

ผลของหมู่แทนที่ตำแหน่งที่ 5 ของไดออกซียูรีดีน โมโนฟอสเฟตต่อปฏิกิริยาการเติมไทโอเลตใน
โพลีดีเลคตินเทสโดยใช้เทคนิคควเอ็ม/เอ็มเอ็ม

นายนพพร ไกยเวช

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาเคมี ภาควิชาเคมี
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2555
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the Graduate School.

EFFECTS OF SUBSTITUENTS ON THE 5-POSITION OF
DEOXYURIDINEMONOPHOSPHATE ON THE THIOLATE ADDITION IN
THYMIDYLATE SYNTHASE USING QM/MM TECHNIQUE

Mr. Nopporn Kaiyawet

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 2012
Copyright of Chulalongkorn University

Thesis Title	EFFECTS OF SUBSTITUENTS ON THE 5-POSITION OF DEOXYURIDINEMONOPHOSPHATE ON THE THIOLATE ADDITION IN THYMIDYLATE SYNTHASE USING QM/MM TECHNIQUE
By	Mr. Nopporn Kaiyawet
Field of Study	Chemistry
Thesis Advisor	Professor Supot Hannongbua, Dr. rer. nat.
Thesis Co-advisor	Professor Adrian J. Mulholland, Ph.D. Thanyada Rungrotmongkol, 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 Supot Hannongbua, Dr. rer. nat.)

THESIS COMMITTEE

..... Chairman
(Assistant Professor Warinthorn Chavasiri, Ph.D.)

..... Thesis Advisor
(Professor Supot Hannongbua, Dr. rer. nat.)

..... Thesis Co-advisor
(Professor Adrian J. Mulholland, Ph.D.)

..... Thesis Co-advisor
(Thanyada Rungrotmongkol, Ph.D.)

..... Examiner
(Associate Professor Pornthep Sompornpisut, Ph.D.)

..... Examiner
(Associate Professor Nongnuj Muangsin, Ph.D.)

..... Examiner
(Assistant Professor Viwat Vchirawongwin, Ph.D.)

..... External Examiner
(Onjira Aruksakunwong, Ph.D.)

นพพร ไกยเวช : ผลของหมู่แทนที่ตำแหน่งที่ 5 ของดีออกซียูริดีน โมโนฟอสเฟตต่อปฏิกิริยาการเติมไทโอเลตในไทมิดิเลตซินเทสโดยใช้เทคนิคควเอ็ม/เอ็มเอ็ม. (EFFECTS OF SUBSTITUENTS ON THE 5-POSITION OF DEOXYURIDINE MONOPHOSPHATE ON THE THIOLATE ADDITION IN THYMIDYLATE SYNTHASE USING QM/MM TECHNIQUE) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ศ.ดร.สุพจน์ หารหนองบัว, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม : ศ.ดร.เอเดรียน มูสฮอลแลนด์, ดร.รัชฎา รุ่งโรจน์มงคล, 91 หน้า.

ไทมิดิเลตซินเทสเป็นเอนไซม์ที่เร่งปฏิกิริยาการเติมหมู่เมทิลแบบรีดิวซ์ของของสารดียูเอ็มพีเพื่อสร้างผลิตภัณฑ์ดีเอ็มพีในกระบวนการสังเคราะห์ไทมินนิวคลีโอไทด์ การเติมแบบไมเคิลและการเกิดสารเชิงซ้อนโควาเลนต์ไทรภาคถูกพิจารณาว่าเป็นขั้นตอนที่มีความสำคัญที่ถูกรวมอยู่ในขั้นกลไกการยับยั้งของยาต่อต้านมะเร็งห้าเอพยู ผลของการแทนที่หมู่ฟังก์ชันที่ตำแหน่งซีห้าของสารดียูเอ็มพีด้วยอะตอมฮาโลเจน (ฟลูออรีน, คลอรีน และ โบรมีน) หมู่ดิงอิลิเกตรอน (ไซยาไนด์และไนโตร) และหมู่ให้อิเล็กตรอน (อะมิโนและไฮดรอกซิล) ต่อความเสถียรของเอนไซม์ไทมิดิเลตซินเทสถูกสืบสวนโดยใช้เทคนิคการจำลองพลวัตเชิงโมเลกุล ผลการจำลองแสดงให้เห็นว่าพันธะไฮโดรเจนมีการปรากฏเพิ่มขึ้นระหว่างหมู่แทนที่ของสารอุปมานดียูเอ็มพี (ในระบบฟลูออรีนดียูเอ็มพี, ไซยาไนด์ดียูเอ็มพี, ไฮดรอกซิลดียูเอ็มพีและอะมิโนดียูเอ็มพี) และหมู่ไฮดรอกซิลของวายเป็นสปีดพลังงานเสรีการยึดจับได้กำหนดการเรียงลำดับของสัมพรรคภาพการยึดจับได้เป็น คลอรีนดียูเอ็มพี ~ ฟลูออรีนดียูเอ็มพี > ดียูเอ็มพี > โบรมีนดียูเอ็มพี สำหรับสารอุปมานฮาโลเจน และ ไซยาไนด์ดียูเอ็มพี ~ ไนโตรดียูเอ็มพี > ไฮดรอกซิลดียูเอ็มพี ~ อะมิโนดียูเอ็มพี สำหรับสารอุปมานแบบดิงอิลิเกตรอนและให้อิเล็กตรอน บนพื้นฐานวิธีไฮบริด B3LYP/6-31+G*-CHARMM การเติมแบบไมเคิลและการเกิดสารเชิงซ้อนโควาเลนต์ไทรภาคระหว่างดียูเอ็มพีและเอ็มทีเอชเอฟสับสเตรตเพื่อเกิดเป็นสารมัธยันตร์สะพานเมทิลินพบว่าเกิดผ่านกลไกแบบคอนเซ็คเต็ด จากการทดสอบเทคนิคควเอ็ม/เอ็มเอ็ม ในหลายลักษณะงานวิจัยนี้ได้เลือกใช้เทคนิค SCS-MP2 ด้วย aug-cc-pVTZ เบซิส เซ็ต สำหรับการคำนวณหาค่าพลังงานการเกิดปฏิกิริยา เนื่องจากสัมพรรคภาพในการยึดจับที่สูงและใช้พลังงานก่อกัมมันต์ใกล้เคียงกับยาต่อต้านมะเร็งฟลูออรีนดียูเอ็มพีดังนั้นเราจึงนำเสนอว่าสารอุปมานคลอรีนดียูเอ็มพีมีศักยภาพที่จะเป็นสารยับยั้งต่อต้านฟังก์ชันของเอนไซม์ไทมิดิเลตซินเทสได้

ภาควิชา.....เคมี.....ลายมือชื่อนิสิต.....
 สาขาวิชา.....เคมี.....ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
 ปีการศึกษา.....2555.....ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....
 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

5273821523 : MAJOR CHEMISTRY

KEYWORDS : Thymidylate synthase / Michael addition / MD simulation / QM-MM / Molecular modeling / Computational chemistry

NOPPORN KAIYAWET: EFFECTS OF SUBSTITUENTS ON THE 5-POSITION OF DEOXYURIDINEMONOPHOSPHATE ON THE THIOLATE ADDITION IN THYMIDYLATE SYNTHASE USING QM/MM TECHNIQUE. ADVISOR: PROF. SUPOT HANNONGBUA, Dr. rer. nat., CO-ADVISOR: PROF. ADRIAN J. MULHOLLAND, Ph.D., THANYADA RUNGROTMONGKOL, Ph.D., 91 pp.

Thymidylate synthase (TS) is an enzyme that catalyses the reductive methylation of dUMP to produce dTMP in the thymine nucleotide synthesis. The Michael addition and ternary covalent complex formation are considered as important steps involved in the inhibition mechanism of a known anticancer drug, 5-FU. The effect of functional group substitution on the C-5 position of dUMP substrate with halogen atoms (F-, Cl- and Br-) and electron withdrawing (CN-, NO₂-) and donating (NH₂-, OH-) groups towards the TS stability is investigated by means of molecular dynamics simulation technique. The simulated results showed that an additionally unique hydrogen bond between the substituted group of dUMP analogues and the hydroxyl group of Y94 was observed in most systems except CldUMP and BrdUMP analogues. The MM/PBSA binding free energy provides the orders of averaged binding affinity are: CldUMP ~ FdUMP > dUMP > BrdUMP for halogen analogues and CNdUMP ~ NO₂dUMP > OHdUMP ~ NH₂dUMP for electron withdrawing and donating analogues. Based on the B3LYP/6-31+G*-CHARMM hybrid method, the Michael addition and ternary covalent complex formation between dUMP and mTHF substrates to form the methylene bridging intermediate is determined via a concerted mechanism. By following the several tests of QM/MM approaches, the SCS-MP2 method with aug-cc-pVTZ basis sets is selected for reaction energy profile evaluation. Due to the close of activation energy with FdUMP anticancer drug and high affinity of the binding in CldUMP analogue, therefore, we proposed that the CldUMP analogue has a potent to be the new candidate inhibitor against thymidylate synthase function.

Department : Chemistry..... Student's Signature.....
 Field of Study : Chemistry..... Advisor's Signature.....
 Academic Year : 2012..... Co-advisor's Signature.....
 Co-advisor's Signature.....

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Prof. Dr. Supot Hannongbua, who gives a great opportunity to study in my Ph. D. program. His understanding in an enzymology not only guides but also inspires me to do and finish this dissertation. I would especially thank Prof. Dr. Adrian J. Mulholland, my foreign co-advisor, for helpful suggestions and valuable experiences in his laboratory, the University of Bristol, UK. I would extend to special thank Dr. Thanyada Rungrotmongkol for her intensive helps and suggestions during my study period and also thank Dr. Richard Lonsdale for his kindly helps about the QM/MM fundamental knowledge and its calculation procedures.

All of my thesis committee, Assist. Prof. Dr. Warinthorn Chavasiri, Assoc. Prof. Dr. Pornthep Sompornpisut, Assoc. Prof. Dr. Nongnuj Muangsin, Assist. Prof. Dr. Viwat Vchirawongwin and Dr. Onjira Aruksakunwong are also acknowledged for their useful suggestion in my examination.

I would acknowledge the Royal Golden Jubilee Ph.D. Program for fully financial support, the Computational Chemistry Unit Cell (CCUC) at Chulalongkorn University and the Centre for Computational Chemistry (CCC) at University of Bristol for computer facilities and software.

Finally, I also thank all of the CCUC and CCC members to share their valuable experiences in both academic and non-academic frameworks during the full period of my study.

CONTENTS

	Page
ABSTRACT IN THAI.....	iv
ABSTRACT IN ENGLISH.....	v
ACKNOWLEDGEMENTS.....	vi
CONTENTS.....	vii
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
LIST OF ABBREVIATIONS.....	xvi
CHAPTER I - INTRODUCTION.....	1
1.1 What is cancer?.....	1
1.2 Cancer staging.....	3
1.3 Cancer chemotherapy.....	4
1.4 Antimetabolites drugs.....	4
1.5 Overall catalysis of thymidylate synthase enzyme.....	6
1.6 Molecular Mechanism of thymidine synthesis.....	7
1.7 Inhibition mechanism of the FdUMP and the C-5 substituted analogues.....	9
1.8 Mechanistic indetermination of methylene bridging intermediate formation.....	10
1.9 Research rationale.....	11
1.10 Objectives.....	11
1.11 Scope of this work.....	12
CHAPTER II – THEORETICAL BACKGROUND.....	13
2.1 Molecular mechanics.....	13
2.1.1 Bond stretching.....	13
2.1.2 Bond bending.....	14
2.1.3 Torsion angle.....	15
2.1.4 Improper.....	15
2.1.5 Electrostatic energy.....	16

	Page
2.1.6 van der Waals interaction	17
2.2 Molecular dynamics	18
2.2.1 The Time Dependent Algorithms	19
2.2.1.1 Verlet Algorithm	19
2.2.1.2 Leap frog algorithm	20
2.2.2 Periodic boundary condition	22
2.2.3 Water models	23
2.2.4 Implicit solvation	24
2.2.4.1 The Poisson Equation	26
2.2.4.2 The Born model	27
2.2.5 Molecular Mechanics Poisson-Boltzmann Surface Area	28
2.3 Quantum mechanics in brief	29
2.3.1 Independent Particle Approximation	30
2.3.2 The Møller-Plesset perturbation theory	30
2.3.3 The Density Functional Theory	32
2.3.3.1 Local Density Approximation	34
2.3.3.2 Generalized Gradient Approximation	35
2.3.3.3 Hybrid functional	36
2.4 Hybrid QM/MM	36
2.4.1 The QM/MM schemes	37
2.4.2 The model of QM/MM boundary treatment	38
CHAPTER III – SYSTEMS PREPARATION	40
3.1 Molecular dynamics simulation details	40
3.2 Quantum mechanics/molecular mechanics details	41
CHAPTER IV – RESULTS AND DISCUSSION	43
4.1 The dUMP halogen analogue systems	43
4.1.1 Overall enzyme and substrates stability	43
4.1.2 Per-residue decomposition energy scanning	47
4.1.3 Enzyme-substrates interactions	49
4.1.4 Reaction coordinate distance analysis	50

	Page
4.1.5 Ligand Binding affinity.....	52
4.2 The dUMP halogen analogue systems.....	54
4.2.1 Overall enzyme and ligands stability.....	54
4.2.2 Per-residue decomposition energy scanning.....	55
4.2.3 Enzyme-ligands interactions.....	56
4.2.4 Inter molecular distances analysis.....	59
4.2.5 Ligand Binding affinity.....	61
4.3 Mechanism identification of Michael addition and covalent complex formation.....	63
4.4 Activation energy barrier requirement for Michael addition and ternary covalent complex formation in various dUMP analogues.....	66
4.4.1 The effect of QM region size.....	66
4.4.2 The effect of DFT methods.....	68
4.4.3 Activation energy barrier from MP2 and SCS-MP2 calculation.....	70
CHAPTER V – CONCLUSIONS.....	77
REFERENCES.....	78
APPENDIX.....	85
BIOGRAPHY.....	89

LIST OF TABLES

	Page
Table 4.1 Averaged binding free energy (kcal/mol) over 100 snapshots taken from the last 10 ns of the ligand(s) binding to TS enzyme in binary and ternary complexes where the standard deviation is shown in bracket.....	53
Table 4.2 The averaged binding free energy (kcal/mol) over 100 snapshots of the two ligands binding to TS enzyme obtained from the last 17 ns of simulation time where the standard deviation is shown in bracket.....	62
Table 4.3 Activation energies barrier responsible to the Michael addition and ternary covalent complex formation on each dUMP analogues obtained from the SCS-MP2/aug-cc-pVTZ method in comparison with the experimental values.....	74
Table 4.4 The averaged critical distances (d1-d5 defined in Fig. 4.23) responsible to Michael addition, ternary complex formation for reactant, transition state and product structures of all dUMP analogues.....	76

LIST OF FIGURES

		Page
Figure 1.1	The comparison of abnormal cell division (neoplasm) between benign and malignant tumors.....	1
Figure 1.2	The schematic of cancer cell development.....	2
Figure 1.3	An example of esophageal tumor staging in Tumor-Node-Metastasis (TNM) classification system.....	3
Figure 1.4	Chemical structures of folic acid and example of folic acid based anticancer drugs.....	5
Figure 1.5	Chemical structures of purine, purine based nucleotide example of purine based anticancer drugs.....	5
Figure 1.6	Chemical structures of pyrimidine, pyrimidine based nucleotide example of pyrimidine based anticancer drugs.....	6
Figure 1.7	The X-ray crystal structure of <i>E. coli</i> thymidylate synthase homodimer in ternary complex of with THF and dUMP molecules bound.....	7
Figure 1.8	Multi-steps reductive methylation mechanism proposed for thymidylate synthase (TS) enzyme by Carreras and Santi.....	8
Figure 1.9	Comparison of the first three steps of reductive methylation between the normal metabolic process of dUMP (A) and FdUMP analogue (B).....	9
Figure 1.10	The two possible schemes of methylene bridging intermediate formation in (A) stepwise and (B) concerted mechanisms.....	10
Figure 2.1	Bond distance identification between any two atoms.....	13
Figure 2.2	Bond angle identification between any three atoms.....	14
Figure 2.3	Torsion angle or dihedral angle identification between four atoms.....	15
Figure 2.4	Improper dihedral angle or out-of-plane angle identification between four atoms.....	15
Figure 2.5	The Lennard-Jones pair potential energy curve.....	17

	Page
Figure 2.6	The potential energy for bond length minimization compared between (A) Molecular mechanics and (B) Molecular dynamics methods..... 18
Figure 2.7	An example of the PBC approximation in two dimensions in which the central, square, shaded box of side L is replicated in both dimensions..... 22
Figure 2.8	Water molecule models generated by 3, 4 and 5 site of point charge..... 23
Figure 2.9	The simplify example of cavity creation model in continuum solvent treatment..... 24
Figure 2.10	The simply representative model comparison between the accessible surface and molecular surface..... 25
Figure 2.11	An example of QM and MM regions identification..... 37
Figure 2.12	The two different approaches (like atom model and GHO model) for bond termination..... 39
Figure 3.1	The QM region used in QM/MM calculations of the Michael addition and ternary covalent complex formation reactions..... 42
Figure 4.1	The RMSD plot for all heavy atoms and protein backbone of TS and ligand heavy atoms versus simulation time of (A) binary complexes and (B) ternary complexes relative to the starting structure..... 43
Figure 4.2	The crystal structure of TS/dUMP/THF ternary complex (1KZI.pdb) and the protein fragments important for binding of dUMP and THF (stick model) shown in different colors..... 44
Figure 4.3	RMSD versus simulation time of each binding domain fragment compared between the absence (A) and presence (B) of mTHF cofactor inside the TS binding pocket..... 45
Figure 4.4	(A) Chemical structures of dUMP and mTHF. Dynamical flexibility on the dUMP and mTHF fragments by means of RMSD calculation in (B) binary and (C) ternary complexes..... 46

Figure 4.5	Decomposition energy scanning on the pairwise per-residue for all TS residues in binding of dUMP and mTHF for binary and ternary complexes.....	47
Figure 4.6	(A) Decomposition binding free energy scanning in pairwise per-residue basis of all protein residues for binding of dUMP and three XdUMPs in binary and ternary complexes (B) The energy components, electrostatic ($E_{ele}+G_{polar}$) and van der Waals ($E_{vdW}+G_{nonpolar}$) terms, for the nine key residues.....	48
Figure 4.7	(A) Percentage occupation of hydrogen bonding for dUMP (or analogues) and mTHF binding to TS residues of the binary (light gray) and ternary (black) complexes. (B) Close up of dUMP and mTHF (green) in the TS catalytic pocket where the dash and solid lines represent the hydrogen bonding and the intermolecular distances, $d1-d3$, involving to active conformation in the Michael addition, respectively.....	49
Figure 4.8	Distribution plot of the $d1-d3$ intermolecular distances, $P(d)$, sampling from the last 10-ns simulation in the binary (dashed line) and ternary (solid line) complexes.....	51
Figure 4.9	The RMSD plots of all heavy atoms, protein backbone and heavy atoms of the two ligands compared between NH ₂ dUMP, OHdUMP, CNdUMP and NO ₂ dUMP analogues.....	54
Figure 4.10	(A) The pairwise per-residue decomposition free energy scanning of all protein residues with individual mTHF and dUMP analogues (B) The electrostatic ($E_{ele}+G_{polar}$) and van der Waals ($E_{vdW}+G_{nonpolar}$) energy components contributed to the active site residues.....	56

Figure 4.11	(A) The percentage of hydrogen bonding occupation for mTHF and dUMP analogues binding with the TS active site residues (B) The binding complex structure of dUMP analogue and mTHF (blue) inside the TS catalytic pocket where the dash and solid lines represent the hydrogen bonding and the intermolecular distances, <i>d1-d6</i> , involving to active conformation in the Michael addition, respectively.....	57
Figure 4.12	The functional group orientation and hydrogen bonding interactions of the four dUMP analogues (OHdUMP, NH ₂ dUMP, CNdUMP and NO ₂ dUMP) with all surrounded residues.....	58
Figure 4.13	Distribution plot of the selected <i>d1-d6</i> distances, <i>P(d)</i> , along the last 17ns of simulation time.....	60
Figure 4.14	Identification of the reaction coordinate distances responsible to Michael addition (RC1) and ternarycovalent complex formation (RC2).....	63
Figure 4.15	(A) Two-dimensional QM/MM (B3LYP/6-31+G(d)/CHARMM) potential energy surface for Michael addition and covalent complex formation. (B) The simplified overall energy profile along the chemical reaction pathway. In both figures, (I), (II) and (III) correspond to the reactant, transition state and product, respectively, and are depicted in Fig. 4.16.....	64
Figure 4.16	QM/MM (B3LYP/6-31+G(d)/CHARMM) optimized structures of the (I) reactant, (II) transition state and (III) product located in the active site of thymidylate synthase during Michael addition and covalent complex.....	65
Figure 4.17	(A) Conformational structure of dUMP, mTHF and the surrounded residues focused on the thiolate anion of C146 (B) The QM/MM contribution energies in terms of QM, QM without dispersion and MM component for the studied system with three different QM regions along the complexed reaction coordinate.....	67

	Page
Figure 4.18 Comparison of QM/MM optimization energy profiles where the QM region are treated by B3LYP, BLYP and BH&HLYP methods at 6-31+G* basis set.....	68
Figure 4.19 Energy profiles comparison at 6-31+G* basis set between the individual DFT optimized and single point energies from BLYP and BH&HLYP methods on the B3LYP based structures.....	69
Figure 4.20 Single point energy profiles of MP2 and SCS-MP2 methods in cc-pVDZ, aug-cc-pVDZ, cc-pVTZ and aug-cc-pVTZ basis sets using the B3LYP/6-31+G* based structures.....	70
Figure 4.21 Single point energy profiles responsible to the Michael addition and ternary covalent complex formation using SCS-MP2/aug-cc-pVTZ method for the three different snapshots of dUMP, FdUMP and CldUMP ternary complexes.....	71
Figure 4.22 Single point energy profiles responsible to the Michael addition and ternary covalent complex formation using SCS-MP2/aug-cc-pVTZ method for the three different snapshots of NH ₂ dUMP, OHdUMP, CNdUMP and NO ₂ dUMP ternary complexes.....	72
Figure 4.23 The critical distances directly and indirectly responsible to the Michael addition, ternary covalent complex formation in the TS/XdUMP/mTHF ternary complex.....	75

LIST OF ABBREVIATION

2D-PES	=	Two Dimensional-Potential Energy Surface
3D	=	Three Dimention
Ala (A)	=	Alanine
Asp (D)	=	Aspartic
Cys (C)	=	Cysteine
DFT	=	Density Functional Theory
DNA	=	Deoxyribonucleic acid
dTMP	=	2-deoxythymidine-5-monophosphate
dUMP	=	2-deoxyuridine-5-monophosphate
FdUMP	=	2-deoxy-5-fluorouridine-5-monophosphate
MD	=	Molecular Dynamics
MM/PBSA	=	Molecular Mechanics/Possion-Boltzmann Surface Area
MP2	=	Møller-Plesset the second order perturbation theory
mTHF	=	5,10-methylene-5,6,7,8-tetrahydrofolate
PDB	=	Protein Data Bank
QM/MM	=	Quantum Mechanics/Molecular Mechanics
RC	=	Reaction Coordinate
RNA	=	Ribonucleic acid
SCS-MP2	=	Spin Component Scaled - Møller-Plesset the second order perturbation theory
TS	=	Thymidylate Synthase
Tyr (Y)	=	Tyrosine
WAT	=	Water

CHAPTER I

INTRODUCTION

The DNA double helix is fundamentally constructed by sequential combination of nucleotide monomers which is sub-classified to four different types (adenine, cytosine, guanine and thymine) depending on nitrogenous bases. Because the thymine starvation leads to cell apoptosis in according to the thymineless death phenomenon, the thymidylate synthase in thymine synthesis becomes an interesting anticancer target for cancer chemotherapy treatment. Due to the thymidylate synthase required for thymine nucleotide synthesis, fundamental information on intermolecular interactions and binding affinity of mTHF, dUMP substrates and the dUMP analogues as well as catalytic mechanism with the enzyme is very useful to develop the new effective inhibitor against the thymidylate synthase function in future.

1.1 What is cancer?

Cancer is characterized by the out of control in cell growth and division^[1]. Due to the high cell reproduction and proliferation rate, the multiplication results to the formation of a cluster cell mass or tumor consequently. The tumor is classified in two majority types, benign and malignant tumors^[2,3] (Fig. 1.1) depending on its spreading ability.

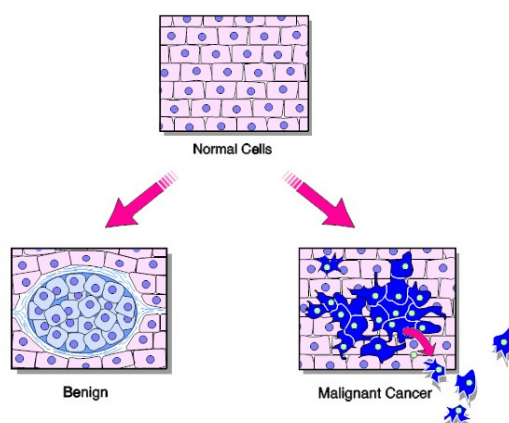


Figure 1.1 The comparison of abnormal cell division (neoplasm) between benign and malignant tumors^[4]

The benign tumor is usually defined as a group of cells which presents the different pattern of cell division from the normal cells (not cancerous). Generally, this tumor does not invade the surrounding tissues or rarely spread out to the other parts of the body. Its border of is well defined by CT scanning and effectively removed by surgery treatment. However, the benign tumor which occurs in a sensitive location, brain tumor for example, also has a serious impact for treatment as well.

Unlike the benign tumor, the malignant tumor (cancerous) invades the surrounding tissues and spreads out to other parts of the body through vascular or lymphatic systems. The process of spreading from one organ to another is called metastasis as simply shown in Fig. 1.2. Although the cancer metastasis has affected to other areas of the body, the initiation organ is still used to identify the type of cancer disease.

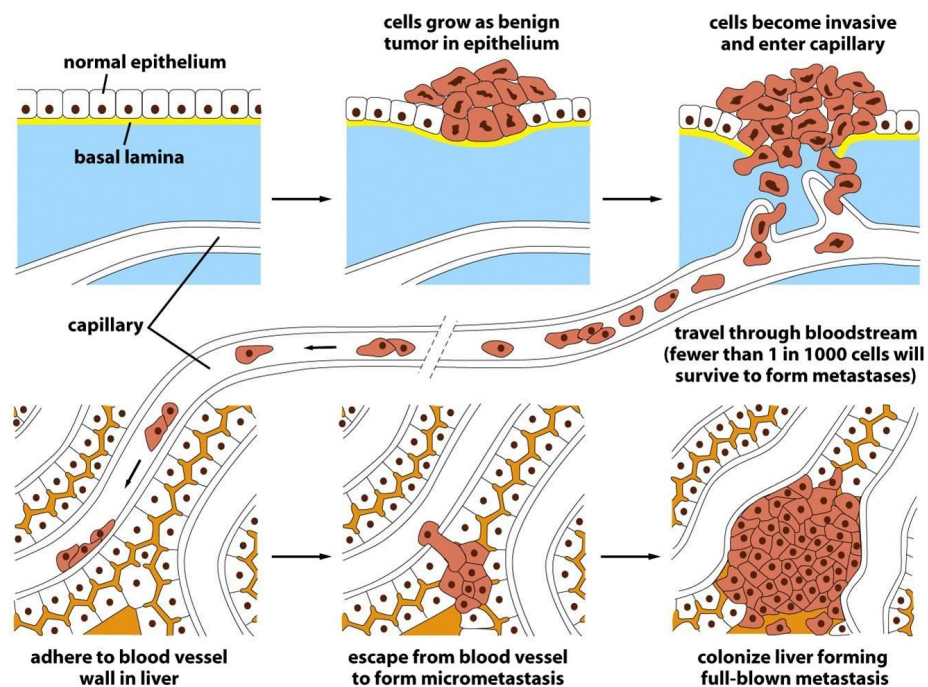


Figure 1.2 The schematic of cancer cell development

1.2 Cancer staging

Cancer staging is the preferably fundamental criteria to identify the severity of cancer disease using location of the primary tumor, tumor size, number of tumors, lymph node involvement and metastasis^[5]. American Joint Committee on Cancer (AJCC)^[6, 7] and International Union for Cancer Control (UICC)^[8] recommends the TNM system for cancer stage guidelines. This system monitors the size of the primary tumor (T), spreading ability to the lymph node (N) and the presence of metastases (M) an important signal for staging identification. The followed number on each letter indicates to the severity level of each parameter (Fig. 1.3).

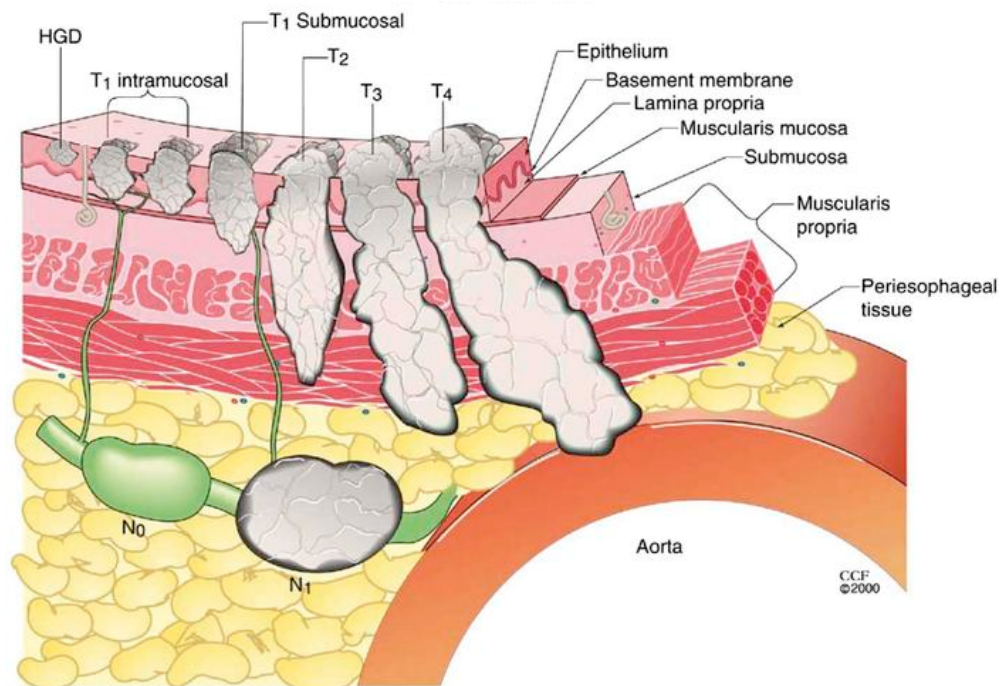


Figure 1.3 An example of esophageal tumor staging in Tumor-Node-Metastasis (TNM) classification system^[9]

1.3 Cancer chemotherapy

Cancer chemotherapy is the administration of cytotoxic agents to control cancer cells where the effectiveness of this treatment depending on the ability of cell division inhibition. Most of the chemotherapeutic drugs effects to some particular phases of the cycle known as the cycle-specific drugs. Therefore, it is necessary to understand the mechanism of anticancer drug inhibition on the specific target in cell division cycle. Till now, anticancer drugs are classified into three different types according to the mechanism mode of inhibition as^[10]

1. DNA building block inhibition (antimetabolite)
2. Directly damage to the DNA
3. Mitosis phase inhibition

Unfortunately, all of the chemotherapy protocol not only kills the cancer cell but it also affects to the healthy cells as well.

1.4 Antimetabolite drug

Antimetabolite agent is the substance that has mimic structures with the natural nucleic acid substrate and co-interferes to the DNA or RNA metabolic pathway during the cell reproduction process^[11]. This similarity allows the antimetabolite agents to serve as a false substrate leading to the un-required product (false nucleobase) and interruption of cell division consequently. In general, antimetabolite agents can be classified in several classes: anti-folate, purine and pyrimidine analogues^[12] where the chemical structures of these analogues are shown in the Fig. 1.4-1.6, respectively.

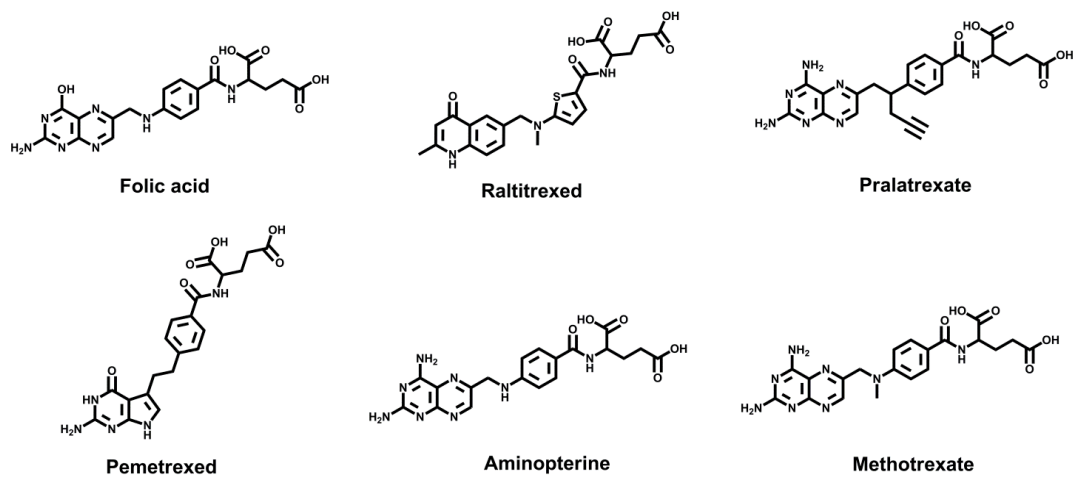


Figure 1.4 Chemical structures of folic acid and example of folic acid based anticancer drugs

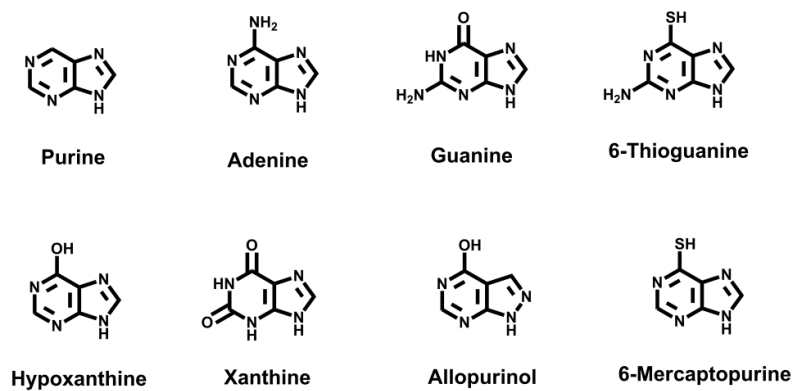


Figure 1.5 Chemical structures of purine, purine based nucleotide example of purine based anticancer drugs

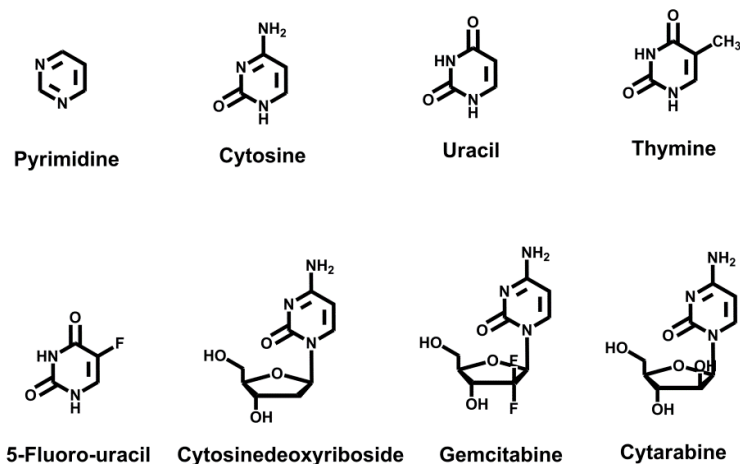


Figure 1.6 Chemical structures of pyrimidine, pyrimidine based nucleotide example of pyrimidine based anticancer drugs

1.5 Overall catalysis of thymidylate synthase enzyme

In DNA replication, the DNA double helix is constructed by sequential combination of deoxynucleotide triphosphate monomer (dATP, dTTP, dCTP and dGTP). Since the thymine starvation results to the thymineless death phenomenon, inhibition of thymine synthesis process has been interested and proposed to be one of the potential targets for cancer controlling. The thymidylate synthase^[13] (TS, Fig. 1.7) has known as an enzyme that catalyzes the reductive methylation of 2-deoxyuridine-5-monophosphate (dUMP) to produce 2-deoxythymidine-5-monophosphate (dTMP) using the 5,10-methylene-5,6,7,8-tetrahydrofolate (mTHF) cofactor^[14-18]. In addition, the relation of H4-folate content and number of dTMP formation was discovered by Humphreys and Greenberg in 1958^[19]. The later studies by Friedkin using tritium isotopic label demonstrated that the tritium-labeled H4-folate was transferred to the methyl group of dTMP product. Therefore, this folate cofactor not only plays as carbon donor but it is confirmed to serve as a reducing agent during the thymine synthesis.

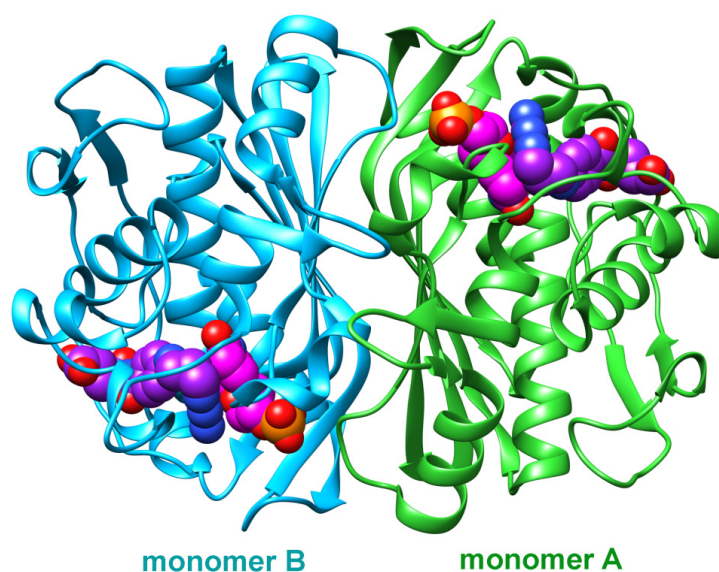


Figure 1.7 The X-ray crystal structure of *E. coli* thymidylate synthase homodimer in ternary complex of with THF and dUMP molecules bound^[20]

1.6 Molecular Mechanism of thymidine synthesis

Although the overall catalyzed reaction is presented by reductive methylation, molecular mechanism of this reaction can be separately considered via multi-step reaction mechanism^[14] (Fig. 1.8). Prior to initiation reaction, the TS enzyme active site is sequentially bound with dUMP and mTHF to generate binary and ternary complexes respectively^[21]. During the non-covalent ternary complex formation, configuration of the mTHF cofactor is substantial adjusted and activated simultaneously with conformational change in the TS binding domain. Interestingly, the mTHF binding into the active site involves the several intrinsic steps which are concomitantly occurred including the non-covalent ternary complex association, enzyme active site conformational closure and mTHF conformational activation^[17, 22, 23].

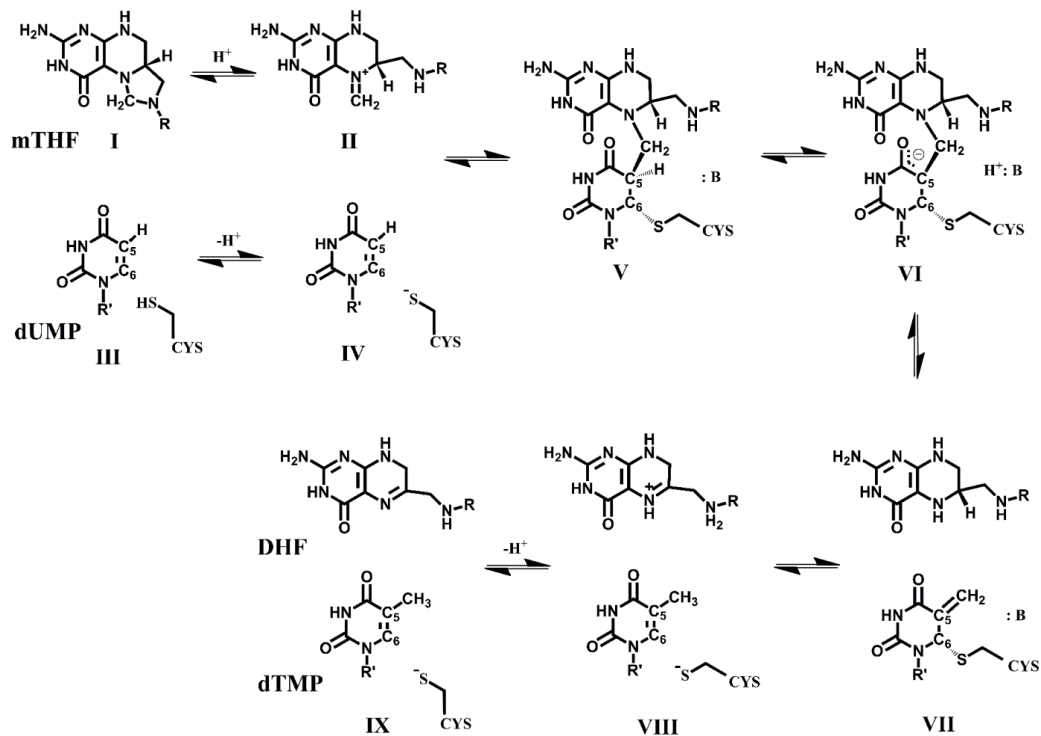


Figure 1.8 Multi-step reductive methylation mechanism proposed for thymidylate synthase (TS) enzyme by Carreras and Santi

It should be noted that the nucleophilic attack (Michael addition) by the reactive cysteine (146) at the C-6 position of dUMP occurs simultaneously with tight binding of the mTHF cofactor into TS protein. Afterwards, the C-5 position of dUMP is activated during Michael addition ongoing which consequently leads to the ternary covalent complex formation between dUMP and mTHF cofactor (intermediate V)^[24]. This step is followed by β -elimination reaction which Y94-proposed base abstracts the C-5 proton of dUMP^[25] (intermediate VI) and generates exocyclic methylene (intermediate VII) consequently. Finally, this intermediate is reduced by a transfer of hydride from the H4-folate to produce dTMP and H2-folate^[24]. This description is widely agreed in the main mechanistic route of overall reaction, however, some of the minor details still open for discussion.

1.7 Inhibition mechanism of the FdUMP and the C-5 substituted analogues

The 5-fluoro-dUMP (FdUMP), a metabolite form of 5-fluorouracil anticancer drug^[26], is an example of the successful dUMP analogue which co-reacts and forms the ternary covalent complex intermediate with the mTHF cofactor in a similar way as the native dUMP substrate^[21, 23] (Fig. 1.9A and B). Differently, the new C-F bonding presents more strength and requires more dissociation energy than the native C-H bonding leading to the termination of β -elimination reaction and its following reaction (Fig. 1.9B). Therefore, the FdUMP molecule is likely trapped in ternary covalent intermediate resulting to an irreversibility of enzyme function finally.

In addition, the fluorine substituent not only interferes to the β -elimination reaction but it also influences to the C-5 atomic charge as well. The change of atomic charge according to the substituent inducement is proposed to influence to the reactivity of Michael addition reaction and ternary covalent complex formation consequently^[27]. Because the ternary covalent intermediate formation is involved by many simultaneous steps, therefore, the description of C-5 substituent effect responsible to the ternary covalent complex formation without any interference by the other intrinsic reactions is limited in an experimental framework study.

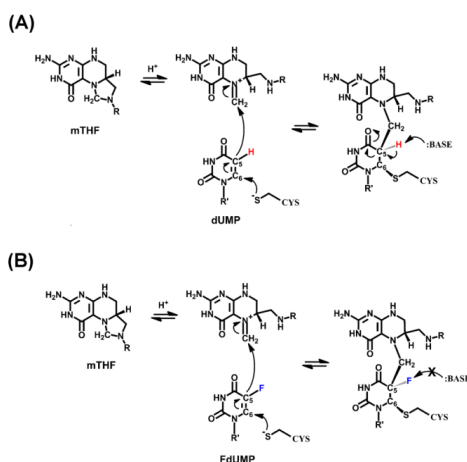


Figure 1.9 Comparison of the first three steps of reductive methylation between the normal metabolic process of dUMP (A) and FdUMP analogue (B)

1.8 Mechanistic indetermination of methylene bridging intermediate formation

Although the methylene bridging intermediate has been isolated and confirmed in experimental investigations^[28, 29], the mechanism for its formation has not been conclusively determined, due to its fast formation^[17, 23, 30]. Therefore, understanding of its reaction mechanistic details are required. After the initiation of the Michael addition step, the thio-ether linkage is formed in concomitantly with the C-5 nucleophilicity increasing which is followed by covalent bonding formation with the methylene of mTHF to form a methylene bridging intermediate. However, catalytic mechanisms for Michael addition and covalent complex formation have been previously proposed in both concerted and stepwise mechanisms^[14, 28, 31-33]. In the stepwise mechanism (Fig 1.10A), it is proposed that a keto-enolate intermediate is formed prior to covalent complex formation, whereas in the concerted mechanism (Fig 1.10B) the keto-enolate intermediate is not formed. Hence, the existence of a keto-enolate intermediate is still undetermined at present.

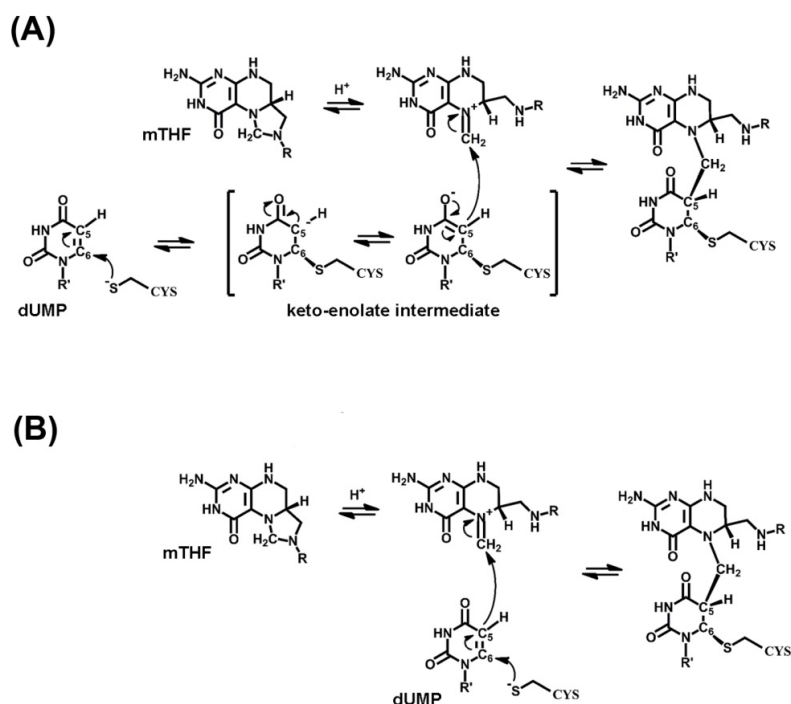


Figure 1.10 The two possible schemes of methylene bridging intermediate formation in (A) stepwise and (B) concerted mechanisms

1.9 Research rationale

As mentioned above, inhibition of the thymine synthesis process becomes an important target in cancer chemotherapy treatment due to the thymineless death phenomenon. The fundamental data of thymidylate synthase and its substrates binding interaction as well as the catalytic mechanism is necessary required for the synthetic chemists and pharmacists who are developing the new effective thymidylate synthase inhibitor. Molecular investigation in an experimental framework is quite limited due to the microscopic problem, therefore, molecular dynamics simulation is the alternative method that is complementary used to investigate molecular basis of binding interaction, binding affinity and thermodynamics stability between the enzyme and substrates as well as its modified analogues. Moreover, the activation energy requirement of each analogue in the critical inhibitory step also provides the kinetic reactivity of reaction according to the change of dUMP functional group in relative to the native dUMP substrate. By following, the analogue which presents the high binding affinity to the enzyme and required the lower activation energy comparing to the native dUMP substrate is potentially proposed as the new effective inhibitor against the thymidylate synthase.

1.10 Objectives

In this work, the acceptable computational chemistry methods such as the molecular dynamics simulation (MD) and hybrid quantum mechanics/molecular mechanics (QM/MM) were selected to investigate the molecular interactions and catalytic mechanism of thymidylate synthase enzyme with mTHF and dUMP substrate as well as the various C-5 functional group analogues of dUMP. The aims of this work cover:

1. To investigate the effect of halogen substitutions (-F, -Cl and -Br atoms) on the C-5 position of dUMP in the complex with and without mTHF (binary and ternary complexes) to non-covalent ternary complex stability prior to the initiation of Michael addition reaction

2. To investigate the effect of electron withdrawing ($-\text{CN}$, $-\text{NO}_2$) and donating ($-\text{OH}$, $-\text{NH}_2$) functional group substitutions on the C-5 position of dUMP to intermolecular interactions and binding affinity toward TS in the presence of mTHF cofactor
3. To determine the molecular mechanism of methylene bridging intermediate formation involved in Michael addition and ternary covalent complex formation reactions
4. To evaluate and compare the required activation energy usage of all dUMP analogues in the Michael addition and ternary covalent complex formation reactions

1.11 Scope of this work

The scope of this study is one-by-one corresponding to the objectives as:

1. MD simulation of the thymidylate synthase enzyme with dUMP substrate compared to its three halogen analogues (FdUMP, CldUMP and BrdUMP) in complex with and without mTHF cofactor
2. MD simulation of the thymidylate synthase enzyme in comparison between the dUMP analogue which is substituted by electron withdrawing ($-\text{CN}$, $-\text{NO}_2$) and donating ($-\text{OH}$, $-\text{NH}_2$) groups
3. Mechanistic identification of the methylene bridging formation in ternary complex by 2-dimensional energy potential surface scanning using B3LYP/6-31+G*-CHARMM hybrid method
4. Activation energy barrier of the Michael addition and ternary covalent complex formation reactions in all dUMP analogues by B3LYP, BLYP and BH&HLYP functional and SCS-MP2 methods

CHAPTER II

THEORITICAL BACKGROUND

2.1 Molecular mechanics

Molecular mechanics (MM) is classical mechanics based model which describes molecular structure by balls-springs connection. The different kind of atoms and bonds are also identified in the meaning of ball size and length respectively. Based on this treatment, molecular properties can be well studied. Molecular mechanics framework assumes the energy of molecule is arisen from geometry changing through bond stretching, bond bending and bond torsional rotation which are simply expressed by classical equation such as harmonic oscillator function. Moreover, the attractive and repulsive inter atomic interactions are also considered by Lennard-Jones and Coulomb equations. All existed constants on these equations are called from database library which is obtained from experimental data or *ab initio* calculation. Therefore, the accuracy of calculated result is mainly depended on the reasonability of database parameters. The total potential energy of a molecule can be written in a summation of all separated interaction terms by this following:

$$E_{total} = E_{bond} + E_{angle} + E_{dihedral} + E_{improper} + E_{ele} + E_{vdW} \quad (2.1)$$

2.1.1 Bond stretching

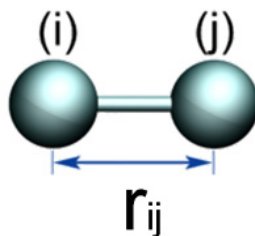


Figure 2.1 Bond distance identification between any two atoms

The bond stretching function, E_{bond} , on each pair atoms can be simply written by followed the Hooke's law for a spring in a harmonic form as:

$$E_{bond} = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 \quad (2.2)$$

Where k_b is the force constant for the bond, r is the actual bond length in the structure between the two atoms defining the bond and r_0 is the equilibrium distance for the bond. The sum runs over all the bonds that have been defined in the system. Because the energy is in harmonic form (Fig 2.2) it means that the energy of the bond will increase steadily without limit as it is distorted from its equilibrium value, r_0 .

2.1.2 Bond bending

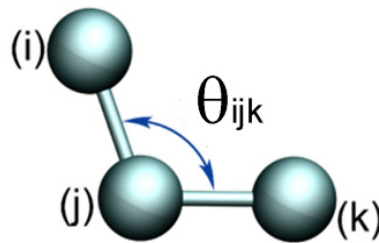


Figure 2.2 Bond angle identification between any three atoms

The E_{angle} term represents the change of energy according to the distortion of bond angle from its equilibrium position, θ_0 . Similar as the bond energy term, the bond angle is also given by the Hookian potential with respect to angle by this following:

$$E_{angle} = \sum_{angle} \frac{1}{2} k_{\theta} (\theta - \theta_0)^2 \quad (2.3)$$

The extra parameter, k_{θ} , is force constant for the angle and the θ_0 is its equilibrium value where the molecule is at rest.

2.1.3 Torsion angle

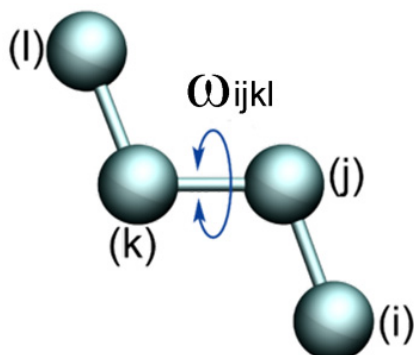


Figure 2.3 Torsion angle or dihedral angle identification between four atoms

The third type of bonding term, $E_{dihedral}$, is the term that describes how the energy of a molecule changes due to the bond rotation. In contrast, harmonic form for dihedral energy is not usually appropriate because the whole range of angles from 0° to 360° may be reproduced in periodic function. Therefore, the dihedral energy can then be written as:

$$E_{dihedral} = \sum_{dihedral} \frac{1}{2} V_n (1 + \cos(n\omega - \delta)) \quad (2.4)$$

2.1.4 Improper

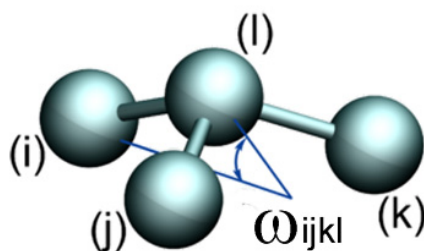


Figure 2.4 Improper dihedral angle or out-of-plane angle identification between four atoms

The fourth term in Equation (2.1) represents the out-of-plane motional energy which is often necessary for planar groups, such as sp² hybridized carbons and aromatic systems due to the planarity maintaining during calculation. The improper dihedral angle is differently defined from dihedral angle because the i, j, k and l atoms are not directly bonded to each other. The calculation of the angle, however, remains similar with the proper dihedral angle.

$$E_{improper} = \sum_{improper} \frac{I}{2} V_n (1 + \cos(n\omega - \delta)) \quad (2.5)$$

2.1.5 Electrostatic energy

An electrostatic energy, E_{elec} , is accounted from electrostatic interactions between the two distributed charge atoms. In molecular mechanics, atomic charge is simply assigned on each individual atom whilst sum of all atomic charges indicates to the total molecular net charge. The electrostatic interactions are described by Coulombic potential function as:

$$E_{elec} = \frac{1}{4\pi \epsilon_0} \sum_{ij} \frac{q_i q_j}{r_{ij}} \quad (2.6)$$

Where q_i and q_j are the fractional charges on atoms i and j and r_{ij} is the distance between the two particles which is run over all pair atoms. The pre-factor term, $1/4\pi\epsilon_0$, is the dielectric constant that will have the value 1 when the system is in vacuum. It should be noted that the partial atomic charge on each atoms are constantly presented during a calculation. Instead of the partial atomic charges, it is possible to define other representations of the charge distribution such as dipole-dipole interaction but the results have not been shown in more significantly accuracy than the simple point-charge model.

2.1.6 van der Waals interaction

The van der Waals interactions are widely used to describe steric interactions of molecule. It is important factor to determine the overall molecular conformation especially in three-dimensional structure of many bio-molecules. Commonly, the Lennard-Jones function which represents sum of repulsive and attractive terms is used for van der Waals evaluation.

$$E_{vdw} = \sum_{i,j \text{ pairs}} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (2.7)$$

The A_{ij} and B_{ij} factors are constant which is depended on the specific type of atom i and j where the repulsive part of the curve is produced by the $1/r_{ij}^{12}$ term and the attractive part by $1/r_{ij}^6$. The shape of the Lennard-Jones potential is example plotted as shown in Fig. 2.5.

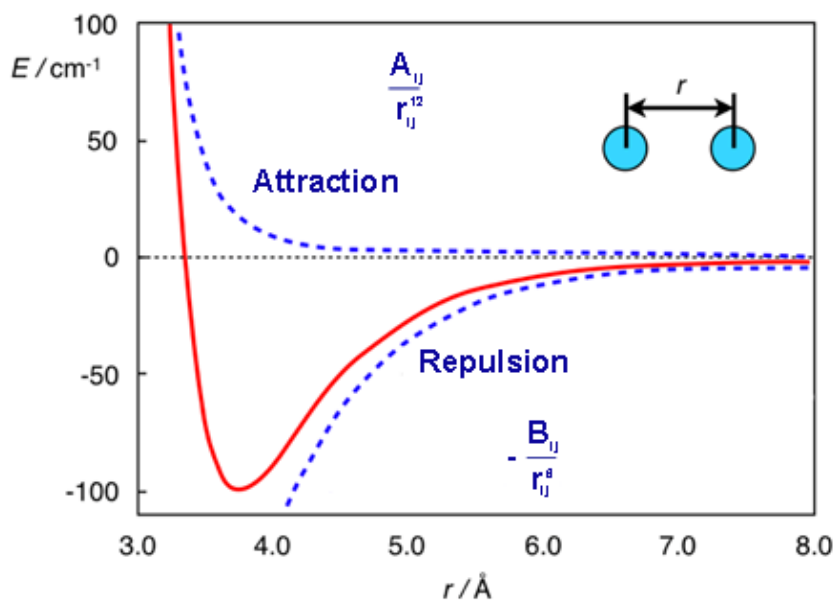


Figure 2.5 The Lennard-Jones pair potential energy curve

2.2 Molecular dynamics

As we have known that molecular motion is essential for all chemical interactions. For example, the binding of substrates to enzymes, the binding of antigens to antibodies, the binding of regulatory proteins to DNA, and the mechanisms of enzyme catalysis are also required the movement of molecule to success its process. Molecular dynamics (MD) simulation is an important method that monitors time-dependent processes in molecular systems in order to study their structural, dynamic, and thermodynamic properties by propagating a set of molecular coordinates and its velocities by numerical solving the Newton's equations of motion.

The difference between molecular mechanics and dynamics is now compared in an example of bond length optimization simply. In molecular mechanics, the bond length would be decreased until its potential energy was reached, Fig. 2.6A. The bond length in molecular dynamics calculation is also decreased continuously until to the equilibrium length. However, the optimum bond length is still oscillated around its equilibrium value instead of static bonding, Fig. 2.6B. In other words, in mechanics the potential energy is minimized, while the kinetic energy of the molecule is ignored. In a dynamics trajectory, both potential and kinetic energy are studied and the total energy is conserved by the motion.

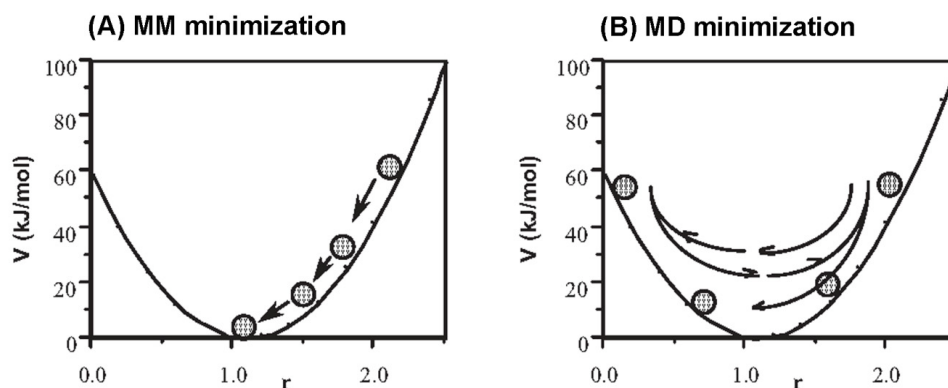


Figure 2.6 The potential energy for bond length minimization compared between (A) Molecular mechanics and (B) Molecular dynamics methods.

2.2.1 The Time Dependent Algorithms

The involved bond stretching, angle bending, dihedral torsions, van der Waals interactions, and electrostatic interactions are described by potential energy function as mentioned in molecular mechanics section. Then, we will further solve molecular motion as a function of time based on the Integrating Newton's Laws of motion concept.

$$F_A = m_A a_A = m_A \frac{dv}{dt} = m_A \frac{d^2 r_A}{dt^2} \quad (2.8)$$

2.2.1.1 Verlet Algorithm

Based on the finite-difference methods, the Taylor expansion of the position at the time $t + \Delta t$ can be written as:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} a(t) \Delta t^2 + \dots \quad (2.9)$$

Where $v(t)$ is the velocity vector and $a(t)$ is the acceleration. If we truncate on the second order and use Taylor expansions corresponding to forward and reverse of the time steps, Δt and $-\Delta t$, we will have:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} a(t) \Delta t^2 \quad (2.10)$$

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{1}{2} a(t) \Delta t^2 \quad (2.11)$$

summation of (2.10) and (2.11):

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2 \quad (2.12)$$

This is known as the Verlet algorithm. If the acceleration is substituted by the force, then:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \left(\frac{F}{m}\right)\Delta t^2 \quad (2.13)$$

These two steps are repeated for every time step for each atom in the molecule. The velocity does not appear in the expression but it can be obtained by subtraction of (2.10) by (2.11).

$$v_A(t) = \frac{r_A(t + \Delta t) - r_A(t - \Delta t)}{2 \Delta t} \quad (2.14)$$

This algorithm is executed in two steps:

1. Use the current position, $r(t)$, to calculate the current force, F , (Eq.2.8)
2. Use the current and previous positions, $r(t)$ and $r(t-\Delta t)$, together with the current force, F , (calculated in step 1) to calculate the next position $r(t+\Delta t)$, (Eq.2.13)

2.1.1.2 Leap frog algorithm

To more improve the Verlet velocity, the leap-frog algorithm is developed by its half-step scheme where the velocities are evaluated at the midpoint of the position and vice versa.

$$v\left(t + \frac{\Delta t}{2}\right) = v(t) + \left(\frac{dv}{dt}\right)\left(\frac{\Delta t}{2}\right) + \frac{1}{2} \left(\frac{d^2v}{dt^2}\right)\left(\frac{\Delta t}{2}\right)^2 \quad (2.15)$$

$$v\left(t - \frac{\Delta t}{2}\right) = v(t) - \left(\frac{dv}{dt}\right)\left(\frac{\Delta t}{2}\right) + \frac{1}{2} \left(\frac{d^2v}{dt^2}\right)\left(\frac{\Delta t}{2}\right)^2 \quad (2.16)$$

Subtract 2.15 by 2.16 and rearrange:

$$v\left(t + \frac{\Delta t}{2}\right) = v\left(t - \frac{\Delta t}{2}\right) + a(t) \Delta t = v\left(t - \frac{\Delta t}{2}\right) + \left(\frac{F}{m}\right) \Delta t \quad (2.17)$$

Using the same procedure for the Taylor expansion of the r_t position at the $t + \Delta t/2$ time point with $\Delta t/2$ time step, we get:

$$r\left(t + \frac{\Delta t}{2} + \frac{\Delta t}{2}\right) = r\left(t + \frac{\Delta t}{2}\right) + v\left(t + \frac{\Delta t}{2}\right)\left(\frac{\Delta t}{2}\right) + \frac{1}{2}a(t) + \frac{\Delta t}{2}\left(\frac{\Delta t}{2}\right)^2 \quad (2.18)$$

$$r\left(t + \frac{\Delta t}{2} - \frac{\Delta t}{2}\right) = r\left(t + \frac{\Delta t}{2}\right) - v\left(t + \frac{\Delta t}{2}\right)\left(\frac{\Delta t}{2}\right) + \frac{1}{2}a(t) + \frac{\Delta t}{2}\left(\frac{\Delta t}{2}\right)^2 \quad (2.19)$$

Subtract 2.18 by 2.19 and rearrange:

$$r(t + \Delta t) = r(t) + v\left(t + \frac{\Delta t}{2}\right)\Delta t \quad (2.20)$$

The (2.17) and (2.20) equations are called the leapfrog algorithm which is reputed to be one of the most accurate and stable techniques for molecular dynamics simulation usage.

In the leapfrog scheme, the atomic position is depended on the velocity which is calculated in the one-half time step out of phase. The velocity are first calculated at time $t + \Delta t/2$ and used to predict the positions of the particles at time $t + \Delta t$. Then, the velocities leap over the positions and the positions leap over the velocities in vice versa.

The leap-frog algorithm is executed in three steps:

1. Use the current position, $r(t)$, to calculate the current force, F , (Eq.2.8)
2. Use the current force, F , and previous mid-step velocity, $v(t-\Delta t/2)$, to calculate the next mid-step velocity, $v(t+\Delta t/2)$, (Eq.2.16)
3. Use the current position, $r(t)$, and the next mid-step velocity, $v(t+\Delta t/2)$, (from step 2) to calculate the position in the next step, $r(t+\Delta t)$, (Eq.2.20)

2.2.2 Periodic boundary condition

One of the most important applications of molecular simulations is the investigated system in condensed phase or solution because most of the chemical and biochemical reaction occurs in solution. Therefore, the system in MD simulation should be extremely solvated by solvent molecules for realistic. Unlike the cluster system, the condensed phase system has an infinite extent. It is very hard to simulate a huge number of solvent molecules in infinite system directly due to the computational limitation. The most widely approximation which is used to represent the environmental effect of the simulated system is periodic boundary conditions (PBC) method. In this approximation, the infinite condensed phase is constructed by periodically repeated by a series of identical finite system itself (Fig. 2.7). Thus, the boundaries region and edge effect is not further found. The most common option which is used for three dimensional extended in PBC approximation is cubic and octahedral boxes.

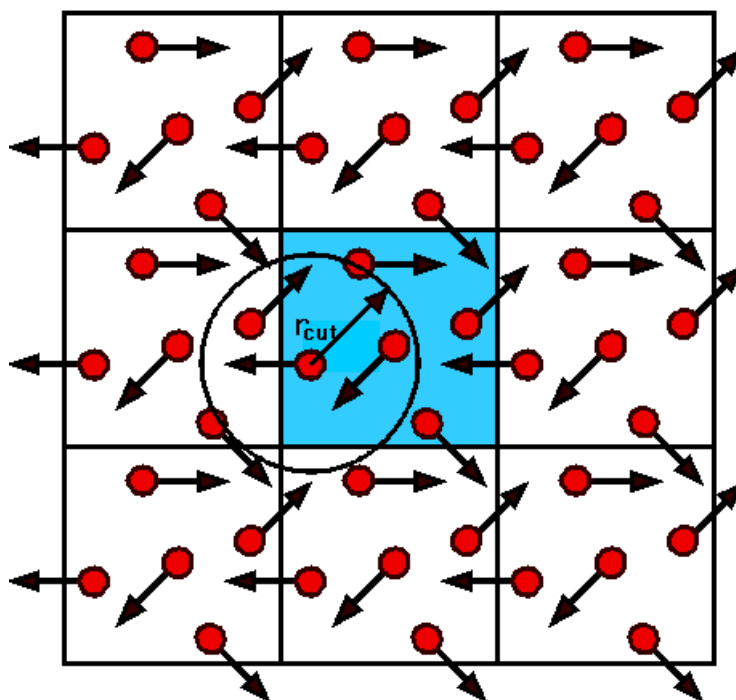


Figure 2.7 An example of the PBC approximation in two dimensions in which the central, square, shaded box of side L is replicated in both dimensions.

2.2.3 Water models

In computational chemistry, an aqueous solution environment is very important for molecular dynamics and statistical mechanics simulations especially for the 3-dimensional hydrogen bonding network formation. The success of these investigations depends on the availability and accuracy of the water-water potential functions resulted to the realistic of aqueous environment representation in our simulation system. In general, all of the proposed water models are developed to fit well with the important parameters of the macroscopic properties such as density, critical parameters and radial distribution function by involving the orientation of electrostatic and Lennard-Jones sites. The simplest water models treat the water molecule as rigid and rely only on non-bonded interactions. The electrostatic interaction is modeled using Coulomb's law and the dispersion and repulsion forces using the Lennard-Jones potential by this following:

$$E_{ab} = \sum_i^{on\ a} \sum_j^{on\ b} \frac{k_c q_i q_j}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{B}{r_{oo}^6} \quad (2.21)$$

where the k_C is electrostatic constant ($332.1 \text{ \AA} \cdot \text{kcal/mol}$); q_i and q_j are the partial charges; r_{ij} is the distance between two atoms or charged sites; and A and B are the Lennard-Jones parameters.

The charged sites may be applied on atoms or dummy sites (such as lone pairs). In most water models, the Lennard-Jones term is only applied for oxygen-oxygen interaction. The figure below shows the general shape of the 3, 4 and 5 site water models (Fig. 2.8). The exact geometric parameters (the O-H distance and the H-O-H angle) are varied and depended on the model individually.

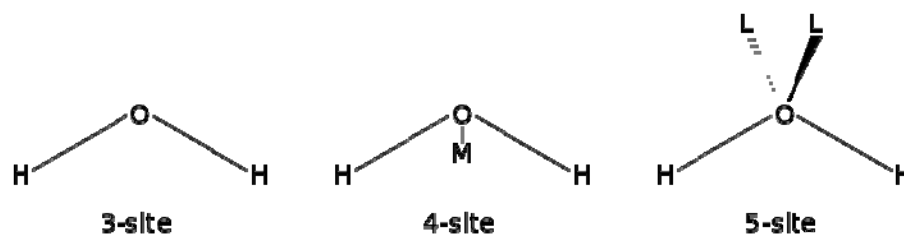


Figure 2.8 Water molecule models generated by 3, 4 and 5 site of point charge

2.2.4 Implicit solvation

The implicit solvation is the simpler model to represent an environmental medium instead of individual “explicit” molecule which is applicable used in MD simulation as well. According to the distributed charge interaction between solute and solvent in an aqueous solution system, therefore, the representing the charge distribution of the solvent explicitly will be simply applied by a continuous electric field during the simulation instead. Then, the energy landscapes obtained with implicit and explicit solvent has possibly difference.

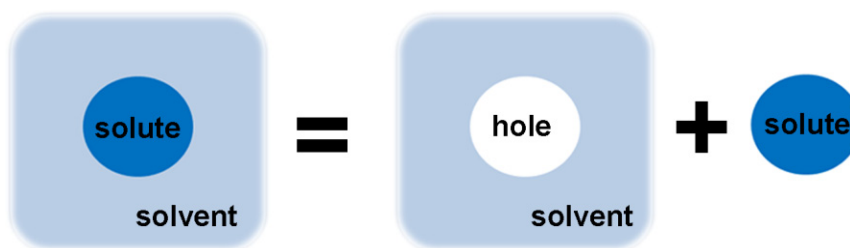


Figure 2.9 The simplify example of cavity creation model in continuum solvent treatment

The continuum model considers all environmental solvents as a uniform medium with dielectric constant, ϵ . The continuity of medium is disrupted (like a hole creation) when the solute is put in (Fig. 2.9). However, creation of the hole requires destabilization energy because the solute may distributes and polarizes to the medium (and in vice versa) through dispersion and electrostatic stabilization. The solvation (free) energy may be written as:

$$\Delta G_{solvation} = \Delta G_{cavity} + \Delta G_{disp} + \Delta G_{elec} \quad (2.22)$$

It should be noted that the dielectric of the medium, ϵ , is normally presented in a constant value depending on the solvent type only (such as: acetone $\epsilon = 20.7$, 1-propanol $\epsilon = 20.1$ or benzene $\epsilon = 2.28$). Therefore, the effect of solute structure to the

solvation is truncated to the cavity surface and dispersion interaction only. The simplest of a cavity model is a spherical shape, however, van der Waals surface of the solute molecule can be considered instead for more realistic. After the surface creation, solvent can solvates around the van der Waals surface of solute. The solvated area is defined by a spherical particle (typical radius of 1.4 Å for water probed model) rolling on its van der Waals surface. This approach is called the Solvent Accessible Surface (SAS) as simply shown in the Fig. 2.10.

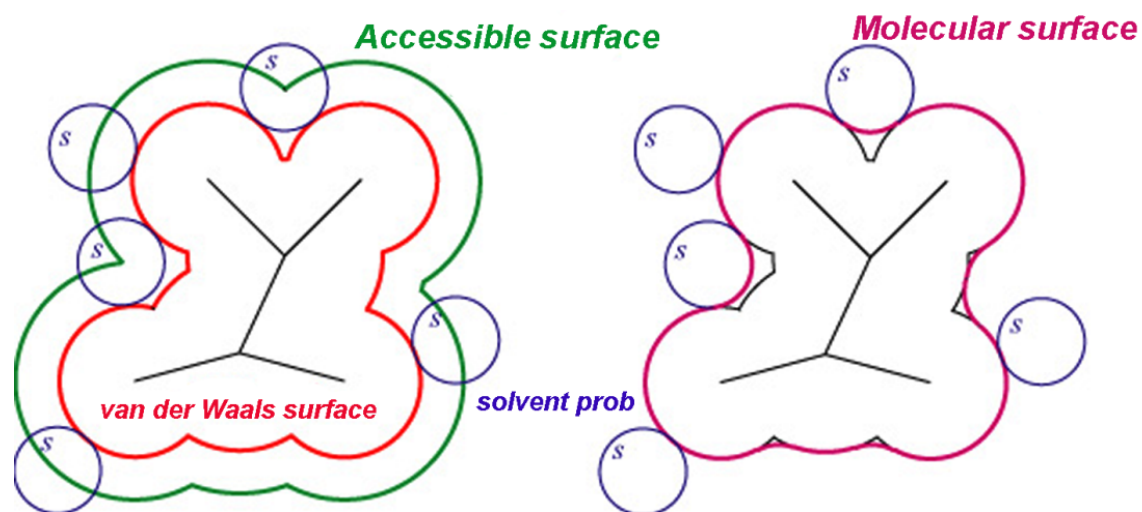


Figure 2.10 The simply representative model comparison between the accessible surface and molecular surface

The required energy for cavity creation is usually assumed to be the proportion of total SAS area creation.

$$\Delta G_{cavity} + \Delta G_{disp} = \gamma SAS + \beta \quad (2.23)$$

2.2.4.1 The Poisson Equation

Most of all continuum solvent models use the Poisson equation to evaluate an electrostatic potential as a function of the charge density and the dielectric constant.

$$\nabla \cdot (\varepsilon(r) \nabla \varphi(r)) = -4\pi \rho(r) \quad (2.24)$$

where ε is the dielectric constant of the medium, $\varphi(r)$ is an electric field at any given position and $\rho(r)$ is charge density of the solute. It should be noted that dielectric in this equation is described as a function of position, however, the electric field in a constant dielectric can be written as:

$$\nabla^2 \varphi(r) = -\frac{4\pi \rho(r)}{\varepsilon} \quad (2.25)$$

In addition, the Poisson equation is valid under conditions of zero ionic strength. The further Poisson–Boltzmann (PB) equation for electrolyte can be applied as:

$$\nabla \cdot (\varepsilon(r) \nabla \varphi(r)) - \kappa^2 \left(\frac{k_B T}{q} \right) \sinh \left[\frac{q \varphi(r)}{k_B T} \right] = -4\pi \rho(r) \quad (2.26)$$

$$\kappa^2(r) = \frac{8\pi q^2 I}{k_B T} \quad (2.27)$$

where q is electrolyte charge, κ^2 is the Debye–Hückel parameter, I is the ionic strength of the solution. Noticeably, all of the equations are represented in a differentiate form which its result describes the electrostatic potential at any point of investigated space. However, the free energy of the electrostatic component can be described by the multiplication of density charge of the solute with the different of electrostatic potential in presence (solution, φ_{sol}) and absence (vacuum, φ_{gas}) of solvent by this following:

$$\Delta G_{ele} = \frac{1}{2} \int \rho(r) (\varphi_{sol} - \varphi_{gas}) dr \quad (2.28)$$

2.2.4.2 The Born model

As we have mentioned that both of the Poisson or Poisson–Boltzmann equations can be used to solve free energy of the electrostatic component clearly. However, the simpler way to calculate free energy of the electrostatic component is spherical cavity model approximation. For the radius spherical cavity of a , net charge of q and dielectric constant of ϵ , the different energy between vacuum and continuum medium based on the Born model is given by

$$\Delta G_{ele}(q) = - \left(1 - \frac{1}{\epsilon} \right) \left(\frac{q^2}{2a} \right) \quad (2.29)$$

The usage of partial charge in Born model equation has been known as generalized Born model (GB). However, it should be questioned from the equation that free energy of the electrostatic component for both positive and negative ions with the same radius are similar in Born model.

2.2.5 Molecular Mechanics Poisson-Boltzmann Surface Area

In biological simulation, the combination of molecular mechanics energy and implicit solvation models, Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA), are widely accepted for binding free energy calculation which obtained from the different energy between the complex and separated free forms.

$$\Delta G_{bind} = \Delta G_{com} - (\Delta G_{rec} + \Delta G_{lig}) \quad (2.30)$$

By following the MM/PBSA (or MM/GBSA) approach, the free energy terms can be more considered as

$$\Delta G_{bind} = \Delta H - T \Delta S \quad (2.31)$$

$$\approx \Delta E_{MM} + \Delta G_{sol} - T \Delta S \quad (2.32)$$

where ΔE_{MM} , ΔG_{sol} and $-T\Delta S$ are the changes of MM energy in vacuum phase, solvation free energy, conformational entropy between the bound and free forms. Furthermore, energy of the MM part can be more separated in several terms followed the molecular mechanics approach.

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} + \Delta E_{improper} + \Delta E_{ele} + \Delta E_{vdW} \quad (2.33)$$

The solvation free energy term can also separated to an electrical and non-polar contribution (cavity) component based on the implicit continuum solvent concept as we have mentioned.

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{SAS} \quad (2.34)$$

2.3 Quantum mechanics in brief

The molecular mechanics method is widely used for a large area of chemical and biological systems by MD simulation approach. However, the limitation of this method is unallowable of an electronic structure change involved in bond-breaking and bond-forming in chemical reaction. Therefore, quantum mechanics (QM) is more required for chemical reaction treatment. By following the Born-Oppenheimer (BO) approximation, electron and nucleus movement can be separately considered according to their high different masses. Then, most of the QM methods have focused on the solving of time-independent Schrödinger equation of electronic structure instead. The typical form of Schrödinger equation with BO approximation can be written by

$$\hat{H}_{el}\Psi_{el} = E_{el}\Psi_{el} \quad (2.35)$$

From this fundamental equation, the total energy (sum of kinetic and potential) is obtained by taking the Hamiltonian operator to the interested molecular wave function. In addition, total energy is contributed from: kinetic of electron and nucleus, the attraction of electron-nucleus, the repulsion of electron-electron and nucleus-nucleus. Together with the Born-Oppenheimer approximation, the Hamiltonian operator is

$$\hat{H}_{el} = \sum_i \left(-\frac{\hbar^2}{2m_e} \nabla^2 \right) - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i < j} \frac{e^2}{r_{ij}} \quad (2.36)$$

where i and j run over electrons, k run over nucleus, \hbar is Planck's constant divided by 2π , m_e is the electron mass, ∇^2 is the Laplacian operator, e is the electron charge unit, Z is an atomic number, and r_{ab} is the distance between any particles a and b . If we work in Cartesian coordinates, the Laplacian has the form

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (2.37)$$

But it can be reduced in an atomic unit as

$$\hat{H}_{el} = \sum_i \left(-\frac{1}{2} \nabla^2 \right) - \sum_i \sum_k \frac{Z_k}{r_{ik}} + \sum_{i < j} \frac{1}{r_{ij}} \quad (2.38)$$

2.3.1 Independent Particle Approximation

The molecular electronic structure can be solved by using Linear Combination of Atomic Orbital model (LCAO) which many-electron system is approximated as the linear combination of independent single electron orbital simply. If Ψ is the molecular orbital function, the LCAO can be written as

$$\Psi = c_1\psi_1 + c_2\psi_2 + c_3\psi_3 + \dots \quad (2.39)$$

Similarly, the Hamiltonian operator of n-electron system is also separate as

$$\hat{H} = \hat{H}(1,2,3,\dots,n) \quad (2.40)$$

By following these approximations, the Schrödinger equation becomes to n-equations of single electron orbital as

$$\hat{H}(i)\psi(i) = \varepsilon(i)\psi(i) \quad (2.41)$$

2.3.2 The Møller-Plesset perturbation theory

Because of an electron-electron correlation is accounted through an average potential in the Hartree-Fock theory, the Møller-Plesset perturbation theory includes that electron-electron correlation by treating the correlation energy as a small perturbation. In principle, the perturbation theory separates Hamiltonian operator into a sum of an exact solution of the corresponding Schrodinger equation and the perturbation term which the later term must be smaller in sense. The Hamiltonian is

$$\hat{H}_\lambda = \hat{H}_0 + \lambda\hat{H}^{(\lambda)} \quad (2.42)$$

where \hat{H}_0 is the approximate Fock-operator in the Hartree-Fock method, $\lambda\hat{H}^{(\lambda)}$ is a small perturbation term, λ is a perturbation parameter and $\hat{H}^{(\lambda)}$ is a perturbation operator. The perturbed wave function, ψ_λ , and energy, E_λ , are expressed in terms of the λ parameter as

$$\psi_\lambda = \psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)} + \lambda^3\psi^{(3)} + \dots \quad (2.43)$$

$$E_\lambda = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \dots \quad (2.44)$$

Because the $\psi_n^{(\lambda)}$ is unknown wave function, then, it is tried by a series of a complete set of unperturbed orthogonal eigenfunction instead.

$$\psi_n^{(1)} = \sum_j a_j \psi_j^{(0)} \quad (2.45)$$

$$\psi_n^{(2)} = \sum_j b_j \psi_j^{(0)} \quad (2.46)$$

In summary, The Hamiltonian operator, wave function and energy state perturbation of the second order ($\lambda = 2$) of the n^{th} state (n^{th} electron) can be finally considered as:

$$H^{(0)}\psi_n^{(2)} + H^{(1)}\psi_n^{(1)} = E_n^{(0)}\psi_n^{(2)} + E_n^{(1)}\psi_n^{(1)} + E_n^{(2)}\psi_n^{(0)} \quad (2.47)$$

$$E_n = E_n^{(0)} + \lambda \langle \psi_n^0 | H^{(1)} | \psi_n^0 \rangle + \lambda^2 \sum_{j \neq n} \frac{|\langle \psi_n^0 | H^{(1)} | \psi_j^0 \rangle|^2}{E_n^{(0)} - E_j^{(0)}} \quad (2.48)$$

$$\psi_n = \psi_n^0 + \lambda \sum_{m \neq n} \frac{\langle \psi_m^0 | H^{(1)} | \psi_n^0 \rangle}{E_n^{(0)} - E_m^{(0)}} \psi_m^0 + \quad (2.49)$$

$$\lambda^2 \sum_{m \neq n} \left\{ \left[\sum_{j \neq n} \frac{\langle \psi_j^0 | H^{(1)} | \psi_n^0 \rangle \langle \psi_m^0 | H^{(1)} | \psi_j^0 \rangle}{(E_n^{(0)} - E_j^{(0)})(E_n^{(0)} - E_m^{(0)})} - \frac{\langle \psi_n^0 | H^{(1)} | \psi_n^0 \rangle \langle \psi_m^0 | H^{(1)} | \psi_n^0 \rangle}{(E_n^{(0)} - E_m^{(0)})^2} \right] \psi_m^0 - \frac{|\langle \psi_m^0 | H^{(1)} | \psi_n^0 \rangle|^2}{2(E_n^{(0)} - E_m^{(0)})^2} \psi_n^0 \right\}$$

2.3.3 The Density Functional Theory

The square of wave function at any position has been known as the probability of electron finding in an infinitesimal volume, dr , or an electron density. Then the total number of electron density can be evaluated by integration of an electron density overall space as this following presented.

$$N = \int \rho(r) dr \quad (2.50)$$

In Hohenberg–Kohn theory, all the properties of a molecule in a ground electronic state are determined by the ground state electron density function. The second is that any obtained electron density function will give higher (or equal) than the true ground state energy. Therefore, the method which will minimize energy of the system must operate on the electronic density. For these concepts, Kohn and Sham have proposed a combination of wave functions and the density approach as

$$E[\rho] = T_0[\rho] + \int [\hat{U}_e(r) + \hat{V}_{ext}(r)] \rho(r) dr + E_{xc}[\rho] \quad (2.51)$$

where $T_0[\rho]$ is an electron kinetic energy in non electron-electron interaction condition. The U_e and V_{ext} terms are electron-electron potential and the external potential energy according to the nucleus field respectively. The $E_{xc}[\rho]$ term is called exchange-correlation energy which includes all of the absent contribution energies such as: Electron correlation and the true kinetic energy of the real $T_e[\rho]$ system.

$$\hat{U}_e(r) = \int \frac{\rho(r')}{|r'-r|} dr' \quad (2.52)$$

$$\hat{V}_{ext}(r) = \sum_a \frac{-Z_a}{|R_a - r|} \quad (2.53)$$

Unfortunately, evaluation of the kinetic term in an electron density is very hard but it is easier done from the wave function. Then, in Kohn-Sham approach, the kinetic term is complementary calculated by wave function based. By following the electron density concept, the kinetic term can be written as:

$$\rho(r) = \sum_{i=1}^N |\varphi_i^{KS}(r)|^2 \quad (2.54)$$

$$T_0[\rho] = \frac{1}{2} \sum_{i=1}^N \langle \varphi_i^{KS} | \nabla_i^2 | \varphi_i^{KS} \rangle \quad (2.55)$$

Furthermore, the exchange-correlation term, $E_{XC}[\rho]$, can also be separated to the exchange, $E_X[\rho]$, and correlation, $E_C[\rho]$, independently.

$$E_{XC}[\rho] = E_X[\rho] + E_C[\rho] = \int \rho(r) \varepsilon_X[\rho(r)] dr + \int \rho(r) \varepsilon_C[\rho(r)] dr \quad (2.56)$$

The various DFT methods are resulted from the different of exchange and correlation function selection which may effects to the different of the calculated energy result consequently. In addition, the treatment of electron-electron correlation between the parallel spin and opposite spin should be considered in the different ways. Therefore, the exchange energy is defined by sum of the α and β spin densities (ρ_α and ρ_β) contribution which the exchange energy is considered on the same spin state only.

$$E_X[\rho] = E_X^\alpha[\rho_\alpha] + E_X^\beta[\rho_\beta] \quad (2.57)$$

$$E_C[\rho] = E_C^{\alpha\alpha}[\rho_\alpha] + E_C^{\beta\beta}[\rho_\beta] + E_C^{\alpha\beta}[\rho_\alpha, \rho_\beta] \quad (2.58)$$

2.3.3.1 Local Density Approximation

The simplest approximation is originated from the *local density approximation, LDA*, based on the uniform electron gas, or equivalently concept which the single local of ρ value is required at every position during the calculation. The exchange energy for a uniform electron gas is given by

$$E_X^{LDA}[\rho] = -C_X \int \rho^{4/3}(r) dr \quad (2.59)$$

The better results of LDA based are obtained by more electron spin assignment on the uniform electron gas model (ρ_α and ρ_β) resulted to unrestricted of LDA model which can be used in an unpaired electron system (like radical) or becoming unpaired electron (electron transfer) systems additionally. This method is called the *local spin density approximation, LSDA*.

$$E_X^{LSDA}[\rho] = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \int (\rho_\alpha^{4/3}(r) + \rho_\beta^{4/3}(r)) dr \quad (2.60)$$

2.3.3.2 Generalized Gradient Approximation

In fact, electron density of a molecule is always varied. Then, the concept of electron density gradient, *generalized-gradient approximation (GGA)*, is included into the LSDA for better results of exchange-correlation term. The most popular GGA exchange function which is related to the LSDA exchange energy is the Beck (B or B88) functional while the gradient corrected correlation is Lee–Yang–Parr, *LYP* functional.

Beck's exchange energy

$$E_X^B[\rho] = E_X^{LDA} - \sum_{\sigma} \int dr \rho_{\sigma}^{4/3} \frac{bX_{\sigma}^2}{1+6bX_{\sigma}\sinh^{-1}X_{\sigma}} \quad (2.61)$$

$$X_{\sigma} = \frac{|\nabla\rho|}{\rho^{4/3}} \quad ; b = 0.0042$$

Lee–Yang–Parr's correlation energy

$$E_C^{LYP}[\rho_{\alpha}, \rho_{\beta}] = -a \int dr \frac{\gamma(r)}{1+d\rho^{-1/3}} \left\{ \rho + 2b\rho^{-5/3} \left[2^{2/3} C_F \rho_{\beta}^{8/3} - \rho t_w + \frac{1}{9} (\rho_{\alpha} t_w^{\alpha} + \rho_{\beta} t_w^{\beta}) \right. \right. \quad (2.62)$$

$$\left. \left. + \frac{1}{18} (\rho_{\alpha} \nabla^2 \rho_{\alpha}) \right] \exp(-c\rho^{-1/3}) \right\}$$

$$\gamma(r) = 2 \left(1 - \frac{\rho_{\alpha}^2(r) + \rho_{\beta}^2(r)}{\rho^2(r)} \right)$$

$$t_w(r) = \frac{1}{8} \frac{|\nabla\rho(r)|^2}{\rho(r)} - \frac{1}{8} \nabla^2 \rho$$

$$C_F = \frac{3}{10} (3\pi^2)^{2/3} \quad ; a = 0.04918, b = 0.132, c = 0.2533, d = 0.349$$

2.3.3.3 Hybrid functional

The most popular hybrid functional (the most popular DFT in fact) is a combination of an exchange-energy of Becke 3 parameters or Becke half-and-half functional and the correlation-energy functional of Lee-Yang-Parr.

$$E_{XC}^{B3LYP} = (1-a)E_X^{LSDA} + aE_X^{HF} + b\Delta E_X^B + (1-c)E_C^{LSDA} + cE_C^{LYP} \quad (2.63)$$

where, a , b , and c are optimized to 0.20, 0.72, and 0.81, respectively.

$$E_{XC}^{H\&HLYP} = \frac{1}{2}E_X^{HF} + \frac{1}{2}E_X^{LSDA} + \frac{1}{2}\Delta E_X^B + E_C^{LYP} \quad (2.64)$$

2.4 Hybrid QM/MM

Usually, computational calculation of the large complex such as the biological system is followed the molecular mechanics model to describe molecular and dynamical properties which evaluates through an empirical force field potential. However, the MM force field is limited to investigate the change of an electronic structure of the molecule during the chemical reaction ongoing, for example, the bond breaking and forming. Therefore, the quantum mechanics is required for a complementary treatment however, calculation with the QM method is also limited to a small system. The algorithm which combines the quantum mechanics and molecular mechanics together (QM/MM) becomes a very useful approach during the past decades. The QM/MM method treats the specific localized region (where chemical reaction is occurred) with QM while the influence of surrounding environment is treated by MM level (Fig. 2.11).

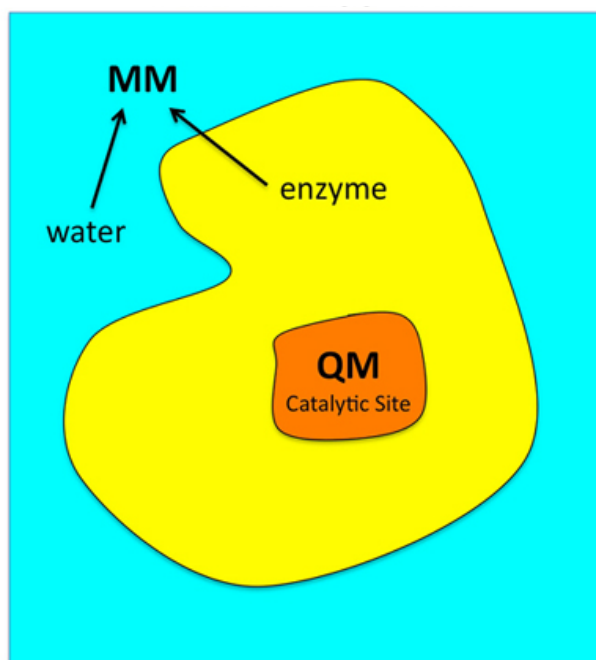


Figure 2.11 An example of QM and MM regions identification

2.4.1 The QM/MM schemes

After the two regions of QM and MM are partitioned, then, energy calculation on these two regions can be employed by using the two alternative schemes: additive and subtractive approaches. In the additive scheme, the total energy of the system is obtained from the summation of each energetic term together.

$$E_{TOTAL} = E_{QM} + E_{MM} + E_{QM-MM} \quad (2.65)$$

From this equation, E_{QM} is energy of the QM region (QM treatment), E_{MM} is the energy of the MM region (force field treatment) and E_{QM-MM} is the interaction between QM and MM region. In addition, the E_{QM-MM} term includes an electrostatic interaction ($E_{ele,QM-MM}$), van der Waals interaction ($E_{vdW,QM-MM}$) and bonded interaction ($E_{bonded,QM-MM}$).

$$E_{QM-MM} = E_{ele,QM-MM} + E_{vdW,QM-MM} \quad (2.66)$$

$$E_{ele,QM-MM} = - \sum_{i \in QM} \sum_{j \in MM} \frac{Q_j}{r_{ij}} + \sum_{A \in QM} \sum_{j \in MM} \frac{Z_A Q_j}{R_{Aj}} \quad (2.67)$$

$$E_{vdW,QM-MM} = \sum_{i \in QM} \sum_{j \in MM} 4\epsilon_{ij} \left\{ \left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^6 \right\} \quad (2.68)$$

In subtractive scheme, the system is divided layers and subtracts double counted energies of the smaller with different energetic calculated method as follows:

$$E_{TOTAL} = E_{QM}^{QM \text{ region}} + E_{MM}^{QM + MM \text{ region}} - E_{MM}^{QM \text{ region}} \quad (2.69)$$

The chosen of either additive or subtractive schemes is not critical whilst the QM region determination should be concerned due to the energy interference between QM and MM regions especially in the case of through bond cutting model.

2.4.2 The model of QM/MM boundary treatment

The boundary treatment between the QM and MM regions is very important in QM/MM method. In addition, the QM region selection can be done by two different ways: with and without covalent bond cutting. In many cases, the QM region is selected by passing through some covalent bonding which is required for some treatment consequently. The link atom method is widely used for QM and MM boundary treatment. Since the covalent bonding is cut in during the QM region selection, then, the “link atom” is introduced as an additional atom which is placed between the cut bonds (Fig. 2.12, left). This link atom is usually used a hydrogen atom due to the simplicity.

The second method for boundary treatment is the localized orbitals usage, the generalized hybrid orbital (GHO) model for example. In GHO, the MM frontier atoms are assigned by four of sp^3 localized hybrid orbitals which the connected frontier QM atom is called the active orbital and the other three orbitals are called auxiliary orbitals (Fig. 2.12, right). Indeed, these four hybrid orbitals are included in the QM calculations but only the active hybrid orbital is participated to the SCF optimization.

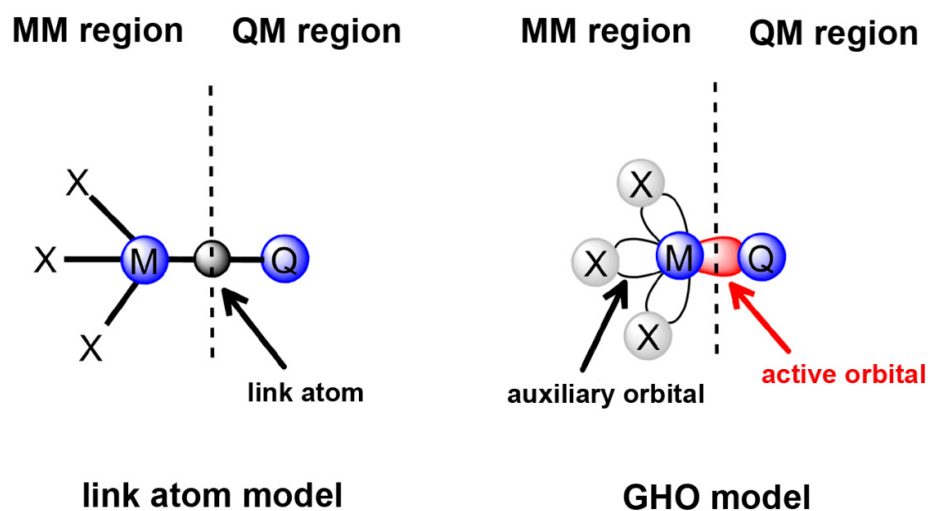


Figure 2.12 The two different approaches (link atom model and GHO model) for bond termination

CHAPTER III

SYSTEMS PREPARATION

To prepare the non-covalent ternary complex of either substrate or dUMP analogues, the crystal structure of TS/dUMP/THF complex obtained from the Protein Data Bank (PDB code: 1KZI)^[20] was selected to be the starting structure. The THF crystal structure was modified to be mTHF by single methylene addition. The three dUMP halogen analogues (FdUMP, CldUMP and BrdUMP) were modeled by single atomic replacement on the C-5 hydrogen of dUMP with fluorine, chlorine and bromine atoms while the electron withdrawing and donating analogues (CNdUMP, NO₂dUMP, OHdUMP and NH₂dUMP) were also modeled by replacement with cyanide, nitro, hydroxyl and amino functional groups respectively. Meanwhile, the mTHF structure in each ternary complex was removed out to generate the relative binary complex.

3.1 Molecular dynamics simulation details

Each system was set up and performed by molecular dynamics simulations in according to our previous works^[34-36]. The ionizable amino acids were assigned at pH 7.0 by PROPKA prediction software^[37, 38]. The lacking hydrogen atoms were added using the LEaP module implemented in AMBER10 software package^[39]. The partial atomic charges of mTHF, dUMP and its analogues were prepared using the standard procedures as follows. The quantum mechanics optimization in Gaussian03^[40] with HF/6-31G(d) basis set was performed for position refinement of all atoms and ESP charge calculation. The obtained charges were then converted to RESP charges using Antechamber in AMBER10. The systems were solvated with TIP3P water molecules^[41] from 10 Å of protein surface in three dimensions and sodium ions were subsequently added in the system for electrical charge neutralization. The periodic boundary condition with isobaric-isothermal (NPT) ensemble under 1 atm of pressure and 310 K of temperature with Berendsen coupling time of 0.2 ps was set. Non-bonded interaction calculation was truncated within a 12 Å residue-base cutoff. The

Particle Mesh Ewald method^[42] was applied to account the long-range electrostatic interactions. The ff03 force-field and GAFF were used for parameterization of protein and ligands, respectively. All of the covalent bonding involving the hydrogen atoms was constrained by 2 fs step size with SHAKE algorithm^[43] along the simulations.

All water molecules were relaxed with 2000 steps of steepest descent (SD) and conjugated gradient (CG) minimization methods whereas the protein and ligand coordinates were fixed. Afterwards, all restrained forces were removed out and whole system was fully minimized by 2000 steps of SD and CG. The system was heated up to 310 K over 200-ps simulation and pre-equilibrated for 300 ps with position restrained on the two ligands and bridging water molecules inside the protein binding pocket with 50 kcal/mol·Å² factors to maintain their positions. The restrained force on the two ligands were reduced to 30, 20, 10, 5, 2 and 0 kcal/mol·Å², and 30, 30, 20, 10, 5 and 2 kcal/mol·Å² for bridging water molecules in every 250 ps. Afterwards, all atoms were freely moved without any restrained force until the 20 ns of simulation time.

3.2 Quantum mechanics/molecular mechanics details

The TS/XdUMP/THF complexed of all dUMP analogues in QM/MM calculation were also prepared in the same procedure with in previous MD system preparation^[44-46]. The systems were solvated using the TIP3P water model and subsequently simulated for 5 ns under stochastic boundary conditions with a truncated sphere of radius 25 Å, centered on the iminium nitrogen of mTHF. Simulations were performed with the CHARMM27 force field^[47, 48] using the CHARMM software package^[49] (version 30b2). The equilibrated structure energy minimized without any reaction coordinate constrain using QM/MM approach where the C146 and A166 side chains, the pteridine fragment of mTHF, and the uracil and deoxyribose fragments of dUMP were included in the QM region (61 atoms, Fig. 3.1). Four H-link atoms^[50] were introduced between the QM and MM boundary. QM calculations were performed using the B3LYP density functional^[51-53] with the 6-31+G(d) basis set using Jaguar 5.5.^[54] Single point energies were calculated for all optimized geometries responsible to the reaction pathway by MP2 and SCS-MP2^[55] at cc-pVTZ

level. The rest of the system was treated with the CHARMM27 force field in the Tinker program^[56]. The QM and MM output created during the optimizations was coupled using QoMMa^[57].

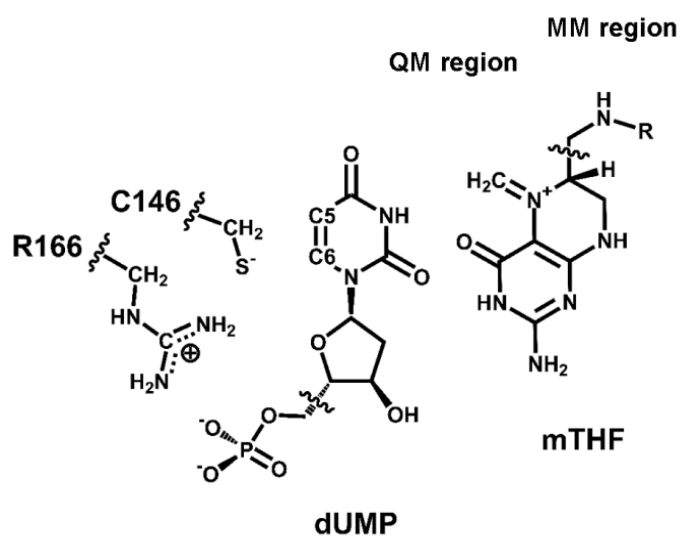


Figure 3.1 The QM region used in QM/MM calculations of the Michael addition and ternary covalent complex formation reactions

CHAPTER IV

RESULTS AND DISCUSSION

4.1 The dUMP halogen analogue systems

4.1.1 Overall enzyme and substrates stability

To monitor the stability of simulated systems, the root mean square displacements (RMSD) along the simulation time of all heavy atoms, protein backbone atoms and ligand heavy atoms is measured and plotted in Fig. 4.1.

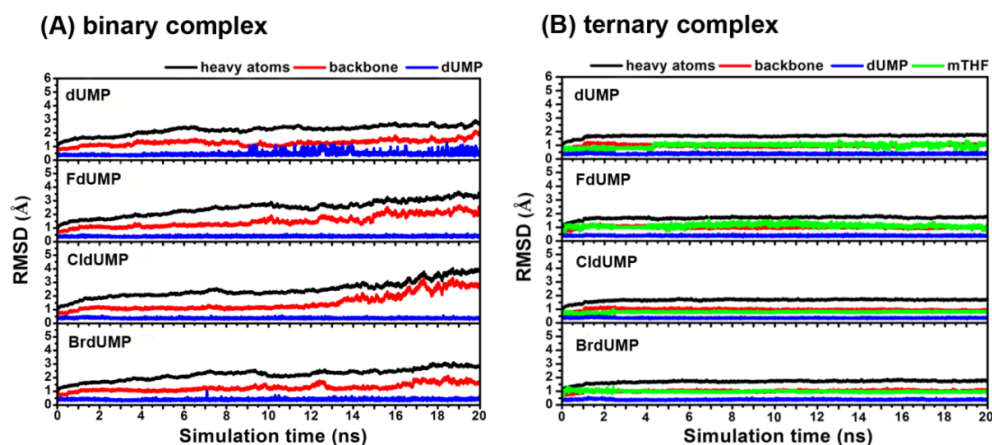


Figure 4.1 The RMSD plot for all heavy atoms and protein backbone of TS and ligand heavy atoms versus simulation time of (A) binary complexes and (B) ternary complexes relative to the starting structure

The RMSD of all heavy atoms and backbone in binary complexes (Fig. 4.1 (A)) are presented in a slightly increased curve in a range of 1.0-2.5 Å during the first 12 ns for all systems. After the 14 ns, the plots likely fluctuate in an extended range especially for the FdUMP and CldUMP complexes. This is contrast for all ternary complexes (Fig. 4.1 (B)), these plots are much less fluctuated and the systems have reached equilibrium after 1ns. Meanwhile, the plots of dUMP and its analogues in both binary and ternary complexes are shown in a very narrow fluctuation range

less than 1.0 Å till the end of simulation. The low fluctuation of mTHF is similar to that of dUMP (only 0.5 Å higher RMSD values). We can summarize from the results that including of the mTHF cofactor into the binding site leads to an increase in the protein stability noticeably. Therefore, dynamical property of the TS enzyme in absence of mTHF cofactor (binary complex) may not be assumed to describe the ternary complex behavior reasonably in contradictory to the previous study^[27].

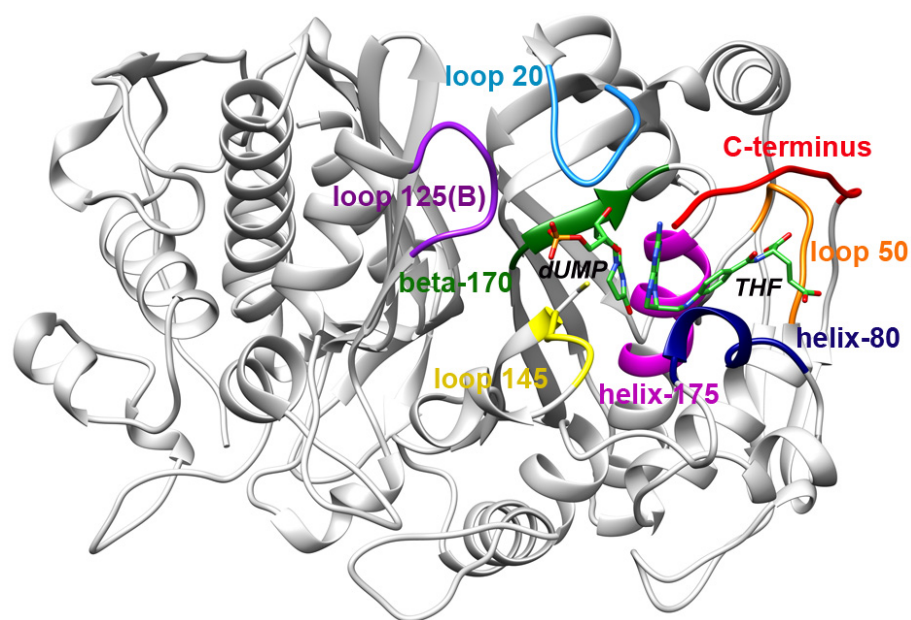


Figure 4.2 The crystal structure of TS/dUMP/THF ternary complex (1KZI.pdb) and the protein fragments important for binding of dUMP and THF (stick model) shown in different colors

Furthermore, the important fragments (five flexible loops, two helices and a beta sheet) involved in the binding sites for dUMP and mTHF are shown in Fig. 4.2. To explore the effect of mTHF binding on TS active site which may involve to the overall protein flexibility, the fluctuation on these fragments considered in term of RMSD plot is compared between binary and ternary complexes in Fig. 4.3.

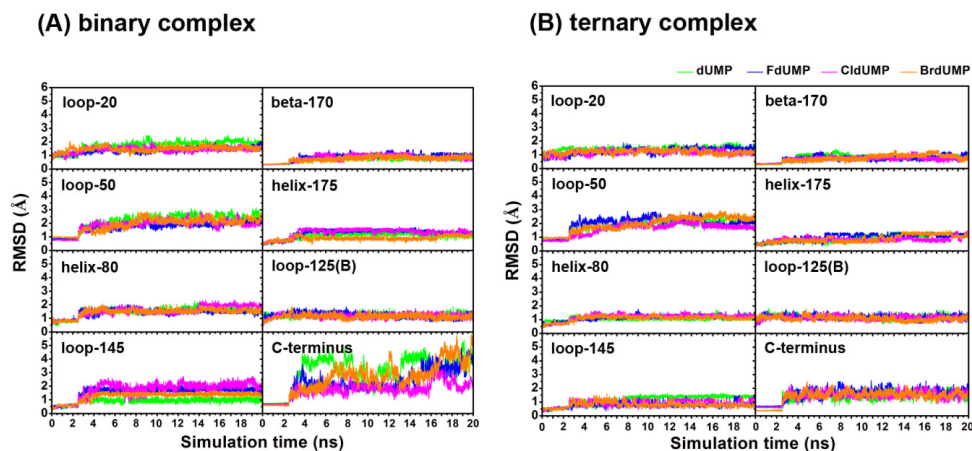


Figure 4.3 RMSD versus simulation time of each binding domain fragment compared between the absence (A) and presence (B) of mTHF cofactor inside the TS binding pocket

In according to the plots, the RMSD profiles can be simply classified in three different patterns depending on the magnitude of fluctuation: low (0-1.5 Å), medium (1.5-2.5 Å) and high (> 2.5 Å) flexibilities. Noticeably, the fluctuation of C-terminal fragment is found in a relatively high flexibility along the fully unrestrained simulation times for all binary complexes, while it is presented in a medium flexibility for ternary complex instead. From the X-ray structure of TS/dUMP/THF ternary complex, the C-terminus is not only specified as the binding gateway to support the incoming substrates but it also plays an important role to stabilize the folate cofactor during the ternary complex formation as well. Therefore, the presence of the mTHF cofactor in the binding pocket can reduce the flexibility of C-terminal fragment significantly compared with the absent systems. The loop-20, helix-80 and loop-145 show a medium flexibility in binary complex but more rigidity in ternary complex, whereas the fluctuation of rest fragments is likely similar in both complexes. Since all systems including ligand binding fragments (Figs. 4.3 (A) and (B)) seem to reach the equilibrium at 10 ns, their MD trajectories from the last 10 ns are collected for intermolecular interaction and binding free energy analysis.

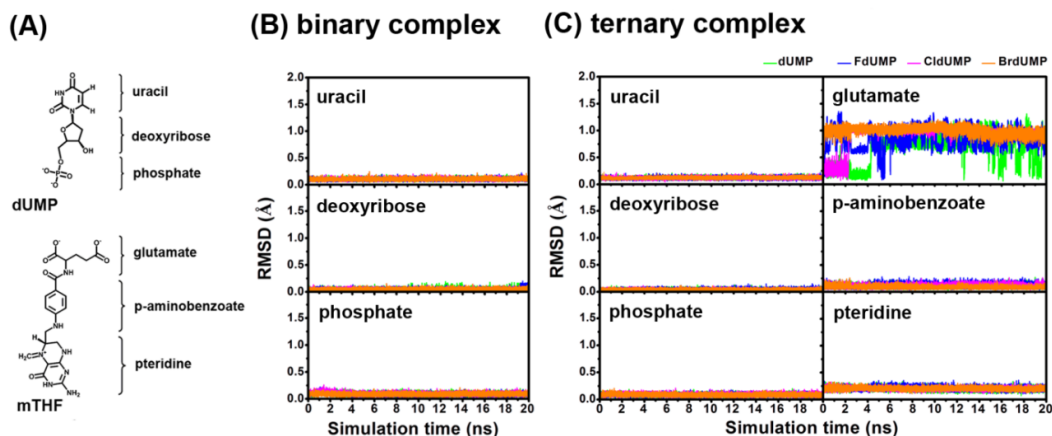


Figure 4.4 (A) Chemical structures of dUMP and mTHF. Dynamical flexibility on the dUMP and mTHF fragments by means of RMSD calculation in (B) binary and (C) ternary complexes

The dUMP contains three main components uracil base, deoxyribose and phosphate while mTHF have pteridine, p-aminobenzoate and glutamate (Fig. 4.4 (A)). The RMSD of their heavy atoms in binary and ternary complexes are plotted and compared in Fig. 4.4 (B) and (C) for investigation on conformational flexibility. It can be seen that fluctuation of all three dUMP components is shown in a very low magnitude in the systems with/without mTHF cofactor binding. For mTHF in ternary complex, high conformational flexibility is dominantly found in the glutamate while the others are in the range with the those of dUMP. This indicates that more structural flexibility of mTHF (Fig. 4.1 (B)) is mostly distributed from glutamate positioned on protein surface (Fig. 4.2) which is fully solvated by waters. Interestingly, the very high flexibility of this component has no influence to the pteridine and PABA conformations. Taken together, the dUMP well occupies inside the binding pocket in both presence and absence of mTHF, while the reactive mTHF pteridine for the methylene bridging formation after Michael addition is unlikely affected by the glutamate fluctuation.

4.1.2 Per-residue decomposition energy scanning

To scan the important residues for ligand binding, the per-residue energy contribution is calculated using MM-GBSA decomposition (DC) energy. The DC energies of all protein residues for ligand binding between binary and ternary complexes are shown in Fig. 4.5.

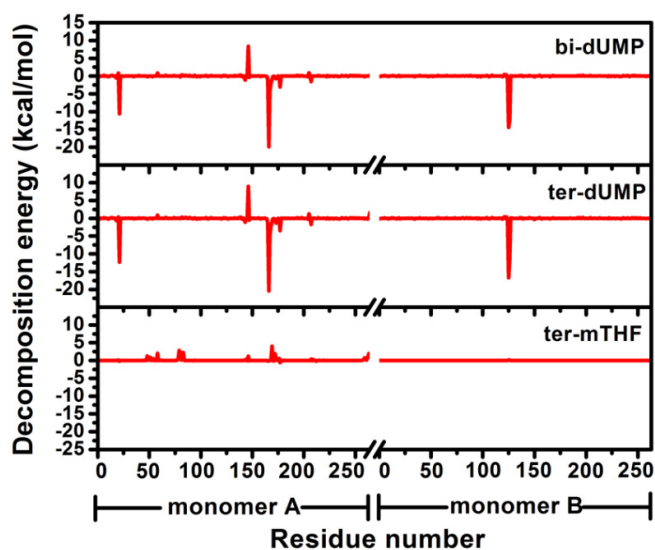


Figure 4.5 Decomposition energy scanning on the pairwise per-residue for all TS residues in binding of dUMP and mTHF for binary and ternary complexes

It can be clearly seen that most of the energy stabilization is found for dUMP binding while no stabilization is detected for mTHF. Therefore, only dUMP (as well as its analogues) is further analyzed and compared in Fig. 4.6 (A).

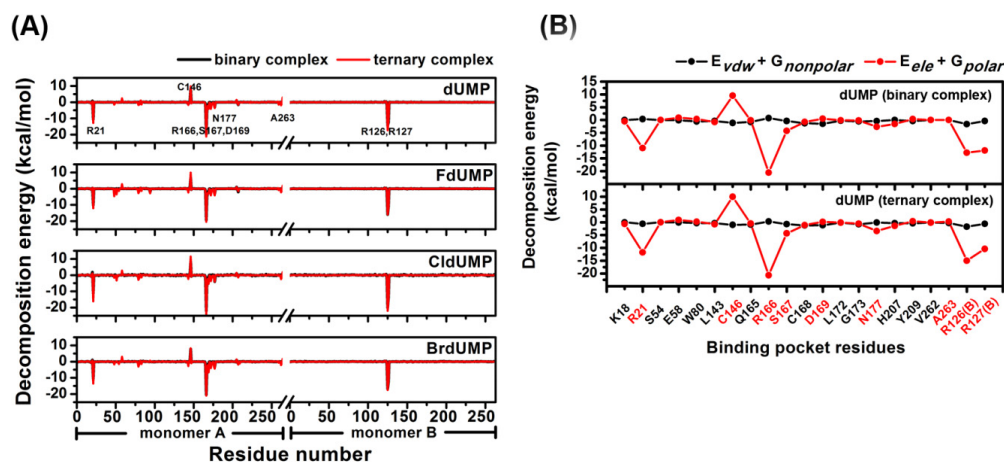


Figure 4.6 (A) Decomposition binding free energy scanning in pairwise per-residue basis of all protein residues for binding of dUMP and three XdUMPs in binary and ternary complexes **(B)** The energy components, electrostatic ($E_{ele} + G_{polar}$) and van der Waals ($E_{vdw} + G_{nonpolar}$) terms, for the nine key residues

The overall per-residue energy pattern for dUMP and XdUMP analogs is likely similar in either binary or ternary complexes, *i.e.*, the main stabilization is contributed from the four arginines (R21, R166, R126(B) and R127(B)) dominantly with < -10 kcal/mol and the S167, D169, N177 and A263 with ~ -5 kcal/mol. In contrast, the reactive C146 residue in negatively charged form provides energy destabilization about ~ 10 kcal/mol due to strong repulsive interaction with phosphate group of dUMP and its analogues. The decomposition energy of these nine important residues towards binding of dUMP as the representative ligand can be considered in terms of electrostatic and van der Waals energies including solvation effect (Fig. 4.6 (B)). As seen from the plots, the dUMP mainly binds to TS enzyme through electrostatic interaction whilst the van der Waals interaction is rather weak.

4.1.3 Enzyme-substrates interactions

Because the electrostatic interaction is highly importance in ligand binding in the focused system and the hydrogen bonding is basically electrostatic interaction, the hydrogen bonding formation between ligand and TS residues is meaningful. To access such H-bonding, the percentage of hydrogen bonding occupation is monitored by following the two criteria: (i) distance between proton donor (D) and acceptor (A) atoms $\leq 3.5 \text{ \AA}$; and (ii) angle of D—H...A $> 120^\circ$. The results for all binary and ternary systems are plotted in Fig. 4.7 (A) where the hydrogen bonding pairs are represented by dash line in Fig. 4.7 (B).

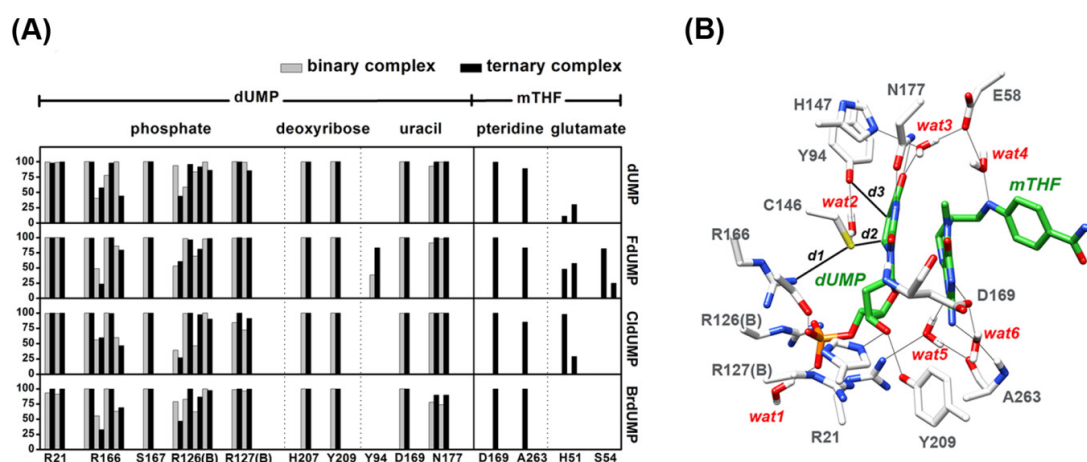


Figure 4.7 (A) Percentage occupation of hydrogen bonding for dUMP (or analogues) and mTHF binding to TS residues of the binary (light gray) and ternary (black) complexes. **(B)** Close up of dUMP and mTHF (green) in the TS catalytic pocket where the dash and solid lines represent the hydrogen bonding and the intermolecular distances, $d1-d3$, involving to active conformation in the Michael addition, respectively

By comparison between dUMP and mTHF in ternary complexes, the number of hydrogen bonding formation is substantially higher detected for dUMP (at least 18 hydrogen bonds with 10 residues). The hydrogen bonding patterns with dUMP are likely similar in both binary and ternary systems and in correspondence with the DC results. The main stabilization is found at its phosphate moiety according to the strong salt-bridge interaction with the four surrounded arginines (R21, R166, R126(B) and R127(B)) and a serine S167. Whereas lower number of hydrogen bond occupation are presented in its deoxyribose (H207 and Y209) and uracil (D169 and N177) moieties. Interestingly, an additional H-bond is formed between the substituted fluorine of FdUMP and Y94 in which is considerably stronger by 40% in ternary complex (to be 80 %).

Among the three fragments (pteridine, PABA and glutamate) of mTHF, no hydrogen bonding is formed at PABA due to its non-polar aromatic ring. The pteridine firmly makes two hydrogen bonding interactions with D169 and A263. Note that the D169 residue strongly forms hydrogen bonding with both dUMP and mTHF (nearly 100 %) leading to the conformational trapping in π - π stacking between the uracil of dUMP and the pteridine of mTHF consequently. The significance of this residue is supported by the deficiency of the ternary complex formation in the mutated systems (D221 numbering in *L. casei*)^[58]. The terminal glutamate of mTHF positioned on the protein surface is partially stabilized by two residues: H51 in all ternary systems except for BrdUMP, and S54 only in FdUMP.

4.1.4 Reaction coordinate distance analysis

The substitution on C5 position of dUMP with electron withdrawing groups was previously proposed by QSAR study to increase the reactivity of Michael addition reaction in the TS enzyme^[27]. To investigate the effect of halogen substitutions at this position on the active conformation of binary and ternary complexes for Michael addition, the distribution of three intermolecular distances: $d1$ [S(C146)-NH2(R166)] between the reactive C146 and the stabilized R166; $d2$ [S(C146)-C6(dUMP)] responsible to the focused reaction; and $d3$ [O(Y94)-

C5(dUMP)] relatively corresponding to a unique hydrogen bonding with Y94, is monitored and plotted in Fig. 4.8 (see Fig. 4.7 (B) for $d1$ - $d3$ definition).

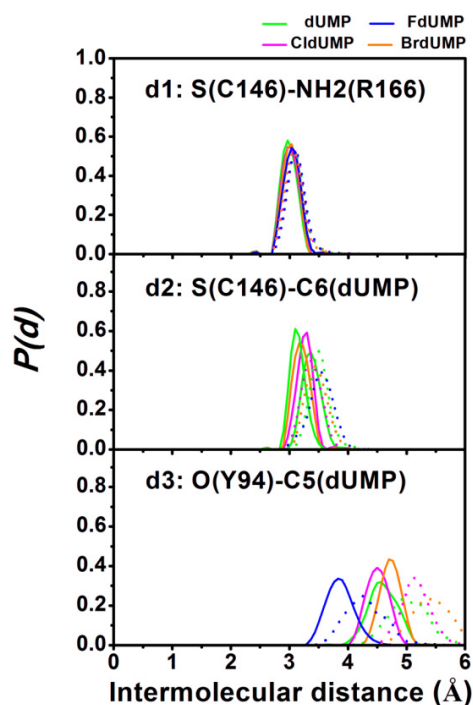


Figure 4.8 Distribution plot of the $d1$ - $d3$ intermolecular distances, $P(d)$, sampling from the last 10-ns simulation in the binary (dashed line) and ternary (solid line) complexes

Prior to Michael addition, the reactive C146 is firstly activated by R166^[59] and on the other hands this arginine plays an role to stabilize the negative C146 via the strong charge-charge interaction. This is well supported by the narrow and sharp peak of $d1$ at ~ 3.2 Å for all systems. From the $d2$ plot which is directly involved to Michael addition, the most probable finding in $d2$ from all ternary complexes is presented in a range of 2.7-3.7 Å (solid lines) whereas the absence of mTHF leads to more movement of dUMP as seen by a slightly shift to 4.0 Å in binary complexes (dashed lines). The similar range of $d2$ values in all systems suggests the Michael addition on the C-6 position of the dUMP analogue is possible but may differ in the activation energy of reaction depended on the nature of the chemical structure. In addition, the

unique hydrogen bonding interaction between the substituted fluorine and Y94 hydroxyl group in the FdUMP systems is confirmed by the shortening of the related distance, $d3$ (blue). This is thus the substitution on the C-5 position of dUMP which can form a strong hydrogen bond with Y94 may lead to an increase in ligand binding affinity with the TS target.

4.1.5 Ligand Binding affinity

The MM/PBSA module implemented in AMBER was used to evaluate the binding strength of individual dUMP/XdUMP and both mTHF and dUMP/XdUMP with TS enzyme in binary and ternary systems, respectively, and the results are given in Table 4.1. The X-ray structure of TS/dUMP/THF ternary complex shows three water molecules bridged between each ligand and residues which may feasibly facilitate the orientation of ligand(s) to be well adjusted in the binding pocket. To reveal this point, the free energy of ligand binding to the TS receptor embedded with/without the bridging waters is calculated and compared in the same table. It can be seen that inclusion of bridging waters can increase the binding free energy of ligand(s) by > 10 kcal/mol in all binary and ternary complexes, suggesting the important role of bridging waters in stabilization of ligand binding. Based on predicted binding free energies (kcal/mol), the order of binding affinity for ternary systems is of: Cl dUMP (-100.3) \sim FdUMP (-96.8) $>$ dUMP (-86.3) $>$ BrdUMP (-77.8). Interestingly, the binding free energy order observing in the binary complexes are presented in a little different pattern (dUMP \sim FdUMP \sim Cl dUMP $>$ BrdUMP) with reduced energies relative to those of ternary complexes. Moreover, the dUMP or its analogues binds dominantly stronger than the mTHF by c.a. 3 times in correspondence with the per-residue decomposition free energy and hydrogen bonding interactions.

Table 4.1 Averaged binding free energy (kcal/mol) over 100 snapshots taken from the last 10 ns of the ligand(s) binding to TS enzyme in binary and ternary complexes where the standard deviation is shown in bracket.

receptor	ligand	binary complex				ternary complex			
		dUMP	FdUMP	CldUMP	BrdUMP	dUMP	FdUMP	CldUMP	BrdUMP
TS	dUMP*, mTHF	-	-	-	-	-71.9 (10.3)	-77.8 (8.9)	-82.8 (11.1)	-65.4 (6.7)
TS, wat1-6	dUMP*, mTHF	-	-	-	-	-86.3 (11.0)	-96.8 (9.6)	-100.3 (10.7)	-77.8 (11.9)
TS	dUMP*	-	-	-	-	-76.4 (11.2)	-87.1 (9.4)	-88.6 (11.7)	-69.9 (10.8)
TS	mTHF	-	-	-	-	-19.8 (6.1)	-28.1 (5.2)	-26.7 (5.4)	-22.3 (6.3)
TS	dUMP*	-59.5 (6.6)	-59.1 (7.4)	-59.8 (7.5)	-53.0 (8.8)	-	-	-	-
TS, wat1-3	dUMP*	-70.9 (6.7)	-72.5 (8.3)	-70.3 (10.1)	-55.5 (8.7)	-	-	-	-

*dUMP or XdUMPs

4.2 dUMP electron withdrawing and donating analogues

4.2.1 Overall enzyme and ligands stability

The dynamical stability of all simulated systems during the simulation is monitored by the root mean square displacements (RMSD) separated in all heavy atoms, protein backbone atoms and heavy atoms of ligand as plotted in Fig. 4.9. Considerable to the all heavy atoms and protein backbone, the curves are noticeably reached to equilibrium after the 2-ns in all systems which are constantly presented in a range of 1.5-2.0 Å and 1.0 Å respectively. The two ligands fluctuation is also found in very narrow range at nearly 1.0 Å and 0.5 Å for mTHF cofactor and dUMP analogues respectively whereas the equilibrium geometries of these two ligands seem to present after the 3-ns instead. Therefore, the MD trajectories since the 3-ns are selected for intermolecular interaction and binding free energy analysis further.

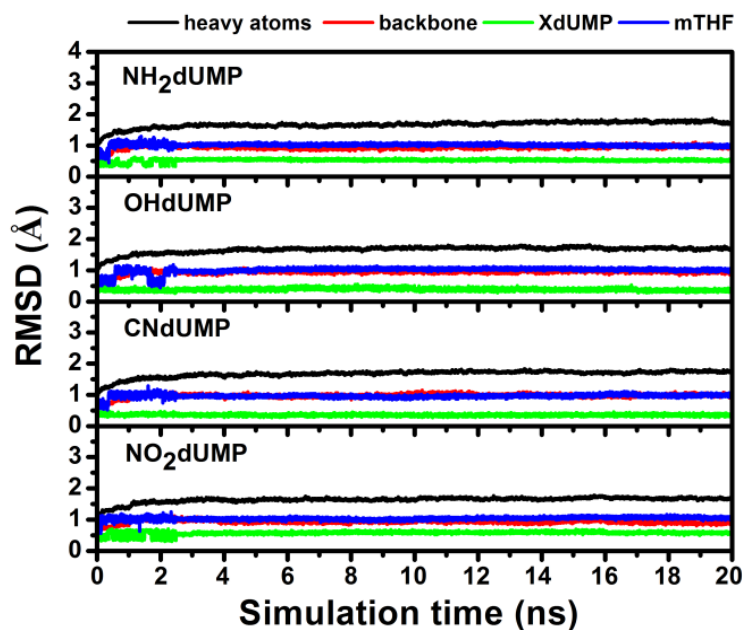


Figure 4.9 The RMSD plots of all heavy atoms, protein backbone and heavy atoms of the two ligands compared between NH₂dUMP, OHdUMP, CNdUMP and NO₂dUMP analogues

4.2.2 Per-residue decomposition energy scanning

All amino acid residues which associate with ligands stabilization are monitored using per-residue MM-GBSA decomposition (DC) energy method. The stabilization energy results on each individual ligands is separately considered as shown in Fig. 4.10 A. Noticeably, the number of stabilized residue is mostly contributed to the dUMP analogue rather mTHF, however, the overall energetic profiles of all analogue systems are founded in the similar pattern. The four arginines residues (R21, R166, R126(B) and R127(B)) are dominantly found to stabilize the dUMP analogue with < -10 kcal/mol while stabilization energy of the S167, D169, N177 and A263 are also detected in -5 kcal/mol approximately. However, destabilization energy is found at the C146 residue (~ 10 kcal/mol) which should be the resulted from strong repulsive interaction between the negative thiolate and its phosphate moiety. Although the lower number of stabilized residues is detected in mTHF, however, some interesting residues such as: the H51, W80, W83, D169 and A263 also associates to the mTHF binding as presented in a range of between -3 and -5 kcal/mol of DC stabilization energy.

The separated energy components in terms of both electrostatic and van der Waals energies with solvation effect focused on the active site residues is further considered as summarized in Fig. 4.10 (B). It is clearly seen from the plots that electrostatic interaction becomes an important role to stabilize the dUMP analogue dominantly. However, the detected stabilization energies on the mTHF molecule are presented in both electrostatic (H51, D169) and van der Waals (W80 and W83) interactions.

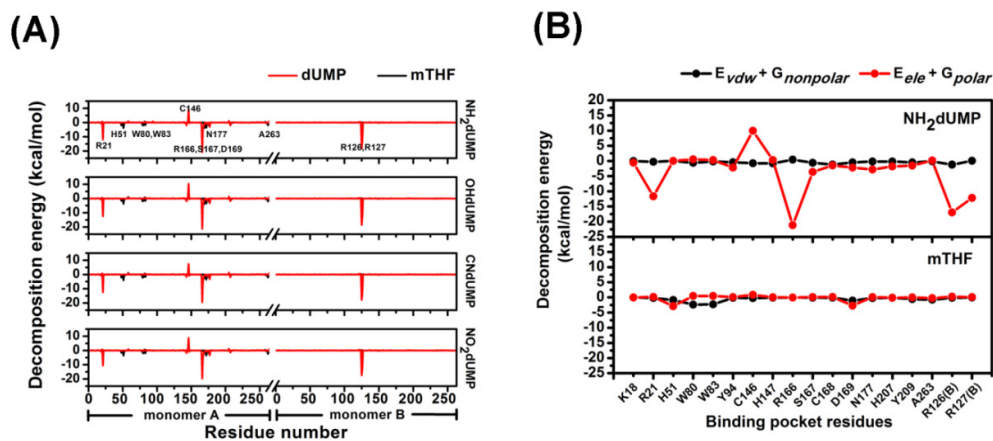


Figure 4.10 (A) The pairwise per-residue decomposition free energy scanning of all protein residues with individual mTHF and dUMP analogues (B) The electrostatic ($E_{ele}+G_{polar}$) and van der Waals ($E_{vdw}+G_{nonpolar}$) energy components contributed to the active site residues

4.2.3 Enzyme-ligands interactions

As indicated from the DC results, electrostatic interaction in the meaning of H-bonding is dominantly interested to represent the enzyme-ligands binding interactions. Based on the two fundamental criteria, (i) distance between proton donor (D) and acceptor (A) atoms ≤ 3.5 Å; and (ii) angle of D—H...A $> 120^\circ$, the percent of H-bonding occupation is determined (Fig 4.11 (A)). Most of the H-bonding interactions are contributed from dUMP analogues (19 interactions with 10 residues) rather mTHF cofactor (3 interactions with 3 residues), however, the overall pattern of interactions in all analogue systems seem to present in the similar way. Interestingly, the main H-bonding interaction founded in dUMP analogues is contributed from the phosphate moiety with serine (S167) and four surrounded arginines (R21, R166, R126(B) and R127(B)) residues through strong charge-charge interactions where the lower number of H-bonding formation are detected at deoxyribose (with H207 and Y209) and uracil (with Y94, D169 and N177) moieties.

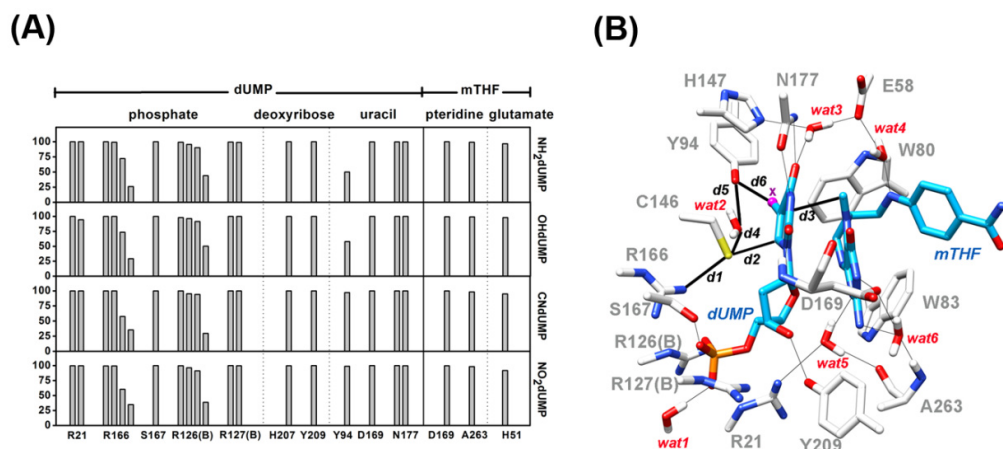


Figure 4.11 (A) The percentage of hydrogen bonding occupation for mTHF and dUMP analogues binding with the TS active site residues **(B)** The binding complex structure of dUMP analogue and mTHF (blue) inside the TS catalytic pocket where the dash and solid lines represent the hydrogen bonding and the intermolecular distances, *d1-d6*, involving to active conformation in the Michael addition, respectively

Focused on the dUMP functional group substituent, H-bonding occupation with Y94 is noticeably presented in the different pattern between electron withdrawing (CN- and NO₂-, ~95%) and donating (OH- and NH₂-, < 50%) analogues. The conformational structures of the substituent groups and nearby residues are more investigated for better understanding (Fig. 4.12). It should be noted from the figure that the hydroxyl hydrogen of Y94 has potentially formed H-bonding with either substituent heteroatom or bridging water molecule. Therefore, the percent of H-bonding detection between Y94 and substituent heteroatom is compensated by the lost of bridging water interaction obviously. As clearly seen from the figure, the heteroatom of all dUMP analogues are positioned nearby the hydroxyl hydrogen of Y94 similarly. Inter atomic distances between the hydroxyl hydrogen of Y94 and the hetero atoms of the cyanide and nitro groups seem to present in shorter length compared to the hydroxyl and amino groups according to indirect contact to the C-5 position.

This difference is the main reason that leads to well support the strong H-bonding formation in CNdUMP and NO₂dUMP analogues rather the OHdUMP and NH₂dUMP consequently (more detailed in distance analysis section). In addition, interaction of the bridging water with the E58, H147 and uracil are strongly conserved in most analogue systems (slightly changes to be E58, H147 and nitro oxygen instead in NO₂dUMP) confirmed the important role of bridging water to stabilize the dUMP substrate (or analogue) in well adjusted configuration inside the TS enzyme pocket.

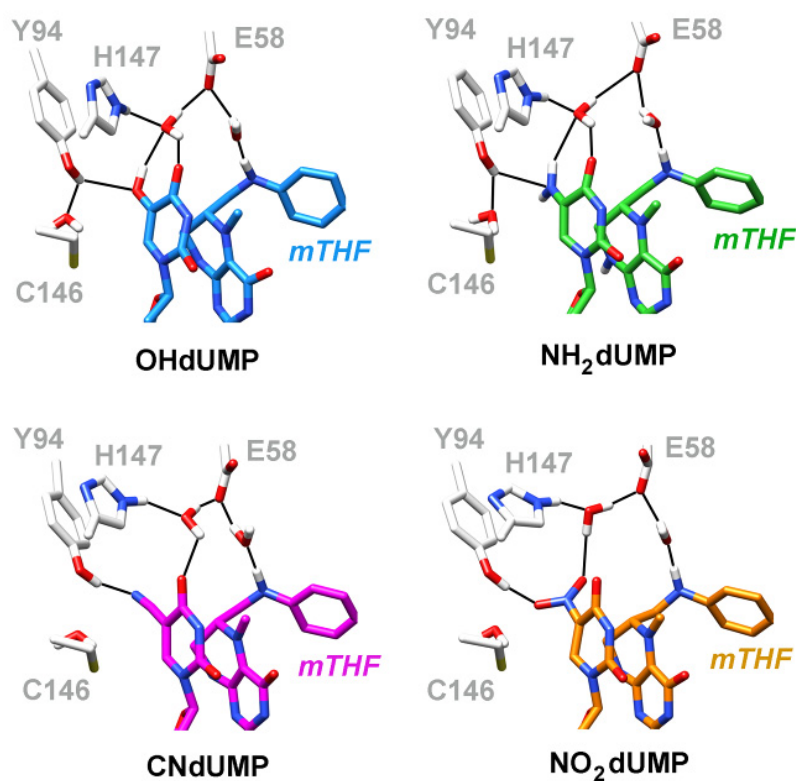


Figure 4.12 The functional group orientation and hydrogen bonding interactions of the four dUMP analogues (OHdUMP, NH₂dUMP, CNdUMP and NO₂dUMP) with all surrounded residues

For the three fragments of mTHF (pteridine, PABA and glutamate), the terminal glutamate positioned on the protein surface is strongly stabilized by H51 with > 90% occupation in all systems. The number of hydrogen bonding is not found at the PABA moiety due to its presented hydrophobic property. Although the number of hydrogen bonding is not found at the PABA moiety due to its hydrophobic property. In pteridine fragment, the two strongly hydrogen bonding interactions are detected at D169 and A263. The two strongly H-bonding interactions with uracil and pteridine fragments indicates to the significant role of D169 residue for substrate conformational trapping in π - π stacking in suitable configuration prior to the initial reaction. In accordance with DC results, structural orientation of the W80 and W83 residues also support the non-bonded interaction through van der Waals interaction with mTHF cofactor as well (Fig. 4.11 (B)).

4.2.4 Inter molecular distances analysis

As previously proposed by QSAR study that the reactivity of Michael addition may be influenced from functional group changing on the C5 position of dUMP, therefore, distribution of the six following distances: $d1$ [NH₂(R166)-S(C146)] between the stabilized R166 and the reactive C146; $d2$ [S(C146)-C6(dUMP)] refer to the Michael addition; $d3$ [C5(dUMP)-CH₂(mTHF)] the C5 atom of dUMP and methylene group of mTHF; $d4$ [S(C146)-O(WAT2)] and $d5$ [OH(Y94)-O(WAT2)] the bridging water; and $d6$ [OH(Y94)-X(dUMP)] the hydroxyl hydrogen and the heteroatom of substituent, are monitored as plotted in Fig. 4.13 (see Fig. 4.11 (B) for $d1$ - $d6$ definition).

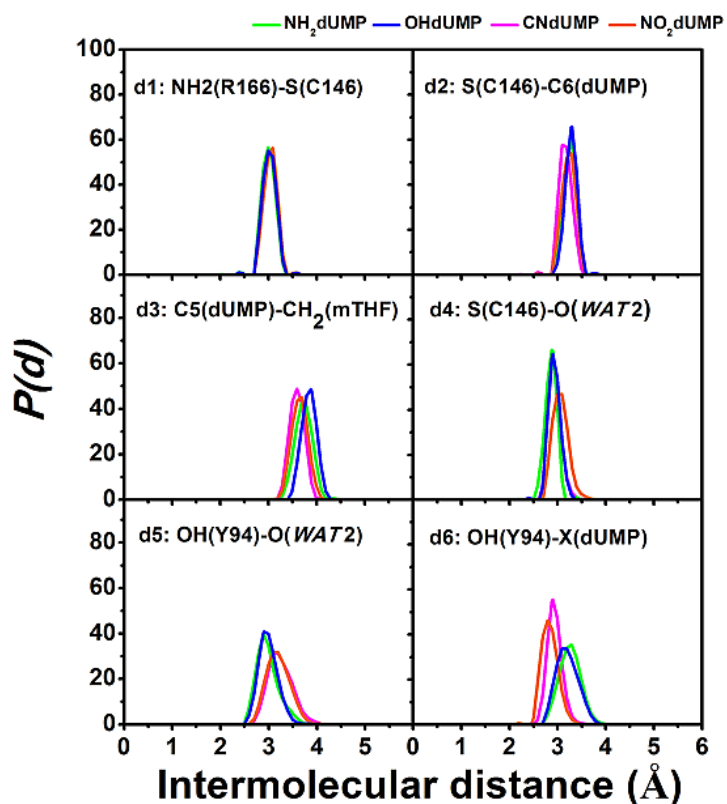


Figure 4.13 Distribution plot of the selected $d1$ - $d6$ distances, $P(d)$, along the last 17ns of simulation time

Before the Michael addition initiation, the R166 is suggested as activator residue to generate the thiolate specie in C146, in the other hands, it also stabilizes this thiolate through strong charge-charge interaction as shown in the narrow and sharp peak of most probable finding $d1$ at ~ 3.0 Å. The shortening and most similar range of $d2$ at 3.2 Å suggests us to propose that the Michael addition reaction on the C-6 position of dUMP analogues is possibly occurred in all systems. Although the nucleophilicity of dUMP C5 position is not actually activated in our model, however, the methylene linkage formation can be implied by the $d3$ distance indirectly. The most values of $d3$ are presented in a range of 3.0-3.2 Å (higher shift to 3.8 Å in OHdUMP) suggesting to the suitable of non-covalent ternary complex conformation prior to the methylene bridging formation afterwards.

The strength of water bridging interactions with C146 and Y94 are directly indicated from the $d4$ and $d5$ distances where the hydrogen bonding formation between the substituted heteroatom and Y94 hydroxyl oxygen is indicated by the $d6$ distance. The bridging water interaction is strongly confirmed in OHdUMP and NH₂dUMP analogues as shown in the low positioned peak at 2.8 and 2.9 Å for $d4$ and $d5$ respectively. In compensate with water bridging strength, the shortening of $d6$ distance is noticeably detected in NO₂dUMP and CNdUMP analogues at 2.7 and 2.9 Å instead (~3.2 Å in both OHdUMP and NH₂dUMP systems). We can summarize from this information that the strength of additional H-bonding interaction between the heteroatom functional group and Y94 is compensated by the lost of Y94-water bridging interaction.

4.2.5 Ligand Binding affinity

The affinity of each individual and both ligands binding with TS enzyme is evaluated using the implemented MM/PBSA module in AMBER. According to the role of bridging waters for ligands conformational adjustment, the ligands binding affinity in embedded with/without the bridging waters is calculated and compared. The results are comparatively summarized as given in Table 4.2. As clearly seen from the table, the important role of bridging waters for ligands stabilization is obviously confirmed by an increasing of binding free energy by > 10 kcal/mol in all systems when these bridging waters are included in the part of receptor. However, the water stabilization seemly effects with the mTHF cofactor binding rather the dUMP analogues as indicated by more stabilized ~5 kcal/mol in most cases.

Interestingly, the binding strength of the dUMP analogues which is substituted by an electron withdrawing group are presented in stronger affinity with the TS enzyme active site compared to the electron donating substitution by this following order: CNdUMP (-110.8) ~ NO₂dUMP (-109.0) > OHdUMP (-102.6) ~ NH₂dUMP (-101.2). The dUMP analogues present more favorable binding with the TS enzyme (~3 times stronger) compared to the mTHF cofactor in well accordance with the per-residue decomposition free energy and H-bonding interactions results.

Table 4.2 The averaged binding free energy (kcal/mol) over 100 snapshots of the two ligands binding to TS enzyme obtained from the last 17 ns of simulation time where the standard deviation is shown in bracket

receptor	ligand	dUMP analogues			
		NH ₂ dUMP	OHdUMP	CNdUMP	NO ₂ dUMP
TS	XdUMP, mTHF	-92.3 (6.9)	-93.9 (5.4)	-99.2 (6.2)	-98.0 (5.8)
TS, wat1-6	XdUMP, mTHF	-101.2 (6.7)	-102.6 (5.2)	-110.8 (6.1)	-109.0 (6.2)
TS	XdUMP	-74.7 (6.6)	-74.4 (6.9)	-80.5 (5.5)	-79.2 (5.6)
TS, wat1-3	XdUMP	-78.5 (6.0)	-81.5 (6.4)	-86.2 (5.1)	-86.6 (5.2)
TS	mTHF	-21.5 (2.9)	-20.4 (3.0)	-20.2 (2.7)	-20.6 (2.8)
TS, wat4-6	mTHF	-29.6 (3.2)	-30.5 (3.1)	-29.8 (3.2)	-29.5 (3.6)

4.3 Mechanism identification of Michael addition and covalent complex formation

As mentioned, the ternary covalent intermediate formation involves nucleophilic attack of C146 at the C-6 position of dUMP and methylene linkage formation between the C-5 atom of dUMP and the reactive methylene of mTHF cofactor. Based on the reaction coordinate driving approach, the potential energy surface responsible with these chemical reactions is calculated as a function of two reaction coordinates: $RC1 = d(S_{C146}-C6_{dUMP})$ and $RC2 = d(C5_{dUMP}-C5M_{mTHF})$ as shown in the Fig. 4.14. Harmonic restraints with a force constant of $3000 \text{ kcal/mol}\cdot\text{\AA}^2$ are used to drive the RC1 and RC2 from 3.5 to 1.6 \AA and 3.4 to 1.4 \AA , respectively, with an interval step of 0.1 \AA .

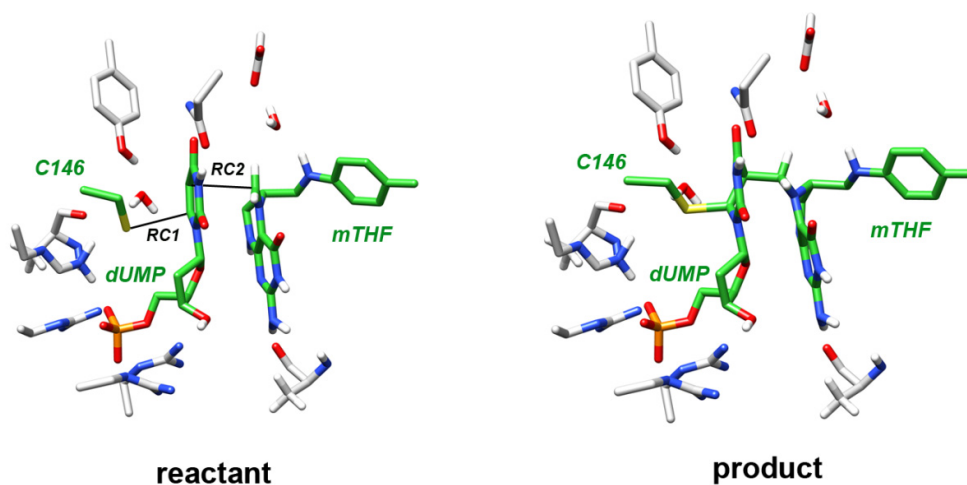


Figure 4.14 Identification of the reaction coordinate distances responsible to Michael addition (RC1) and ternary covalent complex formation (RC2)

In order to establish whether the formation of the ternary covalent intermediate in TS occurs in a concerted or stepwise manner, a two-dimensional potential energy surface (2D-PES) was calculated. The location of the minimum energy pathway along over the 2D-PES will indicate a concerted or stepwise mechanism if the path crosses the center of the surface or not, respectively. The 2D-PES for the ternary covalent intermediate formation is plotted relative to the reactant species in Fig. 4.15 (A). As seen from the PES, two local energy minima corresponding to the (I) reactant and (III) product were detected with a single approximate transition state (TS, II) along the minimum energy pathway. An activation energy barrier of 20.8 kcal/mol is observed, as shown in Fig. 4.15 (B). The minimum energy pathway crosses the middle of the 2D-PES, suggesting that the Michael addition and covalent complex formation steps occur in a concerted manner.

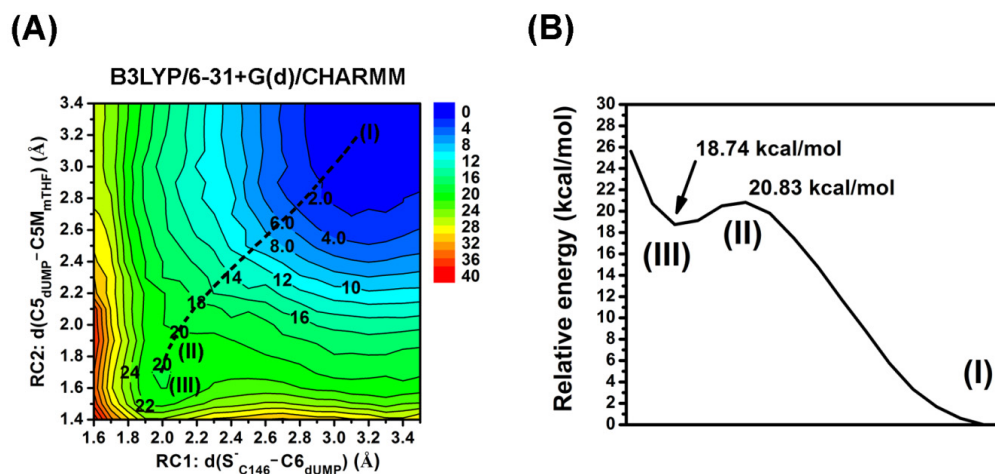


Figure 4.15 (A) Two-dimensional QM/MM (B3LYP/6-31+G(d)/CHARMM) potential energy surface for Michael addition and covalent complex formation. (B) The simplified overall energy profile along the chemical reaction pathway. In both figures, (I), (II) and (III) correspond to the reactant, transition state and product, respectively, and are depicted in Fig. 4.16.

The geometries of the reactant, transition state and product are depicted in Fig. 4.16. According to the figure, hybridization of the C-6 and C-5 atoms is gradually changed (from sp^2 to sp^3), together with the decreasing of two reaction coordinate distances (RC1: 3.5 to 1.6 Å and RC2: 3.4 to 1.4 Å) during the reaction. It should be noted that the hydrogen bonding interaction between the amine hydrogen of R166 and the thiolate sulfur of C146 is broken concomitantly with the Michael addition, however, it is compensated for by the new interaction forming between the amine hydrogen of R166 and the carbonyl backbone of A144, as indicated by the reduced distance of $d(\text{HH}_{21\text{R166}}-\text{O}_{\text{A144}})$ from 2.37 Å (reactant) to 1.69 Å (product). Interestingly, the water molecules surrounding the C146 thiolate not only provide electrostatic stabilization but also retain the strong bridging interaction between C146 and the Y94 hydroxyl group, as indicated by the consistent values of ~ 3.2 Å and ~ 2.7 Å for $(\text{SG}_{\text{C146}}-\text{HO}_{\text{WAT}})$ and $d(\text{O}_{\text{WAT}}-\text{HH}_{\text{Y94}})$, respectively between reactant and product. Because the conformational structure of other environmental residues and bridging waters are not changed significantly during the two reactions, the PES result reasonably represents the expected chemical reaction without any interference from other intrinsic changes.

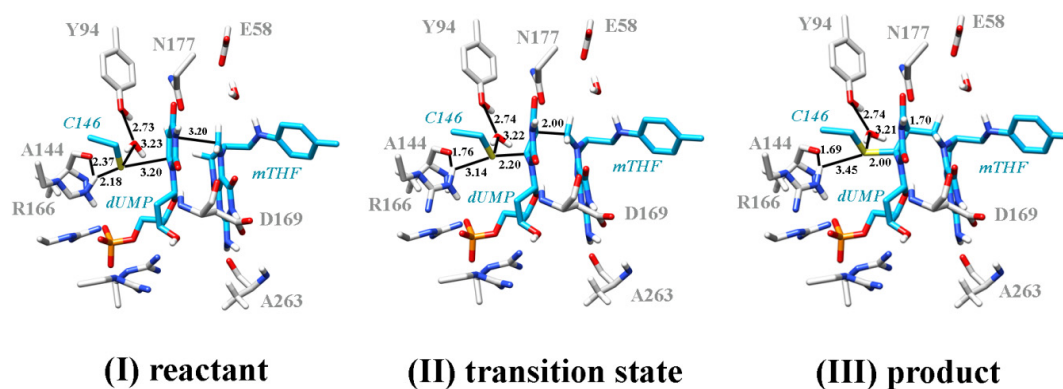


Figure 4.16 QM/MM (B3LYP/6-31+G(d)/CHARMM) optimized structures of the (I) reactant, (II) transition state and (III) product located in the active site of thymidylate synthase during Michael addition and covalent complex

4.4 Activation energy barrier required for Michael addition and ternary covalent complex formation in various dUMP analogues

In consistence with the 2D-PES results, the same two reaction coordinates, $RC1 = d(S_{C146}-C6_{dUMP})$ and $RC2 = d(C5_{dUMP}-C5M_{mTHF})$, are selected in activation energy barrier investigation of the Michael addition and ternary covalent complex formation. However, the combination of these RCs ($RC1+RC2$) is used to reduce the number of calculation points instead of the full individual RCs in 2D-PES scanning. Therefore, in this following case, the complex RC distance which responsible to the Michael addition and ternary covalent complex formation is defined in a reducing range from 6.4 to 3.2 Å with an interval step of 0.2 Å leading to reduce the number of geometries calculation from about 400 to 17 points obviously. Structural geometries along the reaction pathways are mainly obtained by fully optimization with B3LYP/6-31+G*-CHARMM level in an adiabatic mapping procedure similar as the previous 2D-PES calculation. Moreover, the other DFT methods, BLYP and BH&HLYP, as well as the *ab initio* based SCS-MP2 methods are also comparatively tested.

4.4.1 The effect of QM region size

Because the large size of QM region selection leading to the increasing of computational cost significantly, then, the optimum relationship between QM size selection and accuracy of the calculated result should be considered. Focused on the reactive cysteine, the thiolate anion is surrounded by two interesting molecules before the reaction is initiated: the R166 and bridged water (Fig. 4.17 (A)). It may be considerable to additionally include the R166 side chain and water to be QM region further from the minimal reactive C146 and two dUMP and mTHF substrate fragments for more accuracy results. Here, the three types of QM regions are selected: I) the minimal QM size which composed by the C146 side chain, uracil and deoxyribose of dUMP and pteridine of mTHF II) the minimal QM size with R166 side chain and III) the minimal QM size with R166 side chain and the stabilized water. The QM/MM total energy together with its separated components in QM, QM with no dispersion and MM energies along the reaction pathway of each selected model is monitored and compared as shown in the Fig. 4.17 (B).

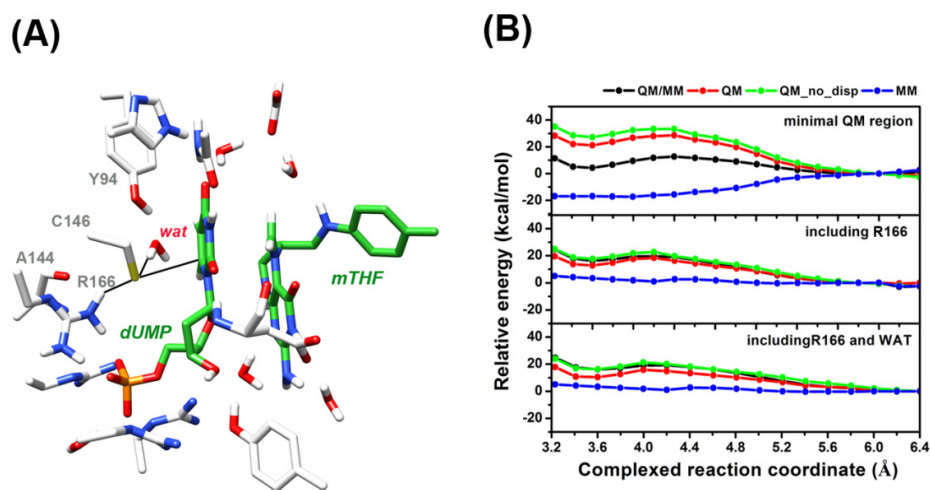


Figure 4.17 (A) Conformational structure of dUMP, mTHF and the surrounded residues focused on the thiolate anion of C146 (B) The QM/MM contribution energies in terms of QM, QM without dispersion and MM component for the studied system with three different QM regions along the complexed reaction coordinate

A single transition state and product species are detected in all different QM size models as indicated in the QM/MM total energy profile (black line). However, the R166 side chain excluded in the QM region (model I) seems to significantly effect to the energetic profile by ~ 20 kcal/mol of MM lowered whereas a nearly constant profile are found in the other two included R166 systems (model II). Without R166 in QM region, its side chain may change conformation and interacts with the other MM residues leading to a more MM energy stabilization during the reaction proceeding. The additional stabilization energy by R166 is accounted to the part of QM energy as clearly seen by the lower magnitude of QM energy component (red line). In the later QM model (III) an including of the stabilized water to the QM part does not effect to the energy profile significantly as indicated by the mostly identical energy components to the system model II. Therefore, the result allows us to summarize that the optimum QM region should be composed of the dUMP and mTHF fragments, reactive C146 side chain and also R166 side chain.

4.4.2 The effect of DFT methods

Although the B3LYP is the popular DFT functional used and accepted in QM/MM calculation, the two other DFT methods, BLYP and BH&HLYP, are also tested. The QM/MM geometry optimization is calculated at 6-31+G*, 6-31++G** and 6-311+G** basis sets on each DFT functional to obtain the reaction energy profiles on this system with the QM model II. The calculated energy profile results are given in the Fig. 4.18.

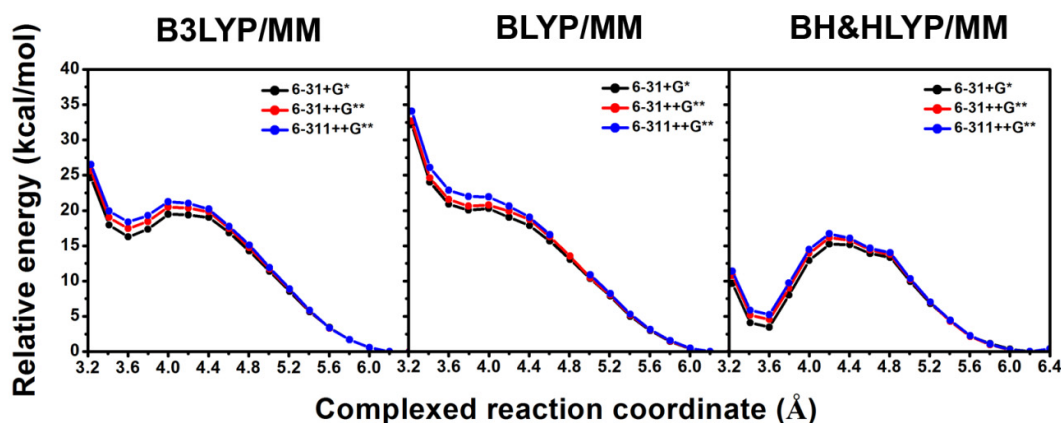


Figure 4.18 Comparison of QM/MM optimization energy profiles where the QM region are treated by B3LYP, BLYP and BH&HLYP methods at 6-31+G* basis set

It can be considered from the figure that the obtained QM/MM energy profiles are presented in the different patterns between DFT methods applied. Noticeably, the BH&HLYP optimized result generates the lower activation energy barrier and more stable methylene bridging intermediate compared to B3LYP and BLYP results. The energy profile difference may be resulted from optimized structural geometries or possibly the particular exchange-correlation function of each functional in energy calculation process.

To clarify this problem, the single point energy calculation by BLYP and BH&HLYP methods at 6-31+G* basis set is performed on the B3LYP optimized structures along the reaction pathway. The obtained results are shown in Fig. 4.19. In addition, the single point energy profile results are relatively compared with their fully optimizations (Fig. 4.18) for better understanding.

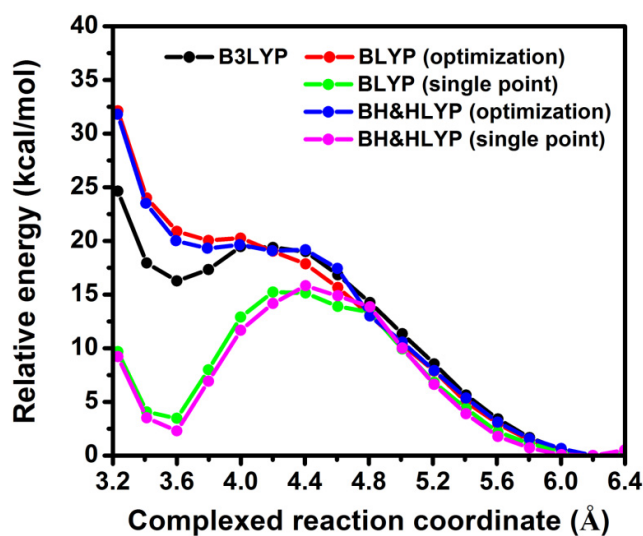


Figure 4.19 Energy profiles comparison at 6-31+G* basis set between the individual DFT optimized and single point energies from BLYP and BH&HLYP methods on the B3LYP based structures

It is clearly seen from the result that single point energy profiles on the B3LYP based structure by BLYP and BH&HLYP methods are presented in the similar pattern with their fully optimization energy profiles. Therefore, we can summarize from this information that the different pattern of energy profiles is caused from its unique calculated procedure (exchange-correlation function, for example) rather than the geometry structure reason. This finding leads us to use the higher accuracy method, such as the SCS-MP2, for determining reaction energy profile further.

4.4.3 Activation energy barrier from MP2 and SCS-MP2 calculation

Due to the different energy profile patterns in DFTs, single point energy calculation (on the B3LYP/6-31+G* based structures) at higher level by MP2 and SCS-MP2 methods with cc-pVDZ, aug-cc-pVDZ, cc-pVTZ and aug-cc-pVTZ basis sets is required for reaction energy profile with higher accuracy where the results are shown in the Fig. 4.20. Noticeably, the MP2 and SCS-MP2 energy profiles are presented in the similar pattern as the previous BH&HLYP calculation.

According to the plots, the basis set effect between aug-cc-pVDZ and aug-cc-pVTZ used in both MP2 and SCS-MP2 methods is found by ~ 4 kcal/mol difference possibly from the size of basis set function in double zeta and triple zeta expansion. Based on the assumption that the larger basis set expansion should give more accuracy calculated result than the smaller one, then, activation energy barrier corresponding to the Michael addition and ternary covalent complex formation between mTHF and dUMP (or its analogues) will be investigated by the SCS-MP2 method with SCS-MP2/aug-cc-pVTZ basis set further.

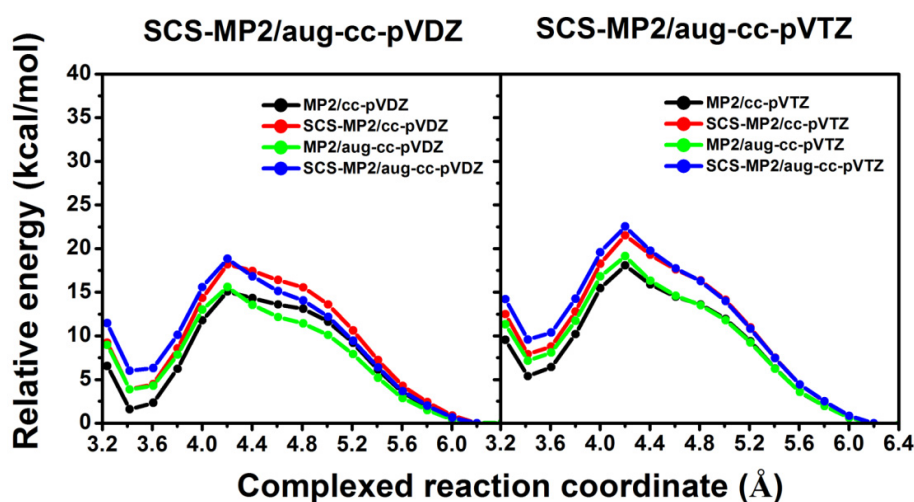


Figure 4.20 Single point energy profiles of MP2 and SCS-MP2 methods in cc-pVDZ, aug-cc-pVDZ, cc-pVTZ and aug-cc-pVTZ basis sets using the B3LYP/6-31+G* based structures

To avoid the snapshot dependent problem, reaction energy profile of each complexed system is calculated in three different structures extracted from the molecular dynamics simulation with a consequence of B3LYP/6-31+G* optimization. In addition, it is well known from the experimental studies that BrdUMP analogue cannot follow the Michael addition and ternary covalent complex formation successfully due to the interference of dehalogenation reaction ($-Br$ and $-I$ elimination itself) during the nucleophilic attack as confirmed by the 340 nm peak in spectrophotometry technique^[60], therefore, investigation of the dUMP halogen analogues is truncated to the FdUMP and CldUMP systems instead.

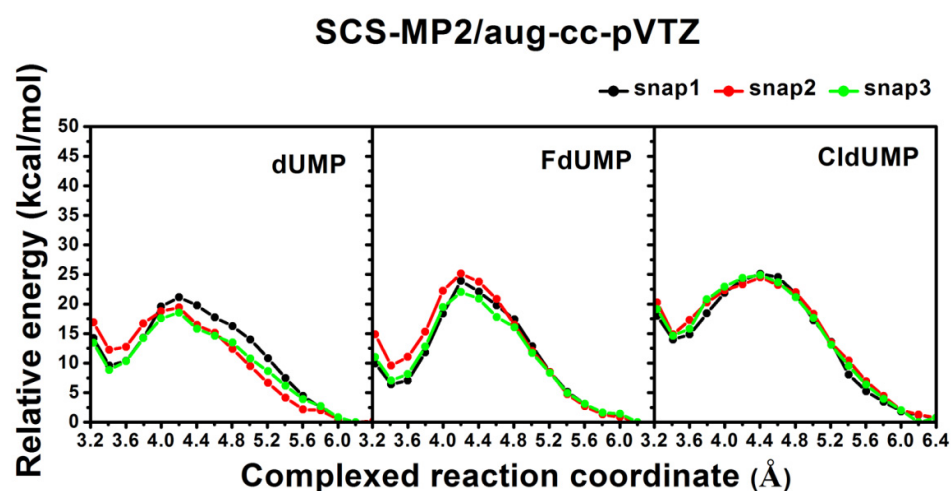


Figure 4.21 Single point energy profiles responsible to the Michael addition and ternary covalent complex formation using SCS-MP2/aug-cc-pVTZ method for the three different snapshots of dUMP, FdUMP and CldUMP ternary complexes

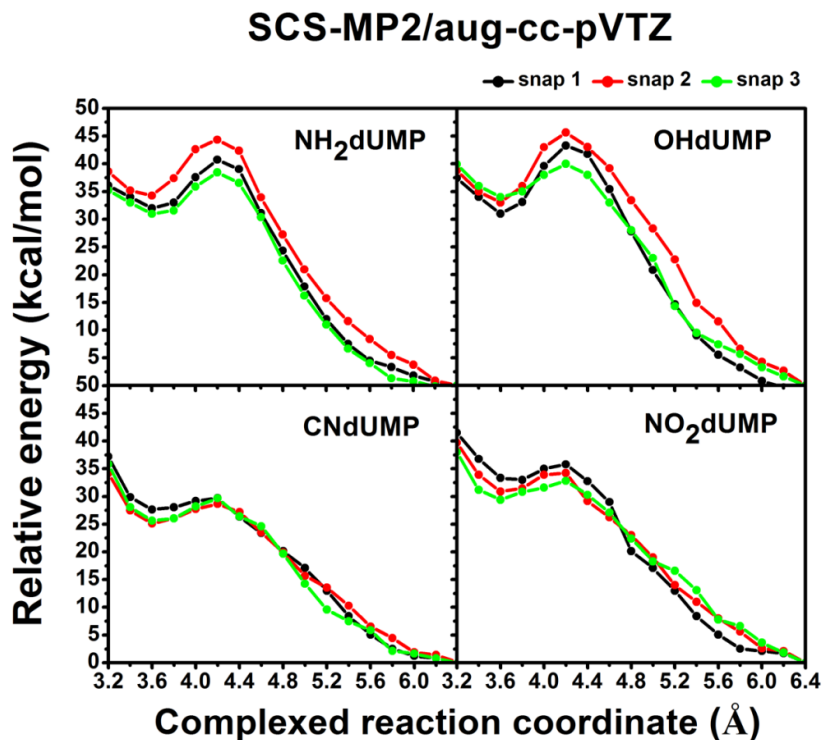


Figure 4.22 Single point energy profiles responsible to the Michael addition and ternary covalent complex formation using SCS-MP2/aug-cc-pVTZ method for the three different snapshots of NH₂dUMP, OHdUMP, CNdUMP and NO₂dUMP ternary complexes

The energy profile results are shown in the Fig. 4.21 for dUMP, FdUMP and CldUMP halogen analogues and in Fig. 4.22 for the NH₂dUMP, OHdUMP, CNdUMP and NO₂dUMP analogues. In addition, the activation energy barrier for all dUMP analogues is summarized in table 4.3 and compared with the experimental values. By considering on the halogen analogue results, the averaged activation energy barrier values of the FdUMP and CldUMP are of 23.7 and 24.8 kcal/mol respectively while the native dUMP substrate requires the lower magnitude at 19.7 kcal/mol. Interestingly, activation energy values of our calculation in dUMP and FdUMP are closely related to the experimental information (18.3 and 19.9 kcal/mol) suggesting the reasonably of selected model and the high accuracy of calculation.

It is well known that FdUMP analogue can form methylene bridging intermediate inside the thymidylate synthase as well as the dUMP substrate. Therefore, the Michael addition and ternary covalent complex formation in CldUMP analogue could be proceed in according to the relatively similar activation energy requirement with FdUMP analogue. Taken together with the tight binding of CldUMP analogue in our previous MD study (section 4.1), the CldUMP analogue is proposed to be the new potent candidate inhibitor against the thymidylate synthase due to the high affinity of the binding and similar range of activation energy value with the conventional FdUMP anticancer drug.

Besides, activation energy barrier of the electron withdrawing and donating dUMP analogues is presented in a significantly higher value compared to the halogen analogues. Interestingly, dUMP analogues substituted by an electron withdrawing group (CNdUMP and NO₂dUMP) require lower activation energy than the donating substituted analogues (NH₂dUMP and OHdUMP) by 10 kcal/mol approximately. The mean activation energy values are of 29.35 and 34.28 kcal/mol for CNdUMP and NO₂dUMP analogues which are 10 and 15 kcal/mol higher than the native dUMP substrate respectively. Unlike the previous QSAR studied, reactivity (also be indicated by activation energy of the reaction) of the Michael addition and ternary covalent complex formation is not possibly increased when C-5 position of dUMP is substituted by an electron withdrawing such as cyanide or nitro functional groups. However, the high activation energy requirement prediction in the donating substituted analogues, like NH₂dUMP and OHdUMP, well supports with the QSAR proposed obviously.

Table 4.3 Activation energy barrier responsible to the Michael addition and ternary covalent complex formation for all dUMP analogues obtained from the SCS-MP2/aug-cc-pVTZ method in comparison with the experimental values

dUMP analogues	$\Delta E_{\text{cal},1}^{\ddagger}$	$\Delta E_{\text{cal},2}^{\ddagger}$	$\Delta E_{\text{cal},3}^{\ddagger}$	$\Delta E_{\text{cal,avg}}^{\ddagger}$	SD	$\Delta E_{\text{exp}}^{\ddagger}$ ^[23]
dUMP	21.14	19.42	18.52	19.70	1.33	18.3
FdUMP	23.90	25.13	22.02	23.69	1.57	19.9
CldUMP	25.07	24.51	24.90	24.83	0.29	<i>N.D.</i>
NH ₂ dUMP	40.73	44.33	38.43	41.16	2.97	<i>N.D.</i>
OHdUMP	43.30	45.65	40.12	43.02	2.78	<i>N.D.</i>
CNdUMP	29.69	28.68	29.67	29.35	0.58	<i>N.D.</i>
NO ₂ dUMP	35.79	34.24	32.80	34.28	1.50	<i>N.D.</i>

**N.D.* = no data

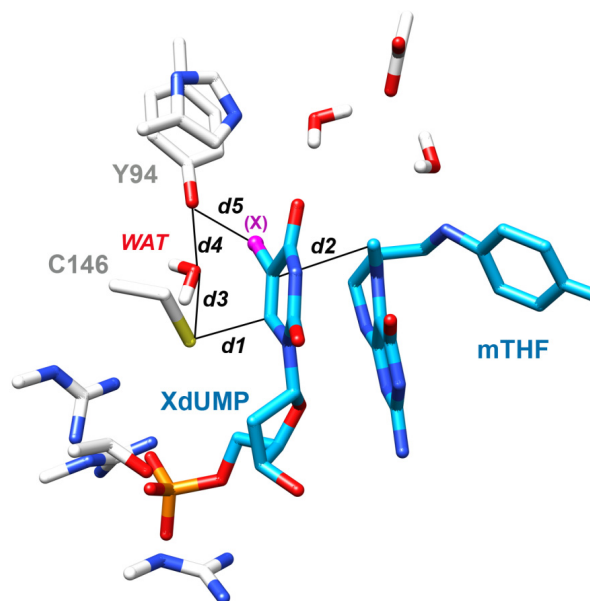


Figure 4.23 The critical distances directly and indirectly responsible to the Michael addition, ternary covalent complex formation in the TS/XdUMP/mTHF ternary complex

The critical distances, d1-d5, which are participated with Michael addition and ternary covalent complex formation reaction are measured and averaged from three snapshots in reactant, transition state and product species of all analogues (table 4.4). As clearly seen from the table, the main difference is found in product in which complex RC value at 3.4 Å but are shifted to 3.6 Å in NH₂dUMP, OHdUMP, CNdUMP and NO₂dUMP analogues. This is due to an increasing of steric hindrance compared to the halogen atom substitution. Interestingly, the bridged water is strongly detected and stabilized to the reactive cysteine (d3) as observed in an almost distance in a range between 3.1-3.3 Å in all systems excepted for the NH₂dUMP analogue (3.9 Å).

Table 4.4 The averaged critical distances (d1-d5 defined in Fig. 4.23) responsible to Michael addition, ternary complex formation for reactant, transition state and product structures of all dUMP analogues.

	reactant	TS	product
dUMP			
d1	3.21	2.29	1.85
d2	3.19	1.91	1.56
d3	3.23	3.20	3.20
d4	2.71	2.72	2.71
d5	3.18	3.22	3.23
FdUMP			
d1	3.13	2.34	1.86
d2	3.26	1.86	1.55
d3	3.29	3.29	3.30
d4	3.19	3.22	3.22
d5	3.08	3.09	3.10
CldUMP			
d1	3.37	2.27	1.85
d2	3.03	1.94	1.56
d3	3.19	3.38	3.45
d4	3.20	3.20	3.21
d5	3.22	3.27	3.29

	reactant	TS	product
NH₂dUMP	d1	3.17	1.97
	d2	3.23	1.64
	d3	3.88	3.92
	d4	2.70	2.71
	d5	3.67	3.81
OHdUMP	d1	3.22	1.95
	d2	3.18	1.65
	d3	3.30	3.33
	d4	2.85	2.91
	d5	3.24	3.35
CNdUMP	d1	3.24	1.94
	d2	3.17	1.67
	d3	3.32	3.35
	d4	2.93	3.03
	d5	2.96	2.97
NO₂dUMP	d1	3.19	1.93
	d2	3.20	1.68
	d3	3.15	3.16
	d4	2.88	2.87
	d5	3.18	3.21

CHAPTER V

CONCLUSIONS

Molecular dynamics simulation was used to investigate molecular behaviors including intermolecular interaction, binding affinity and dynamical property of the mTHF, dUMP and its various analogues with thymidylate synthase enzyme. In complementary with the Quantum Mechanics/Molecular Mechanics approach, reactivity of Michael addition and ternary covalent complex formation for the dUMP substrate and its analogues was studied in terms of reaction energy pathway, activation barrier and the critical structures (reactant, transition state and product).

In molecular dynamics simulation part, the effect of functional substitutions (-F, -Cl, -Br, -NH₂, -OH, -CN and -NO₂) on the C-5 position of dUMP substrate towards the enzyme stability is investigated by means of molecular dynamics simulation technique in ternary complex with mTHF cofactor. Both of dUMP and its analogues were dominantly stabilized by the TS residues through electrostatic interactions including charge-charge and hydrogen bond interactions, whereas the less stabilization was found for mTHF. The number of hydrogen bond formation is highly detected at the phosphate moiety of dUMP/XdUMP by the four surrounded arginines R21, R166, R126(B) and R127(B) like a conserved interaction in all systems. Interestingly, an additionally unique hydrogen bond between the hydroxyl group of Y94 and FdUMP, CNdUMP, NO₂dUMP, OHdUMP and NH₂dUMP was observed in both medium (OHdUMP and NH₂dUMP) or strong (FdUMP, CNdUMP, NO₂dUMP, OHdUMP) interaction strength. Prior to Michael addition, the R166 plays a role to stabilize the negative C146 via the strong charge-charge interaction which is well supported by the narrow and sharp peak of at ~3.2 Å in both dUMP native substrate and its analogues. Moreover, the distance between the S⁻ atom of the reactive C146 and the C6 atom of dUMP related to Michael addition at 2.7-3.7 Å in all systems suggested that, in the meaning of force field based method, the Michael addition on the C-6 position of the dUMP may possibly proceed in all dUMP analogues. Based on the MM/PBSA binding free energy prediction, the dUMP (or its analogues) shows

stronger binding with TS than the mTHF by c.a. 3 times dominantly. The order of averaged binding affinity is: CldUMP (-100.3) ~ FdUMP (-96.8) > dUMP (-86.3) > BrdUMP (-77.8) and CNdUMP (-110.8) ~ NO₂dUMP (-109.0) > OHdUMP (-102.6) ~ NH₂dUMP (-101.2) in kcal/mol.

In quantum mechanics/molecular mechanics part, we have determined that the Michael addition and ternary covalent complex formation between dUMP and mTHF substrates to form the methylene bridging intermediate occur via a concerted mechanism. Therefore the keto-enolate intermediate, which is uniquely proposed in the stepwise mechanism, should not be considered further. Indeed, experimental information such as: *i*) simultaneous occurrence of the several steps (imidazolidine activation, Michael addition and covalent bonding formation) during mTHF cofactor tight binding; and *ii*) unseparable kinetic rate of Michael addition from the ternary covalent complex formation also agree and support the concerted mechanism determination indirectly. Altogether, these findings are very useful for mechanism-based drug design because future attention should be focused on increasing Michael addition reactivity rather than stabilization of the keto-enolate.

The several testing of high-level QM/MM calculations on the ternary complex of dUMP, mTHF and thymidylate synthase provides an important information as follows: the QM region must be composed by the dUMP and mTHF fragments, C146 and R166 side chains at least; and the optimized energy profiles based on different DFT functionals are presented in very dissimilar pattern. The very high level of SCS-MP2 method with aug-cc-pVTZ basis set is selected for exact energy profile determination which is confirmed by the experimental values in the small value of activation energy difference less than 1.5 and 4 kcal/mol for dUMP and FdUMP. Due to the similar activation energy requirement with FdUMP anticancer drug and high affinity of the binding in CldUMP analogue, we could propose the CldUMP analogue to be the newly potent candidate inhibitor against thymidylate synthase.

REFERENCES

- [1] National Cancer Institute [Online]. 2013. Available from : <http://www.cancer.gov/cancertopics/cancerlibrary/what-is-cancer> [2013, March]
- [2] American Brain Tumor Association [Online]. 2013. Available from : <http://www.abta.org/understanding-brain-tumors/diagnosis/benign-and-malignant-tumors.html> [2013, March]
- [3] wiseGEEK [Online]. 2013. Available from : <http://www.wisegeek.com/what-is-a-benign-tumor.htm> [2013, March]
- [4] SmartDraw [Online]. 2013. Available from : http://wcl.smartdraw.com/examples/content/examples/10_healthcare/anatomy_illustrations/normal_benign_&_malignant_tissue_1.jpg [2013, March]
- [5] National Institutes of Health [Online]. 2013. Available from : <http://www.cancer.gov/cancertopics/factsheet/detection/staging> [2013, March]
- [6] American Joint Committee on Cancer [Online]. 2013. Available from : <http://www.cancerstaging.org> [2013, March]
- [7] American Joint Committee on Cancer. AJCC Cancer Staging Manual. 6th ed, New York : Springer, 2002.
- [8] International Union for Cancer Control [Online]. 2013. Available from : <http://www.uicc.org/resources/tnm> [2013, March]
- [9] Center for Continuing Education [Online]. 2013. Available from : <http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/hematology-oncology/esophageal-cancer> [2013, March]
- [10] Durivage, H. J.; Knobf, M. T.; Beaulieu, N. Cancer Chemotherapy Handbook. 2nd ed, Philadelphia : Lippincott Williams & Wilkins, 2001.
- [11] Trigg, M. E.; Flanigan-Minnick, A. Mechanisms of action of commonly used drugs to treat cancer. Commun. Oncology. 8 (2011) : 357-369.
- [12] Marsh, S.; Liu, G. Pharmacokinetics and pharmacogenomics in breast cancer chemotherapy. Adv. Drug Deliv. Rev. 61 (2009) : 381-387.
- [13] Perry, K. M.; Fauman, E. B.; Finer-Moore, J. S.; Montfort, W. R.; Maley, G.

- F.; Maley, F.; Stroud, R. M. Plastic adaptation toward mutations in proteins: Structural comparison of thymidylate synthases. Proteins 8 (1990) : 315-333.
- [14] Carreras, C. W.; Santi, D. V. THE CATALYTIC MECHANISM AND STRUCTURE OF THYMIDYLATE SYNTHASE. Annul. Rev. Biochem. 64 (1995) : 721-762.
- [15] Danenberg, P. V. Thymidylate synthetase - a target enzyme in cancer chemotherapy. Biochem. Biophys. Acta. 473 (1977) : 73-92.
- [16] Santi, D. V. Perspective on the design and biochemical pharmacology of inhibitors of thymidylate synthetase. J. Med. Chem. 23 (1980) : 103-111.
- [17] Spencer, H. T.; Villafranca, J. E.; Appleman, J. R. Kinetic scheme for thymidylate synthase from Escherichia coli: determination from measurements of ligand binding, primary and secondary isotope effects, and pre-steady-state catalysis. Biochemistry 36 (1997) : 4212-4222.
- [18] Stroud, R. M.; Finer-Moore, J. S. Stereochemistry of a multistep/bipartite methyl transfer reaction: thymidylate synthase. FASEB J. 7 (1993) : 671-677.
- [19] Humphreys, G. M.; Greenberg, D. M. Studies on the conversion of deoxyuridylic acid to thymidylic acid by a soluble extract from rat thymus. Arch. Biochem. Biophys. 78 (1958) : 275-287.
- [20] Fritz, T. A.; Liu, L.; Finer-Moore, J. S.; Stroud, R. M. Tryptophan 80 and leucine 143 are critical for the hydride transfer step of thymidylate synthase by controlling active site access. Biochemistry 41 (2002) : 7021-7029.
- [21] Santi, D. V.; Danenberg, P. V. Folates and Pterins. New York : John Wiley & Sons, 1984.
- [22] Matthews, D. A.; Villafranca, J. E.; Janson, C. A.; Smith, W. W.; Welsh, K.; Freer, S. Stereochemical mechanism of action for thymidylate synthase based on the X-ray structure of the covalent inhibitory ternary complex

- with 5-fluoro-2'-deoxyuridylate and 5,10-methylenetetrahydrofolate. J. Mol. Biol. 214 (1990) : 937-948.
- [23] Santi, D. V.; McHenry, C. S.; Raines, R. T.; Ivanetic, K. M. Kinetics and thermodynamics of the interaction of 5-fluoro-2'-deoxyuridylate with thymidylate synthase. Biochemistry 26 (1987) : 8606-8613.
- [24] Barrett, J. E.; Lucero, C. M.; Schultz, P. G. A Model for Hydride Transfer in Thymidylate Synthase Based on Unnatural Amino Acid Mutagenesis. J. Am. Chem. Soc. 121 (1999) : 7965-7966.
- [25] Liu, Y.; Barrett, J. E.; Schul, P. G.; Santi, D. V. Tyrosine 146 of Thymidylate Synthase Assists Proton Abstraction from the 5-Position of 2'-Deoxyuridine 5'-Monophosphate. Biochemistry 38 (1999) : 848-852.
- [26] Parker, W. B.; Cheng, Y. C. Metabolism and mechanism of action of 5-fluorouracil. Pharmacol. Ther. 48 (1990) : 381-395.
- [27] Jarmuła, A.; Cieplak, P.; Krygowski, T. M.; Rode, W. The effect of 5-substitution in the pyrimidine ring of dUMP on the interaction with thymidylate synthase: Molecular modeling and QSAR. Bioorganic & Medicinal Chemistry 15 (2007) : 2346-2358.
- [28] Barrett, J. E.; Maltby, D. A.; Santi, D. V.; Schultz, P. G. Trapping of the C5 Methylene Intermediate in Thymidylate Synthase. J. Am. Chem. Soc. 120 (1998) : 449-450.
- [29] Moore, M. A.; Ahmed, F.; Dunlap, R. B. Trapping and partial characterization of an adduct postulated to be the covalent catalytic ternary complex of thymidylate synthase. Biochemistry 25 (1986) : 3311-3317.
- [30] Finer-Moore, J. S.; Santi, D. V.; Stroud, R. M. Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure. Biochemistry 42 (2003) : 248-256.
- [31] Rode, W.; Leś, A. Molecular mechanism of thymidylate synthase-catalyzed reaction and interaction of the enzyme with 2- and/or 4-substituted analogues of dUMP and 5-fluoro-dUMP. Acta Biochim Pol. 43 (1996) : 133-142.

- [32] Phan, J.; Mahdavian, E.; Nivens, M. C.; Minor, W.; Berger, S.; Spencer, H. T.; Dunlap, R. B.; Lebioda, L. Catalytic Cysteine of Thymidylate Synthase Is Activated upon Substrate Binding. Biochemistry 39 (2000) : 6969-6978.
- [33] Kanaan, N.; Marti, S.; Moliner, V. A QM/MM Study of the catalytic mechanism of the thymidylate synthase. Biochemistry 46 (2007) : 3704-3713.
- [34] Rungrotmongkol, T.; Nunthaboot, N.; Malaisree, M.; Kaiyawet, N.; Yotmanee, P.; Meeprasert, A.; Hannongbua, S. Molecular insight into the specific binding of ADP-ribose to the nsP3 macro domains of chikungunya and Venezuelan equine encephalitis viruses: molecular dynamics simulations and free energy calculations. J. Mol. Graph. Model. 29 (2010) : 347-353.
- [35] Arsawang, U.; Saengsawang, O.; Rungrotmongkol, T.; Sornmee, P.; Wittayanarakul, K.; Remsungnen, T.; Hannongbua, S. How do carbon nanotubes serve as carriers for gemcitabine transport in a drug delivery system? J. Mol. Graph. Model. 29 (2011) : 591-596.
- [36] Sornmee, P.; Rungrotmongkol, T.; Saengsawang, O.; Arsawang, U.; Remsungnen, T.; Hannongbua, S. Understanding the Molecular Properties of Doxorubicin Filling Inside and Wrapping Outside Single-Walled Carbon Nanotubes. J. Comput. Theor. Nanosci. 8 (2011) : 1385-1391.
- [37] Li, H.; Robertson, A. D.; Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. Proteins: Struct. 61 (2005) : 704-721.
- [38] Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very fast prediction and rationalization of pKa values for protein-ligand complexes. Proteins 73 (2008) : 765-783.
- [39] Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, W.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.;

- Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seein, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. AMBER 10 [Software]. 2008.
- [40] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Jr., J. A. M.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03 [Software]. 2004.
- [41] Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics 79 (1983) : 926-935.
- [42] 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. Journal of Chemical Physics 99 (1993) : 8345-8348.
- [43] Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. Journal of Computational Physics 23 (1977) : 327-341.
- [44] Claeysens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.;

- Ranaghan, K. E.; Schutz, M.; Thiel, S.; Thiel, W.; Werner, H. J. High-accuracy computation of reaction barriers in enzymes. Angew. Chem. Int. Ed. 45 (2006) : 6856-9.
- [45] van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. High-level QM/MM modelling predicts an arginine as the acid in the condensation reaction catalysed by citrate synthase. Chem. Commun. 16 (2008) : 1874-1876.
- [46] Lonsdale, R.; Hoyle, S.; Grey, D. T.; Ridder, L.; Mulholland, A. J. Determinants of Reactivity and Selectivity in Soluble Epoxide Hydrolase from QM/MM Modeling. Biochemistry 51 (2012) : 1774-1786.
- [47] MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J. Phys. Chem. B 102 (1998) : 3586-3616.
- [48] Foloppe, N.; MacKerell, A. D. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. J. Comput. Chem. 21 (2000) : 86-104.
- [49] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. 4 (1983) : 187-217.
- [50] Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. J. Comput. Chem. 11 (1990).
- [51] Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. Phys. Rev. B 37 (1988).

- [52] Becke, A. D. Densityfunctional thermochemistry. III. The role of exact exchange. J. Chem. Phys. 98 (1993) : 5648-5652.
- [53] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. J. Phys. Chem. 98 (1994) : 11623-11627.
- [54] Schrödinger. Jaguar 5.5 [Software]. 2003.
- [55] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. J. Chem. Phys. 118 (2003) : 9095-9102.
- [56] University School of Medicine, St. Louis. TINKER: Software Tools for Molecular Design [Software]. 2003.
- [57] Harvey, J. N. Spin-forbidden CO ligand recombination in myoglobin. Faraday Discuss. 127 (2004) : 165-177.
- [58] Chiericatti, G.; Santi, D. V. Aspartate 221 of thymidylate synthase is involved in folate cofactor binding and in catalysis. Biochemistry 37 (1998) : 9038-9042.
- [59] Hardy, L. W.; Finer-Moore, J. S.; Montfort, W. R.; Jones, M. O.; Santi, D. V.; Stroud, R. M. Atomic structure of thymidylate synthase: target for rational drug design. Science 235 (1987) : 448-455.
- [60] Garrett, C.; Wataya, Y.; Santi, D. V. Thymidylate synthetase. Catalysis of dehalogenation of 5-bromo- and 5-iodo-2'-deoxyuridylate. Biochemistry 18 (1979) : 2798-804.

APPENDIX

A. Conferences (Oral and Poster presentation)

1. The 11th Annual National Symposium on Computational Science and Engineering, 28-30 March 2007 at Prince of Songkla University, Thailand
2. The 4th Thai Summer School of Computational Chemistry, 2 - 7 April 2007 at Ubon Ratchathani University, Thailand
3. The 12th Annual National Symposium on Computational Science and Engineering, 27-29 March 2008 at Ubon Ratchathani University, Thailand
4. The 5th Thai Summer School of Computational Chemistry, 13-15 October 2008 at Suranaree University of Technology, Thailand
5. The 34th Congress on Science and Technology of Thailand, 31 October - 2 November 2008 at Queen Sirikit National Convention Centre, Thailand
6. The 6th Thai Summer School of Computational Chemistry, 19-21 October 2009 at Mahasarakham University, Thailand
7. The Second National Seminar and Workshop on Computer Aided Drug Design, 8-11 December 2009 at Universiti Sains Malaysia, Malaysia
8. 4th *Asian Pacific Conference* on Theoretical & Computational Chemistry, 21-23 December 2009 at Port Dickson, Malaysia
9. UK-Malaysia Symposium on Medical Chemistry, 28-29 January 2010 at Sains Malaysia, Malaysia
10. The 14th Annual National Symposium on Computational Science and Engineering, 23-26 March 2010 at Ubon Ratchathani University, Thailand
11. Thailand Conference on Emerging Infectious and Neglected Diseases, 3-4 June 2010 at Amari Orchid Resort&Tower Pattaya Chonburi, Thailand
12. The Thai Student Symposium Progression, School of Chemistry, 8 December 2011 at School of Chemistry, University of Bristol, Bristol, United Kingdom
13. Computational Molecular Science 2012, 24-26 June 2012 at Royal Agricultural College, Cirencester, United Kingdom
14. The 1st CCP-BioSim Conference: Frontiers of Biomolecular Simulation, 27 June 2012 at Royal Agricultural College, Cirencester, United Kingdom
15. The 9th Thai Summer School of Computational Chemistry, 22-25 October 2012 at Department of Chemistry, Faculty of Science, Thaksin University, Phattalung, Thailand

16. The 8th Mathematics and Physical Sciences Graduate Congress, 8-10 December 2012 at Faculty of Science, Chulalongkorn University, Bangkok, Thailand

B. Research Visiting

- November, 2009 - February, 2010** Pharmaceutical Design and Simulation Laboratory, School of Pharmaceutical Sciences, Universiti Sains Malaysia, Malaysia, with Dr. Habibah A. Wahab
- November, 2011 - August, 2012** Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, United Kingdom, with Prof. Dr. Adrian J. Mulholland

C. Scholarship

- 2006** Science Achievement Scholarship of Thailand
- 2009** The Royal Golden Jubilee Ph.D. Program, Thailand Research Fund

D. Publications (* Corresponding Author)

1. P. Intharathep, C. Laohpongspaisan, T. Rungrotmongkol, A. Loisuangsinsin, M. Malaisree, P. Decha, O. Aruksakunwong, K. Chuenpennit, **N. Kaiyawet**, P. Sompornpisut, S. Pianwanit, S. Hannongbua*, "How amantadine and rimantadine inhibit proton transport in the M2 protein channel", *Journal of molecular graphics and modeling*, 27, 2008, 342-348.
2. T. Rungrotmongkol, P. Intharathep, M. Malaisree, N. Nunthaboot, **N. Kaiyawet**, P. Sompornpisut, S. Payungporn, Y. Poovorawan, S. Hannongbua* "Susceptibility of antiviral drugs against 2009 influenza A (H1N1) virus", *Biochemical and Biophysical Research Communications*, 385, 2009, 390-394.
3. N. Nunthaboot, T. Rungrotmongkol, M. Malaisree, P. Decha, **N. Kaiyawet**, P. Intharathep, P. Sompornpisut, Y. Poovorawan, S. Hannongbua*,

- “Molecular insights into human receptor binding to 2009 H1N1 influenza A hemagglutinin”, *Monatshefte Fur Chemie*, 141, 2010, 801-807.
4. N. Nunthaboot, T. Rungrotmongkol, M. Malaisree, **N. Kaiyawet**, P. Decha, P. Sompornpisut, and S. Hannongbua*, Evolution of the human-receptor binding affinity of influenza A (H1N1) 2009 pandemic hemagglutinin from 1918- and 1930-H1N1 influenza viruses, *Journal of Chemical Information and Modeling*, 50, 2010, 1410-1417.
 5. T. Rungrotmongkol*, N. Nunthaboot, M. Malaisree, **N. Kaiyawet**, P. Intharathep, A. Meeprasert and S. Hannongbua, Molecular Insight into the Specific Binding of ADP-ribose to the nsP3 Macro Domains of Chikungunya and Venezuelan Equine Encephalitis Viruses: Molecular Dynamics Simulations and Free Energy Calculations, *Journal of molecular graphics and modeling*, 29, 2010, 347-353.
 6. P. Intharathep, T. Rungrotmongkol, P. Decha, N. Nunthaboot, **N. Kaiyawet**, T. Kerdcharoen, P. Sompornpisut, and S. Hannongbua*, Evaluating how rimantadines control the proton gating of the Influenza A M2-proton port via allosteric binding outside of the M2-channel: MD simulations, *Journal of Enzyme Inhibition and Medicinal Chemistry*, 26, 2011, 162-168.
 7. **N. Kaiyawet**, T. Rungrotmongkol and S. Hannongbua*, Probable polybasic residues inserted into the cleavage site of the highly pathogenic avian influenza A/H5N1 hemagglutinin: speculation of the next outbreak in humans, *International Journal of Quantum Chemistry*, 113, 2013, 569-573.

BIOGRAPHY

Personal Data

Name Nopporn Kaiyawet
Nationality Thai
Date of Birth 14 November 1984
E-mail Address thm_thm_@hotmail.com, n.kaiyawet@gmail.com

Education

2003-2006 B.Sc. in Chemistry
Department of Chemistry, Faculty of Science,
Chulalongkorn University, Bangkok, Thailand

2007-2008 M.Sc. in Chemistry
Department of Chemistry, Faculty of Science,
Chulalongkorn University, Bangkok, Thailand

2009-2012 Ph.D. student in Physical Chemistry, Faculty of Science,
Chulalongkorn University, Bangkok, Thailand