/saquinavir complex in aqueous solution and their contributions to drug resistance: molecular dynamic simulations. *J. Chem. Inf. Comput. Sci.* In press.
- Wlodawer, A., and J. Vondrasek. 1998. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* 27:249–284.
- Yamazaki, T., L. K. Nicholson, D. A. Torchia, P. Wingfield, S. J. Stahl, J. D. Kaufman, C. J. Eyermann, C. N. Hodge, P. Y. S. Lam, Y. Ru, and others. 1994. NMR and x-ray evidence that the HIV protease catalytic aspartyl groups are protonated in the complex formed by the protease and a nonpeptide cyclic urea-based inhibitor. *J. Am. Chem. Soc.* 116:10791–10792.
- Yang, A. S., M. R. Gunner, R. Sampogna, K. Sharp, and B. Honig. 1993. On the calculation of Pk(a)S in proteins. *Proteins.* 15:252–265.
- York, D. M., T. A. Darden, and L. G. Pedersen. 1993a. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. *J. Chem. Phys.* 99:8345–8348.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



Structural and dynamical properties of different protonated states of mutant HIV-1 protease complexed with the saquinavir inhibitor studied by molecular dynamics simulations

Ornjira Aruksakunwong^a, Kitiyaporn Wittayanarakul^a, Pornthep Sompornpisut^a,
 Vannajan Sanghiran^b, Vudthichai Parasuk^a, Supot Hannongbua^{a,*}

^a Department of Chemistry, Faculty of Science, Chulalongkorn University, Pratumwan, Bangkok 10330, Thailand

^b Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

Received 13 July 2005; received in revised form 17 January 2006; accepted 18 January 2006

Abstract

To understand the basis of drug resistance, particularly of the HIV-1 PR, three molecular dynamics (MD) simulations of HIV-1 PR mutant species, G48V, complexed with saquinavir (SQV) in explicit aqueous solution with three protonation states, diprotonation on Asp25 and Asp25' (Di-pro) and monoproteination on each Asp residue (Mono-25 and Mono-25'). For all three states, H-bonds between saquinavir and HIV-1 PR were formed only in the two regions, flap and active site. It was found that conformation of P2 subsite of SQV in the Mono-25 state differs substantially from the other two states. The rotation about 177° from the optimal structure of the wild type was observed, the hydrogen bond between P2 and the flap residue (Val48) was broken and indirect hydrogen bonds with the three residues (Asp29, Gly27, and Asp30) were found instead. In terms of complexation energies, interaction energy of -37.3 kcal/mol for the Mono-25 state is significantly lower than those of -30.7 and -10.7 kcal/mol for the Mono-25' and Di-pro states, respectively. It was found also that protonation at the Asp25 leads to a better arrangement in the catalytic dyad, i.e., the Asp25-Asp25' interaction energy of -8.8 kcal/mol of the Mono-25 is significantly lower than that of -2.6 kcal/mol for the Mono-25' state. The above data suggest us to conclude that interaction in the catalytic area should be used as criteria to enhance capability in drug designing and drug screening instead of using the total inhibitor/enzyme interaction.

© 2006 Elsevier Inc. All rights reserved.

Keywords: HIV-1 protease; Molecular dynamics simulations; Mutation; Protonation state; Saquinavir

1. Introduction

The human immunodeficiency virus type 1 (HIV-1) is the causative agent in acquired immunodeficiency syndrome (AIDS). This disease was recognized in the U.S., around 1981 [1]. There are three essential enzymes involved in the replication cycle of this virus which are reverse transcriptase (RT), protease (PR) and integrase (IN). They are therefore important targets for drug development. Although effective drugs have been developed against HIV-1 PR and HIV-1 RT, however, it was reported that inhibitors at the first target are more potent [2]. Therefore, HIV-1 PR is an attractive target for antiviral therapy. The HIV-1 PR consists of two identical

polypeptides of 99 amino acids (Fig. 1), each chain contains an N-terminal Pro and C-terminal Phe. The active site of PR is formed at the dimer interface containing two conserved catalytic dyad, Asp25 and Asp25' [3]. The substrate binding cleft is composed of equivalent residues from each subunit and is bound on one side by the active site aspartic acids and on the other by the flap region. To date, there are several FDA-approved PR inhibitors in clinical uses and saquinavir is considered to be a highly potent and selective transition state analog inhibitor of the HIV-1 PR [4].

A major problem for the clinical uses of PR inhibitors is the development of drug resistance due to substitutions observed in almost 50% of the residues and over 20 residues associated with resistance to clinically available inhibitors [5]. Recently, the correlation between the inhibitor structure of the HIV-1 PR target and the drug resistance was studied [6]. The kinetic experimental data show that decreased affinity of the drugs for

* Corresponding author. Tel.: +66 2 2187603; fax: +66 2 2187603.

E-mail address: supot.h@chula.ac.th (S. Hannongbua).

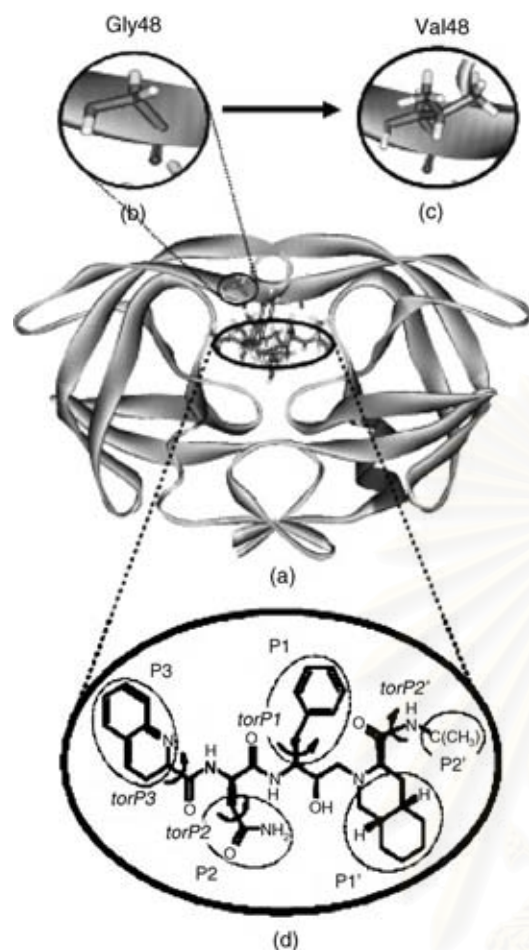


Fig. 1. Schematic representation of the HIV-1 PR complexed with saquinavir (a) in which molecular structure of Gly48 (b) and Val48 (c) in the G48V and of saquinavir (d) were also displayed. Here, P1, P2, P3, P1' and P2' subsites are labeled by circle and torsional angle of each subsite was defined by *tor*.

many mutants is caused primarily by an increase in dissociation rates. The HIV-1 PR mutant species, G48V, is associated in vivo with saquinavir resistance [7]. Residue Gly48 locates in the flap region of the HIV-1 PR and is responsible for the formation of the S2/S2' and S3/S3' binding site—the regions of the enzyme that bind with P2/P2' and P3/P3' of the inhibitor, respectively [8]. In addition, it was found that Gly48 plays a role in shaping the binding pocket of the active site and stabilizing the enzyme–substrate complex [6]. Substitution of this residue by valine introduces a bulky side chain into the S3/S3' binding pocket and results in resistance towards saquinavir leading to an increase of the inhibition constant (K_i) value by 13.5-fold [9]. However, this effect was not observed for other inhibitors.

As an aspartic protease, the protonation state of the catalytic aspartic acids, Asp25 and Asp25', is the key to explain the catalytic mechanism of both wild type and mutant type. Plane wave-based *ab initio* molecular dynamics calculations [10] as well as NMR measurements [11] of the HIV-1 PR complexed with pepstatin lead to the conclusion that the system is, at least, monoprotonated. In addition, the *ab initio* calculation method was performed on the active site of HIV-1 PR and the free energy perturbation (FEP) method was used to determine the

binding free energy of four different protonated states of HIV-1 PR complexed with A74704 by Ky-Youb Nam et al. [12]. The results have the potentially significant implications that the complex is monoprotonated on Asp25.

In this study, molecular dynamics simulations of the G48V HIV-1 PR complexed with saquinavir in explicit aqueous solution were carried out. Due to an unclear detailed mechanism of the reaction catalyzed by HIV-1 PR [13], therefore, the simulations have to be performed for all three possible protonation states of the two aspartic residues, Asp25 and Asp25'.

2. Computational method

2.1. Preparation of the initial structures

In order to investigate the relative dynamics properties among different protonation states of the mutant HIV-1 PR complexed with the inhibitor (saquinavir), three MD simulations were performed, monoprotonate on Asp25 (Mono-25), monoprotonate on Asp25' (Mono-25') and diprotonate on both aspartic acids (Di-pro). The crystal structure of saquinavir bound to wild-type protease was taken from the Protein Data Bank PDB (1HXB) [14] and used as the reference structure. Among the two available crystal structures, the first one which are commonly used in literatures [15] was applied for our simulation. All missing atoms of the protein were added using the LEaP module in the AMBER 7 software package [16]. The protonation state of the ionizable residues, the C- and the N-termini, except for D25/25', was assigned based on the predicted pK_a values at pH 7. The pK_a s of ionizable residues were calculated based on the Poisson–Boltzmann free energy calculations. Details of these calculations are given elsewhere [17]. Hydrogen atoms were then added to the two catalytic aspartic residues in order to generate the Mono-25, Mono-25' and Di-pro states using the LEaP module in AMBER 7.0 software package [16].

It should be noted that the X-ray structure of the double mutant, G48V/L90MSQV complex (1FB7) could be considered as an alternative template [4]. However, the X-ray coordinates of the second monomer of the double mutant are not available. Thus, 1HXB is considered to be more appropriate as a template. The mutant protease enzyme was modeled from this structure replacing glycine by valine at residue 48 (G48V) using the Insight II molecular modeling software.

2.2. Molecular dynamics simulations

Three MD simulations were carried out for the mutant HIV-1 PR complexed with saquinavir in the above mentioned states. The Mono-25, Mono-25' and Di-pro systems were solvated by 9633, 9627 and 9627 TIP3P water molecules [18], respectively. The crystallographic waters were also included in the simulations. The sodium and chloride ions were added to neutralize the system. The corresponding cubical cell for the

three systems was $76 \text{ \AA} \times 79 \text{ \AA} \times 68 \text{ \AA}$. Furthermore, potential parameters for saquinavir are not available in the AMBER package were developed using the following steps. Geometry of the X-ray saquinavir inhibitor was optimized at the Hartree–Fock level with 6-31G** basis functions to adjust the bond-length involving hydrogens. Then, the RESP fitting procedure was employed to calculate partial atomic charges of the inhibitor [19]. Energy minimizations were carried out to relax the model prior to MD runs. A cutoff distance of 12 \AA was applied for non-bonded pair interaction. The Particle Mesh Ewald (PME) method was employed for correcting electrostatic interaction. The SHAKE algorithm [20] was employed to constrain all bonds involving hydrogens. The simulated time step was set to 2 fs. The temperature of the model was gradually raised to 298 K during the first 60 ps, then, kept constant until 1 ns.

The molecular mechanics potential energy minimizations and MD simulations were carried out using the AMBER 7.0 simulation package. Calculations were performed using the parm99 force field. All MD runs reported here were done under an isobaric–isothermal ensemble (NPT) using constant pressure of 1 atm and constant temperature of 298 K [21]. All properties reported and discussed in this study were evaluated after equilibrium was reached, 400 ps.

2.3. Quantum chemical calculations

Quantum chemical calculations were performed to investigate the interaction energy between catalytic residues (Asp25 and Asp25') and saquinavir in the different protonated states of the complexes using the Gaussian 98 program [22]. Initial structures of the complex in the three protonation states were averaged from the whole trajectory, then, the geometry was relaxed using energy minimization from the AMBER 7.0 program. The selected residues, Asp25 and Asp25', were capped by $\text{CH}_3\text{NH-}$ and $-\text{COCH}_3$ groups at the C- and the N-terminal, respectively. The geometry of the hydrogen atoms in the cap methyl group were optimized using semi-empirical calculations at the PM3 level. Then, single point calculation with the density functional theory B3LYP was applied to investigate the total energy of the system. The basis sets used are the 6-31G(d,p) for unprotonated residue and the extended 6-31+G(d,p) for protonated residue. The total interaction energy of the complex (ΔE_{cpx}) was computed as

$$\Delta E_{\text{cpx}} = E_{\text{cpx}} - (E_{\text{SQV}} + E_{\text{Asp25}} + E_{\text{Asp25'}}) \quad (1)$$

where E on the right hand side represents total energy of the systems given in the subscript (cpx = complex and SQV = saquinavir). Evaluations were also performed to understand the contribution of pair interactions, $\Delta E_{\text{Asp25-SQV}}$, $\Delta E_{\text{Asp25'-SQV}}$ and $\Delta E_{\text{Asp25-Asp25'}}$, to the total complexation energy, i.e.,

$$\Delta E_{\text{cpx}} = \Delta E_{\text{Asp25-SQV}} + \Delta E_{\text{Asp25'-SQV}} + \Delta E_{\text{Asp25-Asp25'}} + \Delta E_{3\text{bd}} \quad (2)$$

where the $\Delta E_{3\text{bd}}$ denotes an error due to three-body correction. In other words, the third particle was excluded from the calculation of the pair interactions which are defined as

$$\Delta E_{\text{Asp25-SQV}} = E_{\text{Asp25+SQV}} - (E_{\text{Asp25}} + E_{\text{SQV}}) \quad (3a)$$

$$\Delta E_{\text{Asp25'-SQV}} = E_{\text{Asp25'+SQV}} - (E_{\text{Asp25'}} + E_{\text{SQV}}) \quad (3b)$$

$$\Delta E_{\text{Asp25-Asp25'}} = E_{\text{Asp25+Asp25'}} - (E_{\text{Asp25}} + E_{\text{Asp25'}}). \quad (3c)$$

In addition to the complexation energy shown in Eq. (1), the interaction energies between saquinavir and the catalytic dyad residues ($\Delta E_{\text{SQV-(Asp25+Asp25')}}$) was calculated separately according to the following equation:

$$\Delta E_{\text{SQV-(Asp25+Asp25')}} = E_{\text{cpx}} - (E_{\text{Asp25+Asp25'}} + E_{\text{SQV}}) \quad (4)$$

where $E_{\text{Asp25+Asp25'}}$ is total energy of the Asp25 and Asp25' dimer.

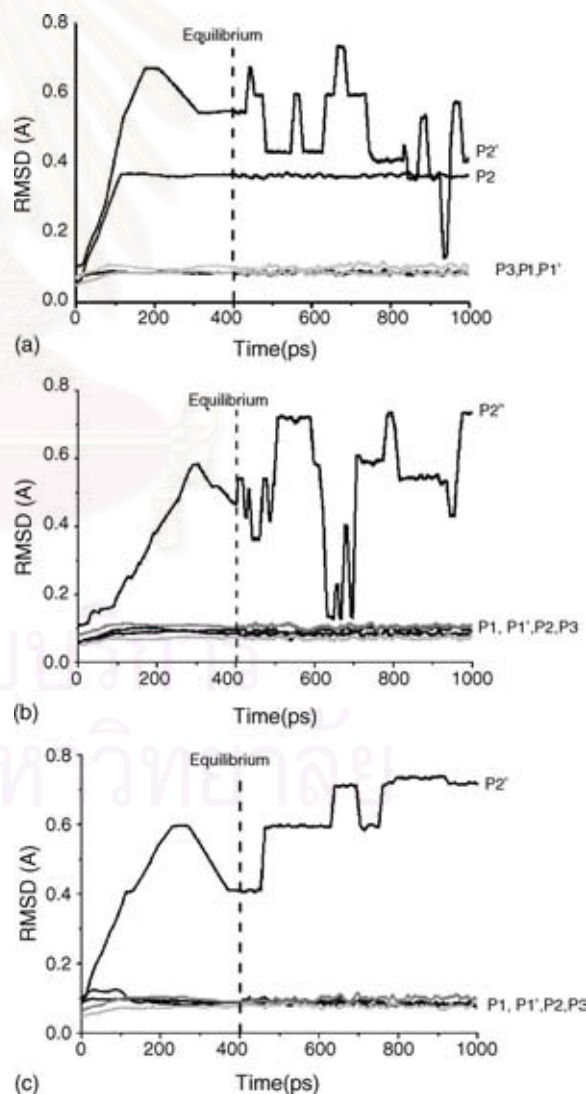


Fig. 2. Root-mean-square displacement (RMSD) of each subsite of saquinavir (defines in Fig. 1d) for the three simulated systems, (a) Mono-25, (b) Mono-25' and (c) Di-pro.

3. Results and discussion

3.1. Flexibility of the saquinavir inhibitor

The structure of saquinavir is known to consist of five subsites (P1, P2, P3, P1' and P2' in Fig. 1d). The root-mean-square displacement (RMSD) of the backbone atoms of each subsite with respect to the starting coordination, the X-ray structure, is plotted in Fig. 2 and the distribution of torsional angle (labeled by *tor* in Fig. 1d) was plotted in Fig. 3. Evaluations were carried out for each subsite of the three protonation states. Note that due to the rigid nature of the planar P1' subsite, therefore, its torsional plot was not taken into consideration. The RMSD plots through the simulation time indicate clearly that conformations of all subsites in the three protonated states are almost unchanged, except for the P2 subsite of Mono-25 and P2' subsite in all states.

As a result of these observations, attention was focused on P2 and P2' subsites. The RMSD plot for P2 subsite for the Mono-25 state (Fig. 2a) is significantly higher than the other two states and the plot for P2' for the three states are highly flexible when compared with the other subsites. The first event indicates a change of P2 subsite of Mono-25 from one (RMSD ~ 0.05 Å) to another (RMSD ~ 0.36 Å) conformation (more details are discussed later). Furthermore, the high flexibility of P2' subsite for all protonation states (Fig. 2a–c) seems to indicate the detection of three preferential conformations for this subsite of saquinavir, e.g., at RMSDs ca. 0.4, 0.6 and 0.7 Å

for Di-pro system, etc. This assumption was confirmed by the plot shown in Fig. 4.

Detailed information on the conformational flexibility of each subsite of saquinavir can be seen in terms of distributions of the torsional angles (Fig. 3) after equilibrium, defined in Fig. 1d. All plots show pronounced and sharp peaks indicating slight flexibility of each subsite in a narrow range. For P2' (Fig. 3c) and P3 (Fig. 3d) subsites, their conformations are not sensitive to the change of protonation states, leading to the most probable torsional angles of 0° and 180° , respectively. Fig. 3b, the torsional angle of P2 subsite of the Mono-25 of $\sim 90^\circ$ (solid line, Fig. 3b) differs substantially from that of $\sim -75^\circ$ for the other two states. The difference between the positions of the two peaks, $\sim 166^\circ$, suggests that the orientation of the P2 subsite of saquinavir in the Mono-25 state complex is almost opposite to that of the other states. This statement is supported by the snapshot shown in an inset of Fig. 3b. Note that single preferential conformation after equilibrium of P2 subsite for the Mono-25 state took place at the torsion angle of $\sim 90^\circ$ (solid line, Fig. 3b) is consistency with the RMSD = 0.36 Å shown in Fig. 2a. Furthermore, the sharp peak at the torsional angles of $\sim -75^\circ$ for the Mono-25' and Di-pro (Fig. 3b) states corresponds to the conformation represented by the RMSDs of ~ 0.08 Å (Fig. 2b and c).

It is interesting to note that, the observed conformation of P2 subsite of saquinavir in the Mono-25 state is similar to the X-ray structure found in the double mutant, G48V/L90M, HIV-1 PR complex [4]. It was also proposed by the X-ray data that

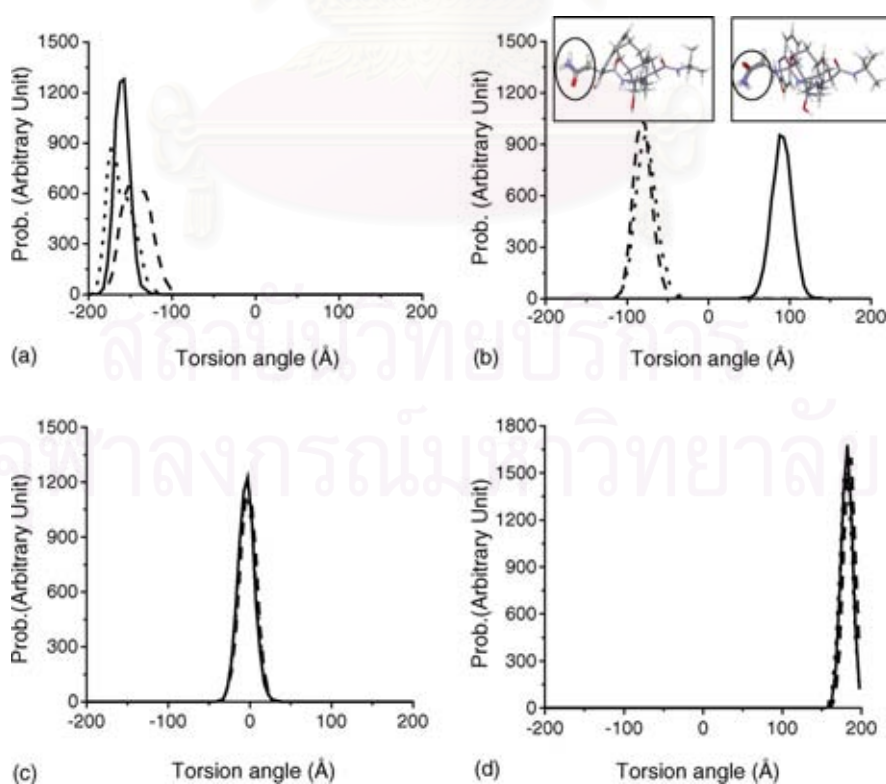


Fig. 3. Changes of torsional angles of four subsites of saquinavir inhibitor: (a) P1 subsite; (b) P2 subsite; (c) P2' subsite; (d) P3 subsite, (defined in Fig. 1d) for the three simulated systems where solid, dashed and dotted lines in the torsional angle plots represent the Mono-25, Mono-25' and Di-pro systems, respectively. The two conformations corresponding to torsional angles of 90.3° and -75.5° of P2 are also given in an inset.

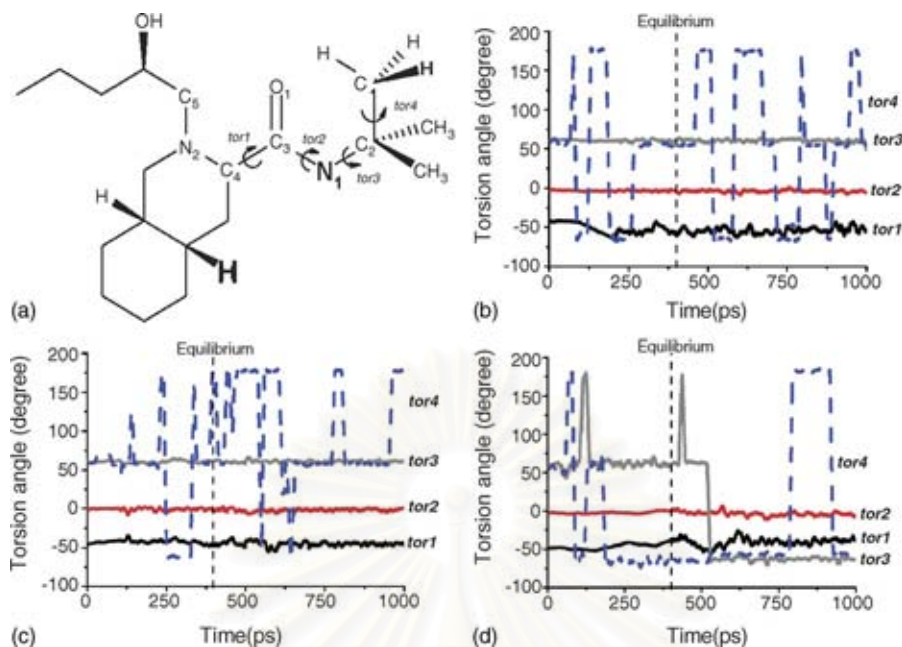


Fig. 4. Definition of torsional angles *tor1*–*tor4* (a) of the P2' subsite of saquinavir inhibitor and their changes as a function of simulation time; (b) Mono-25; (c) Mono-25'; (d) Di-pro.

rotation of the P2 subsite of saquinavir in the G48V/L90M mutation depends only on the position of Gly48 which lies much closer to the inhibitor than the Leu90. However, X-ray structure for the present system (single G48V mutation of HIV-1 PR complexed with saquinavir) are not available.

In addition, the three peaks of P1 subsite for the three states are slightly different (Fig. 3a). A sharper and higher peak at the torsional angle of -157° (Mono-25, solid line) than the other at -172° (Di-pro, dot line) and -142° (Mono-25', dash line) indicate a higher rigidity of P1 subsite in the Mono-25 than the other two states.

To seek more information on the high flexibility of P2' subsite, the four possible torsional angles, *tor1* (N₂-C₄-C₃-O₁), *tor2* (O₁-C₃-N₁-C₂), *tor3* (C₃-N₁-C₂-C₁) and *tor4* (N₁-C₂-C₁-H) were defined (Fig. 4a) and evaluated. The results were plotted in Fig. 4b–d. The *tor1* and *tor2* for all protonated states are highly stable showing preferential orientations at approximately -50° and 0° , respectively. These results suggested that the $-C_3=O_1$ functional group of P2' subsite (see Fig. 4a) was tilted by 50° from the C₃-C₄-N₂ plan while C₂

was detected to locate almost in the O₁-C₃-N₁ plane (*tor2* = 0°). Fluctuations were observed for *tor3* of Di-pro and *tor4* for all states. Particular, *tor4* shows three preferential conformations at -60° , 60° and 180° . This fact can be understood by a free rotation of the $-H_3$ and $-(CH_3)$ functional groups. An appearance of *tor4* at the three torsional angles for all states indicates a staggered conformation among the three hydrogen atoms of the $-CH_3$. The same conclusion can also be made for *tor3* of the diprotonated state, Di-pro, in which the three $-CH_3$ groups are observed to rotate freely with the three preferential states in the staggered conformation.

3.2. Flexibility of enzyme

The RMSD of all atoms with respect to the starting geometry as a function of simulation time was evaluated separately for each chain of the HIV-1 PR. The results were shown in Fig. 5. The plots describe the structure relaxations when the molecule was dissolved in the solution. The RMSD values of all plots increase rapidly for the first 100 ps indicating a change in

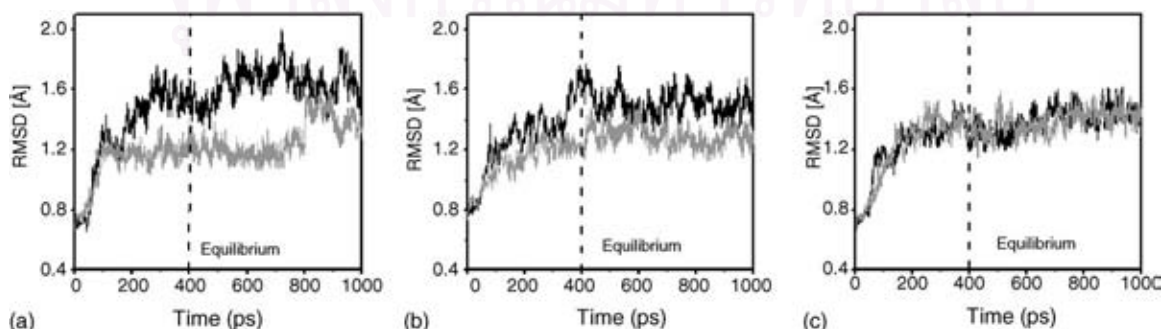


Fig. 5. RMSD as a function of simulation time of the two chains, A (black line) and B (gray line), of the G48V HIV-1 PR mutant-type complexed with saquinavir in the Mono-25 (a), Mono-25' (b) and Di-pro (c) states.

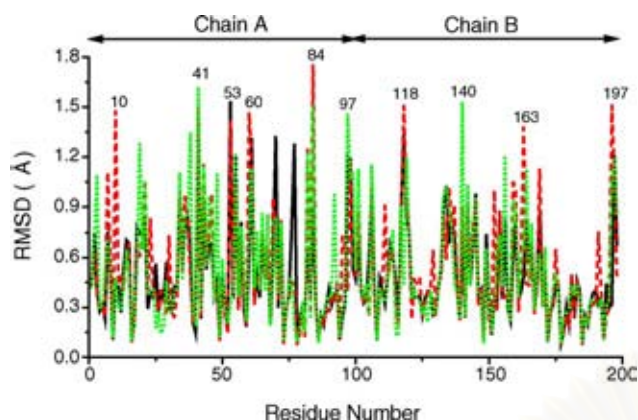


Fig. 6. Root-mean-square deviation for all residues average during 400–1000 ps of the Mono-25 (black), Mono-25' (red) and Di-pro (green) states where the highly flexible residues were labeled.

enzyme geometry relative to the starting X-ray structure. Although the value of almost all plots remain constant after 200 ps, however, the RMSDs for chain A of Mono-25 (Fig. 5a) and both chains of Mono-25' still increase slowly (Fig. 5b). This data suggest that monoprotonated states, Mono-25 and Mono-25', require longer time than the Di-pro state to approach equilibrium. To enhance the reliability of the data, all systems were considered to be in equilibrium after 400 ps. Cooperative effects due to different protonation states and a systematic geometry of the saquinavir inhibitor lead to higher RMSD on chain A (Fig. 5a and b) when compared with chain B of the two monoprotonated states. This difference was not detected for the diprotonated state (Fig. 5c) where both catalytic residues, Asp25 and Asp25' were protonated.

To ascertain for more details on the flexibility of each residue in the enzyme structure, the RMSD of all atoms for each residue were calculated and displayed in Fig. 6. From the plot, the following conclusions can be made (i) for monoprotonated states, the flexibility of chain A is higher than that of chain B Fig. 6. From the plot, the RMSD of almost all residues of the Mono-25 is slightly lower than those of the Mono-25' and the

Di-pro. The exception was found only for the residues 70 and 71 in which their RMSDs are significantly higher than the other two states. In addition, the highly flexible residues were labeled in Fig. 6. Note that these residues, except residue 84 in which mutation usually takes place, were found to be located on the solvent accessible surface of the enzyme.

3.3. Enzyme–saquinavir interaction

It is well known that inhibitor or substrate was held in the active site of enzyme via the intermolecular forces, such as electrostatic, hydrophobic, dispersion forces, etc. For the investigated system where the hydrophilic OH group of saquinavir binds directly to the catalytic residues, the catalytic region contains residues 25–30: Asp25-Thr26-Gly27-Ala28-Asp29-Asp30 and some water molecules. It appears that the electrostatic forces in the catalytic pocket play a key role of the complex. Therefore, attention was focused on the hydrogen bonding between inhibitor and enzyme. In addition, interaction between saquinavir and the catalytic dyad was also investigated.

Hydrogen bonding between HIV-1 PR and saquinavir was determined based on the Carnel module in the AMBER 7.0 using the following criteria: (i) the distance between proton donor (D) and acceptor (A) atoms was less than or equal to 3.5 Å and (ii) the D–H...A angle was greater than or equal to 120°. Analysis was carried out cover the entire trajectory after equilibration. The detected hydrogen bonds with the H...A distances were shown in Fig. 7a–c. It is interesting to note that hydrogen bonds between saquinavir and HIV-1 PR were formed only in the two regions that are flap and active site region. The number of hydrogen bonds were found 4, 5 and 6 for the Mono-25, Mono25' and Di-pro states, respectively. In the flap region, hydrogen bonds were formed only with the Val48 of the HIV-1 PR. In the Mono-25 state (Fig. 7a), the deformation of *d5* bond (5.75 Å) was observed. This observation is due to the rotation of the NH₂ functional group of P2 subsite of saquinavir (inset of Fig. 3b). In addition, the

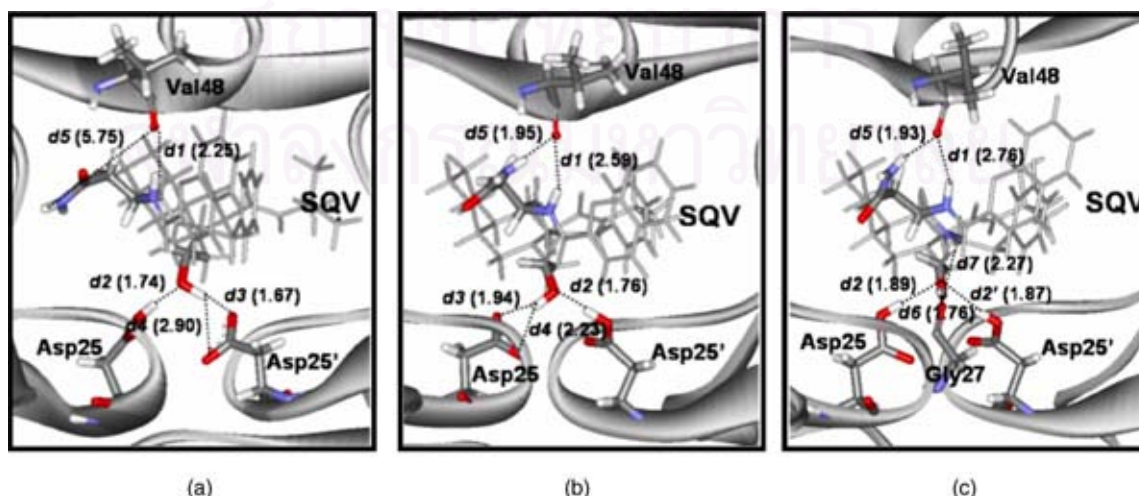


Fig. 7. Detected hydrogen bonds (dash line) in the HIV-1 PR enzyme–saquinavir complex in (a) Mono-25, (b) Mono-25' and (c) Di-pro states, where the average hydrogen bond distances (H...A distance) in Å are numbered.

remaining hydrogen bond (*d1*) at the average distance of 2.25 Å is slightly stronger than those of 2.59 and 2.76 Å for the Mono-25' and Di-pro states, respectively. For the active site region of the two monoprotonated states, the OH functional group of saquinavir was found to form three hydrogen bonds (*d2–d4*) with the two catalytic residues in which, one hydrogen bond with the protonated and the other two with the unprotonated carbonyl group. An exception was found for the Di-pro state where, the unprotonated group is not available, therefore, the other bond was formed with Gly27 (*d6* and *d7* in Fig. 7c).

Furthermore, indirect binding between saquinavir and the three residues, Gly27, Asp29 and Asp30, of the HIV-1 PR via the hydrogen bonds through water molecules was detected and displayed in Fig. 8. In the Mono-25 state (Fig. 8a), the conformation of P2 subsite was found to rotate and indirectly bind to the three residues of enzyme. The oxygen atom and hydrogen atom of P2 subsite form hydrogen bonds with Asp29 and Gly27 via two water molecules and with Asp30 via one water molecule (Fig. 8a), respectively. Even though those hydrogen bonds were not detected for Mono-25'. However, the nitrogen atom of the NH₂ group of P2 subsite for Mono-25' was found to form a hydrogen bond with Asp30 via one water molecule (Fig. 8b), instead.

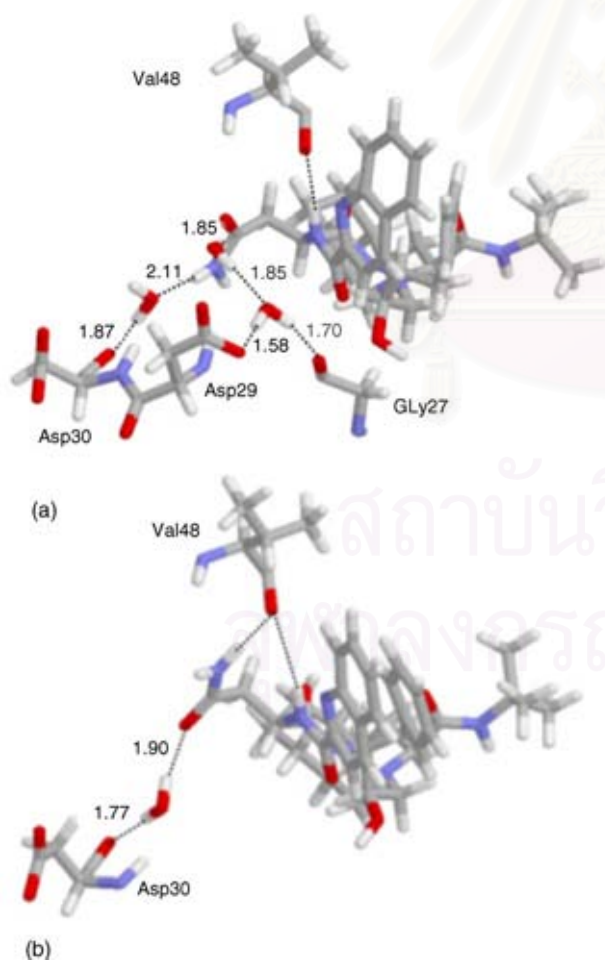


Fig. 8. Indirect binding between HIV-1 PR and saquinavir via hydrogen bonds (dash line) through water molecules in the (a) Mono25 and (b) Mono-25' states where the average distance (in Å) are numbered.

To evaluate the interaction between saquinavir and the catalytic residues, quantitatively, DFT calculations were performed. The total complexation energy (ΔE_{cpx}), interaction between saquinavir and catalytic dyad ($\Delta E_{\text{SQV}-(\text{Asp25}+\text{Asp25}')}$) and pair interactions as well as error due to three-body effects ($\Delta E_{3\text{bd}}$) are summarized in Table 1a (see method of calculation for more details).

The average RMSDs of the ASP25 and ASP25' residues for the three investigated systems with the deviations of less than 0.16 Å were summarized in Table 1b. This indicates reliability of the DFT energy calculations in which the MD snapshots can be represented by the average structure.

In terms of complexation energies, the obtained order of the ΔE_{cpx} is Mono-25 > Mono-25' > Di-pro, i.e., interaction energy of -37.3 kcal/mol for the Mono-25 state is significantly lower than those of -30.7 and -10.7 kcal/mol for the Mono-25' and Di-pro states, respectively. This evidence is in good agreement with that previously reported for saquinavir complexed with wild-type HIV-1 PR in which only Asp25 monoprotonated state is in the reasonable binding affinity [17,23,24]. In addition, quantum mechanics and molecular mechanics calculations [12] for the HIV-1 PR-74704 inhibitor complex leads also to the conclusion that the protonated state of the active site of the complex is monoprotonated state on Asp25.

In Table 1a, the total complexation energy (ΔE_{cpx}) can be separated into three terms of pair interaction and additional three-body correction terms. As expected, the main contribution to complexation energy is due to the electrostatic interactions between saquinavir and non-protonated state aspartic acids which amount to -19.2 and -18.0 kcal/mol for the Mono-25 and Mono-25', respectively. In addition, interactions with the two protonated residues are -8.0 kcal/mol for both Mono-25 and Mono-25'. Summation of the two terms of the pair interaction, $\Delta E_{\text{SQV}-\text{Asp25}}$ and $\Delta E_{\text{SQV}-\text{Asp25}'}$, yields the total interaction of -28.4 kcal/mol for Mono-25 and -28.0 kcal/mol for Mono-25'. These two values are almost consistent with the total interaction between saquinavir and the two aspartic residues calculated from $\Delta E_{\text{SQV}-(\text{Asp25}+\text{Asp25}')}$. The difference is due to many-body correction term, Asp25' and Asp25 were excluded in the calculation of the $\Delta E_{\text{SQV}-\text{Asp25}}$ and $\Delta E_{\text{SQV}-\text{Asp25}'}$, respectively. Taking into account the data summarizing above, no significant difference was found among the interaction between saquinavir and the catalytic residues of the HIV-1 PR of the two mono-protonated states.

Table 1a

Interaction between saquinavir (SQV) and the catalytic residues, Asp25 and Asp25', of HIV-1 PR for the three protonated states yielded from ab initio calculations (see text for more details)

Interaction energy (kcal/mol)	Mono-25	Mono-25'	Di-pro
ΔE_{cpx}	-37.3	-30.7	-10.7
$\Delta E_{\text{SQV}-(\text{Asp25}+\text{Asp25}')}$	-28.4	-28.0	-11.5
$\Delta E_{\text{Asp25}-\text{SQV}}$	-8.0	-18.0	-6.2
$\Delta E_{\text{Asp25}'-\text{SQV}}$	-19.2	-8.0	-7.9
$\Delta E_{\text{Asp25}-\text{Asp25}'}$	-8.8	-2.6	$+0.8$
$\Delta E_{3\text{bd}}$	-1.2	-2.1	$+2.6$

Table 1b

Average RMSDs as well as their deviations for the Asp25 and Asp25' of the Mono-25, Mono-25' and Di-Pro systems evaluated after equilibration, 400–1000 ps

Res. no.	Mono-25		Mono-25'		Di-pro	
	RMSD	Deviation	RMSD	Deviation	RMSD	Deviation
Asp25	0.56	0.10	0.38	0.08	0.18	0.05
Asp25'	0.37	0.06	0.47	0.12	0.35	0.16

Interest is focused on the Asp25–Asp25' interaction which was observed to be the key contribution to the ΔE_{cpx} , leading to a much lower ΔE_{cpx} in the Mono-25 than that of the Mono-25'. A significantly lower Asp25–Asp25' interaction energy of -8.8 kcal/mol for the Mono-25 than that of -2.6 kcal/mol for the Mono-25' suggests us to make a clear conclusion that protonation at the Asp25 leads to a better arrangement in the catalytic dyad. Interestingly, such rearrangement does not make any difference on the $\Delta E_{\text{SQV-Asp25}}$ or $\Delta E_{\text{SQV-Asp25'}}$ interactions (Table 1a) in comparison between the two monoprotonated residues. This finding also agrees well with the hydrogen bond formation between saquinavir and catalytic residues of the Mono-25 and Mono-25' where three hydrogen bonds were observed and the hydrogen bond distance for the two systems are almost equivalent, d_2 , d_3 and d_4 in Fig. 7.

Furthermore, rearrangement of the catalytic dyad was supposed to be the source of conformation change of the P2 subsite of saquinavir in the Mono-25 state (Fig. 3) and of the vanishing of one hydrogen bond between Val48 and saquinavir in the flap region (Fig. 7a).

The above data suggest us to conclude that interaction in the catalytic site should be used as criteria to enhance capability in drug designing and drug screening instead of using the total inhibitor/enzyme interaction. In addition, interaction between the inhibitor and the catalytic region of the enzyme is supposed to relate directly to the activity of the inhibitor in the catalytic process.

4. Conclusions

We have reported a result of 1 ns molecular dynamic (MD) simulation of three protonation states (Mono-25, Mono-25' and Di-pro) for G48V mutant HIV-1 PR complexed with the saquinavir inhibitor. Detailed analyses of structure and dynamics characters among the three protonation states of G48V mutant were given in terms of the flexibility, hydrogen bonding and conformational changes as well as binding energy between inhibitor and enzyme. The results indicate that the three complexes display a distinct dynamical behavior. A major structure change via complexation was found at P2 subsite of saquinavir which affected the hydrogen bonding for enzyme–inhibitor interaction and water molecule. The interaction data yielded from quantum chemical calculations supports the previous conclusion which state that the protonated state of the active site of the mutant-type HIV-1 PR complexed with saquinavir is a monoprotonated state on Asp25 (Mono-25). The energy data indicate also that protonation at the Asp25 lead to a

better arrangement in the catalytic dyad. In addition, conformation of the P2 subsite in the Mono-25 is consistent with that obtained from the X-ray structure of G48V/L90M mutant HIV-1 PR complexed with saquinavir.

Acknowledgments

We are very grateful to the Computational Chemistry Unit Cell and the Austrian-Thai Center for Computer Assisted Chemical Education and research (ATC), Department of Chemistry, Faculty of Science, Chulalongkorn University, Thailand for the use of computer resources. This work was supported by Thailand Research Fund Senior Scholar, Grant No. RTA468008.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2006.01.004.

References

- [1] CDC, First report of AIDS, MMWR 30 (1981) 250–252.
- [2] B. Mahalingam, J.M. Louis, J. Hung, R.W. Harrison, I.T. Weber, Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes, Proteins 43 (2001) 455–464.
- [3] T. Skalova, J. Hasek, J. Dohnalek, H. Petrokova, E. Buchtelova, J. Duskova, M. Soucek, P. Majer, T. Uhlíkova, J. Konvalinka, An ethylenamine inhibitor binds tightly to both wild type and mutant HIV-1 proteases. Structure and energy study, J. Med. Chem. 46 (2003) 1636–1644.
- [4] L. Hong, X.C. Zhang, J.A. Hartsuck, J. Tang, Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance, Protein Sci. 9 (2000) 1898–1904.
- [5] R.W. Shafer, P. Hsu, A.K. Patick, C. Craig, V. Brendel, Identification of biased amino acid substitution patterns in human immunodeficiency virus type 1 isolates from patients treated with protease inhibitors, J. Virol. 73 (1999) 6197–6202.
- [6] C.F. Shuman, P.-O. Markgren, M. Hämäläinen, U.H. Danielson, Elucidation of HIV-1 protease resistance by characterization of interaction kinetics between inhibitors and enzyme variants, Antiviral Res. 58 (2003) 235–242.
- [7] R. Kantor, W.J. Fessel, A.R. Zolopa, D. Israelski, N. Shulman, J.G. Montoya, M. Harbour, J.M. Schapiro, W.R. Shafer, Evolution of primary protease inhibitor resistance mutations during protease inhibitor salvage therapy, Antimicrob. Agents Chemother. 46 (2002) 1086–1092.
- [8] A. Wlodawer, J.W. Erickson, Structure-based inhibitors of HIV-1 protease, Annu. Rev. Biochem. 62 (1993) 543–585.
- [9] J. Ermolieff, X. Lin, J. Tang, Kinetic properties of saquinavir-resistant mutants of human immunodeficiency virus type 1 protease and their implications in drug resistance in vivo, Biochemistry 36 (1997) 12364–12370.
- [10] S. Piana, D. Sebastiani, P. Carloni, M. Parrinello, Ab initio molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site, J. Am. Chem. Soc. 123 (2001) 8730–8737.
- [11] R. Smith, I.M. Brereton, R.Y. Chai, S.B. Kent, Ionization states of the catalytic residues in HIV-1 protease, Nat. Struct. Biol. 3 (1996) 946–950.
- [12] K.-Y. Nam, B.H. Chang, C.K. Han, S.G. Ahn, K.T. No, Investigation of the protonated state of HIV-1 protease active site, Bull. Korean Chem. Soc. 24 (2003) 817–823.
- [13] Z. Zhu, D.I. Schuster, M.E. Tuckerman, Molecular dynamics study of the connection between flap closing and binding of fullerene-based inhibitors of the HIV-1 protease, Biochemistry 42 (2003) 1326–1333.

- [14] A. Krohn, S. Redshaw, J.C. Ritchie, B.J. Graves, M.H. Hatada, Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (*R*)-hydroxyethylamine isostere, *J. Med. Chem.* 34 (1991) 3340–3342.
- [15] W.A.K.P.A. Wang, Computational study of protein specificity: the molecular basis of HIV-1 protease drug resistance, *PNAS* 98 (2001) 14937–14942.
- [16] D.A. Case, J.C.D. Pearlman, T. Cheatham III, J. Wang, W. Ross, C. Simmerling, T.D.T. Merz, R. Stanton, A. Cheng, J. Vincent, M. Crowley, V.T.H. Gohlke, R. Radmer, Y. Duan, J. Pitera, I. Massova, G. Seibel, U.C.S.P. Weiner, P.A. Kollman, AMBER 7, University of California, San Francisco, CA, 2002.
- [17] K. Wittayanarakul, O. Aruksakunwong, S. Saen-oon, W. Chantratita, V. Parasuk, P. Sompornpisut, S. Hannongbua, Insights into saquinavir resistance in the G48V HIV-1 protease: quantum calculations and molecular dynamic simulations, *Biophys. J.* 88 (2005) 867–879.
- [18] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–935.
- [19] C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model, *J. Phys. Chem.* 97 (1993) 10269.
- [20] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, *J. Comput. Phys.* 23 (1977) 327–341.
- [21] H.J.C. Berendsen, J.P.M. Postma, W.F.V. Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81 (1984) 3684–3690.
- [22] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G.J.A.M. Zakrzewski Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian 98 Program (Revision A.7). Journal, 1998.
- [23] Y. Won, Binding free energy simulations of the HIV-1 protease and hydroxyethylene isostere inhibitors, *Bull. Korean Chem. Soc.* 21 (2000) 1207–1212.
- [24] K. Wittayanarakul, O. Aruksakunwong, P. Sompornpisut, V. Sanghiran-Lee, V. Parasuk, S. Pinitglang, S. Hannongbua, Structure, dynamics and solvation of HIV-1 protease/saquinavir complex in aqueous solution and their contributions to drug resistance: molecular dynamic simulations, *J. Chem. Inf. Model.* 45 (2005) 300–308.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Molecular Dynamic and Free Energy Studies of Primary Resistance Mutations in HIV-1 Protease–Ritonavir Complexes

Ornjira Aruksakunwong,[†] Peter Wolschann,[‡] Supot Hannongbua,[†] and Pornthep Sompornpisut^{*,†}

Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10400, Thailand, and
Institute for Theoretical Chemistry, University of Vienna, Vienna 1090, Austria

Received March 15, 2006

To understand the basis of drug resistance of the HIV-1 protease, molecular dynamic (MD) and free energy calculations of the wild-type and three primary resistance mutants, V82F, I84V, and V82F/I84V, of HIV-1 protease complexed with ritonavir were carried out. Analysis of the MD trajectories revealed overall structures of the protein and the hydrogen bonding of the catalytic residues to ritonavir were similar in all four complexes. Substantial differences were also found near the catalytic binding domain, of which the double mutant complex has the greatest impact on conformational changes of the protein and the inhibitor. The tip of the HIV-1 protease flap of the double mutant has the greater degree of opening with respect to that of the others. Additionally, the phenyl ring of Phe82 moves away from the binding pocket S1', and the conformational change of ritonavir subsite P1' consequently affects the cavity size of the protein and the conformational energy of the inhibitor. Calculations of binding free energy using the solvent continuum model were able to reproduce the same trend of the experimental inhibition constant. The results show that the resistance mutants require hydrophobic residues to maintain the interactions in the binding pocket. Changes of the cavity volume correlate well with free energy penalties due to the mutation and are responsible for the loss of drug susceptibility.

INTRODUCTION

The human immunodeficiency virus type 1 protease (HIV-1 PR) is one of the major targets for the treatment of AIDS. The enzyme is essential for processing the precursor viral polypeptide into a functional protein that is essential for viral assembly and the mature infectious virus.^{1–3} HIV-1 PR is a homodimeric protein consisting of two identical polypeptides of 99 amino acids. Many crystal structures of HIV-1 protease complexed with various inhibitors have been solved. The catalytic active site located at the dimer interface (Figure 1) is formed by the catalytic triad residues Asp25-(25')–Thr26(26')–Gly27(27').^{4–6} Additionally, the binding site, which is a large hydrophobic cavity, is stabilized by hydrophobic interactions and hydrogen bonding.

The first inhibitors of HIV-1 PR were based on the transition-state structure of the peptidic substrate by replacing the scissile peptide bond with the uncleavable isostere.⁷ Recently, there have been nine Food and Drug Administration (FDA)-approved protease inhibitors in clinical uses (indinavir, saquinavir, nelfinavir, ritonavir (RTV), lopinavir, atazanavir, fosamprenavir, amprenavir, and tipranavir). These HIV-1 PR inhibitors bind to the substrate binding site.⁸ Ritonavir is a very potent HIV-1 PR inhibitor of wild-type (wt) HIV-1 PR and has high oral bioavailability in humans. It is normally used in combination with other drugs.^{5,9}

Unfortunately, the virus evolves under the selective pressure of these inhibitors. They often become clinically resistant, during the first year of monotherapy. On the basis

of genotypic and phenotypic resistance testing, more than 87 mutations of HIV-1 PR have been observed in at least 49 positions. Many of them have shown cross-resistance to more than one drug.⁸ In particular, the double mutation V82F/I84V has shown high resistance against ritonavir, saquinavir, nelfinavir, indinavir, and amprenavir.¹⁰ The single mutations, V82F and I84V, show major determinants of phenotypic and clinical resistance to ritonavir. These two mutations were classified as primary as they directly confer reduced drug susceptibility. Three-dimensional structures of the V82F/I84V and V82F mutants complexed with symmetric protease inhibitors have been determined by X-ray crystallography.¹¹ Nevertheless, the crystal structure of the mutant–ritonavir complex is not yet known. Experimental thermodynamic studies of the binding of the drug to the wild-type and to the V82F/I84V HIV-1 protease were investigated by Todd et al.¹² The study indicated that the V82F/I84V mutant shows a decrease of 700-fold of the binding affinity to ritonavir. Additionally, the V82F/I84V mutant reduced dimer stability relative to the autolysis-resistant mutant at pH 7.0. The reduction in drug affinity may be due to the combined effects of mutations on both dimer stability and inhibitor binding.¹³ The position of the mutated residues is located at the edges of the active site. It makes direct contact with the inhibitors, particularly, the subsites P1 and P1'. Substitution between hydrophobic residues with different bulkinesses of the side chain may distort the wild-type geometry without changing its polarity or chemical composition. Therefore, the mutations represent an excellent candidate for a study of the structure in the binding sites and their changes leading to drug resistance. A structural study of the protein will provide a

* Corresponding author tel.: 662-2187603; fax: 662-2187603; e-mail: pornthep.s@chula.ac.th.

[†] Chulalongkorn University.

[‡] University of Vienna.

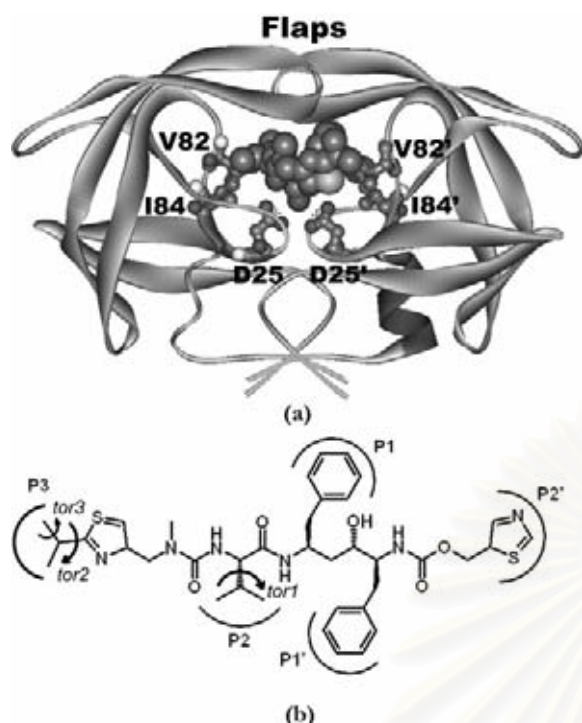


Figure 1. Schematic representation of wild-type HIV-1 protease complex (a) and ritonavir (b). In a, the catalytic residues Asp25/Asp25' and the mutated residues V82/V82' and I84/I84' are shown by the ball-and-stick model, and ritonavir is shown by the space-filled model. The inhibitor side chains denoted by P1, P2, P3, P1', and P2' correspond respectively to the substrate binding pockets denoted by S1, S2, S3, S1', and S2'.

basis for understanding molecular mechanisms of HIV drug resistance, which will be useful for designing anti-HIV inhibitors to combat AIDS.

In this study, we investigated structural and dynamics properties of the primary HIV-1 resistance against ritonavir using molecular dynamic (MD) simulations. The simulations were carried out for the wild-type, V82F, I84V, and V82F/I84V HIV-1 protease complexed with ritonavir in an explicit aqueous solution. The MD approach has provided insightful information on the enzyme–substrate interactions and binding conformations. In addition, the free energy of binding was performed using the molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) approach. The MM/PBSA method offers an efficient computation in calculating the binding free energy of biomolecular systems. It has been extensively used to explore many receptor–drug complexes.

METHODS

Starting Structure. The structures of the HIV-1 protease–RTV complex include the wt, two single mutants (V82F and I84V), and a double mutant (V82F/I84V). The 1.8 Å resolution crystallographic structure of the wt HIV-1 PR–RTV complex (PDB entry 1HXW) was used as the initial structure.⁵ All missing atoms and hydrogens of the enzyme were added using the LEaP module in the AMBER 7 package.¹⁴ Because the X-ray structure of the mutants is unknown, the comparative models of V82F, I84V, and V82F/I84V were therefore constructed using the wt model as the template. According to the pK_a calculation, the ionization states of all ionizable residues computed by the University of Houston Brownian Dynamics program¹⁵ were assigned

to their common charge, that is, Glu⁻¹, Asp⁻¹, Lys⁻¹, and Arg⁺¹, except for the catalytic Asp25.^{16,17} The single protonation at the Asp25 residue has experimentally and theoretically been observed for a number of the HIV-1 protease–inhibitor complexes, particularly, the inhibitor containing the hydroxyl–ethylene isostere.^{18,19} In this study, one of the two aspartyl residues Asp25/25' of the enzyme was assigned as a neutral state. Thus, the four simulated systems were in the monoprotection state.

The force-field parameters for RTV, which are not available in the AMBER package, were obtained as follows. Hydrogen atoms were added to the X-ray structure (1HXW) containing the RTV atomic coordinates. Then, partial geometric optimization of the hydrogen was performed at the Hartree–Fock level with the 6-31G** basis function using the Gaussian 98 program.²⁰ Subsequently, electrostatic potentials were generated on the optimized structure. Atomic charges of the inhibitor were calculated by the RESP fitting method.²¹ Partial atomic charges and force-field parameters for the inhibitor were generated by the Antechamber suite.²²

Each complex was immersed in a 10 Å radius of the TIP3P water model.²³ The crystallographic water molecules were kept in the system. Counterions were added to neutralize the system using the LEaP module. In all systems, the total number of water molecules in the $83 \times 72 \times 64 \text{ \AA}^3$ periodic box was 9027 molecules.

Molecular Dynamics Simulations. Energy-minimization and MD simulations were performed using the SANDER module of AMBER 7.0¹⁴ with the Cornell force field.²⁴ First, energy minimization was applied for water molecules using 500 steps of the steepest descents and then 500 steps of the conjugate gradients. Subsequently, the whole systems were subjected to energy minimization with 1000 steps of the steepest descents and 1500 steps of the conjugate gradients.

The MD simulation was performed employing the periodic boundary condition with the NPT ensemble. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and pressure of the systems.²⁵ The SHAKE algorithm²⁶ was employed to constrain all bonds involving hydrogens. A simulation time step of 2 fs was used. All MD simulations were run with a 12 Å residue-based cutoff for nonbonded interactions, and the particle-mesh Ewald method was used for an adequate treatment of long-range electrostatic interactions.²⁷

The MD simulation consists of thermalization and equilibration. During the 0–60 ps timeframe, the temperature of the system increased from 0 to 298 K using constant number of particles, volume, and temperature (NVT) conditions. Then, the temperature was maintained at 298 K during equilibration employing constant number of particles, pressure, and temperature (NPT) conditions. The reference pressure and temperature were set to 1 atm and 298 K, respectively. The MD trajectories were collected every 0.1 ps. The convergence of energies, temperatures, and pressures of the systems, and the atomic root-mean-square deviation of the enzyme and the inhibitor (RMSD), were used to verify the stability of the systems. The series of snapshots between 1.5 and 3 ns of the equilibrium phase was used for free energy calculations and structure evaluation.

Calculation of the Binding Free Energy (ΔG_{bind}). In general, the binding free energy of a protein–ligand complex (ΔG_{bind}) is defined as

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{receptor}} + G_{\text{ligand}}) \quad (1)$$

where G_{complex} , G_{receptor} , and G_{ligand} are the free energies of the complex, the protein, and the ligand, respectively.

For each system, the free energy can be estimated in terms of the molecular mechanic potential energy (E^{MM}), the solvation free energy (G_{solv}), and the entropic contribution (TS) as follows:

$$G_{\text{bind}} = E^{\text{MM}} - \text{TS} + G_{\text{solv}} \quad (2)$$

$$E^{\text{MM}} = E_{\text{internal}} + E_{\text{elec}} + E_{\text{vdW}} \quad (3)$$

$$G_{\text{solv}} = G_{\text{elec,solv}} + G_{\text{nonpolar,solv}} \quad (4)$$

where E_{internal} includes the bond, angle, and torsional angle energies; E_{elec} and E_{vdW} are intermolecular electrostatic and vdW energies, respectively. Therefore, E^{MM} is associated with the enthalpic changes in the gas phase upon binding. G_{solv} is composed of electrostatic $G_{\text{elec,solv}}$ and nonpolar $G_{\text{nonpolar,solv}}$ contributions to solvation. The former can be obtained by solving the Poisson–Boltzmann equation,²⁸ while the latter is approximated on the basis of the solvent-accessible surface area using the hard-sphere atomic model.²⁹

The calculation of ΔG_{bind} values of HIV-1 protease–ritonavir complexes was carried out using the MM/PBSA module in AMBER 7. E^{MM} was computed with no cutoff. The electrostatic solvation energy was computed using the Poisson–Boltzmann solver implemented in the Delphi 4 program.³⁰ The grid spacing was set to $1/3$ Å. The Debye–Huckel potentials were employed. The dielectric constants of protein and water were set to 4.0 and 80.0, respectively. The atomic charges of the protein were taken from the Cornell force field.²⁴ The nonpolar solvation contribution is estimated as

$$G_{\text{nonpolar,solv}}^S = \gamma A + b \quad (5)$$

where A is the solvent-accessible surface area and the solvation parameters, γ and b , are 0.00542 kcal/mol Å² and 0.92 kcal/mol, respectively. The probe radius of the solvent was set to 1.4 Å. The atomic radii of the solute were taken from the PARSE parameter set.²⁹

Considering the entropic effects, structural reorganization and solvent entropy are important contributions in enzyme–ligand bindings. However, a number of studies published by Kollman and Kuhn and Warshel’s group showed that this effect is minor for some protein complex systems.^{31,32} In the case of HIV-1 protease, a thermodynamic study of four FDA-approved drugs, including indinavir, nelfinavir, saquinavir, and ritonavir, showed that the experimental $T\Delta S$ values of V82F/I84V are ~ 0.7 kcal/mol less favorable than that of the wild type.¹² On the basis of the information, the entropy difference between the wild type and the three mutants complexed with the same inhibitor should be very small. Normal-mode analysis (NMA) is useful to estimate entropic changes of the solute molecule, but the calculation is considered problematic and time-consuming. In addition, this approach does not take the solvent entropy into consideration. For the above reasons, NMA would not greatly improve the correlation between the experimental K_i and the binding free

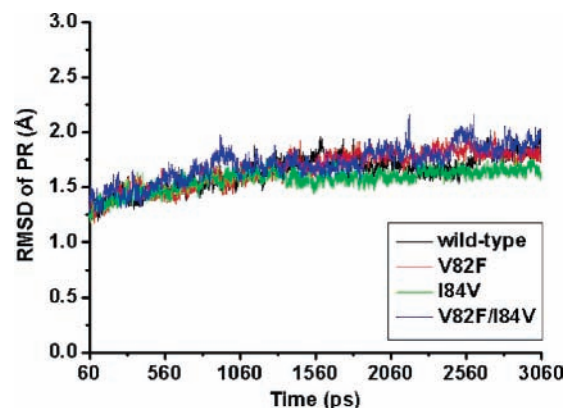


Figure 2. Plot of RMSDs versus the time for the wild-type, V82F, I84V, and V82F/I84V complexes. The obtained RMSD was computed using the structure at $t = 0$ as a reference.

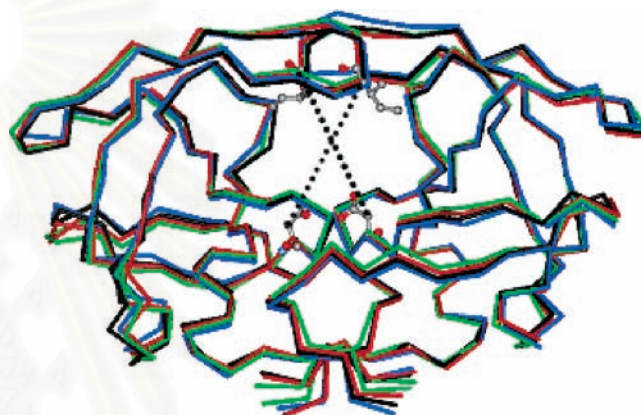


Figure 3. C α trace of the wild type (black) and V82F (red), I84V (green), and V82F/I84V (blue) mutants. Asp25, Asp25', Ile50, and Ile50' are displayed in the ball-and-stick. Dashed lines demonstrate the measured distances of C α –Ile50/C β –Asp25 and of C α –Ile50'/C β –Asp25'. A small panel shows the average RMSD per residue of the mutant structures compared to the wt structure.

energy. Thus, the calculation of the solute $T\Delta S$ term has been omitted in the study.

RESULTS AND DISCUSSION

The 3 ns MD trajectories of four systems consisting of the wild-type, V82F, I84V, and V82F/I84V HIV-1 PR complexed with ritonavir were generated. The atomic RMSD of the protein structure is shown in Figure 2. The small RMSD fluctuation and the convergence of energies, temperatures, and pressures of the systems during a course of the simulations indicated well-behaved systems.

Structural Features of the HIV-1 Protease. Structures of the wt and all three mutants taken from a snapshot of the MD trajectories are almost identical (Figure 3). From a RMSD comparison, the global tertiary structures of all four systems are similar to each other. The structure of the catalytic triad residues of all three mutants is essentially identical to that of the wt with RMSD < 0.5 Å (data not shown). However, a large RMSD difference can be seen on the HIV flaps of chain B (49'–54') in the V82F/I84V mutant.

The degree of the flap change is simply demonstrated by a measure of the inter-residue distances of C α –Ile50/C β –Asp25 (chain A) and of C α –Ile50'/C β –Asp25' (chain B). The distance separation was calculated by taking structures from the MD trajectory and then statistically analyzed as a

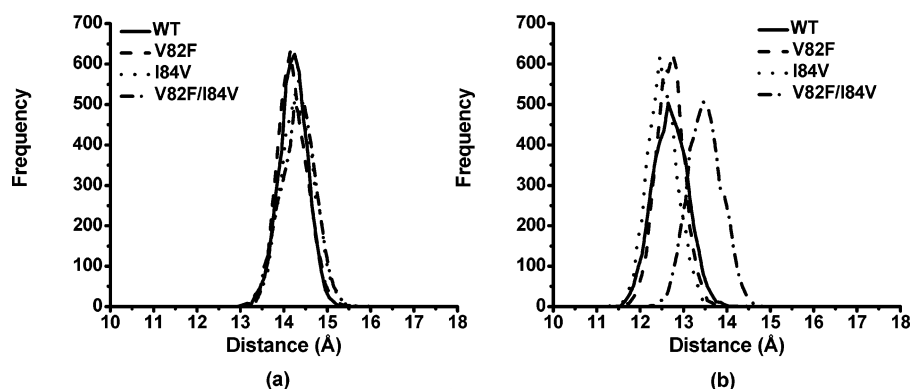


Figure 4. Distributions of the $C\alpha$ -Ile50/ $C\beta$ -Asp25 (a) and the $C\alpha$ -Ile50'/ $C\beta$ -Asp25' (b) distances (for details, see text and labels in Figure 3) sampled during the 1.5–3 ns timeframe.

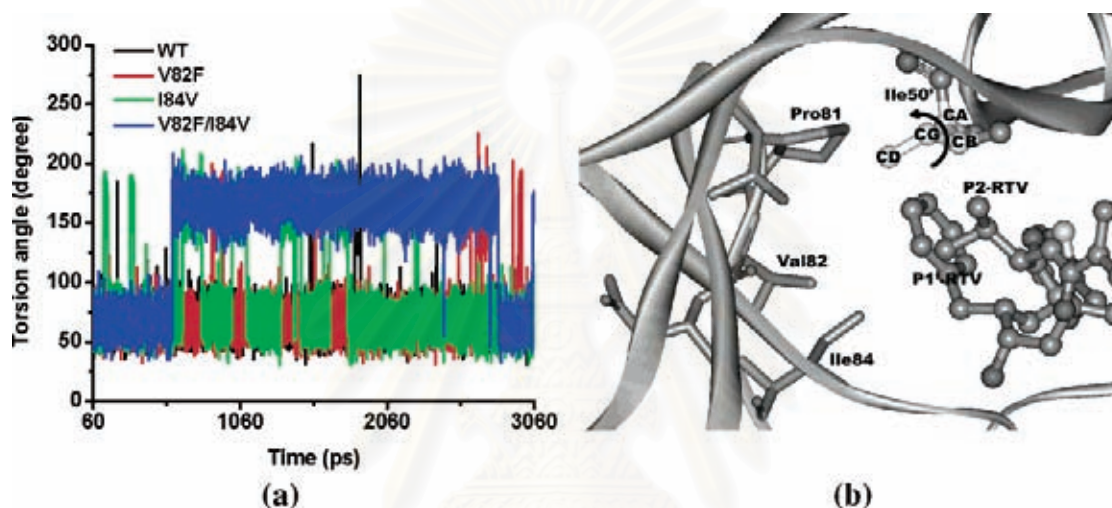


Figure 5. (a) Trajectory of Ile50' torsion defined by the $C\alpha$ - $C\beta$ bond (arrow in b). (b) Ile50', ritonavir, and residues 81–84 of the wild-type HIV-1 PR.

population frequency. From Figure 4, distance distribution peaks of the wt and the mutants are mostly overlying for chain A, while those for chain B are not (Figure 4). A slight shift of the inter-residue distance of chain B was observed for the single mutants, but a greater change was found for the double mutant within a range of 1.0–1.5 Å. This suggests the flap structure may have changed because of the mutation. The most shifted distance was found for the double mutant, and the change of the flap structure of chain A is not equal to that of chain B. It has been known that the inhibitor is not symmetric. It is likely that the flap tip of the double mutant has a greater degree of opening than that of the others. This may relate to a loss of interactions of the drug in the binding pocket of HIV-1 V82F/I84V protease as compared with that of the wt and the V82F and I84V mutants.

A number of MD studies published previously have reported large conformational changes of the flap.³³ The MD simulations of the apo wild type and the apo V82F/I84V mutant of HIV-1 protease observed more frequent and more rapid curling of the flap tip in the mutant than in the wild type. Moreover, the mutant's flap also opened farther than the wild type's, implying more flexibility of the mutant. The flexibility of the flap is highly dependent on the availability and type of ligands.

It appears that the Ile50' side chain of chain B is in close contact with P1' and P2' subsites of RTV and residues 81–84 of chain A (Figure 5b). One should note that the

interactions in this binding pocket are formed by the hydrophobic packing of side chains. Thus, the shape and size of the substituted nonpolar side chains of V82F and I84V are responsible for the binding of the drug. Detailed attention was focused on residues 50', 82, and 84 as well as on the inhibitor subsites. The trajectories of the side chain torsions χ_1 (Ile50') and χ_1 (Phe82) defined by the $C\alpha$ - $C\beta$ bond are shown in Figure 5. For Ile50', the majority of the χ_1 torsion is in a range of 50–90° for the wt and V82F proteases. This suggested no substantial change of Ile50' side chain. On the other hand, the χ_1 (Ile50') in the I84V and V82F/I84V mutants produces two distinct values with a difference of $\sim 120^\circ$, implying two preferable conformations of Ile50'. The two phenyl rings of Phe82 and Phe82' in the V82F/I84V mutant have similar χ_1 torsions, while they are significantly different for the V82F mutant. This discrepant side chain orientation of Phe82 implies different interactions between the V82F and the V82F/I84V mutants and the inhibitor.

Conformational Flexibility of Ritonavir. From the simulations, a fluctuation of the RTV structure compared to the X-ray structure exhibits RMSDs of 1.25 ± 0.24 , 1.43 ± 0.21 , 1.28 ± 0.18 , and 1.78 ± 0.13 Å for the wt, V82F, I84V, and V82F/I84V, respectively. The RMSD fluctuation of RTV subsites P1, P2, P3, P1', and P2' is shown in Figure 6.

The RMSD plots show that the phenyl ring of P1 (green) and P1' (blue) and the thiazol ring of P2' (sky-blue) do not

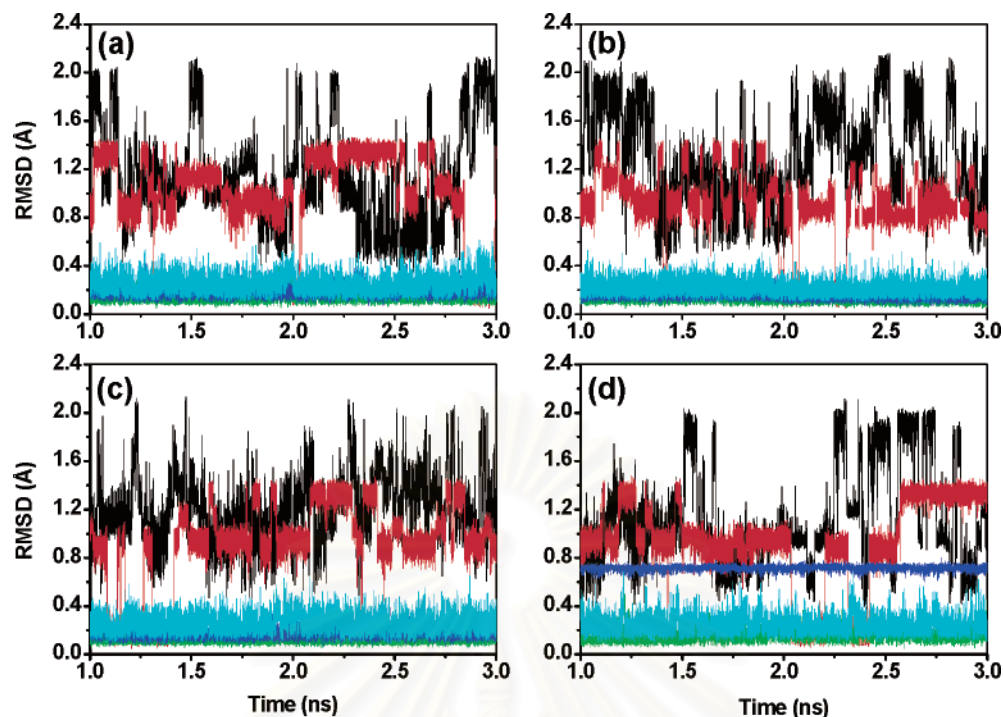


Figure 6. RMSDs of each ritonavir subsite with respect to the starting structure versus the simulation time for the wild type (a) and V82F (b), I84V (c), and V82F/I84V (d) mutants. P3, P2, P1, P1', and P2' subsites are given in black, red, green, blue, and sky-blue lines, respectively.

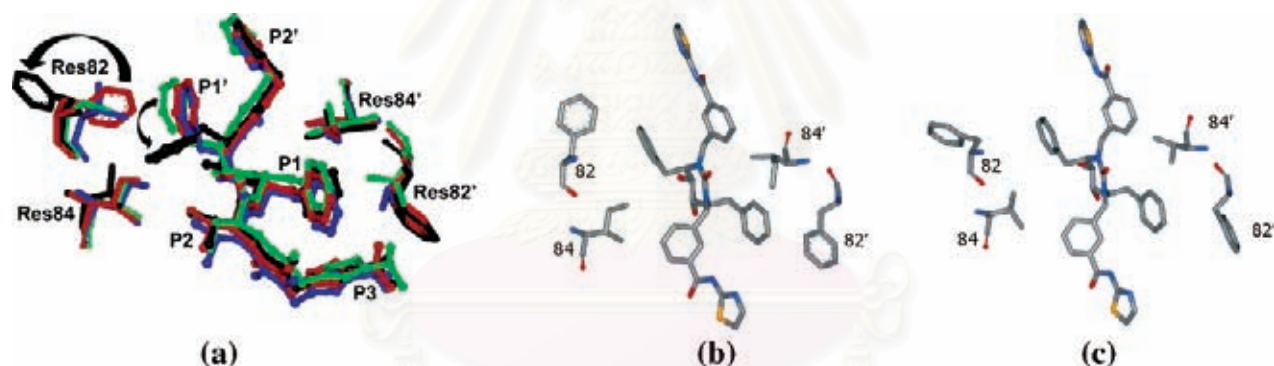


Figure 7. Structure comparison of residues 82 and 84 and ritonavir in HIV-1 protease: (a) the MD structures of the wild type (blue), V82F (red), I84V (green), and V82F/I84V (black); (b and c) the X-ray structures of the V82F (1BV7) and the V82F/I84V mutants (1BWA), respectively. Arrows show different orientations of the phenyl ring of the P1' and of Phe82 between the single and double mutants.

change their conformation significantly during the simulations, except for the P1' of the V82F/I84V system. This indicated that the inhibitor side chains of the bound RTV in the binding pocket of the enzyme are considerably rigid. Meanwhile, P3 (black) and P2 (red) side chains show a larger degree of atomic displacements. These are due to the rotation of the two methyl groups of the isopropyl chain. Thus, the flexibility of these subsites is relatively higher upon binding. It appears that the most significant difference between the wt and the mutants was found at the P1' subsite of RTV. The RMSD change in the double mutant is higher than that in the wt and the two single mutants. This indicates that there is a greater degree of conformational change of the P1' phenyl ring taking place on the V82F/I84V complex only.

Conformational changes of the P1' phenyl ring in the double mutant are illustrated in Figure 7a. The P1' rings in the wt and in the two single mutants remain largely similar to each other, while that in the double mutant is oriented differently. Additionally, the ring orientation of Phe82 of

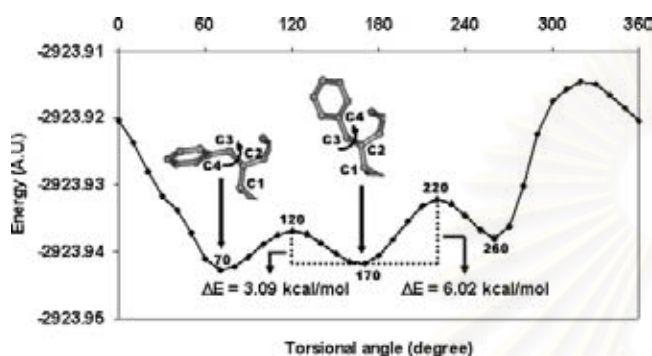
V82F is apparently different from that of the V82F/I84V mutant. Nevertheless, an arrangement of Phe82 side chain was found to be similar to the crystal structures of the single and double mutations of HIV-1 protease (Figure 7b,c).

It should be noted that the ring rearrangement of the P1' subsite in the V82F/I84V complex has consequently affected the binding pocket S1' of the protein and the conformational energy of the inhibitor. In the V82F/I84V complex, the side-chain rearrangements of Phe82 coupled with the smaller side chain of Val84 create a larger gap of the S1' pocket (Figure 7). The volumes of the binding cavity calculated using the CAST program³⁴ are shown in Table 1. The order of the cavity volumes for all four complexes is V82F/I84V > I84V > wild type > V82F. Apparently, the mutation in V82F/I84V has the largest impact on the volume change in the binding cavity of the protein. This may change direct contact between RTV and the residues in the binding pocket S1'. On the contrary, the inhibitor-binding site of the single V82F mutant becomes more restricted as the size of the cavity shrinks because of the bulkiness of the group. This

Table 1. Calculated Volume in the Binding Cavity of the Protease Complexes

complexes	cavity ^a (Å ³) volume (Å ³)	ΔV^b (Å ³)
wild type	1257 ± 80	
V82F	1241 ± 84	-16
I84V	1317 ± 59	60
V82F/I84V	1472 ± 55	215

^a Because the cavity volume was obtained from the MD trajectory, the average value and standard deviation are reported. ^b Change of the cavity volumes (V) calculated from $V(\text{wild type}) - V(\text{mutant})$.

**Figure 8.** HF/6-31G potential energy versus RTV conformations defined by the C2–C3 torsion of the P1' side chain.

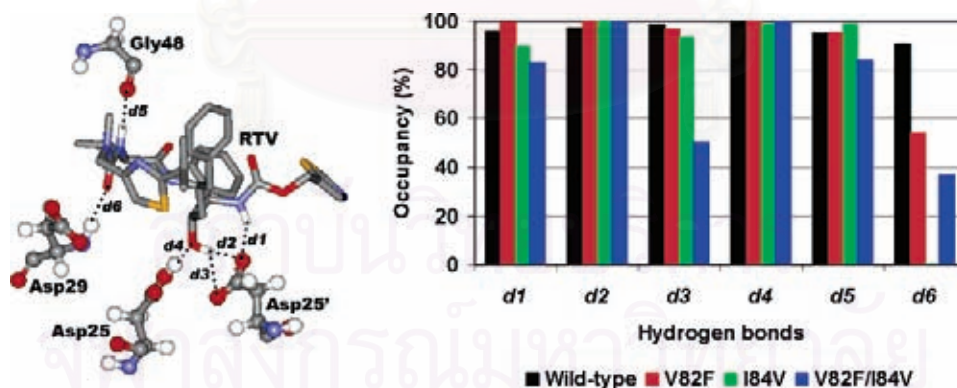
possibly causes the P1' ring to remain unchanged in the V82F complex.

A question may arise whether residue 82' in another subunit of the double mutant and the P1 ring of the inhibitor undergo similar conformational changes. The fact is that RTV is an asymmetric inhibitor; thus, the distinct binding mode in the second monomer is not uncommon. From the simulations, the substituted residues of the protein and the P1' change of the inhibitor did not significantly alter the

conformation of Phe82' and the P1 ring. In conclusion, the conformation perturbation of residue 82 took place only on one subunit of the HIV-1 protease.

The conformational energy of RTV was investigated by calculating the intramolecular potential energy with various conformations of the P1' subsite. We began with a starting structure of RTV taken from the MD snapshot, and then, the C1–C2–C3–C4 torsion defined for the conformation of the P1' phenyl ring (Figure 8) was subjected to change from 0 to 360°. The single-point energy calculations to the whole RTV structure were performed on each generated conformation at the HF/6-31G level. A plot of the calculated potential energy versus RTV conformations at the scanned torsional angles (Figure 8) indicates three minimum potential energies with torsions of 70, 170, and 260°. The results are in good agreement with the P1' torsions determined from the MD simulations (Figure 7). The lowest total energy corresponds to the RTV conformer in the V82F/I84V complex, while the minimum at 170° is close to the conformer found in the wt and the two single mutants. It should be noted that in the simulations the starting RTV conformer for all systems was ~170° corresponding to the second energy minimum. The rearrangement of the P1' ring in the double mutant can overcome the energy barrier of ~3 kcal/mol to reach an energetically more favorable state; although, the lowest energy of the conformation is about 0.5 kcal/mol lower than the second minimum. This conformation change might relax the strain around the P1' subsite due to the steric conflict of the V82F/I84V mutant.

Ritonavir–HIV-1 Protease Binding Energy. Table 2 shows the estimated binding free energy (ΔG_{bind}) averaged over the trajectory for each complex. Our calculations show a good correlation with experimental results. A net loss of ΔG_{bind} in the mutant complexes with respect to the wt corresponds to a decrease in the inhibition potency. The

**Figure 9.** Hydrogen bonding (d1–d6) between HIV-1 PR and ritonavir and the percent occurrence of the corresponding hydrogen bonds calculated from the MD trajectories.**Table 2.** Averaged Energy Contributions (kcal/mol) to the Binding Free Energy for Wild-Type and Mutant HIV-1 Protease–Ritonavir Complexes

energy	wild-type	V82F	I84V	V82F/I84V
ΔE_{vdw}^a	-75.83 ± 3.81 ^b	-76.28 ± 4.91	-69.81 ± 4.44	-68.89 ± 4.00
ΔE_{elec}^c	-52.59 ± 4.76	-49.35 ± 5.09	-47.93 ± 4.67	-44.07 ± 5.26
$\Delta G_{\text{elec,solv}}^d$	23.63 ± 1.67	24.52 ± 1.96	23.40 ± 1.63	20.89 ± 1.05
$\Delta G_{\text{nonpolar,solv}}^e$	-6.29 ± 0.18	-6.31 ± 0.28	-5.93 ± 0.36	-5.82 ± 0.41
ΔG_{bind}^f	-111.08 ± 4.89	-107.43 ± 6.11	-100.27 ± 4.21	-97.90 ± 2.40
$K_i^{\text{experiment}}(\text{nM})^g$	0.17	0.14	1.90	119

^a Van der Waals energy. ^b Standard error of the mean. ^c Electrostatic energy. ^d Electrostatic contribution to solvation. ^e Nonpolar contribution to solvation. ^f Binding free energy in the absence of entropic contribution. ^g The experimental inhibition constant.

calculated difference in free energy of binding between the wt and each mutant is -3.65 kcal/mol for V82F, -10.81 kcal/mol for I84V, and -13.18 kcal/mol for V82F/I84V. This suggested that the binding affinity of RTV is strongest for the native protein and becomes less favorable in the mutants. Particularly, biochemical data indicated that RTV has about 700 times lower of an inhibition constant (K_i) value with the double mutant. All complexes are correctly rank-ordered, although the magnitudes of ΔG_{bind} are overestimated.

A detailed analysis suggested that there is a decrease in both van der Waals and electrostatic contributions in the mutants upon binding. It appears that for all complexes van der Waals interactions are the largest contribution to the binding free energy ($\sim 70\%$). This explains an essence of the substituted hydrophobic residues in the primary resistance mutation. In addition, it is in accordance with the fact that the binding pocket of the HIV-1 protease is considerably hydrophobic. However, the magnitude of ΔE_{elec} (Table 2) should also be taken into consideration. Differences in the electrostatic components between the wt and each mutant show an energy loss of -1.10 kcal/mol for V82F, -4.41 kcal/mol for I84V, and -8.19 kcal/mol for V82F/I84V, suggesting that unfavorable interactions arose from the mutations. In terms of the magnitude, the loss of the nonbonded energy changes correlates well with the volume change of the interior cavity (ΔV in Table 1) induced by the mutation.

Effect of Mutation to the Binding in the Catalytic Domain. Residues in the binding domain of HIV-1 PR such as the catalytic triad and the flap region allow for several important hydrogen bonds to the main-chain RTV to be formed. The hydrogen bonding was determined using the following criteria: (i) a distance between donor and acceptor heavy atoms ≤ 3.5 Å, (ii) an angle of donor-H-acceptor $\geq 120^\circ$, and (iii) a percentage of occurrence in the simulations. From the histogram in Figure 9, six hydrogen bonds of the wt complex exist over a 90% occurrence frequency, whereas the single and the double mutants have five and four hydrogen bonds, respectively. Hydrogen bonding between the catalytic residues Asp25/Asp25' and the OH of RTV denoted by d2, d3, and d4 is apparently strong. Particularly, d2 and d4 fluctuate in a range of 2.6–2.8 Å for the distance and 164 – 172° for the angle in all four MD trajectories. Moreover, hydrogen bonds d1 and d5 have the same magnitude in terms of the observed quantities. The significant difference may be depicted by d6, which is a hydrogen bond between D29 and the P3 subsite. The hydrogen bond is considerably weak and appears only in the wt complex.

CONCLUSION

We have studied the structure, dynamics, and free energies of binding of wild-type and primary resistance mutants, V82F, I84V, and V82F/I84V, HIV-1 protease complexed with ritonavir. Overall, structures of the enzyme and the hydrogen bonding of the catalytic residues to ritonavir are similar in all four complexes. On the basis of the MD data, the tip of the HIV-1 protease flap of the double mutant has the greater degree of opening with respect to that of the wild type and the two single mutants. The greatest conformational changes in the catalytic binding domain also took place on

the double mutant complex. The substituted residues of the resistance mutants require a hydrophobic property to maintain the interactions in the binding pocket. The computed free energy of binding using the solvent continuum model is in good agreement with the experimental inhibition constant. Those changes induced by the mutations are free of energy penalties for the binding of the drug and are responsible for the loss of the drug's susceptibility.

ACKNOWLEDGMENT

Financial support by the Thailand Research Fund and the generous supply of computer time by Computational Chemistry Unit Cell (CCUC) and the Faculty Facility are gratefully acknowledged.

REFERENCES AND NOTES

- (1) Hellen, C. U.; Krausslich, H. G.; Wimmer, E. Proteolytic Processing of Polyproteins in the Replication of RNA Viruses. *Biochemistry* **1989**, *28*, 9881–9890.
- (2) Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A.; Scolnick, E. M.; Sigal, I. S. Active Human Immunodeficiency Virus Protease Is Required for Viral Infectivity. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686–4690.
- (3) Swanstrom, R.; Erona, J. Human Immunodeficiency Virus Type-1 Protease Inhibitors: Therapeutic Successes and Failures, Suppression and Resistance. *Pharmacol. Ther.* **2000**, *86*, 145–170.
- (4) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. Conserved Folding in Retroviral Proteases: Crystal Structure of a Synthetic HIV-1 Protease. *Science* **1989**, *245*, 616–621.
- (5) Kempf, D. J.; Marsh, K. C.; Denissen, J. F.; McDonald, E.; Vasavanonda, S.; Flentge, C. A.; Green, B. E.; Fino, L.; Park, C. H.; Kong, X. P. ABT-538 is a Potent Inhibitor of Human Immunodeficiency Virus Protease and Has High Oral Bioavailability in Humans. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2484–2488.
- (6) Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A. Substrate Shape Determines Specificity of Recognition for HIV-1 Protease: Analysis of Crystal Structures of Six Substrate Complexes. *Structure* **2002**, *10*, 369–381.
- (7) Lepsik, M.; Kriz, Z.; Havlas, Z. Efficiency of a Second-Generation HIV-1 Protease Inhibitor Studied by Molecular Dynamics and Absolute Binding Free Energy Calculations. *Proteins* **2004**, *57*, 279–293.
- (8) Ohtaka, H.; Freire, E. Adaptive Inhibitors of the HIV-1 Protease. *Prog. Biophys. Mol. Biol.* **2005**, *88*, 193–208.
- (9) Nair, A. C.; Bonin, I.; Tossi, A.; Wels, W. J.; Miertus, S. Computational Studies of the Resistance Patterns of Mutant HIV-1 Aspartic Proteases towards ABT-538 (Ritonavir) and Design of New Derivatives. *J. Mol. Graphics Modell.* **2002**, *21*, 171–179.
- (10) Velazquez-Campoy, A.; Kiso, Y.; Freire, E. The Binding Energetics of First- and Second-Generation HIV-1 Protease Inhibitors: Implications for Drug Design. *Arch. Biochem. Biophys.* **2001**, *390*, 169–175.
- (11) Ala, P. J.; Huston, E. E.; Klabe, R. M.; Jadhav, P. K.; Lam, P. Y.; Chang, C. H. Counteracting HIV-1 Protease Drug Resistance: Structural Analysis of Mutant Proteases Complexed with XV638 and SD146, Cyclic Urea Amides with Broad Specificities. *Biochemistry* **1998**, *37*, 15042–15049.
- (12) Todd, M. J.; Luque, I.; Velazquez-Campoy, A.; Freire, E. Thermodynamic Basis of Resistance to HIV-1 Protease Inhibition: Calorimetric Analysis of the V82F/I84V Active Site Resistant Mutant. *Biochemistry* **2000**, *39*, 11876–11883.
- (13) Xie, D.; Gulnik, S.; Gustchina, E.; Yu, B.; Shao, W.; Qoronfleh, W.; Nathan, A.; Erickson, J. W. Drug Resistance Mutations Can Effect Dimer Stability of HIV-1 Protease at Neutral pH. *Protein Sci.* **1999**, *8*, 1702–1707.
- (14) Case, D. A.; Pearlman, J. W.; Caldwell, T. E.; Cheatham, J., III; Wang, W. S.; Ross, C. L.; Simmerling, T. A.; Darden, K. M.; Merz, R. V.; Stanton, A. L.; Cheng, J. J.; Vincent, M.; Crowley, V.; Tsui, H.; Gohlke, R. J.; Radmer, Y.; Duan, J.; Pitera, I.; Massova, G. L.; Seibel, U. C.; Singh, P. K.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, CA, 2002.
- (15) Davis, M. E.; Madura, J. D.; Luty, B. A.; Mccammon, J. A. Electrostatics and Diffusion of Molecules in Solution – Simulations

- with the University-of-Houston-Brownian Dynamics Program. *Comput. Phys. Commun.* **1991**, *62*, 187–197.
- (16) Wittayanarakul, K.; Aruksakunwong, O.; Sompornpisut, P.; Sanghiran-Lee, V.; Parasuk, V.; Pinitglang, S.; Hannongbua, S. Structure, Dynamics and Solvation of HIV-1 Protease/Saquinavir Complex in Aqueous Solution and Their Contributions to Drug Resistance: Molecular Dynamic Simulations. *J. Chem. Inf. Model.* **2005**, *45*, 300–308.
- (17) Wittayanarakul, K.; Aruksakunwong, O.; Saen-oon, S.; Chantratita, W.; Parasuk, V.; Sompornpisut, P.; Hannongbua, S. Insights into Saquinavir Resistance in the G48V HIV-1 Protease: Quantum Calculations and Molecular Dynamic Simulations. *Biophys. J.* **2005**, *88*, 867–879.
- (18) Nam, K. Y.; Chang, B. H.; Han, C. K.; Ahn, S. G.; No, K. T. Investigation of the Protonated State of HIV-1 Protease Active Site. *Bull. Korean Chem. Soc.* **2003**, *24*, 817–823.
- (19) Won, Y. Binding Free Energy Simulations of the HIV-1 Protease and Hydroxyethylene Isostere Inhibitors. *Bull. Korean Chem. Soc.* **2000**, *21*, 1207–1212.
- (20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11; Gaussian, Inc.: Pittsburgh, PA, 2001.
- (21) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of Resp Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.
- (22) Wang, J. M.; Wang, W.; Kollman, P. A. Antechamber: An Accessory Software Package for Molecular Mechanical Calculations. *Abstr. Pap. Am. Chem. Soc.* **2001**, *222*, U403.
- (23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (24) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (25) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (26) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (27) York, D. M.; Darden, T. A.; Pedersen, L. G. The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals – A Comparison of the Ewald and Truncated List Methods. *J. Chem. Phys.* **1993**, *99*, 8345–8348.
- (28) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the Electrostatic Potential of Molecules in Solution: Method and Error Assessment. *J. Comput. Chem.* **1987**, *9*, 327–335.
- (29) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- (30) Rocchia, W.; Alexov, E.; Honig, B. Extending the Applicability of the Nonlinear Poisson–Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions. *J. Phys. Chem. B* **2001**, *105*, 6507–6514.
- (31) Villa, J.; Strajbl, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. How Important Are Entropic Contributions to Enzyme Catalysis? *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11899–11904.
- (32) Kuhn, B.; Kollman, P. A. Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models. *J. Med. Chem.* **2000**, *43*, 3786–3791.
- (33) Perryman, A. L.; Lin, J. H.; McCammon, J. A. HIV-1 Protease Molecular Dynamics of a Wild-Type and of the V82F/184V Mutant: Possible Contributions to Drug Resistance and a Potential New Target Site for Drugs. *Protein Sci.* **2004**, *13*, 1108–1123.
- (34) Liang, J.; Edelsbrunner, H.; Woodward, C. Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design. *Protein Sci.* **1998**, *7*, 1884–1897.

CI060090C

VITAE

Name: Ornjira Aruksakunwong

Date of birth: July 15, 1979

Birth place: Rayong, Thailand

Education:

Secondary school: Rayongwitthayakhom School, Muang, Rayong, Thailand, 1998

Bachelor's degree: Chulalongkorn University, Bangkok, Thailand, 2002

Grants Received:

- TRF senior Scholar of The Thailand Research Fund, lead by Prof. Dr. Supot Hannongbua.
- North-South Dialogue Scholarship for 9 months (1/10/04 – 31/6/05) in Austria, working with Prof. Dr. Peter Wolschann, University of Vienna.

Publications:

1. Lee, V. S., Aruksakunwong, O., et.al. ScienceAsia 29 (2003): 181-188.
2. Wittayanarakul, K., Aruksakunwong, O., et.al. J. Bio. Phys. 88 (2005): 867-879.
3. Wittayanarakul, K., Aruksakunwong, O., et.al. J. Chem. Inf. Model. 45 (2005): 300-308.
4. Aruksakunwong, O., Wittayanarakul, K., et. al. JMGM. 25 (2006): 324–332.
5. Aruksakunwong, O., Hannongbua, S., et. al. J. Mol. Structure. 790 (2006): 174-182.
6. Aruksakunwong, O., Promsri, S., et. al. Current Computer-Aided Drug Design (2006) accepted.
7. Lee, V. S., Aruksakunwong, O., et. al. J. Mol. Structure: THEOCHEM. (2006) revised.
8. Lee, V. S., Aruksakunwong, O., et. al. J. Bio. Phys. (2006) revised.
9. Aruksakunwong, O., Wolschann, P., et. al. J. Chem. Inf. Model. (2006) accepted.
10. Saen-oon. S., Aruksakunwong, O., et. al. (2006) submitted.
11. Aruksakunwong, O., Malaisree, M., et. al. J. Bio. Phys. (2006) revised.

* 5 Oral and 5 poster presentations in the national and international conferences.