

ตรีศิวเอสเออาร์ และ โมเลกุลาร์ไดนามิกส์ซิมูเลชันของเอชไอวี-1 อินทิเกรต

และสารประกอบเชิงซ้อนกับตัวรับยั้ง



นางสาวณัฐฐานนตร นันทบุตร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาเคมี ภาควิชาเคมี

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2549

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย



4 5 7 3 8 1 1 2 2 3

3D-QSAR AND MOLECULAR DYNAMICS SIMULATIONS OF HIV-1 INTEGRASE
AND ITS COMPLEX WITH INHIBITORS

MISS NADTANET NUNTHABOOT

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

Copyright of Chulalongkorn University

492294

Thesis Title 3D-QSAR AND MOLECULAR DYNAMICS SIMULATIONS
 OF HIV-1 INTEGRASE AND ITS COMPLEX WITH
 INHIBITORS

By Miss Nadtanet Nunthaboot


Field of Study Chemistry

Thesis Advisor Associate Professor Sirirat Kokpol, Ph. D.

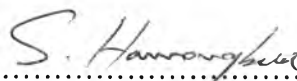
Thesis Co-advisor Professor Peter Wolschann, Dr. rer. nat.

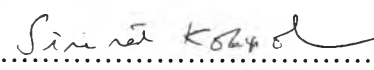
Thesis Co-advisor Associate Professor Vudhichai Parasuk, Ph. D.

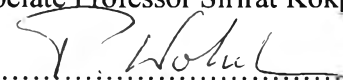
Accepted by the Faculty of Science, Chulalongkorn University in Partial
Fulfillment of the Requirements for the Doctoral Degree


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

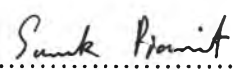
THESIS COMMITTEE

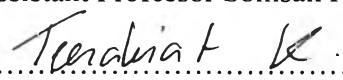

..... Chairman
(Professor Supot Hannongbua, Ph. D.)

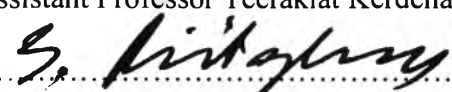

..... Thesis Advisor
(Associate Professor Sirirat Kokpol, Ph. D.)


..... Thesis Co-advisor
(Professor Peter Wolschann, Dr. rer. nat.)


..... Thesis Co-advisor
(Associate Professor Vudhichai Parasuk, Ph.D.)

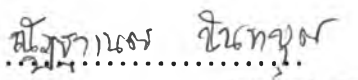
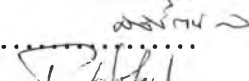
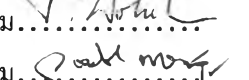
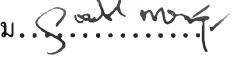

..... Member
(Assistant Professor Somsak Pianwanit, Ph. D.)


..... Member
(Assistant Professor Teerakiat Kerdcharoen, Ph.D.)


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

ณัฐฐานทร นันทบุตร: ทรีดีคิวเอสเออาร์ และโมเลคิวลาร์ไดนามิกส์ซิมูเลชันของเอชไอวี-1 อินทิเกรส และ สารประ- กอบเชิงซ้อนกับตัวยับยั้ง (3D-QSAR AND MOLECULAR DYNAMICS SIMULATIONS OF HIV-1 INTEGRASE AND ITS COMPLEX WITH INHIBITORS) อ. ที่ปรึกษา: รศ. ดร. ศิริรัตน์ ก๊กผล, อ. ที่ปรึกษาร่วม: ศ. ดร. ปีเตอร์ โวลเซน, อ. ที่ปรึกษาร่วม: รศ. ดร. วุฒิชัย พาราสุข, 190 หน้า.

ได้ใช้เทคนิคทรีดีคิวเอสเออาร์ คือ คอมฟาและคอมเซีย หาคความสัมพันธ์ระหว่างสมบัติทางโครงสร้างและฤทธิ์ทางชีวภาพของสารยับยั้งเอนไซม์เอชไอวี-1 อินทิเกรส ที่มีโครงสร้างหลากหลาย 11 กลุ่มจำนวน 89 สาร และ 10 กลุ่มจำนวน 84 สาร สำหรับการยับยั้งกระบวนการ 3'-processing และ strand transfer ตามลำดับแบบจำลองเดี่ยวที่ได้จากคอมฟามีค่า r^2_{cv} เป็น 0.698 สำหรับ 3'-processing และเป็น 0.720 สำหรับ strand transfer ส่วนแบบจำลองเดี่ยวที่ได้จากคอมเซียมีค่า r^2_{cv} เป็น 0.724 สำหรับ 3'-processing และเป็น 0.656 สำหรับ strand transfer ผลการวิเคราะห์ที่ได้แสดงให้เห็นว่าหมู่แทนที่ขนาดใหญ่ใกล้ระนาบของหมู่อินโดล และ หมู่แทนที่ขนาดเล็กใกล้หมู่เตตระโซลของ 5CITEP จะทำให้สารมีประสิทธิภาพในการออกฤทธิ์ทางชีวภาพได้สูงขึ้น และสารควรจะมีหมู่แทนที่ที่มีความหนาแน่นของอิเล็กตรอนสูง เนื่องจากหมู่แทนที่ดังกล่าวสามารถเกิดอันตรกิริยาได้ดีกับโลหะและกรดอะมิโนที่มีประจุบวกเช่นไลซีนลำดับที่ 156 และ 159 ในตำแหน่งกัมมันต์ของเอนไซม์อินทิเกรส นอกจากนี้ยังได้คำนวณโมเลคิวลาร์ไดนามิกส์ซิมูเลชันแบบดั้งเดิมและแบบผสมระหว่างกลศาสตร์ควอนตัมและกลศาสตร์เชิงกล (QM/MM) ของสารประกอบเชิงซ้อนของเอนไซม์เอชไอวี-1 อินทิเกรสกับ 5CITEP (HIV-1 IN-5CITEP) และอนุพันธ์ของ 5CITEP คือ DKA (HIV-1 IN-DKA) เพื่อให้เข้าใจถึงพฤติกรรมทางโครงสร้างและพฤติกรรมทางพลวัตของสารเชิงซ้อนระหว่างโปรตีนกับลิแกนด์ โดยในบริเวณที่เป็นกลศาสตร์ควอนตัมได้ใช้การคำนวณด้วยวิธี PM3 ความแตกต่างที่สังเกตได้ชัดเจนระหว่างสองเทคนิคข้างต้นคือ บริเวณกรดอะมิโนลำดับที่ 116 ถึง 119 กรดอะมิโนลำดับที่ 140 ถึง 149 แอลฟาเฮลิคัลลำดับที่ 4 และอันตรกิริยาแบบไอออนิก นอกจากนี้ยังพบว่าในกรณีที่ใช้เทคนิคแบบดั้งเดิม โคออร์ดิเนชันของเมกนีเซียมไอออนกับลิแกนด์และตัวทำละลายน้ำจะเป็นออกตะฮีดรัลเกือบสมบูรณ์ ทำการเปรียบเทียบผลการคำนวณที่ได้จากการใช้เทคนิคแบบ QM/MM ระหว่าง HIV-1 IN-5CITEP และ HIV-1 IN-DKA เพื่ออธิบายความแตกต่างของกลไกการยับยั้งของสารทั้งสองตัว ข้อแตกต่างที่สำคัญของทั้งสองระบบคือบริเวณกรดอะมิโนลำดับที่ 116 ถึง 119 กรดอะมิโนลำดับที่ 140 ถึง 149 โคออร์ดิเนชันของเมกนีเซียมไอออน และอันตรกิริยาแบบไอออนิก เมกนีเซียมไอออนมีโคออร์ดิเนชันแบบออกตะฮีดรัลที่บิดเบี้ยวกับออกซิเจน 6 และ 7 อะตอม สำหรับ HIV-1 IN-5CITEP และ HIV-1 IN-DKA ตามลำดับ หมู่คาร์บอกซิเลตของ DKA เกิดอันตรกิริยาแบบไอออนิกกับไลซีนลำดับที่ 159 ได้ดีกว่าในกรณีของหมู่เตตระโซลของ 5CITEP นอกจากนี้ค่าพลังงานเสรีการยึดเหนี่ยวที่คำนวณได้ของ HIV-1 IN-DKA มีความเสถียรกว่าของ HIV-1 IN-5CITEP ผลที่ได้สนับสนุนค่าฤทธิ์ทางชีวภาพที่ดีกว่าของ DKA เมื่อเทียบกับ 5-CITEP

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

4573811223: MAJOR CHEMISTRY

KEY WORDS: HIV-1 INTEGRASE/CoMFA /CoMSIA /QUANTUM MECHANICAL/
MOLECULAR MECHANICAL

NADTANET NUNTHABOOT: 3D-QSAR AND MOLECULAR DYNAMICS
SIMULATIONS OF HIV-1 INTEGRASE AND ITS COMPLEX WITH
INHIBITORS. THESIS ADVISOR: ASSOC. PROF. SIRIRAT KOKPOL, Ph. D..
THESIS COADVISOR: PROF. PETER WOLSCHANN, Dr. rer. nat. THESIS
COADVISOR: ASSOC. PROF. VUDHICHAH PARASUK, Ph. D., 190 pp.

Three dimensional quantitative structure activity relationship (3D-QSAR) techniques including CoMFA and CoMSIA were employed to determine the relationship between structural properties and biological activities of 11 and 10 diverse structural classes of HIV-1 integrase (IN) inhibitors against 3'-processing (89 compounds) and strand transfer (84 compounds) reactions, respectively. The single CoMFA model gave $r^2_{cv} = 0.698$ for 3'-processing activity and $= 0.720$ for strand transfer activity. The single CoMSIA model showed $r^2_{cv} = 0.724$ and $= 0.656$ for 3'-processing and strand transfer activities, respectively. Result analyses suggest that large substituents in the plane of the indole ring and the small bulky groups at the tetrazole ring of 5CITEP are required to increase the inhibitory potency of IN inhibitors. The groups of high electron density on ligand are also necessary for better interaction with the metal ion or the positive charge amino acids such as Lys156 and Lys159 in the active site of IN enzyme. To understand the structural and dynamical behaviors of protein-ligand complex, classical and hybrid quantum mechanical/molecular mechanism (QM/MM) molecular dynamics (MD) simulations were conducted for systems of HIV-1 IN complexed with 5CITEP (HIV-1 IN-5CITEP) and with its derivative, DKA (HIV-1 IN-DKA). The PM3 method was used to describe the QM region. Substantial differences between the two levels of calculations were observed at residues 116-119, residues 140-149, α 4-helix, and the salt link interaction. In addition, the coordination of Mg^{2+} ion with ligand and water solvent in a near perfect octahedral complex was noticed in classical MD simulation. The QM/MM results between HIV-1 IN-5CITEP and HIV-1 IN-DKA were compared in order to explain the difference of inhibition mechanisms of the two inhibitors. The main findings of the differences between both systems are residues 116-119, the flexible loop residues 140-149, Mg^{2+} coordination, and salt link interaction. Mg^{2+} forms distorted octahedral complexes with 6 and 7 oxygen atoms as observed in HIV-1 IN-5CITEP and HIV-1 IN-DKA, respectively. The carboxylate moiety of DKA forms a stronger salt bridge interaction with Lys159 than the tetrazole ring of 5CITEP does. Moreover, the calculated binding free energy of HIV-1 IN-DKA is energetically more favorable than that of HIV-1 IN-5CITEP. The results supported the higher inhibitory potency of DKA in comparison with 5CITEP.

Department..... Chemistry.....
Field of Study..... Chemistry.....
Academic Year..... 2006.....

Student's Signature..... *Nadtanet Nunthaboot*
Advisor's Signature..... *Sirirat Kokpol*
Co-advisor's Signature..... *Peter Wolschann*
Co-advisor's Signature..... *Vudhichai Parasuk*

ACKNOWLEDGEMENTS

First, I wish to express my gratitude to my supervisor Assoc. Prof. Dr. Sirirat Kokpol for providing the opportunity to undertake this work, for her guidance, valuable advices and suggestions.

I also would like to thank Assoc. Prof. Dr. Vudhichai Parasuk and Assist. Prof. Dr. Somsak Pianwanit for their helpful suggestions. Without their encouragement and constant guidance, I may not complete this dissertation.

I would like to extend my sincere thanks to Prof. Dr. Peter Wolschann for the very helpful discussions, to help me correct the writing of this dissertation and for his taking care of me during my visit in Vienna.

Prof. Dr. James M. Briggs is gratefully acknowledged for his critical remarks and his hospitality in Houston, where part of this work was completed.

Thesis committee, Prof. Dr. Supot Hannongbua, Assist. Prof. Dr. Teerakiat Kerdcharoen and Assist. Prof. Dr. Surapong Pinitglang, are acknowledged.

I recognize Dr. Christian Hensen and Dr. Jerry O. Ebalunode for their kind CHARMM assistance.

Special thanks go to all present and the former working colleagues of the Computational Chemistry Unit Cell (CCUC) for their constant and reliable willingness to help, for their patience, and for providing the friendly working atmosphere.

I gratefully acknowledge the CCUC at Chulalongkorn University, the Schrödinger cluster at University of Vienna, the TLC² at University of Houston and the National Electronics and Computer Technology (NECTEC) for providing computational time and all facilities.

This dissertation was smoothly completed with the funding supported by the Royal Golden Jubilee Program (3.C.CU/45/S.1), the Thailand Research Fund.

The Fulbright Foundation is gratefully acknowledged for an awarding of the Fulbright visiting scholarship.

Finally, this work would not have been completed without unconditional support and encouragement from my parents who always cheer me up.

CONTENTS

	Page
ABSTRACT IN THAI.....	iv
ABSTRACT IN ENGLISH.....	v
ACKNOWLEDGEMENTS.....	vi
CONTENTS.....	vii
LIST OF TABLES.....	xii
LIST OF FIGURES.....	xiii
LIST OF SCHEMES.....	xviii
NOTATIONS.....	xix
CHAPTER I INTRODUCTION.....	1
1.1 Acquired Immune Deficiency Syndrome.....	1
1.1.1 History.....	1
1.1.2 The global AIDS epidemic today.....	1
1.1.3 AIDS epidemic in Thailand.....	2
1.2 Human Immune Deficiency Virus (HIV).....	3
1.2.1 HIV-1 Structure.....	3
1.2.2 HIV-1 Replication.....	4
1.2.3 HIV-1 Treatment.....	6
1.3 HIV-1 Integrase (IN).....	9
1.3.1 HIV-1 IN Structure.....	9
1.3.1.1 N-terminal domain.....	10
1.3.1.2 Catalytic core domain.....	10
1.3.1.3 C-terminal domain.....	11
1.3.1.4 The catalytic core domain complexed with inhibitors.....	11
1.3.2 HIV-1 IN Mechanism and the possible drug targets.....	12
1.3.2.1 Binding of IN to viral DNA.....	13
1.3.2.2 First catalytic step: 3'-processing.....	13
1.3.2.3 Second catalytic step: strand transfer.....	13
1.3.2.4 Gap repair and ligation of viral DNA to host DNA.....	14
1.3.3 Progress of HIV-1 IN inhibitors design.....	15

	Page
1.3.4 Outlook and Future challenges of HIV-1 IN.....	19
1.4 Research Inspiration.....	19
1.5 Research Goals.....	20
CHAPTER II LITERATURE REVIEWS ON MOLECULAR MODELING OF HIV-1 IN.....	22
2.1 Molecular Modeling.....	22
2.2 Molecular Modeling of HIV-1 IN.....	22
2.2.1 Two/Three Dimensional-Quantitative Structure Activity Relationship (2/3D-QSAR).....	23
2.2.2 Molecular Dynamics (MD) simulation.....	26
CHAPTER III THEORY.....	32
3.1. Quantitative Structure Activity Relationship (QSAR).....	32
3.1.1 Comparative Molecular Field Analysis (CoMFA).....	33
3.1.2 Comparative Molecular Similarity Indices Analysis (CoMSIA).....	34
3.1.3 3D-QSAR set up.....	35
3.1.3.1 Compound selection.....	35
3.1.3.2 Ligand 3D structure generation.....	36
3.1.3.3 Molecule alignment.....	36
3.1.3.4 Molecular field calculation.....	36
3.1.3.5 Statistical Analysis.....	37
3.1.3.6 Result Interpretation.....	38
3.2 Molecular dynamics simulation.....	39
3.3 Langevin dynamics.....	41
3.4 Stochastic boundary.....	41
3.5 Quantum Mechanical /Molecular Mechanical (QM/MM).....	42
3.5.1 The QM/MM approach.....	42
3.5.2 Frontier bonds in QM/MM methods.....	44
3.6 Free energy MM-PB (GB) SA.....	46
CHAPTER IV COMPARATIVE MOLECULAR FIELD ANALYSIS (CoMFA) AND COMPARATIVE MOLECULAR SIMILARITY INDICES ANALYSIS (CoMSIA).....	48

	Page
4.1 Computational Methods.....	48
4.1.1 Biological data.....	48
4.1.2 Compound generation.....	74
4.1.3 Template selection.....	74
4.1.4 Alignment procedure.....	74
4.1.4.1 The atom-based root means square (RMS) fit method.....	75
4.1.4.2 The flexible fitting (multi-fit method).....	75
4.1.4.3 The rigid body field fit.....	75
4.1.5 Field calculation for CoMFA.....	77
4.1.6 Field calculation for CoMSIA.....	77
4.1.7 Region focusing.....	77
4.1.8 Derivation of models	78
4.2 Results and Discussion.....	78
4.2.1 3'-processing activity.....	78
4.2.1.1 CoMFA Statistical Parameters.....	78
4.2.1.2 CoMFA Contour.....	84
4.2.1.3 CoMSIA Statistical Parameters.....	87
4.2.1.4 CoMSIA Contour.....	91
4.2.2 ST activity.....	94
4.2.2.1 CoMFA Statistical Parameters.....	94
4.2.2.2 CoMFA Contour.....	100
4.2.2.3 CoMSIA Statistical Parameters.....	103
4.2.2.4 CoMSIA Contour.....	107
4.3 Conclusion.....	110
CHAPTER V Hybrid Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of HIV-1 Integrase/Inhibitor Complexes.....	112
5.1 Initial structure of protein-ligand complex	112
5.1.1 Metal-ligand system	114
5.1.1.1 Computational Methods.....	114
5.1.1.2 Results and Discussion.....	117

	Page
5.1.2 Metal-amino acid system.....	118
5.1.2.1 Computational methods.....	118
5.1.2.2 Results and Discussion.....	119
5.1.3 Hybrid of metal-ligand-amino acid system.....	120
5.1.3.1 Computational methods.....	120
5.1.3.2 Results and Discussion.....	121
5.1.4 Pair and total interaction energies calculations.....	122
5.1.4.1 Computational Methods.....	122
5.1.4.2 Results and Discussion.....	125
5.2 MD simulation.....	129
5.2.1 Computational methods.....	129
5.2.1.1 System set-up and Conventional MD simulation.....	129
5.2.1.2 QM/MM set up.....	132
5.2.2 Results and Discussion.....	132
A. Comparison of HIV-1 IN-5CITEP using MM and QM/MM force fields.....	132
A.1 IN-5CITEP.....	132
A.1.1 Structural dynamics of HIV-1 IN in HV-1 IN-5CITEP	132
A.1.2 Mg ²⁺ coordination and 5CITEP binding mode.....	137
A.2 IN-DKA.....	141
A.2.1 Structural dynamics of HIV-1 IN in HIV-1 IN-DKA.....	141
A.2.2 Mg ²⁺ coordination and DKA binding mode.....	146
B. Comparison of QM/MM results between HIV-1 IN-DKA and HIV-1 IN- 5CITEP.....	150
B.1 Stability of MD trajectories.....	150
B.2 Residue Fluctuation.....	150
B.3 Structural difference between IN-DKA and IN-5CITEP.....	151
B.4 Inhibitor binding mode/structural features of inhibitor.....	155
B.5 Mg ²⁺ coordination.....	157
B.6 Salt bridge analysis.....	158
B.7 Role of water molecules.....	160

	Page
B.8 Residue contributions.....	161
B.9 Interaction energy decomposition.....	163
B.10 Atomic charge polarization.....	165
B.11 Protein-ligand binding free energy.....	167
5.3 Conclusion.....	168
CHAPTER VI CONCLUSION.....	171
6.1 3D-QSAR.....	171
6.2 Classical and hybrid QM/MM MD simulations.....	172
6.3 Suggestions for future works.....	173
REFERENCES.....	174
APPENDICES.....	188
BIOGRAPHY.....	190

LIST OF TABLES

	Page
Table 1.1 Global summary of the AIDS epidemic in 2006 (data as of December 2006), the number is an average while the number in parentheses is for range. The data were taken from the jointed United Nations programme on HIV/AIDS (UNAIDS) 2006	2
Table 1.2 Antiretroviral drugs currently approved by FDA.....	8
Table 1.3 HIV-1 IN inhibitors: Phase of preclinical and clinical development.....	16
Table 4.1 Summary of CoMFA results of 3'-processing mechanism.....	79
Table 4.2 Actual and calculated activities for 3'-processing mechanism of compounds used in training and test sets, obtained from CoMFA and CoMSIA.....	80
Table 4.3 Summary of CoMSIA results of 3'-processing activity.....	89
Table 4.4 CoMFA statistical parameters for ST activity.....	96
Table 4.5 Actual and calculated activities for ST mechanism of compounds used in training and test sets, obtained from CoMFA and CoMSIA.....	96
Table 4.6 CoMSIA statistical parameters for ST mechanism.....	105
Table 5.1 Interaction energies of complexed Mg ²⁺ -ligand models.....	118
Table 5.2 Interaction energies of Mg ²⁺ -amino acids models.....	120
Table 5.3 Interaction energies of hybrid system.....	122
Table 5.4 Description of models and list of amino acids within 8 Å from inhibitor...	123
Table 5.5 Pair and total interaction energies (kcal/mol) of HIV-1 IN-5CITEP.....	127
Table 5.6 Pair and total interaction energies (kcal/mol) of HIV-1 IN-DKA.....	128
Table 5.7 Standard deviation of amino acid decomposition analysis.....	163
Table 5.8 Electrostatic interaction energies and their components (kcal/mol) of QM subsystem of HIV-1 IN-DKA and HIV-1 IN-5CITEP.....	165
Table 5.9 Average partial atomic charges and polarizations. Polarization is defined as the difference in averaged charged in atomic units on a given atom in condense phase and in gas phase.....	166
Table 5.10 Binding free energy (kcal/mol) of HIV-1 IN-DKA and HIV-1 IN-5CITEP, calculated using the MM/GBSA approach.....	167

LIST OF FIGURES

	Page
Figure 1.1	Structure of HIV-1 particle..... 4
Figure 1.2	HIV-1 replication cycle..... 6
Figure 1.3	HIV provirus gene structure and IN domains..... 9
Figure 1.4	HIV-1 IN structure..... 11
Figure 1.5	(a) X-ray co-crystal of HIV-1 IN-5CITEP and (b) schematic drawing of hydrogen bonds between 5CITEP and surrounding amino acids..... 12
Figure 1.6	HIV-1 integration process and potential drug targets for inhibition..... 15
Figure 1.7	Chemical structure of HIV-1 IN inhibitors..... 18
Figure 2.1	Chemical structure of compounds used in references [59-63]..... 26
Figure 2.2	Representation of protein-protein and protein-solvent interactions seen in HIV-1 IN active site (a) without Mg^{2+} and (b) with Mg^{2+} 27
Figure 2.3	Orientation of L-731, 988 in the (a) bound native, (b) ion bound double mutant and (c) the bound double mutant. 5CITEP is shown by dark grey ball and stick model..... 30
Figure 3.1	Schematic representations of the Lennard-Jones and Coulomb attractive and repulsive potentials describing steric and electrostatic contributions to CoMFA fields..... 34
Figure 3.2	A Gaussian type function used in CoMSIA calculations..... 35
Figure 3.3	Illustration of QSAR table generated by SYBYL..... 37
Figure 3.4	Schematic illustration of the partition of region in stochastic boundary molecular dynamics approach..... 42
Figure 3.5	Illustration for the QM/MM method in the enzyme system. The active center is treated at the QM level and the surrounding is treated at the MM level..... 43
Figure 3.6	Partition between QM and MM regions in link-atom approach..... 45
Figure 4.1	Distribution plots of activity values of the training set compounds against (a) 3'-processing and (b) ST mechanism..... 49
Figure 4.2	Structure of 5CITEP showing the C1-C5 atoms used for superimposition. 74

	Page	
Figure 4.3	3D-view of aligned molecules (training and test sets) based on RMS fit (a = 3'-processing, b = ST), multi-fit (c = 3'-processing, d = ST), and field fit (e = 3'-processing, f = ST).....	76
Figure 4.4	Plot of observed and calculated activities against 3'-processing mechanism of (a) training set and (b) test set using CoMFA.....	83
Figure 4.5	CoMFA StDev \times coeff steric contour map for 3'-processing activity.....	85
Figure 4.6	CoMFA StDev \times coeff electrostatic contour map for 3'-processing activity	87
Figure 4.7	Plot of observed and calculated activities against 3'-processing mechanism of (a) training and (b) test sets using CoMSIA	90
Figure 4.8	CoMSIA StDev \times coeff steric contour map for 3'-processing activity.....	91
Figure 4.9	Mapping of CoMSIA StDev \times coeff hydrogen bond donor contour plots within the active site of the complex structure of HIV-1 IN/5CITEP for 3'-processing activity.....	93
Figure 4.10	Mapping of CoMSIA StDev \times coeff hydrogen bond acceptor contour plots within the active site of the complex structure of HIV-1 IN/5CITEP for 3'-processing activity.....	94
Figure 4.11	Plot of observed and calculated activities against ST mechanism of (a) training and (b) test sets using CoMFA	99
Figure 4.12	CoMFA StDev \times coeff steric contour map for ST activity.....	101
Figure 4.13	CoMFA StDev \times coeff electrostatic contour map for ST activity.....	103
Figure 4.14	Plot of observed and calculated activities against ST mechanism of (a) training and (b) test sets using CoMSIA.....	106
Figure 4.15	CoMSIA StDev \times coeff steric contour map for ST activity.....	107
Figure 4.16	Mapping of CoMSIA StDev*coeff hydrogen bond donor contour plots within the active site of the complex structure of HIV-1 IN/5CITEP for ST activity.....	108

	Page
Figure 4.17 Mapping of CoMSIA StDev*coeff hydrogen bond acceptor contour plots within the active site of the complex structure of HIV-1 IN/5CITEP for ST activity.....	110
Figure 5.1 Chemical structure and atomic definition of (a) 5CITEP and (b) DKA.....	113
Figure 5.2 Proposed models of metal-ligand chelation.....	115
Figure 5.3 Proposed models of metal-amino acids chelation.....	119
Figure 5.4 Hybrid chelation of metal, ligand and amino acids.....	121
Figure 5.5 Schematic representation of the active site of HIV-1 IN and the amino acid residues involved in 3'-processing and strand transfer reactions.....	129
Figure 5.6 Schematic drawing of proposed metal-ligand binding of HIV-1 IN-inhibitor complexes. A QM regions used in QM/MM calculation are encircled. HQ refers to hydrogen link atom HQ type which is added between the alpha carbon and beta carbon of both Asp64 and Asp116.....	130
Figure 5.7 Time-dependent heavy atom RMSDs with respect to the average structure versus simulation time of HIV-1 IN-5CITEP which were obtained from conventional MD and QM/MM MD simulations.....	133
Figure 5.8 Average RMSF of protein backbone atoms as a function of residue number of HIV-1 IN-5CITEP. The protein structures were taken from classical and hybrid QM/MM MD simulations.....	134
Figure 5.9 (a) Superimposition of the protein conformation of the average structure obtained from conventional MD (pink) and combined QM/MM MD calculation (ice blue). Residues Asp64, Asp116 and inhibitors are shown in stick models while Mg ²⁺ ion and the catalytic water molecules are displayed in balls. (b) and (c) display hydrogen-bonds between amino acids located in 116-119 obtained from conventional and QM/MM MD simulations of HIV-1 IN-5CITEP complex, respectively.....	136
Figure 5.10 Stereo views of the flexible loop region (residues 140-149) and the α -helix of HIV-1 IN-5CITEP. Selected residues and inhibitors are displayed in stick models. The distances between ligand and Lys159 are shown in dotted lines.....	137

	Page	
Figure 5.11	Time evolution of the inter-atomic distances between Mg^{2+} and its coordinating atoms obtained from (a) conventional MD of HIV-1 IN-5CITEP and (b) QM/MM of HIV-1 IN-5CITEP.....	138
Figure 5.12	Plots of the salt link distances between the negatively charged group of ligand and the ammonium group of Lys59. The distance between tetrazolate ring of 5CITEP and the ammonium nitrogen of Lys159 taken from the conventional MD is shown in panel (a) while that from QM/MM is given in panel (b).....	139
Figure 5.13	Comparison of the distributions of torsional angles (a) <i>Tor1</i> and (b) <i>Tor2</i> of 5CITEP using MM and QM/MM.....	141
Figure 5.14	Time-dependent heavy atom RMSDs with respect to the average structure versus simulation time of HIV-1 IN-DKA which were obtained from conventional MD and QM/MM MD simulations.....	142
Figure 5.15	Average RMSF of protein backbone atoms as a function of residue number of IN-DKA. The protein structures were taken from classical and hybrid QM/MM MD simulations.....	143
Figure 5.16	(a) Superimposition of the protein conformation of the average structure obtained from conventional MD (pink) and taken from a combined QM/MM MD calculation (cyan). Residues Asp64, Asp116 and inhibitors are shown as stick models while Mg^{2+} ion and the catalytic water molecules are displayed in balls. (b) displays the orientations of amino acid residues 64-69 obtained from conventional and QM/MM MD simulations of IN-DKA complex.....	144
Figure 5.17	Stereo views of the flexible loop region (residues 140-149) and the α -helix of HIV-1 IN-DKA. Selected residues and inhibitors are displayed in stick models. The distances between ligand and Lys159 are shown in dotted lines.	145
Figure 5.18	Time evolution of the inter-atomic distances between Mg^{2+} and its coordinating atoms obtained from (a) conventional MD of HIV-1 IN-DKA and (b) QM/MM of HIV-1 IN-DKA	146

Figure 5.19	Plot of salt link distances between the negatively charged group of DKA and the ammonium group of Lys59. The distance between carboxylate group of DKA and the ammonium nitrogen of Lys159 taken from the conventional MD is shown in panel (a) while that from QM/MM is given in panel (b).	147
Figure 5.20	Distribution of torsional angles (a) <i>Tor1</i> , (b) <i>Tor2</i> of DKA using MM and QM/MM.....	149
Figure 5.21	Heavy atom RMSDs between IN-DKA and IN-5CITEP taken from a hybrid QM/MM MD simulations.....	150
Figure 5.22	RMSF of protein backbones of HIV-1 IN-DKA and HIV-1 IN-5CITEP complexes obtained from QM/MM MD simulations.....	151
Figure 5.23	(a) Superimposition of average structure obtained from QM/MM MD simulations between HIV-1 IN-DKA (cyan) and HIV-1 IN-5CITEP (ice blue). Binding modes of DKA (b) and 5CITEP (c) in the HIV-1 IN active site. Mg^{2+} ion is displayed as a green ball. The coordinating distances between Mg^{2+} and its chelating atoms are represented in dotted lines.....	153
Figure 5.24	Comparison of the distributions of torsional angles (a) <i>Tor1</i> and (b) <i>Tor2</i> of both DKA and 5CITEP taken from QM/MM MD simulation.....	156
Figure 5.25	Snapshots of the active site of HIV-1 IN-DKA complex (a) and HIV-1 IN-5CITEP complex (b). Hydrogen-bonds and salt bridges are indicated by dotted lines. All distances are in Angstrom unit.....	159
Figure 5.26	Amino acid decomposition analysis showing contributions of each individual MM amino acid to the QM subsystem. Stabilizing and destabilizing interactions are represented as positive and negative energy values, respectively. Only residues exhibiting an energy greater than 1 kcal/mol are shown.....	162

LIST OF SCHEMES

	Page
Scheme 4.1 Structures of compounds used in the training set (1-61) and test set (62-89) of 3'-processing process. The labeled numbers correspond to atoms C1-C5 of 5CITEP (template) used in superposition process.....	50
Scheme 4.2 Structures of compounds used in the training set (1-64) and test set (65-84) of ST process. The labeled numbers correspond to atoms C1-C5 of 5CITEP (template) used in superposition process.....	63

NOTATIONS

ABNR	Adopted basis Newton Raphson
ASV	Avain Rous Sarcoma Virus
AIDS	Acquired immune deficiency syndrome
CDC	The Centers for Disease Control and Prevention
CD4	Cluster of differentiation 4
CoMFA	Comparative Molecular Field Analysis
CoMSIA	Comparative Molecular Similarity Indices Analysis
DKAs	Diketo acids
DNA	Deoxyribonucleic acid
FDA	The US Food and Drug Administration
GB	Generalize Born
HAART	Highly active antiretroviral treatment
HIV	Human immunodeficiency virus
HIV-1	Human immunodeficiency virus type I
HIV-2	Human immunodeficiency virus type II
IN	Integrase
KS	Kaposi's Sarcoma
LTR	Long terminal repeat regions
MD	Molecular Dynamics
MM	Molecular Mechanics
mRNA	Message ribonucleic acid
NMR	Nuclear Magnetic Resonance
ONC	Optimum number of component
PB	Possion-Boltzman
PCP	Pneumocystis carinii pneumonia
PIC	Preintegration complex
PLS	Partial least square
PLWHA	People living with HIV alive
PR	Protease

PRESS	Sum of squared prediction error called the predictive residual sum of squares
QM/MM	Quantum Mechanical/Molecular Mechanical
r^2	Squared correlation coefficient
r^2_{cv}	Cross-validated correlation coefficient
RMS	Root Means Square
RNA	Ribonucleic acid
RT	Reverse transcriptase
SARs	Structure activity relationships
SASA	Solvent accessible surface area
SEP	Standard error of estimate
SD	Standard deviation
SH3	Src homology domains
SIV	Simian immunodeficiency virus
SLBOs	Strictly localized bond orbitals
RMSD	Root mean square deviations
RMSF	Root mean square fluctuation
ST	Strand transfer
UNAIDS	The jointed United Nations programme on HIV/AIDS
3D-QSAR	Three-dimensional Quantitative Structure Activity Relationship
5CITEP	1-(5-chloroindole-3-yl)-3-hydroxy-3(2H-tetrazol-5-yl)-propanone