การยึดเหนี่ยวและพลวัตของสารประกอบเชิงซ้อนของนิวรามินิเดสแบบชนิดย่อยเอ็นหนึ่งกับสาร
ยับยั้งและกับซับสเตรตด้วยการจำลองพลวัตเชิงโมเลกุลและการจำลองคิวเอ็ม/เอ็มเอ็ม เอ็มดี

นางสาวมธุรส มาลัยศรี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาเคมี   ภาควิชาเคมี
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553

BINDING AND DYNAMICS OF NEURAMINIDASE SUBTYPE N1
COMPLEXED WITH INHIBITORS AND WITH SUBSTRATE BY MOLECULAR
DYNAMICS AND QM/MM MD SIMULATIONS
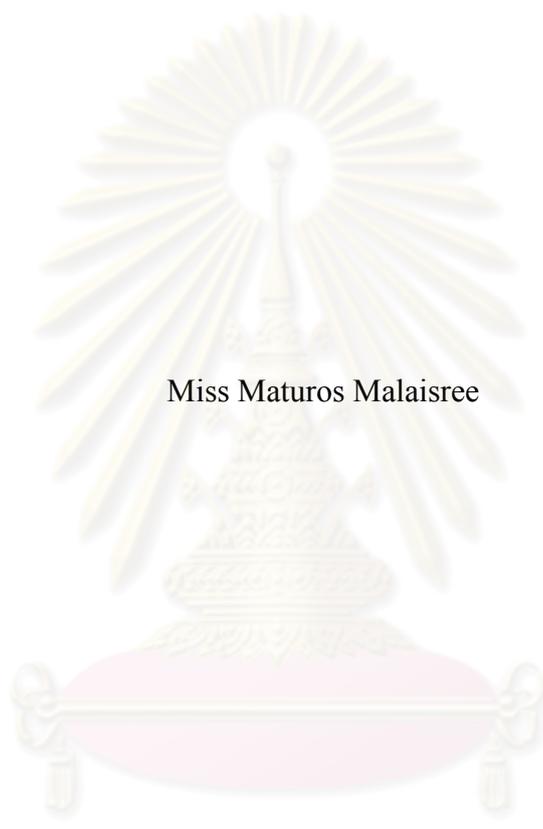
Miss Maturos Malaisree

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 2010

Thesis Title   BINDING AND DYNAMICS OF NEURAMINIDASE
                   SUBTYPE N1 COMPLEXED WITH INHIBITORS AND
                   WITH SUBSTRATE BY MOLECULAR DYNAMICS
                   AND QM/MM MD SIMULATIONS

By   Miss Maturos Malaisree

Field of Study   Chemistry

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

Thesis Co-Advisor Assistant Professor Pornthep Sompornpisut, Ph.D.

                   Professor Adrian Mulholland, 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

(Associate Professor Sirirat Kokpol, Ph.D.)

……………………………………………. Thesis Advisor

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

…………………………………………….. Thesis Co-Advisor

(Assistant Professor Pornthep Sompornpisut, Ph.D.)

…………………………………………….. Thesis Co-Advisor

(Professor Adrian Mulholland, Ph.D.)

…………………………………………….. Examiner

(Associate Professor Vudhichai Parasuk, Ph.D.)

………………………………………….. Examiner

(Assistant Professor Somsak Pianwanit, Ph.D.)

………………………………………….. External Examiner

(Assistant Professor Surapong Pinitglang, Ph.D.)

มธุรส มาลัยศรี: การยึดเหนี่ยวและพลวัตของสารประกอบเชิงซ้อนของนิวรามินิเดสแบบ ชนิดย่อยเอ็นหนึ่งกับสารยับยั้งและกับซับสเตรตด้วยการจำลองพลวัตเชิงโมเลกุลและการ จำลองคิวเอ็ม/เอ็มเอ็ม เอ็มดี. (BINDING AND DYNAMICS OF NEURAMINIDASE SUBTYPE N1 COMPLEXED WITH INHIBITORS AND WITH SUBSTRATE BY MOLECULAR DYNAMICS AND QM/MM MD SIMULATIONS) อ. ที่ปรึกษา วิทยานิพนธ์หลัก: ศ. ดร. สุพจน์ หารหนองบัว, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม: ผศ. ดร. พรเทพ สมพรพิสุทธิ์, ศ. ดร. เอเดรียน มูลฮอลแลนด์ 112 หน้า.

การอุบัติของไวรัสไข้หวัดใหญ่ชนิดเอสายพันธุ์เอชห้าเอ็นหนึ่งและเอชหนึ่งเอ็นหนึ่ง 2009 ได้สร้างความวิตกกังวลไปทั่วโลกถึงการแพร่ระบาดของโรคไข้หวัดใหญ่ เทคนิคต่างๆทางเคมี คอมพิวเตอร์จึงได้ถูกนำมาใช้ศึกษาสมบัติทางโครงสร้าง อันตรกิริยาระหว่างยากับเป้าหมายและ พลังงานเสรีการยึดจับของสารยับยั้งนิวรามินิเดส งานวิจัยนี้ได้ใช้วิธีจำลองพลวัตเชิงโมเลกุลศึกษา สารประกอบเชิงซ้อนของยาที่ใช้ในปัจจุบันสามชนิดคือ โอเซลทามิเวียร์ ซานามิเวียร์และพีรามิเวียร์ กับเอ็นหนึ่ง พบว่าหมู่คาร์บอกซิลิกและกัวนิดิเนียมของพีรามิเวียร์สร้างพันธะไฮโดรเจนกับกรดอะมิ โนรอบข้างได้มากกว่ายาชนิดอื่นโดยเฉพาะกับกรดอะมิโนD151 ซึ่งอยู่ตรงช่วงลูบที่150 ส่วนสายโซ่ ขนาดใหญ่ของยาทั้งสามชนิด พบพันธะไฮโดรเจนที่หมู่ไฮโดรฟิลิกของซานามิเวียร์เท่านั้น ส่วนหมู่ ไฮโดรโฟบิกของโอเซลทามิเวียร์มีขนาดใหญ่เกินไปกับบริเวณเร่งจึงนำไปสู่การลดประสิทธิภาพ ของยาต่อเอ็นหนึ่ง ส่วนการกลายพันธุ์ที่ H274Y พบว่าสาเหตุของการดื้อยาโอเซลทามิเวียร์เกิดจาก การลดขนาดของโพรงไฮโดรโฟบิก ทำให้พลังงานเสรีการยึดจับลดลงจาก -14.6 ± 4.3 เป็น -9.9 ± 6.4 กิโลแคลอรี่ต่อโมล ส่วนการกลายพันธุ์ที่เป็นไปได้ของไวรัสไข้หวัดใหญ่สายพันธุ์ใหม่ปี2009 ที่ R292K, E119V, H274Y และ N294S การลดประสิทธิภาพยาเกิดจากการลดลงของอันตรกิริยาของยา กับเอ็นหนึ่ง อันได้แก่พันธะไฮโดรเจน แรงระหว่างประจุ และแรงแวนเดอร์วาลล์ นอกจากนี้ได้ใช้ วิธีการจำลองทางคิวเอ็ม/เอ็มเอ็ม เอ็มดี ศึกษากลไกขั้นแรกของปฏิกิริยาการตัดซับสเตรตของมนุษย์ กับนิวรามินิเดสสายพันธุ์ต่างๆ ได้แก่ N1-1918, N1-2005, N1-2009, N2-1967 และ N8-1963 พบว่า สายพันธุ์เอ็นหนึ่งและเอ็นสองเท่านั้นที่สามารถเปลี่ยนรูปร่างของซับสเตรตจากรูปแชร์เป็นรูปทวิต โบท ซึ่งถูกทำให้เสถียรด้วยN/Q347และK431 ในขั้นตอนสุดท้ายได้พัฒนาวิธีใหม่ในการทำนายการ ยึดจับชื่อว่า Water-swap reaction coordinate (WSRC) โดยพลังงานเสรีสมบูรณ์ของการยึดจับคำนวณ ได้จากการเปลี่ยนแปลงพลังงานเสรีของการเปลี่ยนกลุ่มของน้ำกับลิแกนด์ในบริเวณเร่งของโปรตีน พบว่าผลที่ได้สอดคล้องกับค่าการทดลอง ดังนั้นWSRC จึงเป็นแนวทางใหม่ในการทำนายการยึดจับ ของโปรตีนกับลิแกนด์ซึ่งสามารถนำไปเป็นเครื่องมือสำคัญในกระบวนการพัฒนายา

ภาควิชา.......เคมี................... ลายมือชื่อนิสิต...

สาขาวิชา....เคมี................... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.......................

ปีการศึกษา..2553................. ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม......................

ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม ....

4873838523 : MAJOR CHEMISTRY

KEYWORDS : Neuraminidase / Influenza virus / Molecular Dynamics Simulations / QM/MM MD simulations / Water-swap reaction coordinate (WSRC)

MATUROS MALAISREE: BINDING AND DYNAMICS OF NEURAMINIDASE SUBTYPE N1 COMPLEXED WITH INHIBITORS AND WITH SUBSTRATE BY MOLECULAR DYNAMICS AND QM/MM MD SIMULATIONS. THESIS ADVISOR: PROF. SUPOT HANNONGBUA, Dr. rer. nat., THESIS CO-ADVISOR: ASST. PROF. PORNTHEP SOMPORNPISUT, Ph.D., PROF. ADRIAN MULHOLLAND, Ph.D., 112 pp.

The emergence of influenza A virus subtypes H5N1 and H1N1-2009 has raised global concerns of a flu pandemic. Computational chemistry approaches were carried out to study structural properties, drug-target interactions and binding free energies of neuraminidase (NA) inhibitors. Three available agents, oseltamivir, zanamivir and peramivir, complexed with N1 were studied using molecular dynamics (MD) simulations. The carboxylate and guanidinium groups of peramivir were found to form many more hydrogen bonds with the surrounding residues compared to the other drugs, especially D151 located in the 150-loop. For the bulky side chain of the three drugs, hydrogen bonds were detected only with the hydrophilic group of zanamivir while the hydrophobic group of oseltamivir was slightly too big to fit within the active site. This leads to the lower efficacy of oseltamivir against the N1 strain. The investigation was extended to study the H274Y mutation. It was found that the source of oseltamivir resistance was due to the reduction of the hydrophobic pocket size which led to a decrease in $\Delta G_{binding}$ from -14.6 ± 4.3 to -9.9 ± 6.4 kcal mol^{-1}. For the novel H1N1-2009 strain, the probable mutations, R292K, E119V, H274Y and N294S, were modelled to predict oseltamivir binding affinity. Reduction in oseltamivir-N1 interaction energies was observed in terms of lower hydrogen bonds, electrostatic and van der Waals interactions. To understand the first step of the NA cleavage mechanism, the natural human substrate (SA-α-2,6-GAL) bound to different NA subtypes such as N1-1918, N1-2005, N1-2009, N2-1967 and N8-1963 was studied using QM/MM MD simulations. SA-α-2,6-GAL in both the N1 and N2 was found to change its conformation from the chair to twist-boat forms. This conformational change was found to be stabilised by the N/Q347 and K431 residues. In the last part, a novel method for binding prediction, called the water-swap reaction coordinate (WSRC), was developed. The free energy change of swapping between an identified water cluster in bulk water and a ligand in a protein active site is, directly, the absolute binding free energy. The agreement of the results with experiment is encouraging. The final results show that WSRC provides a promising new direction of predicting protein-ligand binding, which could be a powerful tool for drug development process.

Department : .....CHEMISTRY...... Student's Signature .....Maturos Malaisree.....

Field of Study : ...CHEMISTRY...... Advisor's Signature .....S. Hannongbua.....

Academic Year : ...2010...... Co-Advisor's Signature ....

Co-Advisor's Signature..................

ACKNOWLEDMENTS

CONTENTS

LIST OF TABLES

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

LIST OF FIGURES

LIST OF ABBREVIATION

3D	=	Three Dimension
ABNR	=	Adopted Basis Newton Raphson
Ala (A)	=	Alanine
Arg (R)	=	Arginine
Asn (N)	=	Asparagine
Cys (C)	=	Cysteine
FDA	=	The US Food and Drug Administration
FDTI	=	Finite-Difference Thermodynamic Integration
FEP	=	Free Energy Perturbation
GAL	=	Galactose
Gln (Q)	=	Glutamine
Glu (E)	=	Glutamic Acid
Gly (G)	=	Glycine
HA	=	Hemagglutinin
His (H)	=	Histidine
Ile (I)	=	Isoleucine
Ki	=	Inhibition Constant
Leu (L)	=	Leucine
LIE	=	Linear Interaction Energy
LJ	=	Lennard-Jones
Lys (K)	=	Lysine
MC	=	Monte Carlo
MD	=	Molecular Dynamics
Met (M)	=	Methionine
MM/PBSA	=	Molecular Mechanic/Possion-Boltzmann Surface Area
NA	=	Neuraminidase
OTV	=	Oseltamivir
PDB	=	Protein Data Bank
Phe (F)	=	Phenylalanine
PMF	=	Potential of Mean Force
Pro (P)	=	Proline

PRV	=	Peramivir
QM/MM	=	Quantum Mechanical/Molecular Mechanical
RDF	=	Radial Distribution Functions
RETI	=	Replica Exchange Thermodynamic Integration
RMSD	=	Root Mean Square Deviation
SASA	=	Solvent Accessible Surface Area
SD	=	Steepest Descents
Ser (S)	=	Serine
SA	=	Sialic acid
TI	=	Thermodynamic Integration
Thr (T)	=	Threonine
Trp (W)	=	Tryptophan
Tyr (Y)	=	Tyrosine
Val (V)	=	Valine
WHO	=	World Health Organization
WSRC	=	Water-Swap Reaction Coordinate
ZNV	=	Zanamivir

CHAPTER I

INTRODUCTION

1.1 IN THIS THESIS

The outbreak of influenza virus subtypes H5N1 and the novel H1N1 (H1N1-2009) has raised concern about a potential global flu pandemic. Among the three known targets determining the virus life's cycle, namely Hemagglutinin (HA), Neuraminidase (NA) and M2 channel, the last step of viral replication is mediated by NA in releasing the virion from the host cell receptor [1]. Currently, three NA inhibitors, zanamivir, oseltamivir and peramivir, are available commercially for treatment of infected patients. However, mutations of the virus that lead to resistance to the available drugs, especially oseltamivir, have been widely reported [2-5]. Understanding of known drug-target interactions in both wild-type (I) and mutant strains (II) as well as NA-substrate interaction (III) is the key to success in designing and discovering new potent inhibitors that fit better into the active site of NA. Therefore, I-III become the rational goal of this work, as studied by molecular modelling approaches. Using the inhibitor-NA as a case study, the calculations were extended to develop a new method that provides a promising route to predict absolute protein-ligand binding free energies (IV).
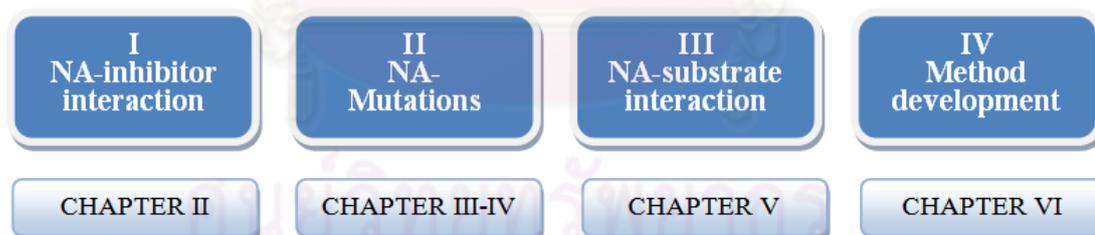
Figure 1.1 Schematic representation of the four parts (I-IV) of this studies in which detailed calculations and investigations are given in chapters II-VI, respectively.

Schematic representation of the four parts (I-IV) is illustrated in Figure 1.1 in which detailed calculations, investigations, results and discussion are given in terms of research articles [6-8] and submitted manuscripts [9,10] in chapters II-VI, respectively. All articles are part of this dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry at Chulalongkorn University.

The aim of this research is to provide the molecular level knowledge of the drug-target interactions between NA and the available commercial agents. Molecular dynamics (MD) simulations and Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) calculations were applied to estimate binding free energies in both the wild-type and the H274Y mutant H5N1 strain. However, predictions of binding free energy using the MM/PBSA method had a large random error. Therefore, the linear interaction energy (LIE) method, which has higher precision and accuracy, was used to predict the efficiency of oseltamivir against the probable H1N1-2009 mutant strains: R292K, E119V, H274Y and N294S. To understand how NAs cleave the terminal sialic acid in the first step of the cleavage mechanism, QM/MM MD simulations of different NA subtypes complexed with the natural substrate were applied. As it is a limitation that the existing methods cannot reliably predict binding free energies when there are large differences in free and bound conformations of protein and ligand complex. To overcome this problem, a new method, called the water-swap reaction coordinate (WSRC), has been developed. This provides a promising new route to predict absolute protein-ligand binding free energies.

1.2 RESEARCH BACKGROUND AND RESEARCH RATIONALE

1.2.1 History

Three major influenza pandemics of human infection were caused by the H1N1 (Spanish flu - 1918), H2N2 (Asian flu - 1957) and H3N2 (Hong Kong flu - 1968) subtypes, leading to numerous deaths around the world [1]. Recently, two new strains of influenza virus, H5N1 and H1N1-2009, have caused wide-spread global infections in both animal and human populations. The avian influenza virus subtype H5N1 is highly pathogenic causing many human deaths (approximately 300 cases between 1997 and 2010 [11]) in Asia and Europe. The first outbreak appeared in Hong Kong and this subtype has been found to be capable of crossing species from avian to human [12]. Subsequently, the first influenza pandemic of the 21st century was caused by a subtype of H1N1 that emerged in 2009. The H1N1-2009 infection has rapidly spread by human-to-human transmission throughout the world, especially in the United States and Mexico [13]. The only publically available oral drug, oseltamivir, has been used widely. However, this widespread use of oseltamivir has

led to the development of drug resistance in both the H5N1 and H1N1 subtypes [2-5, 14, 15]. This has led to an urgent need to understand influenza virus function in order to develop new potent antiviral agents against H5N1 and H1N1-2009.

1.2.2 Influenza A virus

The avian (H5N1) and the novel (H1N1-2009) influenza viruses belong to the genus of influenza virus type A in the *Orthomyxoviridae* family [16]. The surface membrane of influenza virus contains three important proteins (see also Figure 1.2): Hemagglutinin (HA: H1-H16), the M2 protein channel and neuraminidase (NA: N1-N9) [1]. HA plays a role in attaching the receptor, sialic acid, to the host cells and allowing penetration into the target cell. Consequently, the M2 protein acts as an ion channel allowing proton transport which adjusts the internal *pH* of the virus. After virus replication, NA cleaves the terminal sialic acid from the host cells and releases the new viruses to infect other cells [17]. Influenza viruses are classified into different subtypes based on the antigenic relationship of HA and NA proteins. Only HA subtypes H1-H3 and NA subtypes N1-N2 have been found to infect humans. The ability of influenza to infect a host is influenced by the different types of linkage between sialic acid and galactose on the host cell surface [18]. Human influenza virus attaches to host substrate via a SA-α-2,6-GAL linkage, while the avian influenza virus binds preferentially to the SA-α-2,3-GAL linkage [19]. Until now, there are two types of antiviral agents for clinical treatment, targeting the M2 and NA proteins. The first set of drugs, amantadine and rimantadine, are M2 inhibitors, which act by preventing proton transfer through the M2 channel. They are limited by side effects and drug resistance due to several mutations [17]. The second set of commercial drugs, zanamivir, oseltamivir and peramivir, were designed to block NA function. Because of the important role in viral replication, NA has become the main target for drug development against influenza virus [20].
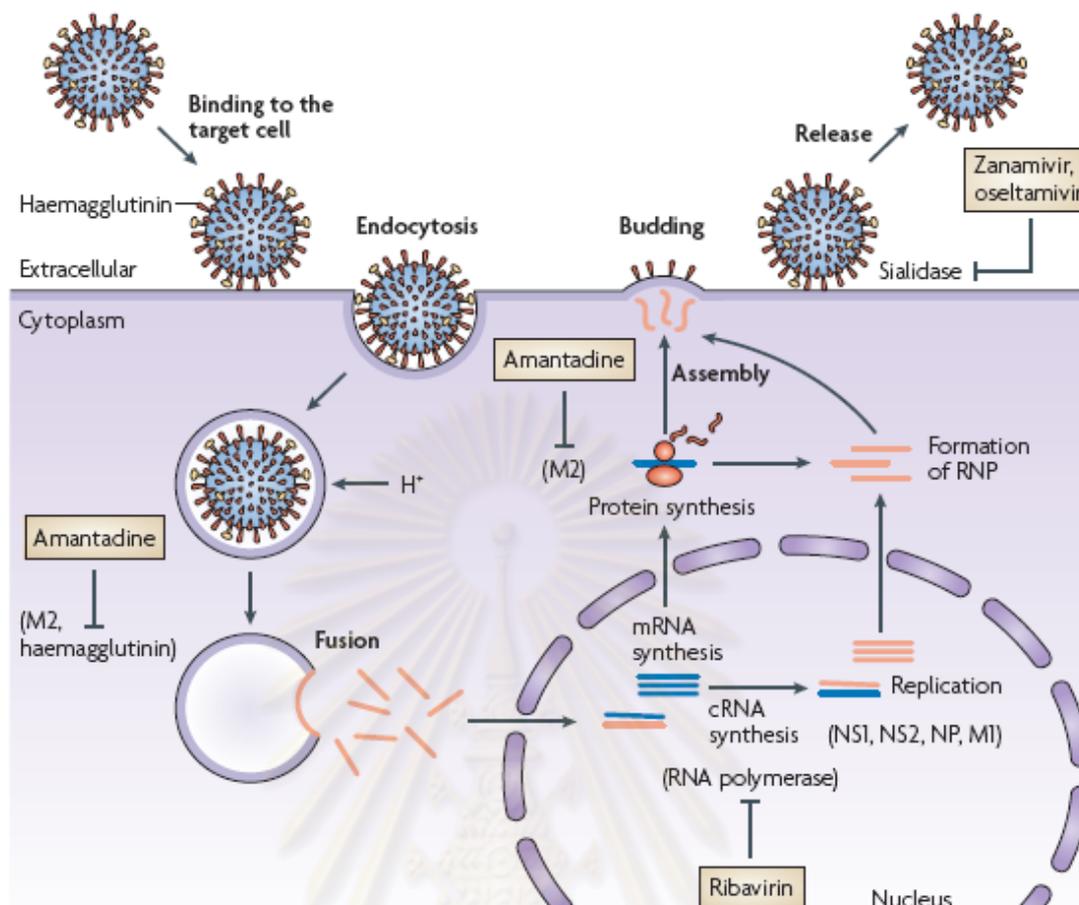
Figure 1.2 Life cycle of influenza virus and target of antiviral agents [1]

1.2.3 Neuraminidase

Neuraminidase (NA) is a membrane glycoprotein of influenza virus that cleaves the terminal sialic acid from the host cells and releases the new viruses to infect the other cells. NA is classified into two groups: group-1 (N1, N4, N5 and N8) and group-2 (N2, N3, N6, N7 and N9) [21]. After the crystal structures of group-2 influenza NA complexed with sialic acid were determined [22], many inhibitors were designed based on the active site of group-2 NA and have been used for all subtypes. However, the large cavity adjacent to the active site on the 150-loop (residues 147-152) in group-1 NA was determined by Russell *et al.* in 2006 [23]. This leads to the different formation of open and closed conformations in group-1 NA due to the position of the D151 and E119 residues. Based on the crystal structures in both groups, the important amino acid residues in the NA active site are conserved in all subtypes. The catalytic sites (R118, D151, D152, R224, E276, R292, R371 and Y406) were found to interact with the sialic acid while the framework sites (E119, R156,

W178, S179, D198, I222, E227, H274, E277, N294, and E425) (in N2 numbering) were stabilized the active site structure [24]. Although the active site of both groups is highly conserved, the activity of the current inhibitors against the N1 subtype has significantly lower efficiency than the other subtypes [25, 26]. This evidence calls for an urgent need to design new selective inhibitors for the N1 subtype.

1.2.4 NA inhibitors

NA inhibitors have been designed and synthesized based on the transition state analog of sialic acid. Three NA inhibitors, zanamivir (Relenza), oseltamivir (Tamiflu), and Peramevir, have been approved by the FDA for the treatment of influenza virus [27]. Zanamivir is administered by oral inhalation while oseltamivir is the only oral-active drug that is commercially available. The third drug, peramivir, is a potent and selective inhibitor against NA function but it failed to demonstrate efficacy in clinical trials when administered orally [17]. Recently, peramivir has been developed for intramuscular/intravenous injection and it has been authorized for treatment patients against H1N1-2009 subtype [27]. Moreover, the other NA inhibitors, pyrrolidine analogue (A-315675) [28], R-125489 [29], glycomonomers of sialic acid [30], phospha-oseltamivir (Tamiphosphor) monoester [31] and sialoglycoconjugates analog [32] are in development and have been found to provide a new potent structural template against NA in several subtypes in both wild-type and mutant strains.

1.2.5 Mutations

High resistance to available drugs is the major concern in the current treatment of infected patients with influenza viruses in both H5N1 and H1N1-2009 strains [2-5, 14, 15]. The different mutations in drug resistance profiles of zanamivir, oseltamivir and peramivir were characterized by different substitution groups of the inhibitors in binding with the active site residues [33]. Due to an increase in medical use for clinical treatment, the emergence of oseltamivir resistance and several mutations have been reported with the reduction of oseltamivir sensitivity [14, 15]. The mutated framework residues H274Y and N294S are indentified in the N1 subtype [33-35], while mutations on the binding residues, E119V and R292K, were detected in the N2 and N9 subtypes after treatment with infected patients [36, 37]. For H5N1 and H1N1-

2009 strains, the H274Y mutation leads to a high-level resistance with a 300– to 1,700–fold reduction to oseltamivir susceptibility [38-40]. The H274Y is resistant to peramivir due to the same bulky hydrophobic substitution with oseltamivir but zanamivir retains its susceptibility with this mutation. Because the structure of zanamivir is similar with the natural substrate in comparison with oseltamivir and peramivir, zanamivir could bind with amino acid residues in NA active site without disrupting their conformation [41]. However, zanamivir resistance has been detected with the E119/G/A/D mutations in the N2 and N9 subtypes [3, 33]. Based on the increase of resistance to commercial NA drugs, this evidence calls for an urgent need to design new selective inhibitors for the N1 subtype. In order to achieve a better chance of improving and developing new NA inhibitors, an understanding of current drug-target interactions and the mechanism of drug-resistance at the molecular level may provide information that would be useful for rational drug design.

1.2.6 Research rationale

As stated above, the outbreak of the influenza virus subtypes H5N1 and the novel H1N1-2009 as well as the increase of resistance to commercial NA drugs call for an urgent need to design new selective inhibitors for the N1 subtype. Understanding of known NA-drug and the mechanism of drug-resistance as well as NA-substrate interactions at the molecular level are the keys of success. Molecular dynamics simulation is the method of choice in studying microscopic properties, such as, structural, dynamics, thermodynamics as well as solvation data of the biological systems, which is of great fundamental and practical importance in structural based drug designed. In addition, free energy calculation measures the strength of binding between a ligand and a protein, and an algorithm that would allow its accurate prediction would be a powerful tool for drug development.

1.3 OBJECTIVES

In this study, several computational chemistry methods, based on Molecular Dynamic (MD), Monte Carlo (MC) and hybrid Quantum Mechanical/Molecular Mechanical (QM/MM) models were used to provide a molecular level insight into the interactions between NA and commercially available drugs and its natural substrate. The aims of these simulations were;

1. To understand and provide detailed information about the drug-target interactions, structure, solvation and dynamic properties of NA subtype H5N1 complexed with the commercial drugs, zanamivir, oseltamivir and peramivir.

2. To find the primary source of oseltamivir resistance due to the H274Y mutation in terms of conformational changes, intra- and intermolecular interactions of the H274Y residue and binding free energy of the complex, focused on the catalytic pocket of NA.

3. To predict the efficiency of oseltamivir against the probable mutants in the novel H1N1-2009 strain: R292K, E119V, H274Y and N294S.

4. To study the catalytic mechanism of the natural substrate in different NA subtypes and to investigate the function of amino acid residues in the binding pocket for the first step of the cleavage mechanism.

5. To develop a new method of calculating absolute binding free energies using a single simulation.

1.4 SCOPE OF THE DISSERTATION

This dissertation was carried out in order to study and investigate structure, binding, interactions, solvation and dynamic properties of NA. As shown in Figure 1.1, the scope of this study is one-by-one corresponding to the research objectives:

1. NA-inhibitor interactions: Three NA inhibitors, oseltamivir, zanamivir, and peramivir, embedded in the catalytic site of NA subtype H5N1 were studied using MD simulations and calculated binding free energies using MM/PBSA method.

2. NA-mutation (H274Y): Oseltamivir bound to the H274Y mutant of the H5N1 strain was modelled using MD simulations. The estimated binding free energy was calculated via the MM/PBSA method.

3. NA-mutation (H1N1-2009): Prediction of oseltamivir efficiency in probable mutants, R292K, E119V, H274Y and N294S, was calculated using the LIE method.

4. NA-substrate interactions: Substrate-enzyme interactions between the natural substrate, SA-α-2,6-GAL, and different NA subtypes: N1-1918, N1-2005, N1-2009, N2-1967 and N8-1963, were studied using QM/MM MD simulations.

5. Method development: Water-swap reaction coordinate (WSRC), a new method for calculating the absolute protein-ligand binding free energies was developed based on first principles using a single simulation.

1.5 EXPECTED RESULTS

The main goal of this study is to understand NA function using the known drug-target interactions from the commercial agents, zanamivir, oseltamivir and peramivir. The obtained information at the molecular level would be useful for drug development against NA. In order to overcome the oseltamivir-resistance problems caused by several mutations of NA, understanding of the primary source of drug resistance in N1 subtypes, H5N1 and H1N1-2009, could provide a better chance for designing the new potent inhibitors. This was followed by a study of the mechanism of substrate binding. The molecular details in the first step of the NA reaction calculated using high accuracy QM/MM MD simulations was used to understand the function of amino acid residues in the binding pocket. The obtained results can directly help in the design of highly effective inhibitors that can fit better into the NA active site in both the wild-type and mutant strains. Finally, the accurate prediction of absolute protein-ligand binding free energies was enabled by developing a new method, called the water-swap reaction coordinate (WSRC). The application of WSRC method will be able to use for binding calculation with different inhibitors and different NA subtypes as well as the other protein-ligand systems in subsequent work. The WSRC method will be implemented in Sire: a complete molecular simulation framework (http://siremol.org/).

CHAPTER II

NA-INHIBITOR INTERACTIONS

Understanding of known drug-target interactions in the catalytic pocket
of neuraminidase subtype N1

Maturos Malaisree[1], Thanyada Rungrotmongkol[1], Panita Decha[1],
Pathumwadee Intharathep[1], Ornjira Aruksakunwong[2] and Supot Hannongbua[1]

---

[1]*Department of Chemistry, Faculty of Science, Chulalongkorn University, Phayathai
Road, Patumwan, Bangkok 10330, Thailand*

[2]*Department of Chemistry, Faculty of Science, Rangsit University, Pathumtani, 12000
Thailand*

2.1 ABSTRACT

To provide detailed information and insight into the drug-target interaction, structure, solvation, dynamic and thermodynamic properties, the three known-neuraminidase (NA) inhibitors, oseltamivir (OTV), zanamivir (ZNV) and peramivir (PRV), embedded in the catalytic site of NA subtype N1 were studied using molecular dynamics simulations. In terms of ligand conformation, there were major differences in the structures of the guanidinium and the bulky groups. The atoms of the guanidinium group of PRV were observed to form many more hydrogen bonds with the surrounded residues and were much less solvated by water molecules, in comparison with the other two inhibitors. Consequently, D151 lying on the 150-loop (residues 147-152) of group-1 NA (N1, N4, N5 and N8) was considerably shifted to form direct hydrogen bonds with the –OH group of the PRV, which was located rather far from the 150-loop. For the bulky group, direct hydrogen bonds were detected only between the hydrophilic side chain of ZNV and residues R224, E276 and E277 of N1 with rather weak binding, 20-70% occupation. This is not the case for OTV and PRV, in which flexibility and steric effects due to the hydrophobic side chain lead to the rearrangement of the surrounded residues, *i.e.*, the negatively charged side chain of E276 was shifted and rotated to form hydrogen bonds with the positively charged moiety of R224. Taking into account all the ligand-enzyme interaction data, the gas phase MM interaction energy of -282.2 kcal mol^{-1} as well as the binding free energy ($\Delta G_{binding}$) of -227.4 kcal mol^{-1} for the PRV-N1 are significantly lower than those of the other inhibitors. The ordering of $\Delta G_{binding}$ of PRV < ZNV < OTV agrees well with the ordering of experimental IC_{50} value.

2.2 INTRODUCTION

The avian influenza subtype H5N1 is a virulent disease with results in significant outbreaks throughout the world, causing numerous human and animal deaths [42]. High resistance to commercially available drugs is the most important risk in dealing with this viral subtype. Understanding of known drug-target interactions is the key to success in designing and discovering new inhibitors that fit better to the catalytic sites of the targets in both wild and mutant types.

The surface membrane of the influenza virus is known to contain three important proteins, including 16 subtypes of hemagglutinin (HA: H1-H16), 9 subtypes of neuraminidase (NA: N1-N9) and the M2 protein channel [12]. HA plays a role in attachment to sialic acid, a receptor on the cell surface, allowing penetration into the target cell. Consequently, M2 functions as an ion channel which regulates the internal *pH* of the virus. After virus replication, NA cleaves the terminal sialic acid from the receptors and releases the progeny virus to infect other cells.

According to its phylogenetic tree, NA was classified into two groups: group-1 (N1, N4, N5 and N8) and group-2 (N2, N3, N6, N7 and N9) [21]. Recently, crystal structures of some group-1 proteins (N1, N4 and N8) were published in 2006 by Russell *et al.* [23]. In contrast to group-2 proteins, a large cavity adjacent to the active site on the 150-loop (residues 147-152) of group-1 was determined. This leads consequently to the formation of both closed and open conformations in group-1 proteins in which both forms can bind with the inhibitors. It was also proposed that the difference among the two groups is due to the position and orientation of the D151 and E119 residues.

There are two classes of anti-influenza virus inhibitors that could possibly prevent an influenza pandemic, attacking the M2 channel and NA target. Amantadine and rimantadine are the two inhibitors that act by hindering proton transfer through the M2 channel. They are limited by side effects, and hence, emergence of viral resistance. Interest is then focused on the NA target, aiming to prevent viral replication of NA protein [17]. Due to the lack of the 3D structure of group-1 NA, progress and development of this type of inhibitor was based on the structure of the group-2 NA complexed with its transition state analog of the substrate, sialic acid [20, 22]. To date, two NA inhibitors, zanamivir and oseltamivir, are available commercially while a third, peramivir (BCX-1812) fails to show statistically significant inhibition due to the relatively low blood levels that were obtained following oral administration [43, 44]. However, high resistance and mutation to the available drugs, especially oseltamivir, have been widely reported [40, 45, 46]. The N1 subtype treated with oseltamivir and zanamivir demonstrates significantly increased IC_{50} compared with peramivir. The chemical structures of the three NA inhibitors are shown in Figure 2.1.

To understand and provide detailed information and insight into the drug-target interaction, structure, solvation, dynamic and thermodynamic properties, the three NA inhibitors, oseltamivir (OTV), zanamivir (ZNV) and peramivir (PRV), embedded in the catalytic site of neuraminidase subtype N1, were studied using molecular dynamics simulations.
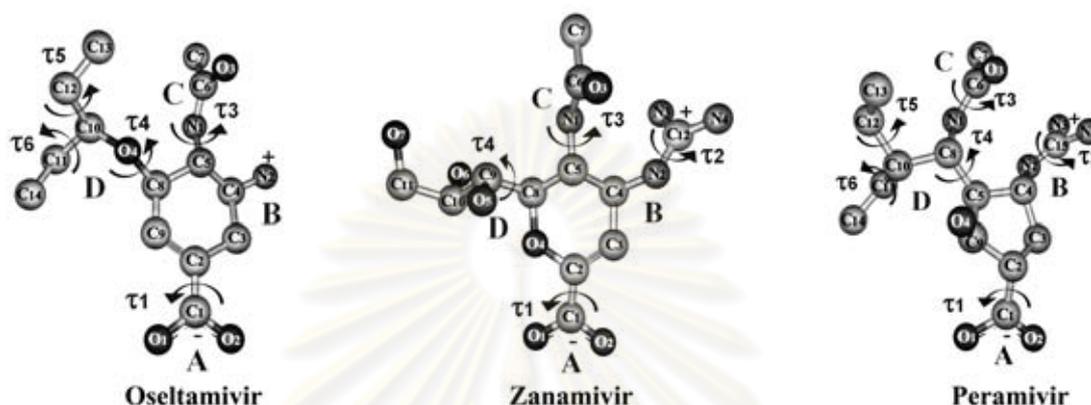
Figure 2.1 Three-dimensional structures of the three neuraminidase inhibitors, oseltamivir, zanamivir and peramivir, where the four side chains are labeled as A-D and torsional angles were defined.

2.3 METHODS

2.3.1 Starting structure and protein preparation

The X-ray structure of neuraminidase subtype N1 complexed with oseltamivir, OTV-N1, was obtained from the Protein Data Bank (PDB), code 2HU4. The atomic coordinates of ZNV and PRV inhibitors were taken from N9 (1NNC) and N8 (2HTU) cocrystal complexes, respectively. To prepare the ZNV-N1 and PRV-N1 complexes, superpositions of ZNV-N9 and PRV-N8 with N1 were performed. The N9 and N8 enzyme coordinates were then removed and ZNV-N1 and PRV-N1 were obtained. All missing hydrogen atoms of the proteins were added using the LEaP module in the AMBER 7 software package [47]. All ionizable side chains of amino acid residues in the protein were configured in their characteristic ionization states at *pH* 7.0 using the LEaP module of AMBER. For the histidine residue, mono-protonation on the nitrogen atom was obtained.

The starting structures and force field parameters for the three inhibitors were obtained as follows. Hydrogen atoms were added to the X-ray coordinates of the three ligands by taking into account the hybridization of the covalent bonds. To obtain minimized geometries for the electrostatic potential calculations, ligand geometries were optimized using Gaussian03 [48] with the HF/6-31G* basis set to adjust bond lengths and angles involving hydrogen atoms. Single-point calculations with Gaussian03 were then carried out to compute the electrostatic potential around each compound using the same basis set and level of theory as in the optimization step. The RESP charges were generated from the HF/6-31G* quantum mechanical data by fitting with the RESP module of AMBER [49]. Three inhibitors in zwitterion form (see Figure 2.1) were given a formal charge of 0. Partial charge generation and assignment of the force field were performed using the Antechamber suite [50].

To incorporate the solvent and counterions, each system was solvated with a TIP3P [51] water box with the minimum distance of 10 Å from the protein surface to the edge of the simulation box. Then the solvated box dimensions were set to 78 Å X 80 Å X 81 Å. All neutralization by the counterions was treated using the LEaP module. The total atoms were 41,313, 41,013, and 41,018 for the OTV-N1, ZNV-N1, and PRV-N1 systems, respectively.

2.3.2 Molecular dynamics simulations

Energy minimization and MD simulations were performed using the SANDER module of AMBER [47]. An all-atom representation of the system was used, employing the Cornell *et al.* [52] force field to assign parameters for the standard amino acids. The periodic boundary condition with the NPT ensemble was applied. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and standard pressure of the system [53]. The SHAKE algorithm [54] was applied to constrain all bonds involving hydrogen atoms and the simulation time step of 2 fs was used. All MD simulations were run with a 10 Å residue-based cutoff for non-bonding interactions and the particle mesh Ewald method was used for an adequate treatment of long-range electrostatic interactions [55]. The simulation steps consist of thermalization, equilibration and production phases. Initially, the temperature of the system was gradually increased from 0 K to 298 K during the first 60 ps. Then, the system was maintained at 298 K until the MD simulations reached equilibration

phase. Finally, the production phases were from 1.5 ns to 3.0 ns for OTV-N1 and PRV-N1 and 2.0 ns to 3.5 ns for ZNV-N1. The convergence of energies, temperature, pressure and global RMSD was monitored to verify the stability of the system. The MD trajectory was collected every 0.2 ps and analyzed in terms of RMSD, torsion angles, hydrogen bonds, and solvation and binding free energies using the CARNAL, Ptraj and MM-PBSA modules of AMBER.

2.3.3 Binding free energy calculations

The free energy of binding, $\Delta G_{binding}$, was calculated according to Eq. 1 from the free energy of the receptor-ligand complex (G_{cpx}) with respect to the unbound receptor (G_{rec}) and ligand (G_{lig}):

$$\Delta G_{binding} \quad = \quad G_{cpx} - (G_{rec} + G_{lig}) \tag{1}.$$

The MM-PBSA (Molecular Mechanics - Poisson-Boltzmann/Surface Area) methodology allows the calculation of the complete binding reaction energy, including the desolvation of the ligand and the unbound protein, on the basis of a thermodynamic cycle. Therefore, Eq. 1 can be approximated as

$$\Delta G_{binding} \quad = \quad \Delta E_{MM} - T\Delta S + \Delta G_{sol} \tag{2}$$

$$\Delta E_{MM} \quad = \quad \Delta E_{ele} + \Delta E_{vdw} \tag{3}.$$

All energies represented in the above equations were averaged over the course of the molecular dynamics trajectories. In Eq. 3, ΔE_{MM} is the molecular mechanical energy obtained from the electrostatic (ΔE_{ele}) and the van der waals (ΔE_{vdw}) interactions within the system. Here, $T\Delta S$ is the solute entropic contribution at temperature T (Kelvin) and the solvation free energy (ΔG_{sol}) represents the electrostatic and nonpolar free energy of solvation, and therefore can be expressed as

$$\Delta G_{sol} = \Delta G_{sol}^{ele} + \Delta G_{sol}^{nonpolar} \tag{4}$$

where ΔG_{sol}^{ele} is the polar contribution to solvation and $\Delta G_{sol}^{nonpolar}$ is the nonpolar solvation term. The former component was calculated using the PB calculation, whereas the latter term is determined using Eq. 5:

$$\Delta G_{sol}^{nonpolar} = \gamma SASA + b \tag{5}$$

where SASA is the solvent-accessible surface area ($Å^2$) and $γ$ and b represent experimental solvation parameters, which are 0.00542 kcal $mol^{-1}Å^{-2}$ and 0.92 kcal mol^{-1}, respectively [56]. The $ΔG_{binding}$ value was obtained using the MM-PBSA module in the program AMBER 7, which interfaces the program DelPhi 4 [57]. To calculate the electrostatic free energy of solvation, the grid resolution of 0.33 Å and the boundary conditions of Debye-Hückel potentials were employed. Atomic charges were taken from the Cornell force field [52]. The water dielectric value was set to 80. The protein dielectric value of 4, normally used for recent MM/PBSA studies in protein structure prediction [58], was used for our study. As an estimation of the entropic effect from normal mode analysis requires high computational demands, based on an assumption that the entropic difference among the three systems should be very small because all system models, OTV, ZNV and PRV complexed with N1, are quite similar. Therefore, the contribution of the entropy ($TΔS$) was not included in this study. In addition, interpretation will focus only on the relative values of the binding free energy.

2.4 RESULTS AND DISCUSSION

2.4.1 Overall enzyme and inhibitor structure

The root mean square displacement (RMSD) of the overall structure and three inhibitors with respect to the initial configuration were evaluated and plotted in Figure 2.2. Substantial changes were observed for the OTV-N1 and ZNV-N1 systems while no significant difference was found for PRV-N1. The OTV-N1 and PRV-N1 systems were found to reach equilibrium at 1.5 ns while ZNV-N1 reached equilibrium at 2 ns.
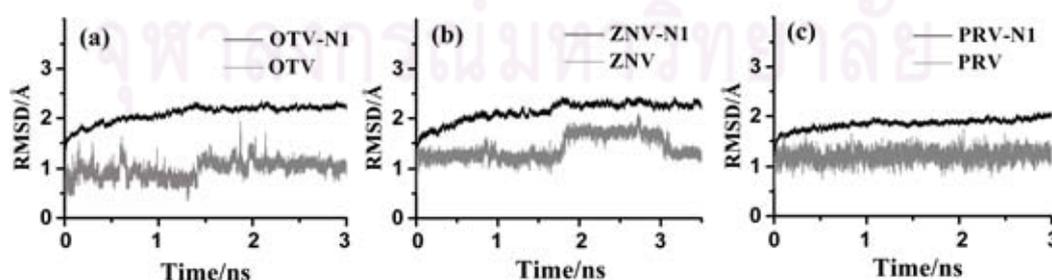
Figure 2.2 RMSDs relative to the initial structure for all atoms of the complexes (black) and inhibitors, oseltamivir (OTV), zanamivir (ZNV) and peramivir (PRV) (grey) for the three systems: (a) OTV-N1, (b) ZNV-N1 and (c) PRV-N1.

2.4.2 Changes of the inhibitor structure

To assess the structural flexibility of the three inhibitors embedded in the N1 pocket, the torsional angles of their side chains were monitored. Six torsional angles, $\tau 1$-$\tau 6$, are described by sets of four atoms as defined in Table 2.1 and indicated by arrows in Figure 2.1. The plots were evaluated over the last 1.5 ns of MD trajectories of the three different systems and are compared in Figure 2.3.

Considering the carboxylate group (side chain A in Figure 2.1), all $\tau 1$ torsional angles for the N1 complexes are found at almost 0°. This means that the –COO⁻ group was observed to lie parallel to the ligand's ring. This is similar to that observed for the rotation of the –NHAc group (side chain C in Figure 2.1), where $\tau 3$ angles were detected in the range between -90° and -150°.

Dramatically structural variations were observed for $\tau 2$, $\tau 4$, $\tau 5$ and $\tau 6$. The $\tau 2$ torsional angle of the guanidinium group (side chain B in Figure 2.1) in the ZNV-N1 and PRV-N1 complexes appears at -127.5° and 127.5°, representing the out-of- and in-plane conformations, respectively. The positively charged guanidinium group of peramivir was originally designed to interact better with the residues in the binding pocket, structurally unlike the orientation of zanamivir [59]. This information was intensively analyzed and explained in the solvation section.

Table 2.1 Definition of torsional angles ($\tau 1$-$\tau 6$, shown in Figure 2.1) for the three inhibitors, oseltamivir (OTV), zanamivir (ZNV) and peramivir (PRV).

τ	OTV	ZNV	PRV
1	C_9-C_2-C_1-O_1	O_4-C_2-C_1-O_1	C_9-C_2-C_1-O_1
2	-	C_4-N_2-C_{12}-N_4	C_4-N_2-C_{15}-N_4
3	C_4-C_5-N_1-C_6	C_4-C_5-N_1-C_6	C_5-C_8-N_1-C_6
4	C_5-C_8-O_4-C_{10}	C_5-C_8-C_9-C_{10}	C_4-C_5-C_8-C_{10}
5	O_4-C_{10}-C_{12}-C_{13}	-	C_8-C_{10}-C_{12}-C_{13}
6	O_4-C_{10}-C_{11}-C_{14}	-	C_8-C_{10}-C_{11}-C_{14}
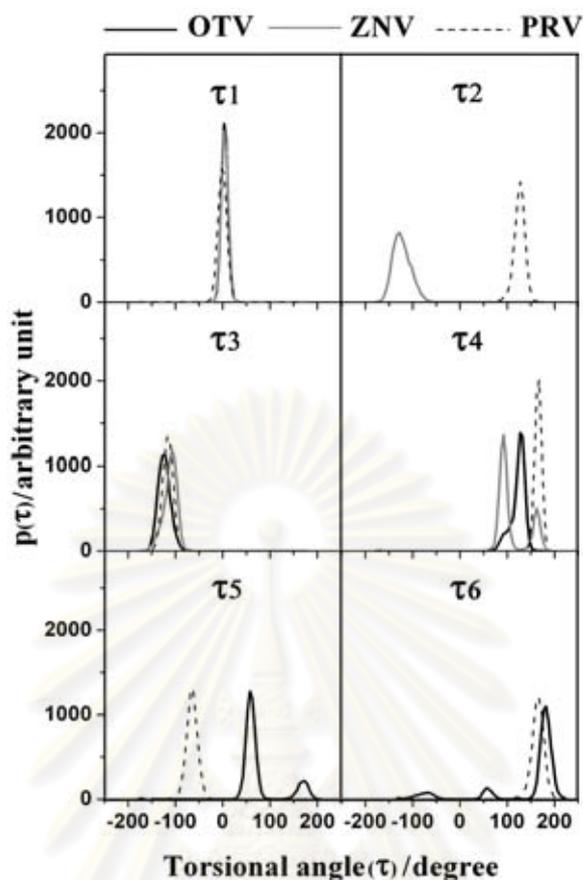
Figure 2.3 Plot of torsional angle distribution sampling over the last 1.5 ns of MD simulations of the OTV-N1, ZNV-N1 and PRV-N1 complexes

Regarding the bulky hydrophobic group of OTV in the OTV-N1 complex, $\tau 4$ was found to exhibit a maximum at 127.5° with a broad peak ranging from 57.5° to 162.5°. The appearance of the shoulder is due to the rotation of this side chain in the N1 binding pocket. There are two separated peaks for the ZNV-N1 system at 92.5° and 162.5°, where the population ratio indicates a preferential conformation, *i.e.*, the two angles represent the orientation of the C_8-C_9 bond lying perpendicular or parallel to the ligand's ring. In contrast, PRV-N1 shows a single preferential orientation represented by $\tau 4$ = 167.5°. The detected sharp and narrow peak implies the rigidity of this side chain.

Regarding $\tau 5$ and $\tau 6$, while side chain D of PRV shows only one preferential orientation, the equivalently ended branches of OTV can rotate rather freely. There is remarkable flexibility of OTV (peaks in $\tau 5$ at 57.5° and 172.5°; $\tau 6$ at -67.5°, 57.5° and 182.5°) in comparison with PRV (peaks in $\tau 5$ at -62.5° and $\tau 6$ at 162.5°).

In summary, conformations of the three inhibitors embedded in the neuraminidase subtype N1 were compared in terms of torsional angles. There were major differences in the structures of the guanidinium ($\tau 2$) and the bulky ($\tau 4$ - $\tau 6$) groups. This is due to the fact that the first side chain lies within the cavity formed by the 150-loop (residues 147-152), which exists only for group-1 NA (N1, N4, N5 and N8). In addition, the conformation of the bulky hydrophobic side chain relates directly to the position and orientation of residue E276 of N1, which was proposed to determine the susceptibility of the ligand in inhibiting the N1 enzyme [40]. Detailed investigations and discussions are reported in the following sections.

2.4.3 Inhibitor-enzyme hydrogen bonding

As mentioned previously, changes in the inhibitor conformation can directly affect the ligand-enzyme interaction. To analyze such effects, the percentage and number of hydrogen bonding between the inhibitor and binding pocket residues were determined based on the following criteria: (*i*) proton donor-acceptor distance ≤ 3.5 Å, and (*ii*) donor-H-acceptor bond angle $\geq 120°$. The results are given in Figure 2.4.

In the three complexes, a high number of strong hydrogen bonds (% occupation ≥ 80) is measured between the –COO⁻ group (side chain A in Figure 2.1), and its nearest residues. Eight hydrogen bonds with R118, R292 and R371 were firmly observed for PRV. These strong hydrogen bonds with the arginine triad were also found in simulations of the N-DANA/NA subtype N9 complex [60]. For OTV and ZNV, only six bonds were formed, so interactions with R118 and R292 were lost. Referring to the existing experimental data, the simulation results for OTV are unlikely to be found in the crystal structure of the OTV-N1 complex, where position of OTV side chain A is surrounded and well fitted by all three arginines [23]. A high number and percentage of hydrogen bond occupation at side chain B was measured for PRV, in comparison with OTV and ZNV. The latter two had comparable hydrogen bond occupations (Figure 2.4). The ammonium group of OTV is stabilized by hydrogen bonding interactions with the two negatively charged residues (E119 and D151) and tryptophan (W178). Referring to Masukawa *et al.* [61], hydrogen bonds with D151 were also found for the N9/OTV and N9/ZNV complexes. In addition, this bond cannot be formed for sialic acid and DANA because of the lack of positive charge on side chain B (Figure 2.1).
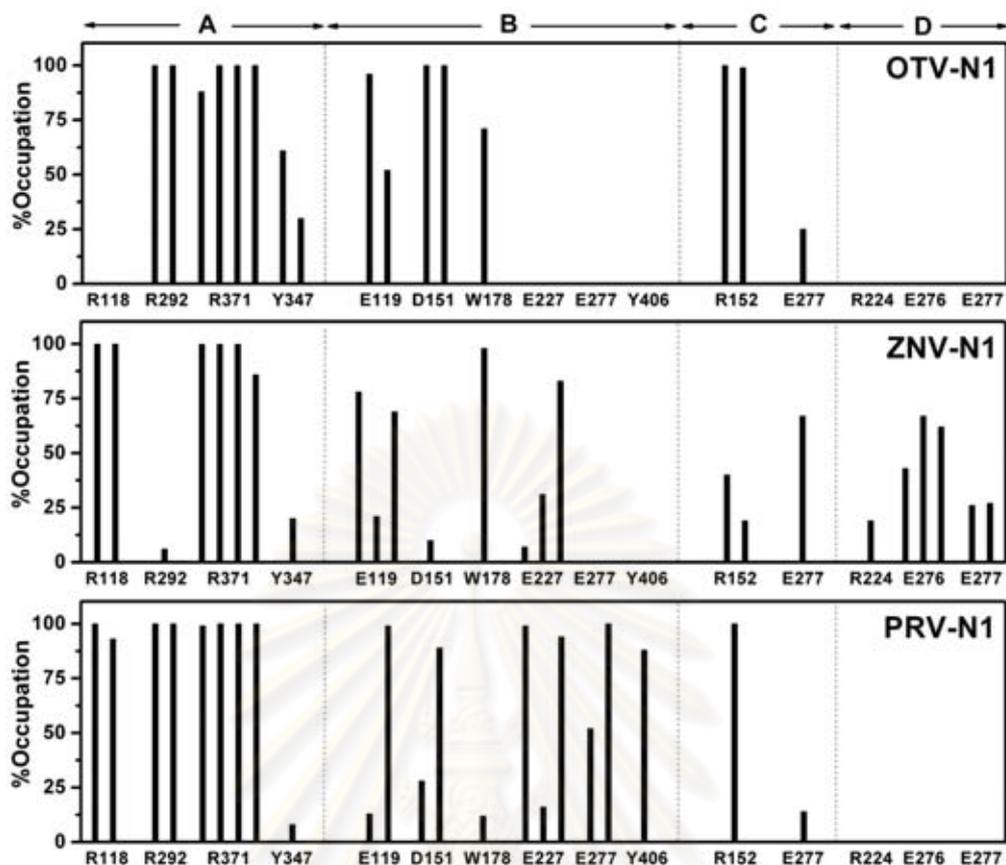
Figure 2.4 Percent occupation pattern of hydrogen bonds between residues in the binding pocket (label given along x-axis) and the four side chains of the inhibitors for the three simulated systems, OTV-N1, ZNV-N1 and PRV-N1.

In consistent with the X-ray structures and our MD simulations of the ZNV and PRV inhibitors in that the orientation of the equivalent side chain (guanidinium group in terms of $\tau 2$ as shown in Figure 2.3) is significantly different. This directly affects its binding with the neighboring residues in the pocket of the NA in which the guanidinium group of PRV is observed to interact much more tightly to enzyme residues. This could be the reason why the guanidinium side chain of PRV rotated only within a narrow range at the position around its preferential configuration (see $\tau 2$ in Figure 2.3). This is different from ZNV, where $\tau 2$ shows a broad distribution. Detailed investigations of the hydrogen bonding between side chain B and the 150-loop residues are reported in the next section.

For side chain C, the inhibitor-enzyme interactions in the three systems are comparable with 2-3 hydrogen bonds with R152 and E227 residues. This is in contrast with side chain D, in which hydrogen bonds were detected only for ZNV, between its

hydrophilic side chain and R224, E276 and E277. However, the percentage of hydrogen bond occupation varies from 25 up to 70. This relates directly to the flexibility of this chain (shown by the distribution of the $\tau4$ angle in Figure 2.3) as discussed previously. The obtained results are similar to those found from the X-ray structure of group-2 NA complexed with ZNV [62] and sialic acid [22], in which hydrogen bonds were observed between the hydrophilic side chain and the carboxylate group of E276. In contrast, there are no hydrogen bonds formed in the lipophilic pocket (side chain D in Figure 2.1) of the OTV and PRV inhibitors. Consequently, the negatively charged side chain of E276 was shifted and rotated to form hydrogen bonds with the positively charged moiety of R224 to form a pocket for the bulky hydrophobic side chain. This is consistent with previous reports [40, 63].

2.4.4 Changes in the active site adjacent to the 150-loop

To provide detailed information on the cavity site near the 150-loop of group-1 neuraminidase, including subtypes N1, N4, N5 and N8 [23], electrostatic potentials of N1 complexed with the three inhibitors (using the last MD snapshot) and N2 complexed with sialic acid (using the X-ray structure, PDB code: 2BAT [22]) were calculated and plotted in Figure 2.5. Hydrogen bonds and the corresponding distances to the 150-loop residues D151 and R152 are also given.

In the results, the characteristics of the hydrogen bonding between the 150-loop and the four structures, three inhibitors and one substrate, are totally different. The 150-loop was observed to bind strongly via D151 and R152 to the two side chains of OTV (Figure 2.5(a)). For ZNV, only one hydrogen bond with 10% occupation was detected between the 150-loop and the side chain of ZNV. Instead, the two side chains were tightly solvated (details given in the previous section). Therefore, the ligand-enzyme interaction for the ZNV-N1 complex was then bridged by water molecules (Figure 2.5(a)). The situation is different for the PRV inhibitor, which was newly designed to better fit the cavity of NA. Although the –OH group of PRV was located far from the 150-loop, in comparison with sialic acid, the hydrogen bonding between them can be firmly formed. This is very similar to the case of sialic acid, in that the D151 and R152 residues for both systems were observed to bind directly to the –OH group and the acetamido (-C=O) group located at side chain C, respectively. Therefore, the 150-loop of the PRV-N1 complex was shifted to move

closer to the ligand, in comparison with the other systems. As can be seen from the 3D plot shown in Figure 2.5, the 150-loop of the PRV-N1 complex (Figure 2.5(c)) covers the ligand plane up to almost half of the 5-membered ring. This is not the case for the other two simulated systems.
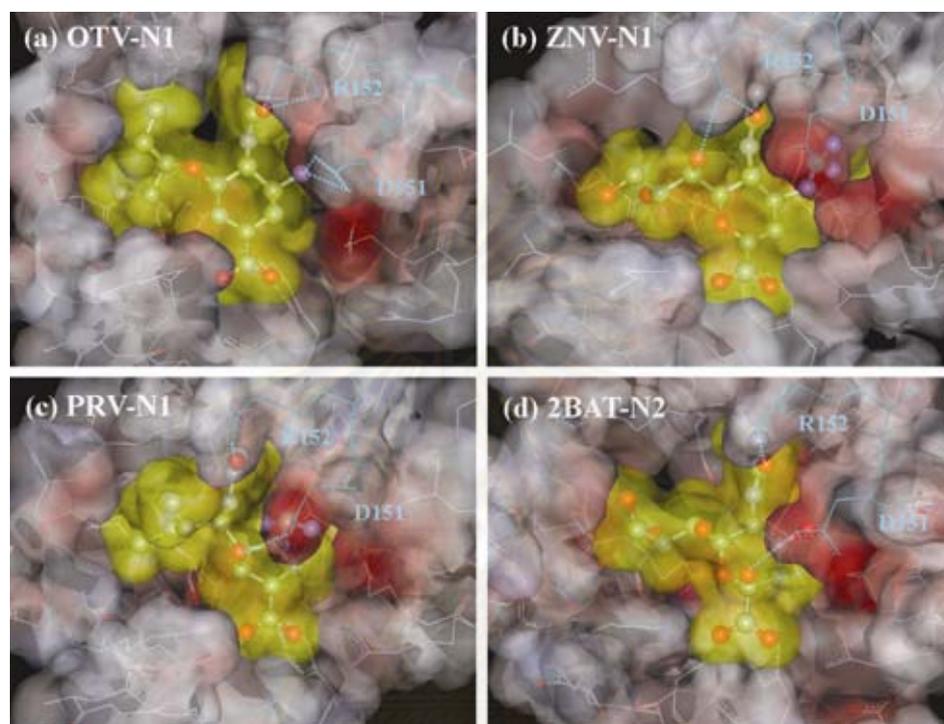
Figure 2.5 Electrostatic potential of the neuraminidase N1 complexed with the three inhibitors and its substrate (sialic acid) including the van der Waals cavity of the inhibitors and substrate (in yellow) and hydrogen bonds to the 150-loop residues, D151 and R152 (negative regions are in red and positive regions are in blue).

2.4.5 Enzyme-inhibitor interactions

The OTV-N1, ZNV-N1 and PRV-N1 interactions were analyzed in terms of binding free energies, computed using MM-PBSA calculations in AMBER. The results are shown in Table 2.2. The gas phase MM interaction energy (ΔE_{MM}) of -282.2 kcal mol^{-1} as well as the binding free energy ($\Delta G_{binding}$) of -227.4 ± 7.6 kcal mol^{-1} for PRV-N1 indicates that PRV fits tightly in the enzyme cavity, especially in comparison with the other two inhibitors. Among the three complexes, the calculated order of the total binding free energy of PRV < ZNV < OTV agrees completely with that of the experimental IC_{50}. Note that the experimental IC_{50} as shown in Table 2.2 depends strongly on the method used as well as the subtype of influenza virus.

Table 2.2 Calculated binding free energy and its components (kcal mol^{-1}) as well as the experimental IC_{50} (in nM) of the three neuraminidase complexes, OTV-N1, ZNV-N1 and PRV-N1.

	OTV-N1	ZNV-N1	PRV-N1
$\varDelta E_{ele}$	-191.9 ± 10.5	-223.2 ± 15.0	-254.9 ± 11.1
$\varDelta E_{vdw}$	-28.8 ± 3.6	-27.0 ± 4.8	-27.4 ± 4.5
$\boldsymbol{\varDelta E_{MM}}$	**-220.6 ± 9.7**	**-250.2 ± 14.6**	**-282.2 ± 9.3**
$\varDelta G_{sol}^{nonpolar}$	-3.5 ± 0.1	-3.5 ± 0.3	-3.9 ± 0.2
$\varDelta G_{sol}^{ele}$	48.9 ± 2.0	56.6 ± 3.1	58.8 ± 2.0
$\varDelta G_{sol}$	45.4 ± 2.0	53.1 ± 3.0	54.9 ± 2.0
$\varDelta G_{sol}^{ele} + \varDelta E_{ele}$	-143.0 ± 8.8	-166.6 ±12.4	-196.1 ± 9.6
$\boldsymbol{\varDelta G_{binding}}$	**-175.2 ± 8.0**	**-197.1 ± 11.9**	**-227.4 ± 7.6**
IC_{50} H1N1 [65]	53.20 ± 10.30	5.80 ± 2.10	3.40 ± 1.00
IC_{50} H1N1 [25]	17.80	4.53	1.11
IC_{50} H1N1 [26]	0.45	0.95	0.34
IC_{50} H1N1 [66]	0.90	1.14	0.27
IC_{50} H1N1 [66]	1.53	0.65	0.41
IC_{50} H1N1 [67]	0.559 ± 0.150	0.126 ± 0.018	0.045 ± 0.009
IC_{50} H1N1 [68]	1.40	0.90	0.40
IC_{50} H1N1 [68]	2.00	0.80	0.30
IC_{50} H1N1 [69]	1.10 ± 0.20	1.35 ± 0.18	0.11 ± 0.01
IC_{50} H5N1 [70]	0.33 ± 0.27	0.57 ± 0.46	0.37 ± 0.26
IC_{50} H6N1 [65]	36.10 ± 8.30	19.70 ± 4.20	2.60 ± 0.40

Some comments could be made concerning the neglect of an entropic term, $T\varDelta S$, in the binding free energy calculation. This assumption is based on the fact that the three systems are very similar, and our interpretations as well as discussion are

based on only relative comparisons among the three complexes. Regarding the HIV-1 protease complexed with saquinavir, ritonavir, nelfinavir and indinavir, although sizes of the 4 drugs are different, the calculated $T\Delta S$ contributions of -14.0, -11.2, -14.2 and -15.7 kcal mol^{-1}, respectively [64], are negligibly different.

2.4.6 Solvation structure

Hydrophobicity and hydrophilicity, especially in the catalytic pocket, are known to play an essential role in the catalytic mechanism of the enzyme. To seek for such detailed information, the radial distribution functions (RDF) $-g_{ij}(r)$: the probability of finding a particle of type i in a spherical radius, r, around the particle of type $j-$ centered on the selected atoms (see Figure 2.1 for atomic labels) of the three inhibitors were calculated. C_{13}, which is not a H-bond donor or acceptor, was included into consideration because it is located on the flexible side chain where its interactions with solvent molecules and surrounding residues have to be intensively investigated. Evaluation was classified into two categories, in which the central atom types for the OTV, ZNV and PRV ligands are topologically equivalent and different. The results are given in Figure 2.6.

By considering the solvation of the –COO$^-$ group (side chain A in Figure 2.1) of the three complexes, the OTV and ZNV ligands were observed to be solvated by the same number of water molecules, i.e., the sharp first peak centered on O_1 and O_2 with the maximum at ~3.0 Å indicates strong hydrogen bonding between water molecules and ligands, while the integration number up to the first minimum of 2 denotes the number of water molecules (coordination number, CN) that form hydrogen bonds with each central atom (Figure 2.6(a)). Interestingly, the two symmetric oxygen atoms of the –COO$^-$ groups for the OTV and ZNV inhibitors are unsymmetrically solvated with a different manner, i.e., only O_1 of ZNV or O_2 of OTV was solvated by 2 water molecules. This can be understood in terms of the number and/or orientation of the enzyme residues around the –COO$^-$ side chains of the two inhibitors. The situation is different for the solvation of this side chain of the PRV, where the small sharp first peak centered at ~3.0 Å indicates the hydrogen bond formation between O_2 and a water molecule with the percentage occupation of 50%, $CN = 0.5$.

For side chain B, the RDFs centered on N_2, N_3 and N_4 atoms were evaluated. Clear solvation, shown by the first peak at a distance of ~3.0 Å, was detected only for N_2 and N_3 of the ZNV inhibitor. The other RDFs, which show the first peak appearing at a distance longer than 4 Å, reveal that direct binding with solvent molecule was not detected, *i.e.*, their central atoms insert and fit well to the enzyme subsite. This is especially true for the PRV inhibitor, in which water molecules start to be detected at the distances of 4.6 Å, 6.6 Å and 5.3 Å, respectively.

Solvation of N_1 and N_2 are for the three inhibitors is quite similar in that only N_1 of ZNV can be accessed by 0.5 water molecules. The N_1---O distance of 4.7 Å, the maximum of the N_1 RDF of the OTV inhibitor, is too far for the two water molecules lying under this peak to be considered as its solvation shell. In addition, O_3 atoms of both OTV and ZNV were solvated by 0.5 and 1.0 water molecules, respectively.

RDFs in Figure 2.6(b) shown for the atoms of the inhibitors that are not comparable to each other. In terms of *CN*, O_4 of ZNV (*CN* = 3) was better solvated than O_4 of PRV and OTV (*CN* = 2 and 1, respectively). However, water can approach closer (indicated by the distance of maximum RDF) to O_4 of PRV and OTV than to that of ZNV. Then O_5 to O_7 of ZNV can be accessed by 3, 1 and 0.5 water molecules, respectively.

Taking into account all the data and the discussion given above, the PRV atoms are accessed much less by water molecules than the other two inhibitors, considered in terms of both coordination number and distance to solvent molecules. This implies that there is less space available between the PRV atoms and the enzyme residues. This conclusion was strongly supported by the hydrogen bond formation (Figure 2.4) and ligand-enzyme interaction data (ΔE_{MM} in Table 2.2), *i.e.*, among the three inhibitors, PRV shows the highest interaction with the enzyme (-227.4 ± 7.6 kcal mol^{-1}) in which the main contribution was the number as well as percentage occupation of hydrogen bonds shown in Figure 2.4. As PRV displays lowest experimental IC_{50}, the simulation evidence suggests a clear conclusion that activity of the inhibitor is directly determined by a direct interaction between the inhibitor and enzyme.
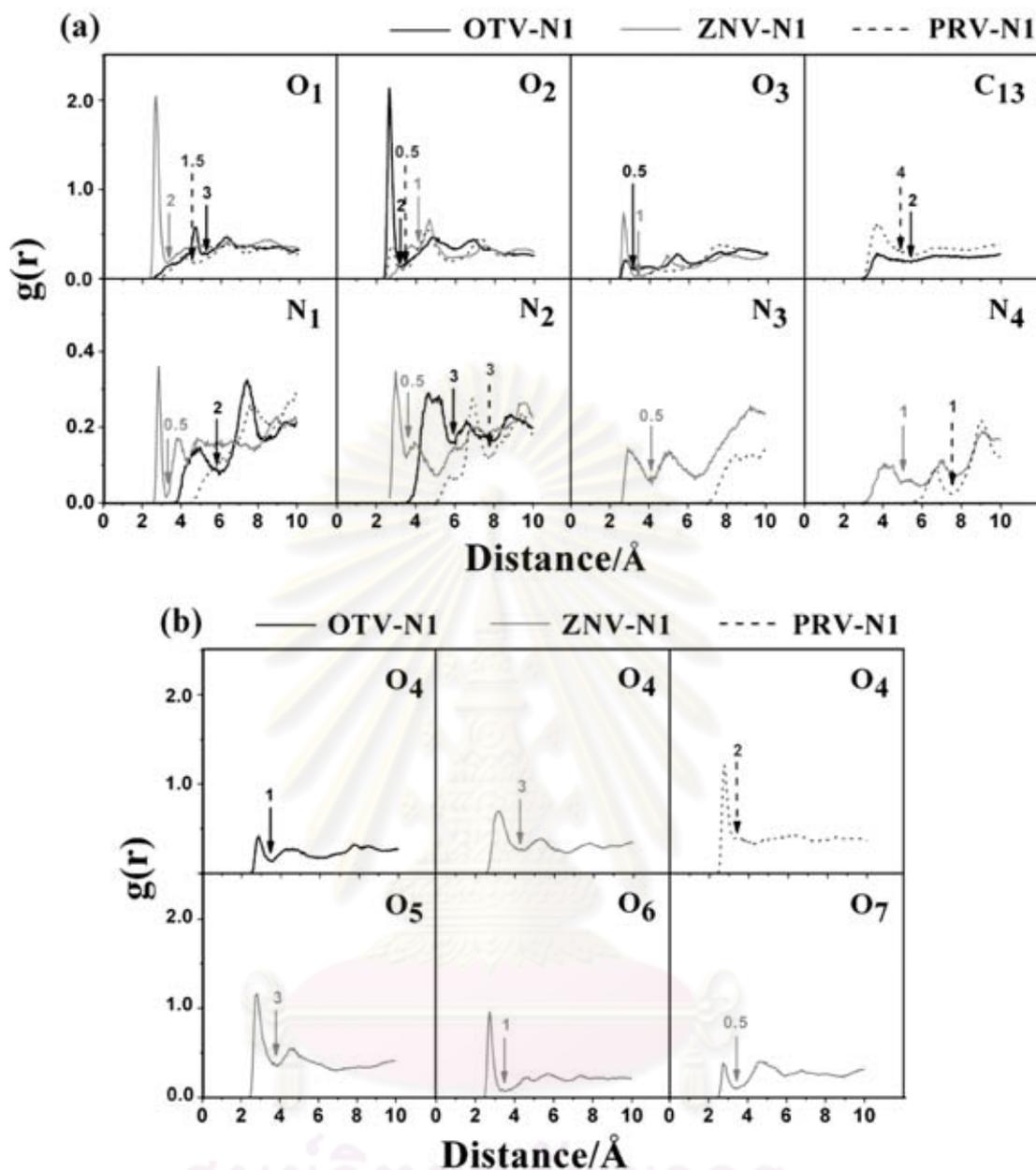
Figure 2.6 Radial distribution functions, *g(r)*, centered on the inhibitor atoms (see Figure 2.1 for definitions) to oxygen atoms of modelled water of the three complexes, OTV-N1, ZNV-N1 and PRV-N1, including the running integration number up to the first minimum (marked by arrows). Here, N_3, N_4, O_5, O_6, O_7 and C_{13} are not available for all inhibitors. The plots are classified into two categories where the central atom types for the three ligands are topologically (a) equivalent and (b) different.

2.5 CONCLUSIONS

The technique of molecular dynamics simulations was applied to seek detailed information on the dynamic nature of the three inhibitors, OTV, ZNV and PRV, embedded in N1 neuraminidase. In conjunction with the Molecular Mechanics—Poisson-Boltzmann/Surface Area (MM-PBSA), binding free energy of the complexes was calculated and found, on a relative scale, to agree very well with the experimental IC_{50} value.

The structural properties, position and conformation of PRV and its side chains are uniformly preferential, *i.e.*, its conformation fits very well with the N1 active site. This leads consequently to the following conclusions about PRV, in comparison with the OTV and ZNV inhibitors: *(i)* strong direct ligand-enzyme hydrogen bonding, both in terms of number and %occupation, with a very slight effect on the surrounding residues; *(ii)* less space available in the N1 pocket, and hence, less access by water molecules; and *(iii)* PRV interacts tightly, via its –OH group, with the D151 residue located in the 150-loop region. In good agreement with the structural nature of PRV, which was designed to fit better, compared with the commercially available inhibitors, to the NA pocket, the PRV-N1 complex shows the lowest binding free energy. In addition, a high number of strong hydrogen bonds (% occupation ≥ 80) is detected between the –COO⁻ group of the three inhibitors and the nearest residues, indicating that it is the essential side chain for the NA inhibitors. In contrast, flexibility and rotation of the bulky side chain of the three ligands can be a primary source of lower susceptibility of the NA inhibitors.

CHAPTER III

NA-MUTATION (H274Y)

Source of oseltamivir resistance in avian influenza H5N1 virus with the H274Y mutation

Maturos Malaisree[1,a], Thanyada Rungrotmongkol[1,a], Nadtanet Nunthaboot[2], Ornjira Aruksakunwong[3], Pathumwadee Intharathep[1], Panita Decha[1], Pornthep Sompornpisut[1] and Supot Hannongbua[1]

[1]*Department of Chemistry, Faculty of Science, Chulalongkorn University, Patumwan, Bangkok 10330, Thailand.*

[2]*Department of Chemistry, Faculty of Science, Mahasarakham University, Mahasarakham, 44150, Thailand.*

[3]*Department of Chemistry, Faculty of Science, Rangsit University, Pathumtani, 12000 Thailand.*

[a]These authors contributed equally to this work.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.1 ABSTRACT

To provide detailed information on the oseltamivir-resistance resulting from the H274Y mutation in neuraminidase (NA) of avian influenza H5N1 viruses, molecular dynamics simulations were carried out for the mutant oseltamivir-NA complex. In contrast with a previous proposal, the H274Y mutation does not prevent E276 and R224 from forming the hydrophobic pocket for the oseltamivir bulky group. Instead, reduction of the hydrophobicity and size of pocket in the area around an ethyl moiety at this bulky group were found to be the source of the oseltamivir-resistance. These changes were primarily due to the dramatic rotation of the hydrophilic $-COO^-$ group of E276 toward the ethyl moiety. In addition, hydrogen bonding interactions with N1 residues at the $-NH_3^+$ and $-NHAc$ groups of oseltamivir were replaced by a water molecule. The calculated binding affinity of oseltamivir to NA was significantly reduced from -14.6 kcal mol^{-1} in the wild-type to -9.9 kcal mol^{-1} in the mutant type.

3.2 INTRODUCTION

The outbreak of avian influenza A (H5N1) viruses has raised global concern for animal and human health. Recently, isolates of the virus with amino acid changes in neuraminidase (NA) that likely confer the reduction in susceptibility to oseltamivir have been reported [2, 3, 33]. In order to have a better chance of overcoming these and predicting future new resistance problems, an understanding of the mechanism of oseltamivir resistance at the molecular level is required.

The influenza viral NA is a receptor-destroying enzyme that cleaves a terminal sialic acid and releases viral progeny from infected cells, and has been separated into two groups by genetic and structural relationships: group-1 NA includes N1, N4, N5 and N8 whereas group-2 NA includes N2, N3, N6, N7 and N9. Based on the crystal structures, group-1 NA differs from group-2 NA by having a large cavity close to the active site on the 150-loop, at residues 147-152 [23]. However, the active site in all subtypes is conserved. The catalytic sites (R118, D151, D152, R224, E276, R292, R371 and Y406) were found to interact with the sialic acid while the framework sites (E119, R156, W178, S179, D198, I222, E227, H274, E277, N294 and E425) (in N2 numbering) were supposed to stabilize the active site structure [24].

NA has been targeted for improved anti-influenza drugs. The approved NA inhibitors were designed on a basis of the transition state analog of sialic acid complexed with group-2 NA [17]. Two marketable NA inhibitors, zanamivir and oseltamivir, are available for treatment of influenza virus infections [41]. To date, high resistance and mutations against oseltamivir have been widely reported, not only in vitro and in vivo but also in clinical treatment. Mutations in group-1 NA (N1) were detected at H274Y and N294S [33-35, 38-40] and in group-2 NA (N2 and N9) at E119V, I222V and R292K [36, 37]. In H5N1 subtypes of the virus, the H274Y mutation in the NA active site leads to a high-level resistance to oseltamivir, typically with a 300-1,700 fold reduction in susceptibility [33-35, 38-40]. The proposed mechanism of oseltamivir resistance caused by the H274Y mutation is that the hydrophobic pocket around the bulky oseltamivir group, which is generally created by the rotation of the E276 side chain to bind with the R224 side chain, cannot be formed [1, 37, 40]. However, detailed information on the results of such proposed changes has not been identified in molecular level.

To provide detailed information on the primary source of oseltamivir-resistance due to the H274Y mutation, molecular dynamics (MD) simulations have been carried out for the mutant type of NA complexed with oseltamivir. The results were analyzed in terms of conformational changes, intra- and intermolecular hydrogen bonds, ligand solvation and binding free energy of the complex, focused on the catalytic pocket of NA. These structural and molecular properties were compared and discussed extensively with our previous study on the wild-type N1 complexed with oseltamivir [6].

3.3 MATERIALS AND METHODS

3.3.1 Molecular dynamics simulations

The crystal structure of the NA subtype N1 complexed with oseltamivir (2HU4.pdb) [23] was used as the initial structure. To prepare the H274Y mutant for inhibitor binding (OTV-MT), the N1 residue 274 was changed from histidine to tyrosine using the LEaP module of the AMBER software package [47]. All missing hydrogen atoms of the protein and inhibitor were added with standard bond lengths and angles. The ionization states of amino acids with electrically charged side chains

were assigned using the PROPKA program [71]. The system was solvated with a TIP3P water box [51] and neutralized by the counterions. The total number of atoms in the periodic box of 78 Å x 80 Å x 81 Å was 41,302.

All calculations were performed using AMBER software package [47]. The Cornell force field [52] was applied for the amino acids while the partial atomic charges and force-field parameter of the inhibitor were taken from our previous calculations in wild-type system. A time step of 2 fs was used with a cut-off radius of 12 Å for the non-bonded interactions. The particle mesh Ewald method was used for calculating the long-range electrostatic interactions [55]. The periodic boundary MD simulations were performed with the NPT ensemble. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and pressure of the system [53] The SHAKE algorithm [54] was employed to constrain all bonds involving hydrogen.

To remove bad contacts, locations of hydrogen atoms and the added water molecules were initially optimized, and then energy minimization of the entire system was carried out. Afterwards, the MD simulations were performed over three periods: a thermalization phase from 0 K to 298 K over 60 ps, an equilibration phase at 300 K for 1.5 ns, and the final production period for 1.5 ns. Only the snapshots taken from the production phase were used for analysis. The convergence of energies, temperature, pressure, and RMSD were used for verification of the system stability.

3.3.2 Binding free energy calculations

To estimate the binding affinity between protein (NA) and ligand (oseltamivir), the free energy was calculated based on the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) methodology [72-74], and implemented in AMBER 7. One hundred MD snapshots were extracted (at every 75 ps) from the production phase and were used as a structural ensemble to evaluate the MM/PBSA binding free energy. After all the water molecules and counterions were striped, the binding free energy of the protein-ligand complex (ΔG_{bind}) was computed by:

$$\Delta G_{bind} = \Delta G_{complex} - \left[\Delta G_{protein} + \Delta G_{ligand} \right]$$

(1)

where $\Delta G_{complex}$, $\Delta G_{protein}$ and ΔG_{ligand} refer to the absolute free energy of the complex, protein and ligand, respectively. Typically, the total free energy of a given conformational state i contains the enthalpy and entropy contributions and is expressed by:

$$\Delta G_i = \Delta H_i - T\Delta S_i \tag{2}$$

where ΔH of a system is composed of the enthalpy changes in the gas phase upon complex formation (ΔE_{MM}) and the solvated free energy contribution (ΔG_{sol}) and $-T\Delta S$ refers to the entropy contribution to the binding. Therefore, Eq. (2) can be approximated as:

$$\Delta G_i = \Delta E_{(MM)i} + \Delta G_{(sol)i} - T\Delta S_i \tag{3}$$

$\Delta E_{(MM)i}$ is further divided into van der Waals (ΔE_{vdW}) and electrostatic (ΔE_{ele}) interaction energies between the ligand and its receptor (Eq. 4). These interaction energies were calculated using the SANDER module in AMBER 7.

$$\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele} \tag{4}$$

$\Delta E_{(sol)i}$ is composed of electrostatic (ΔG_{sol}^{ele}) and non-polar ($\Delta G_{sol}^{nonpolar}$) contributions (Eq. 5).

$$\Delta G_{sol} = \Delta G_{sol}^{ele} + \Delta G_{sol}^{nonpolar} \tag{5}$$

The Delphi4 program suite [57] was used to evaluate the electrostatic free energy of solvation. A grid spacing of 0.33 Å with the boundary condition of Debye-Hückel potentials was applied. Atomic charges were taken from the Cornell force field [52]. The dielectric values applied for water and protein were set to 80 and 1, respectively. The non-polar contribution to the solvation free energy is calculated according to the linear function of the solvent accessible surface area (SASA) (Eq. 6).

$$\Delta G_{sol}^{nonpolar} = \gamma SASA + b \tag{6}$$

where $\gamma = 0.00542$ kcal mol^{-1}Å^2, and $b = 0.92$ kcal mol^{-1}. In this study, the same set of coordinates for complex, protein, and ligand were used, and the entropy contribution was ignored as reported and justified in previous works [75-78].

3.4 RESULTS AND DISCUSSION

To evaluate the stability of the simulated systems, RMSDs of OTV-WT and OTV-MT complexes obtained during the 3 ns MD simulations relative to the initial structure were calculated and plotted in Figure 3.1. As can be seen from the RMSD plots, the two systems were found to reach equilibrium at 1.5 ns and, therefore, the MD trajectories extracted from the 1.5 - 3.0 ns simulations were used for the following analysis.
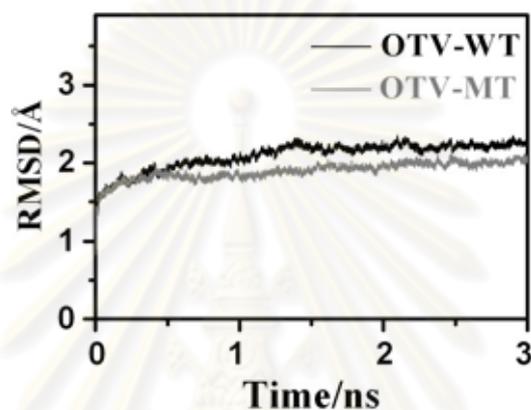
Figure 3.1 RMSD plots of all atoms of OTV-WT and OTV-MT systems along the simulation time.

3.4.1 Conformational changes in the neighbourhood of H274Y

In order to determine the effect of the H274Y mutation on the interactions between oseltamivir and the N1 active site, changes of the molecular conformations within the neighborhood of H274Y were monitored. Interpretations were focused on oseltamivir's 1-ethylproxy (–OCHEt$_2$), and the Y274, E276 and R224 side chains. The results are shown in Figure 3.2(b) in terms of the torsional angles, *Tor-H274Y, Tor-E276, Tor-R224, Tor-X, Tor-Y and Tor-Z*, defined by sets of four atoms (arrows in Figure 3.2(a)). Here, the corresponding data from our previous oseltamivir-wild type (OTV-WT) simulations [4] are also given for comparison.

Considering the torsional angle of the side chain of residue 274 before and after mutation (Figure 3.2(b)), *Tor-H274Y* of the histidine's immidazole ring in the OTV-WT complex exhibits a sharp peak at 107.5° (black line), which was altered to 62.5° for the *Tor-H274Y* of the tyrosine's phenyl ring in OTV-MT (grey line). This

rotation and an apparent size increase of the side chain in changing from H274 to Y274 is supposed to lead to a rearrangement of the neighboring residues, R224 and E276 (Figure 3.2(a)). Note that the formation of the hydrophobic pocket for the oseltamivir's bulky $-OCHEt_2$ side chain in the NA is determined by these two residues [1, 37, 40].

The meaningful changes after the H274Y mutation were found to control the new conformation of the E276's carboxylate group, where *Tor-E276* was noticeably shifted by 115° from -47.5° to 67.5° (Figure 3.2(b)). This is due to the bulky side chain of Y274 where the molecular refractivity (parameter of bulkiness) of 31.83 for tyrosine is significantly larger than that of 23.79 for histidine [79]. Changes of the E276 conformations were found to strongly influence the rotation and flexibility of the $-OCHEt_2$ group of oseltamivir.

This fact is indicated by the appearance of a single sharp and narrow peak in *Tor-X* of the $-OCHEt_2$ side chain at the X-terminal in the OTV-MT complex at 57.5° (grey line in Figure 3.2(b)). This is different from what is found for OTV-WT. The flexibility and rotation of this terminal is shown by the three peaks at -77.5°, 57.5° and 182.5° (black line Figure 3.2(b)). A dramatic change in the preferential conformation of the X-terminal by 125° (from 182.5° for OTV-WT to 57.5° for OTV-MT) is clearly due to the rotation of *Tor-E276* (by 115° as described above). This means that the rotation brings the hydrophilic $-COO^-$ group of the E276 side chain to approach the X-terminal (see Figure 3.2(a)) of the $-OCHEt_2$ group of oseltamivir and leads consequently to the reduction of the hydrophobic pocket of the N1 mutant (see Figure 3.3). In other words, rotation of the $-CH_2CH_3$ moiety at the $-OCHEt_2$ group of oseltamivir was found to be induced by the reduction of the hydrophobicity of the catalytic pocket of N1.
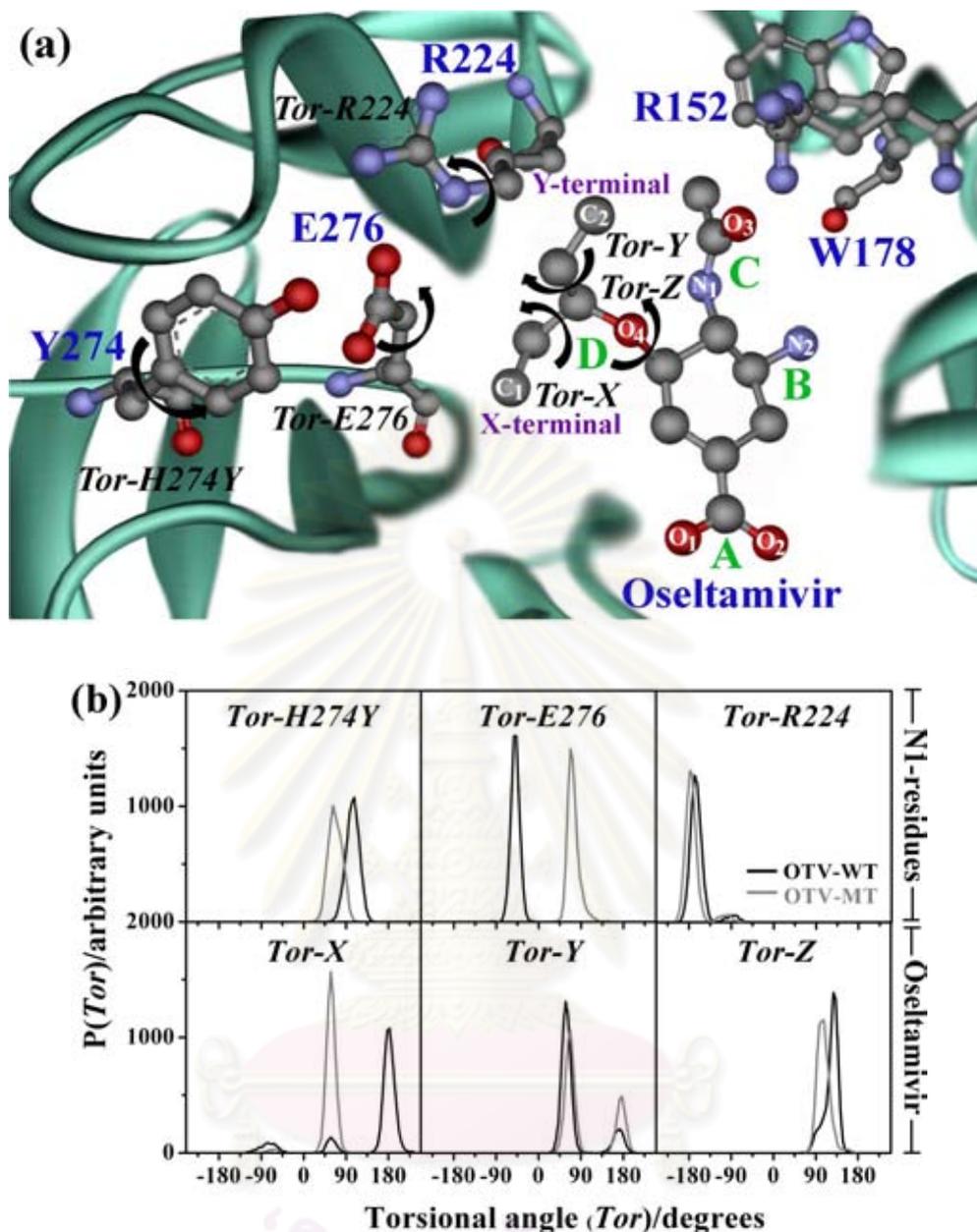
Figure 3.2 (a) Model structure of the oseltamivir bound to the N1-H274Y mutation where R224, Y274 and E276 were shown in the ball and stick model, while locations of R152 and W178 were also given. The torsional angles of the side chains of residues H274Y, E276 and R224 were defined by *Tor-H274Y*, *Tor-E276* and *Tor-R224*, respectively, while those of the hydrophobic side chain of oseltamivir were described by *Tor-X*, *Tor-Y* and *Tor-Z*. (b) Plot of torsional angle distribution (black for wild-type and grey for mutant type).
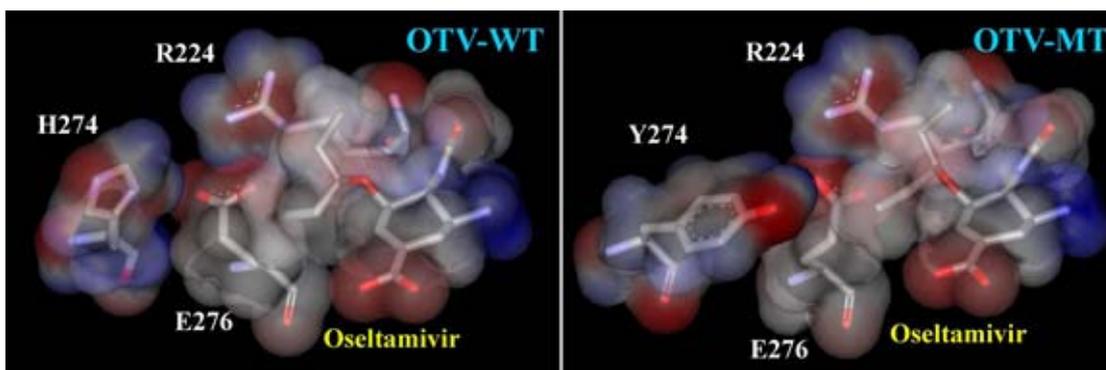
Figure 3.3 Van der Waals surface of the wild-type (OTV-WT) and mutant N1 (OTV-MT) complexed with oseltamivir including the electrostatic potential of the inhibitor, mutated residue 274 and the two residues that principally contribute to the hydrophobic pocket (R224 and E276). The negative and positive regions are shown in red and blue, respectively.

To clarify this point, structural alignment of the MD snapshots of both OTV-WT and OTV-MT complexes sampling every 0.5 ns after equilibration, was plotted in Figure 3.4(c). It can be seen that the E276 side chain was noticeably shifted toward the bulky moiety of oseltamivir while displacement of oseltamivir was also slightly exhibited. Consequently, this fact was found to induce some changes of the positions of two binding residues, R152 and W178, leading to the loss of their hydrogen bonds with oseltamivir (details in the next section).

Interestingly, the remarkable changes mentioned above do not cause significant changes to the –OCHEt$_2$ moiety at the Y-terminal and its contacted residue R224, suggesting that the conformation of this region is considerably conserved. In the results; *Tor-Y* and *Tor-R224,* which determine the rotations of the Y-terminal and the R224 side chain, respectively, for both OTV-WT and OTV-MT complexes were found at almost the same angles (see Figure 3.2(a) and 3.2(b)). In addition, a slight shift of the maximum of *Tor-Z* by 20° from 127.5° (for OTV-WT) to 107.5° (for OTV-MT) indicates a strong influence of the H274Y mutation, which can establish a rotation around the principal axis of the –OCHEt$_2$ group of oseltamivir.
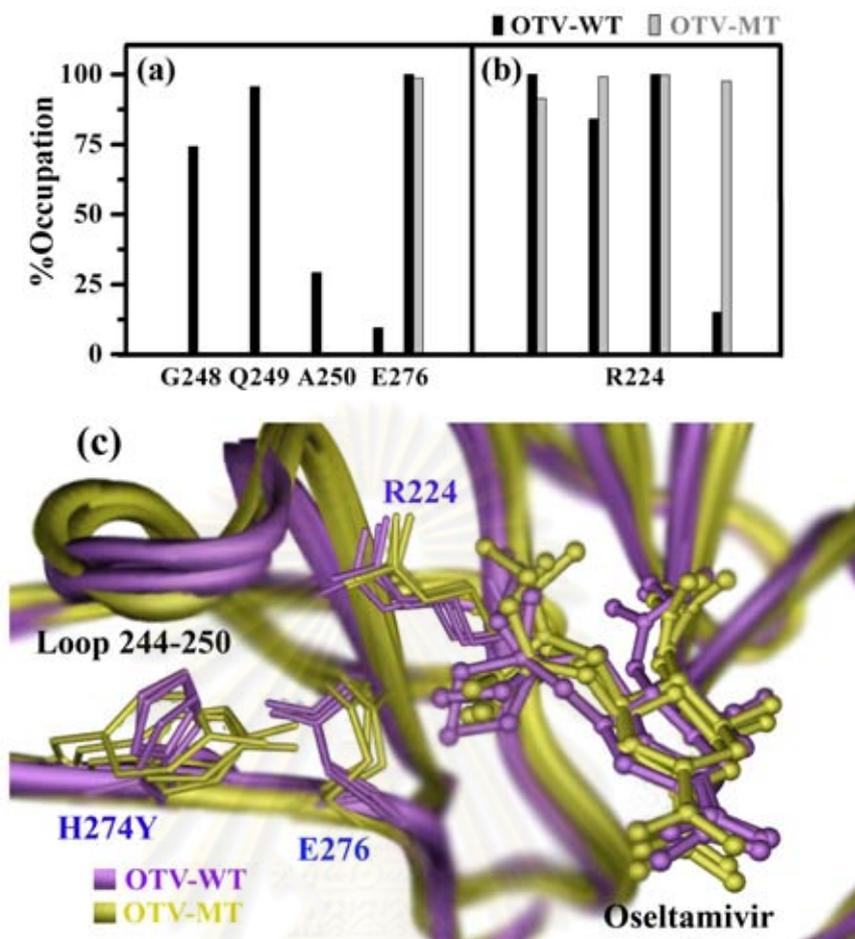
Figure 3.4 Percentage occupation of hydrogen bonding of (a) the mutated residue 274 (H274Y) to its surrounding residues, and (b) between the carboxylate group of E276 and the guanidinium group of R224 (c) Structural alignment between the three MD snapshots of OTV-WT and OTV-MT sampling every 0.5 ns after equilibration phase. Close-up of oseltamivir, the two important residues R224 and E276, and the loop 244-250 located in the neighborhood of the mutated residue 274 are all shown.

3.4.2 Loss of intramolecular hydrogen bonds of the enzyme

To understand the effects of the H274Y mutation on the intramolecular interactions, percentage and number of hydrogen bonds between H274Y and its neighboring residues and between the two charged residues (E276 and R224) were determined, based on the following criteria: *(i)* the distance between proton donor (D) and acceptor (A) atoms was less than or equal to 3.5 Å and *(ii)* the D-H..A angle was greater than or equal to 120°. The results were evaluated and given in Figure 3.4 whereas details of hydrogen bond type were shown in Table 3.1.

It is shown from the plots that the immidazole ring of H274 strongly interacts with the four residues G248, Q249, A250 and E276 through a hydrogen bonding framework (black columns in Figure 3.4(a)) in which this ring acted as the centre of interactions. This is in contrast with what is found for OTV-MT, in which only one hydrogen bond with E276 was retained. Rotation of the *Tor-H274Y* by 62.5° in changing from the histidine's immidazole ring (OTV-WT: black line in Figure 3.2(b)) to the tyrosine's phenyl ring (OTV-MT: grey line in Figure 3.2(b)), as discussed above, was found to be the source for the loss of these hydrogen bonds. In addition, the results show that the loop containing residues 244-250 is considerably conserved (see Figure 3.4(c)). Loss of the hydrogen bond framework in the region close to the bulky group of oseltamivir is likely to reduce the stability of the hydrophobic pocket and subsequently to lower the capacity of the oseltamivir binding to the N1 active site.

Surprisingly, four hydrogen bonds formed between E276 and R224 were observed for both OTV-WT and OTV-MT complexes (Figure 3.4(b)), suggesting that the H274Y mutation does not reduce the attractive interaction between these two residues. This finding contradicts the previous hypothesis [40] which stated that the H274Y mutation inhibits the formation of the hydrophobic pocket around the bulky side chain of oseltamivir*, i.e.,* hydrogen bonding between E276 and R224 cannot be formed in the H274Y complex. From our results, the H274Y mutation does not prevent the E276 and R224 from forming a pocket. As discussed above, reduction of the hydrophobicity of the catalytic pocket of N1 in the area around the X-terminal of the $-OCHEt_2$ moiety of oseltamivir is the primary source of the oseltamivir resistance.

Table 3.1 Description of direct hydrogen bonding interactions between oseltamivir (OTV) and particular residues of the N1 enzyme

Residues	Type	OTV-WT	OTV-MT
H274Y---G248	H274_NE_2---H_N_G248	74.2	-
H274Y---Q249	H274_NE_2---H_N_Q249	95.6	-
H274Y---A250	H274_NE_2---H_N_A250	29.2	-
H274Y---E276	H274_ND_1_HD_1---OE_1_E276 Y274_OH_HH---OE_1_E276	9.5	-
H274Y---E276	H274_ND_1_HD_1---OE_2_E276 Y274_OH_HH---OE_2_E276	100.0	98.7
E276---R224	E276_OE_1---H_{11}_NH_1_R224	100.0	91.4
E276---R224	E276_OE_2---H_{11}_NH_1_R224	84.1	99.2
E276---R224	E276_OE_1---HE_NE_R224	100.0	99.8
E276---R224	E276_OE_2---HE_NE_R224	15.0	97.6
OTV---R292	OTV_O_1---HH_{12}_NH_1_R292	100.0	100.0
OTV---R292	OTV_O_1---HH_{22}_NH_2_R292	100.0	99.8
OTV---R371	OTV_O_1---H_{22}_NH_2_R371	87.8	88.9
OTV---R371	OTV_O_2---H_{22}_NH_2_R371	100.0	100.0
OTV---R371	OTV_O_1---H_{12}_NH_1_R371	100.0	100.0
OTV---R371	OTV_O_2---H_{12}_NH_1_R371	99.9	99.9
OTV---Y347	OTV_O_1---HH_OH_Y347	61.2	39.1
OTV---Y347	OTV_O_2---HH_OH_Y347	30.3	12.5
OTV---E119	OTV_N_2_$H_{21,22,23}$---OE_1_E119	98.0	100.0
OTV---E119	OTV_N_2_$H_{21,22,23}$---OE_2_E119	52.3	62.3
OTV---D151	OTV_N_2_$H_{21,22,23}$---OD_1_D151	100.0	86.2
OTV---D151	OTV_N_2_$H_{21,22,23}$---OD_2_D151	100.0	100.0
OTV---W178	OTV_N_2_$H_{21,22,23}$---O_W178	71.1	-
OTV---R152	OTV_O_3---HH_{22}_NH_2_R152	99.0	59.4
OTV---R152	OTV_O_3---HH_{21}_NH_1_R152	100.0	-
OTV---E277	OTV_N_1_H_{24}---OE_2_E277	24.8	37.0

3.4.3 Loss of inhibitor-enzyme hydrogen bond

Hydrogen bonding analysis was extended to determine the particular interactions between the inhibitor and its binding pocket residues, using the same criteria as mentioned above. From the histograms shown in Figure 3.5(a), numbers of hydrogen bonds are highly detected in the wild- and mutant-type, formed by R292, R371, Y347, E119, D151, W178, R152 and E277. Loss of inhibitor-enzyme hydrogen bonds, due to the H274Y mutation, takes place at W178 and R152 residues.
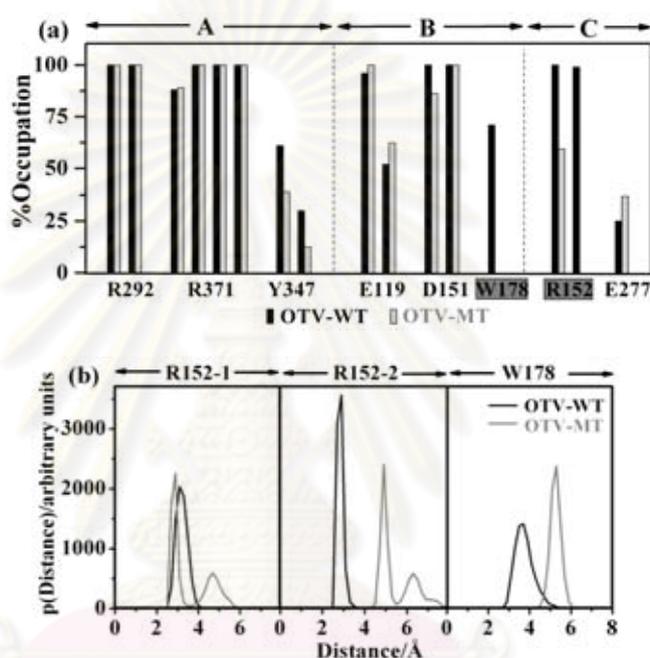
Figure 3.5 (a) Percent occupation pattern of hydrogen bonds between residues in the binding pocket (labeled on the *x*-axis) and side chains A, B and C (defined in Figure 3.2(a)) of the inhibitors for the two simulated systems, OTV-WT and OTV-MT. (b) The probability distributions of distance between specific atoms of the inhibitors and NA residues, R152 and W178, located in the active site of the oseltamivir.

The results in Figure 3.5 lead to a clear conclusion that the interaction with W178 and the one hydrogen bond with R152 (see Figure 3.2(a) for their locations) were lost in the OTV-MT complex (see grey lines). This observation can be understood by the distribution plots of the related distances shown in Figure 3.5(b). With the distances of ~ 3.0 Å (maxima of the first two black-line peaks), the nitrogen atoms of R152's guanidinium group in OTV-WT can form two strong hydrogen bonds with the carbonyl oxygen (O_3) of oseltamivir's –NHAc group (side C in Figure 3.2(a)). This is different for the orientation of R152 in the OTV-MT complex, where

the guanidinium group was flipped; one of the two nitrogen atoms was turned to be located ~ 5.0 Å from the O_3-oxygen (see grey line in Figure 3.5(b)) and, therefore, only one hydrogen bond remained. For the W178 residue, the loss of the hydrogen bond in OTV-MT is a consequence in the increase in the distance between the ammonium nitrogen (N_2) of oseltamivir and the backbone oxygen (O) of W178, from 3.7Å in OTV-WT to 5.3 Å in OTV-MT. Question arose whether loss of the hydrogen bonds between oseltamivir and these two residues, R152 and W178, is due to the displacement of the oseltamivir. It can be clearly seen from Figure 3.4(c) that oseltamivir in the mutant system does not move toward its $-OCHEt_2$ groups (side D in Figure 3.2(a)).

3.4.4 Oseltamivir's solvation

The ligand solvation was monitored in terms of atom-atom radial distribution functions (RDFs, $g_{xy}(r)$ —the probability of finding a particle of type y within a sphere radius r around the particle of type x). The results as well as the running coordination numbers integrated up to the first minimum (marked by an arrow) of the corresponding RDF are summarized in Figure 3.6. No change was found in terms of both number and distribution of water molecules around the O_1- and O_2-oxygens on the carboxylate group (side A) and the O_4-oxygen on the bulky group (side D) of oseltamivir (atomic labels shown in Figure 3.2(a)). Interest is focused on the N_1, N_2 and O_3 atoms on the $-NH_3^+$ (side B) and $-NHAc$ (side C) groups of oseltamivir where solvation in the wild-type and mutant systems is considerably different. The three RDFs of OTV-MT show sharp first peaks centered at ~3.0 Å, with the corresponding coordination numbers of 1, 1 and 0.5 water molecules for the N_1, N_2 and O_3 atoms, respectively. Based on detailed analysis of the simulated trajectories, the three coordination numbers are found to results from one water molecule located in the vicinity of the $-NH_3^+$ and $-NHAc$ side chains of oseltamivir. The situation is different for the OTV-WT complex, where water was found at distances longer than 4.0 Å from the two nitrogen atoms while the coordination number of 0.2 was observed for O_3. The simulation results lead us to conclude that these three atoms in OTV-MT were increasingly solvated. This can be a clear reason why sides B and C of oseltamivir lose their interactions with the R152 and W178 residues of the N1 enzyme (Figure 3.5(a)).
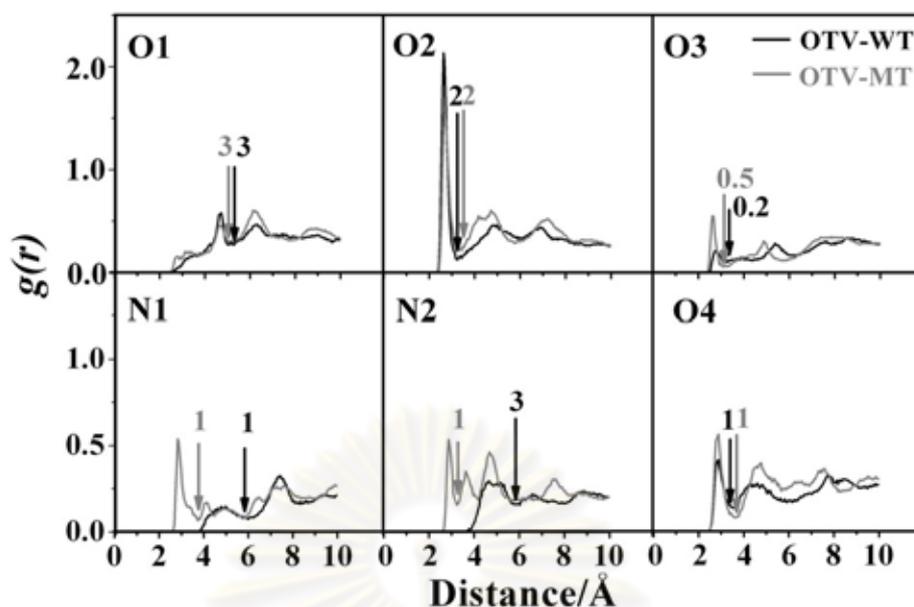
Figure 3.6 Radial distribution functions, *g(r)*, centered on oseltamivir atoms, to water oxygen atoms of the OTV-WT and OTV-MT complexes. The corresponding running coordination numbers integrated up to their first minima (marked by arrows) are also given.

3.4.5 Binding affinity of oseltamivir

The MM/PBSA-based binding free energies for the OTV-WT and OTV-MT complexes were calculated and separated into electrostatic (ΔE_{ele}), van der Waals, (ΔE_{vdw}), non-polar solvation ($\Delta G_{sol}^{nonpolar}$) and electrostatic solvation (ΔG_{sol}^{ele}) components. The results are summarized in Table 3.2. As expected, the electrostatic energy, part of the ΔE_{MM}, makes a major contribution to the oseltamivir-enzyme binding. The H274Y mutation leads to a decrease in the $\Delta G_{binding}$ from -14.6 ± 4.3 kcal mol^{-1} to -9.9 ± 6.4 kcal mol^{-1}. The major contribution to the weaker binding affinity in the MT complex comes from the electrostatic solvation free energy term, 214.3 ± 10.3 kcal mol^{-1} compared with 206.0 ± 8.8 kcal mol^{-1} for OTV-WT. This is in agreement with the solvation data where more water molecules were found in the vicinity of oseltamivir in OTV-MT than that in OTV-WT. The observed trends, as well as the absolute values of the predicted binding free energies, are consistent with experimental studies [34] indicating the lower inhibitory potency of OTV to mutant NA compared with wild-type NA.

Table 3.2 Calculated binding free energy and its components (kcal mol^{-1}) for the calculated OTV-WT and OTV-MT complexes.

	OTV-WT	OTV-MT
ΔE_{ele}	-191.9 ± 10.5	-195.7 ± 11.2
ΔE_{vdw}	-28.8 ± 3.6	-28.5 ± 3.4
ΔE_{MM}	**-220.6 ± 9.7**	**-224.2 ± 9.8**
$\Delta G_{sol}^{nonpolar}$	-3.5 ± 0.1	-3.2 ± 0.4
ΔG_{sol}^{ele}	209.5 ± 8.8	217.5 ± 10.3
ΔG_{sol}	206.0 ± 8.7	214.3 ± 10.3
$\Delta G_{sol}^{ele} + \Delta E_{ele}$	17.6 ± 5.3	21.8 ± 7.5
$\Delta G_{binding}$	**-14.6 ± 4.3**	**-9.9 ± 6.4**
$\Delta G_{experiment}$	**-13.0 ± 0.2**	**-8.6 ± 0.2**

3.5 CONCLUSIONS

We have studied and compared the molecular and structural properties of the oseltamivir binding to both wild-type and mutant of avian influenza (H5N1) NA by using the molecular dynamics simulation approach. The extensive analysis of OTV-MT and OTV-WT complexes was carried out with the simulated data obtained from the present study and previously calculated simulations [6], respectively. The results provide detailed information on the oseltamivir-resistance caused by the H274Y mutation at the molecular level. Within the N1 active site of OTV-MT, bulky side chain of the Y274 phenol ring drives the E276 carboxylate group to rotate away in a direction toward an ethyl moiety of oseltamivir's bulky group, resulting in a rather small hydrophobic pocket that is unable to accommodate this bulky group. This is found to be the primary source of the oseltamivir-resistance associated with the H274Y mutation, which is in contrast with the proposed mechanism of oseltamivir-resistance to this particular mutation in H5N1 NA [37, 40]. The E276 side chain rotation confers the reduction of hydrophobicity at the bulky group. In addition, the −NH$_3^+$ and −NHAc groups of oseltamivir lose their hydrogen bonding interactions with the N1 residues, which was compensated by greater accessibility to water molecule at these two regions. This basic knowledge will be useful for the development and refinement of new antiviral inhibitors for a high potency against both wild-type and drug-resistant strains of neuraminidase N1.

CHAPTER IV

NA-MUTATION (H1N1-2009)

Molecular prediction of oseltamivir efficiency against probable influenza
A (H1N1-2009) mutants: molecular modelling approach

Thanyada Rungrotmongkol[1,2,a], Maturos Malaisree[1,a], Nadtanet Nunthaboot[3],
Pornthep Sompornpisut[1] and Supot Hannongbua[1]

---

[1]*Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science,
Chulalongkorn University, Bangkok, 10330, Thailand.*

[2]*Center of Innovative Nanotechnology, Chulalongkorn University, Bangkok, 10330,
Thailand.*

[3]*Department of Chemistry, Faculty of Science, Mahasarakham University,
Mahasarakham, 44150, Thailand.*

[a]These authors contributed equally to this work.

4.1 ABSTRACT

To predict the susceptibility of the probable 2009 influenza A (H1N1-2009) mutant strains to oseltamivir, MD/LIE approach was applied to oseltamivir complexed with the most frequent drug-resistant strains of neuraminidase subtypes N1 and N2: two mutations on the framework residues (N294S and H274Y) and the two others on the direct binding residues (E119V and R292K) of oseltamivir. Relative to those of the wild-type, loss of drug-target interaction energies especially in terms of electrostatic contributions and hydrogen-bonds were dominantly established in the E119V and R292K mutated systems. The inhibitory potencies of oseltamivir towards the wild-type and mutants were predicted in according with the ordering of binding free energies: wild-type (-12.3 kcal mol^{-1}) > N294S (-10.4 kcal mol^{-1}) > H274Y (-9.8 kcal mol^{-1}) > E119V (-9.3 kcal mol^{-1}) > R292K (-7.7 kcal mol^{-1}), suggesting that the H1N1-2009 influenza with R292K substitution perhaps conferred a high level of oseltamivir resistance, while the other mutants revealed moderate resistance levels. This result calls for an urgent need to develop new potent anti-influenza agents against the next pandemic of potentially higher oseltamivir-resistant H1N1-2009 influenza.

4.2 INTRODUCTION

The 2009 influenza A (H1N1) virus has rapidly spread across the world with an evidence of human-to-human transmission. The probable mutation in the neuraminidase (NA) genes could cause resistance to the available drugs, especially oseltamivir. A new drug-resistant strain probably leads to a large-scale outbreak of novel pandemic flu and an increase the national and global public health concerns. Common mutations in N1 (a subtype in NA group-1) are detected at N294S and H274Y, while the E119V and R292K mutations are mostly found in the N2 and N9 subtypes (in NA group-2), with oseltamivir-resistance levels relative to the wild-type of 20 - 80, 700 - 1700, 20 - 1000 and 1500 - 10000 fold higher, respectively [28, 33-40, 80]. To date, oseltamivir has been found to effectively inhibit this new virus (2009 A (H1N1)) due to the following reasons: the N1 of the new H1N1 influenza and N9 share an identical active site [81, 82], and oseltamivir was designed to fit well to the active site of the NA group-2. With an increase in medical use and stockpile of oseltamivir for the recent outbreak, the question arises, and is the main goal of this

study, can we predict the inhibitory activity of oseltamivir with respect to those frequent mutations that take place in the influenza A (H1N1-2009) strain, and thus the potential evolution and spread of resistant strains. The oseltamivir-resistant influenza NA mutants would perhaps serve as the emergence of a potential pandemic strain of the 2009 H1N1 virus.

Oseltamivir is an antiviral drug against NA which functions by preventing viral replication in the last step of the viral life cycle. It was found to directly interact with the catalytic residues of the NA active site, while the framework residues stabilized the enzyme structure (Figure 4.1) [33]. Mutations at the conserved residues of NA appear to associate with oseltamivir resistance in a subtype specific manner. Thus the mutated framework residues H274Y and N294S are regularly indentified in N1, while in the N2 and N9 subtypes, mutations on the binding residues (E119V and R292K) of oseltamivir were detected after treatment in infected patients with high oseltamivir resistance [33-40].

To provide information at the molecular level in order to aid the control and prevention of emerging potential pandemic strains of the 2009 H1N1 influenza, multi-molecular dynamics (MD) simulations in conjunction with the linear interaction energy (LIE) method have been performed on complexes of oseltamivir bound to each of the four most likely 2009 H1N1 mutated strains; that is with the H274Y, N294S, E119V and R292K substitutions. The structural property, drug-target interaction and the binding affinity of oseltamivir against the mutated models are extensively discussed and compared to those recently published for the wild-type strain of the 2009 H1N1 and H5N1 influenza A virus [81, 82].
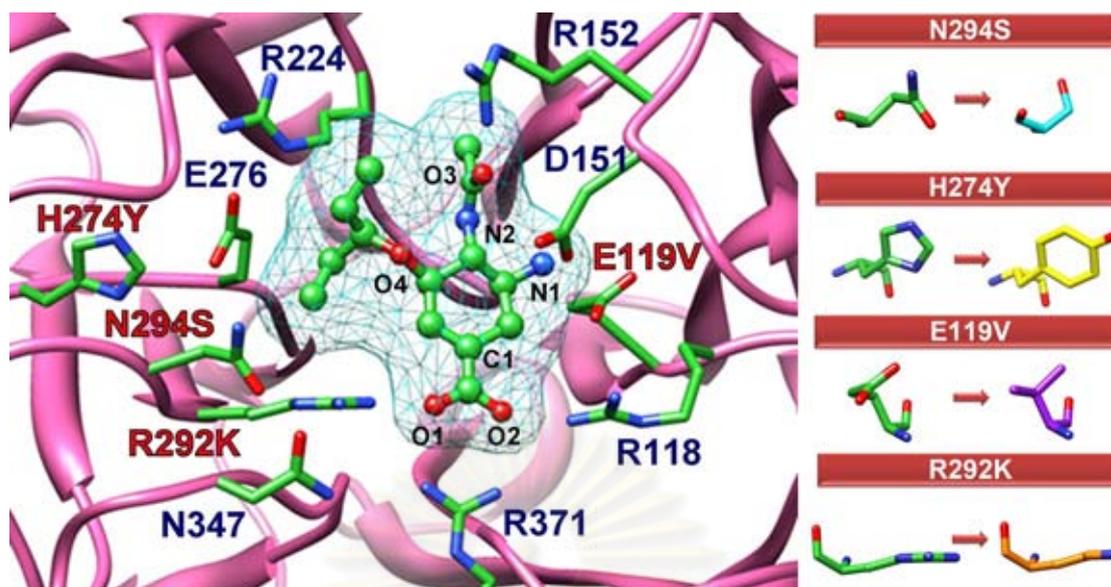
Figure 4.1 Modelled structure of oseltamivir bound to the wild-type strain of 2009 H1N1 influenza NA. Among the labeled residues, four residues coloured in red are singly mutated for investigation in this work: N294S, H274Y, E119V and R292K. The selected atoms of oseltamivir are numbered for simplicity in the discussion.

4.3 MATERIALS AND METHODS

4.3.1 Molecular dynamics simulations

The homology model of oseltamivir bound to the wild-type NA (OTV-WT) of the 2009 H1N1 virus [81] was used as the initial structure for the modelling of the four single mutations: N294S, H274Y, E119V and R292K. To prepare each mutant, the specific residue was changed using the LEaP module of the AMBER 10 program package [83], keeping the backbone and identical sidechain atoms. All mutated NA strains with oseltamivir bound were then set up and treated in accordance with the 20-ns MD simulations for the wild-type novel H1N1 influenza [81], as follows.

Each simulated system was performed by MD simulations with spherical boundary condition under the surface constrained all atom solvent model [84] using the Q-program [85], version 5. The atomic charges of oseltamivir were taken from our previous study [6]. The AMBER force field [83] was applied to the amino acid and inhibitor atoms. To set up the environment for oseltamivir to be the most similar in all simulated systems and to take the conformational change of the oseltamivir into account, the C1 atom of oseltamivir was then chosen to be the center of simulation

and the whole oseltamivir structure was thus considered as ligand. In the simulations, the system was capped by a 25 Å sphere of TIP3P water molecules centered on the C1 atom of oseltamivir (see Figure 4.1 for atomic label). Atoms positioned further than 25 Å from the C1 center were taken into consideration as structural restrains. All acidic and basic side chains of residues lying within a 22 Å sphere were fully charged. In contrast, these ionizable residues positioning between 22 Å and 25 Å distances were neutralized, except for the pairs of charged residues with a probable formation of hydrogen bonding interactions. The rest ionizable residues located outside a 25 Å sphere were considered as uncharged entities. Local reaction field approximation was employed for calculating the long-range electrostatic interactions, with a 10-Å cut-off radius for the non-bonded interactions. The SHAKE algorithm [54] was applied to fix all bonds involving hydrogen atom. The NVT ensemble was performed and a 2-fs time step was used. Initially, locations of the water molecules were simulated by MD simulations at 5 K, keeping all other atoms fixed to their initial positions, and the whole structure was then relaxed by four steps of simulations. Afterwards, the system was heated to 298 K over 300 ps, followed by equilibration phase. At last, the four equilibrated structures were randomly chosen for employing a production phase of 5-ns simulation.

4.3.2 Linear interaction energy

To predict the binding free energies (ΔG_{bind}) of oseltamivir towards the NA mutants the LIE method [86, 87] was used. The total binding free energy, which includes the van der Waals (vdW) (U^{vdW}) and the electrostatic (ES) interaction energies (U^{ES}), of the two simulated states: *(i)* the solvated ligand (free state), and *(ii)* the ligand bound to the solvated protein (bound state), were evaluated using the equation:

$$\Delta G_{bind} = \alpha (<U^{vdW}>_{bound} - <U^{vdW}>_{free}) + \beta (<U^{ES}>_{bound} - <U^{ES}>_{free}) + \gamma \qquad (1)$$

where α and β are the empirical scaling coefficients for the vdW and ES interaction energies, respectively, and γ is a constant. Here, Wall's coefficients ($\alpha = 0.472$, $\beta = 0.122$ and $\gamma = 2.603$), which were efficiently derived from a statistical analysis of the inhibitor sets binding to the relevant NA enzyme [88], were chosen to fit the LIE equation due to the three following reasons. (*i*) Since the outbreak of the novel H1N1 pandemic flu was just arisen in April 2009, the experimental inhibitory activities

required for the construction and validation the LIE model for the training set are not available. *(ii)* This set of coefficients was successfully applied on the avian influenza A (H5N1) virus in prediction the inhibitory activity of oseltamivir against both wild-type and mutant strains [81, 82]. *(iii)* The four single mutated strains of 2009-H1N1 NA virus in the present study were built by a specific mutation on the wild-type strain modelled from the crystal structure of the H5N1 neuraminidase with the sequence identity of 91% [81, 82]. Therefore, both wild-type and mutant NA strains of 2009-H1N1 are relatively similar to the H5N1 NA enzyme.

4.4 RESULTS AND DISCUSSION

4.4.1. Reduced oseltamivir binding to probable H1N1-2009 mutants

To examine oseltamivir susceptibility within the NA pocket of the 2009 H1N1 mutant models, intermolecular hydrogen-bond (H-bond), ES and vdW interactions between the oseltamivir's side chains and the NA residues were evaluated and compared to those of the wild-type [81]. The ES and vdW energetic differences were evaluated using equations (2a) and (2b), and are summarized in Table 4.1:

$$<\Delta\Delta U^{ES}>_{bound} \ = <U^{ES}>_{bound} \text{[mutant]} - <U^{ES}>_{bound} \text{[wild-type]} \qquad (2a)$$

$$<\Delta\Delta U^{vdW}>_{bound} = <U^{vdW}>_{bound} \text{[mutant]} - <U^{vdW}>_{bound} \text{[wild-type]} \qquad (2b).$$

The positive and negative values of the energy components indicate that the selected moiety of oseltamivir in the mutants decreases and increases its binding potency, relative to that of the wild-type, respectively. The H-bonds were calculated according to the two criteria that: *(i)* the proton donor (D) and acceptor (A) distance is ≤ 3.5 Å and *(ii)* the D-H..A angle is $\geq 120°$. The results are shown in Figure 4.2, whereas the descriptions are given in Table 4.2. The schematic views of hydrogen bonds formed between oseltamivir and its binding residues extracted from the simulations were given in Figure 4.3.

Lower oseltamivir binding free energies to the probable H1N1-2009 mutants were observed in terms of the H-bonds, ΔU^{ES} and ΔU^{vdW} energies relative to the wild-type, depending on where the mutation is located. As expected, the ES effect of single mutation at the framework residues (H274Y and N294S) is drastically less than that at the direct binding residues (E119V and R292K), which leads to the ΔU^{ES} reduction being in the range of c.a. 10 - 11 and 24 - 25 kcal mol^{-1}, respectively (Table 4.1(a)).

Table 4.1 Changes in the electrostatic and van der Waals interactions for the four substituent functional groups of oseltamivir in the bound states of the modelled H1N1 mutant relative to those of wild-type. Means and standard deviations are derived from four separate 5 ns simulations.

(a) Electrostatic interactions

Functional group	$<\Delta\Delta U^{ES}>_{bound}$ [kcal mol^{-1}]			
	N294S	**H274Y**	**E119V**	**R292K**
–COO$^-$	**8.3 ± 0.2**	**11.3 ± 0.1**	5.4 ± 0.1	**14.1 ± 0.2**
–NH$_3^+$	-2.7 ± 0.4	-2.8 ± 0.5	**11.1 ± 0.6**	2.3 ± 1.2
–NHAc	3.0 ± 0.1	2.0 ± 0.1	6.0 ± 0.1	4.2 ± 0.2
–OCHEt$_2$	1.3 ± 0.6	0.6 ± 0.6	1.5 ± 0.9	4.3 ± 1.2
Oseltamivir	**10.0 ± 0.6**	**11.1 ± 0.6**	**23.9 ± 0.9**	**24.9 ± 1.2**

(b) Van der Waals interactions

Functional group	$<\Delta\Delta U^{vdW}>_{bound}$ [kcal mol^{-1}]			
	N294S	**H274Y**	**E119V**	**R292K**
–COO$^-$	0.3 ± 0.0	0.3 ± 0.0	0.3 ± 0.0	0.8 ± 0.3
–NH$_3^+$	0.6 ± 0.1	0.6 ± 0.0	-0.3 ± 0.1	0.8 ± 0.1
–NHAc	0.2 ± 0.1	**1.0 ± 0.1**	**1.1 ± 0.0**	-0.7 ± 0.1
–OCHEt$_2$	**1.4 ± 0.1**	**1.5 ± 0.0**	0.1 ± 0.0	**3.2 ± 0.2**
Oseltamivir	2.5 ± 0.0	**3.4 ± 0.0**	1.2 ± 0.0	**4.2 ± 0.2**
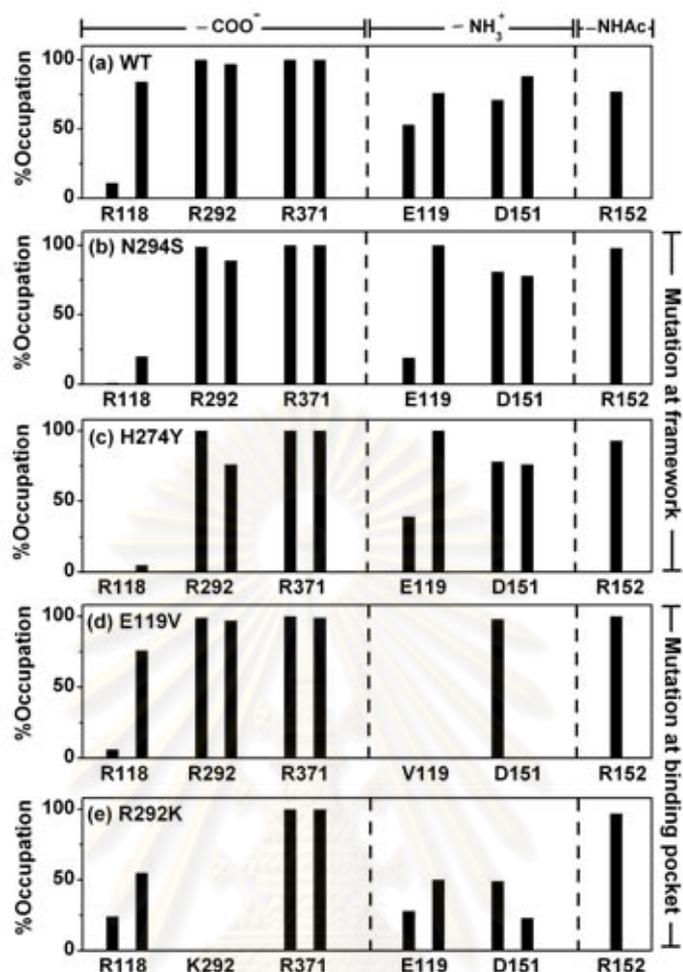
Figure 4.2 Percentage occupation of H-bonds between the functional groups of oseltamivir and the NA residues (see Figure 4.1 for residue positions) in the mutant models with the single substitution at two district regions: the framework residues closed to the hydrophobic pocket (N294S and H274Y), and the direct binding residues (E119V and R292K).

For the H1N1 wild-type, the strong oseltamivir-NA interactions were found via five, three and one H-bonds ($\geq$ 75%) with the $-COO^-$, $-NH_3^+$ and $-NHAc$ moieties, respectively (Figures 4.2(a) and 4.3(a)). In the two framework region mutations, it can be seen that the $-COO^-$ group of oseltamivir has almost lost the H-bonds with R118 (see Figures. 4.2(b) and 4.3(b) for N294S; Figures. 4.2(c) and 4.3(c) for H274Y) in correspondence with the ΔU^{ES} ($-COO^-$) reduction of 11.3 kcal mol^{-1} in H274Y and 8.3 kcal mol^{-1} in N294S (Table 4.1(a)). As expected (Table 4.1(b)), the decreased ΔU^{vdW} (oseltamivir) of 2.5 kcal mol^{-1} by N294S and of 3.4 kcal mol^{-1} by

H274Y were mainly contributed from the loss of vdW interactions at the bulky OCHEt$_2$ group (~1.5 kcal mol^{-1}). Previous theoretical studies on the influenza NA mutants have already explained how the H274Y mutation confers oseltamivir-resistance by a meaningful change of E276's sidechain conformation with a consequent effect upon the shape and size of the hydrophobic pocket for the –OCHEt$_2$ moiety [6, 81, 82] while E276 in the N294S mutant acted as the center of H-bond network between R224 and S294 [82] similar to that found in the crystal structure of the oseltamivir-resistant H5N1 N294S variant [80].

Table 4.2 Hydrogen bond descriptions and interactions detected between oseltamivir (OTV) and particular residues of the N1 enzyme in Figure 4.2 (see Figure 4.1 for atomic label)

Residue	Type	WT	N294S	H274Y	E119V	R292K
R118	OTV_O2---NH1_R118	11	-	-	6	24
R118	OTV_O2---NH2_R118	84	20	5	76	55
R292/K292	OTV_O1---NH1_R292	100	99	100	99	-
R292/K292	OTV_O1---NH2_R292	97	89	76	97	-
R371	OTV_O1---NH1_R371	100	100	100	100	100
R371	OTV_O2---NH2_R371	100	100	100	99	100
E119/V119	OTV_N1---OE1_E119	53	19	39	-	28
E119/V119	OTV_N1---OE2_E119	76	100	100	-	50
D151	OTV_N1---OD1_D151	71	81	78	98	49
D151	OTV_N1---OD2_D151	88	78	76	-	23
R152	OTV_O3---NH2_R152	77	98	93	100	97

With a relatively high reduction in the ES contribution to oseltamivir in the mutations at the binding residues (E119V and R292K, Table 4.1(a)), the mutated residues V119 and K292 showed a complete loss of H-bond interactions with the -NH$_3^+$ and –COO$^-$ moieties of oseltamivir (Figures. 4.2(d) and 4.3(d) for E119V; Figures. 4.2(e) and 4.3(e) for R292K), supported by an increase in the ΔU^{ES} (-NH$_3^+$) by 11.1 kcal mol^{-1} and in the ΔU^{ES} (–COO$^-$) by 14.1 kcal mol^{-1}. In addition, only one

H-bond with D151 in the E119V mutant was maintained, while lower H-bond strengths in the R292K mutant were observed at R118, E119 and D151. Moreover, a reduced vdW interaction of 3.2 kcal mol^{-1} was found at the -OCHEt$_2$ group in the R292K mutant because the side chain of K292 (Figure 4.3(e)) is smaller and shorter than that of R292 (Figure 4.3(a)). The results of the R292K mutation were somewhat comparable to the computational study of the sialic acid analogs binding to the R292K mutated NA subtype N9 [90].
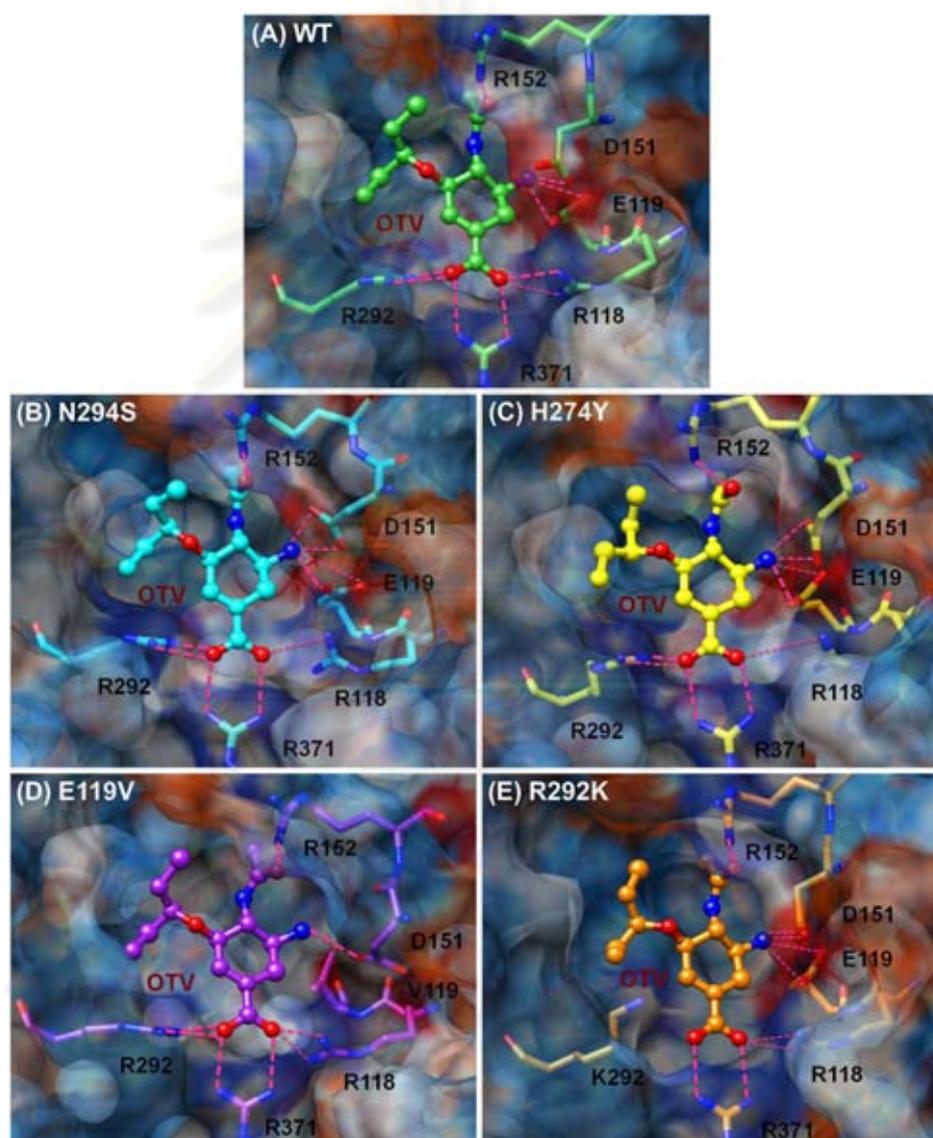
Figure 4.3 Electrostatic potential of five different NA strains complexed with oseltamivir where negative regions are in red and positive regions are in blue: (a) wild-type, (b) N294S, (c) H274Y, (d) E119V and (E) R292K. Close-up of oseltamivir, hydrogen bonds to its binding residues are represented by red dashed line.

4.4.2. Prediction of inhibitory activity against the H1N1 mutated strains

Based upon the MD/LIE approach, the binding affinities of oseltamivir towards different mutant models of the 2009 H1N1 influenza (A/California/04/2009(H1N1)), according to equation (1), were predicted and are summarized in Table 4.3. As expected, oseltamivir's binding free energy against the wild-type (WT) is the most favorable one at -12.8 kcal mol^{-1}. Only moderate binding free energy values were found for the N294S, H274Y and E119V mutated strains, in which the corresponding ΔG_{bind} of -10.4, -9.8 and -9.3 kcal mol^{-1}, respectively. The lowest favorable binding of oseltamivir is found in the R292K mutant with a predicted ΔG_{bind} of -7.7 kcal mol^{-1}. All the calculated binding free energies were found to fall within the ranges of those experimentally determined for various wild-type and mutant strains of the other influenza N1 and N2 subtypes (Table 4.3) [33-40]. Taking all the above data into account, it seems likely that oseltamivir will be significantly less potent an inhibitor for all the modelled mutants of the 2009 H1N1 strains, with the ranked order of: R292K < E119V < H274Y < N294S.

Table 4.3 MD/LIE binding free energies (ΔG_{bind}) of oseltamivir towards the 2009 H1N1 influenza NA (A/California/04/2009(H1N1)) for the wild-type (WT) and the probable single mutations: N294S, H274Y, E119V and R292K. Means and standard deviations are derived from four separate 5 ns simulations. The experimental ΔG_{bind} for different strains of N1 and N2, converted from the K_I inhibitory and IC_{50} values, are also given for comparison.

NA strain	ΔG_{bind} [kcal mol^{-1}]				
	WT	N294S	H274Y	E119V	R292K
Predictive					
A/California/04/2009(H1N1)	-12.8 ± 0.9	-10.4 ± 0.9	-9.8 ± 1.0	-9.3 ± 0.8	-7.7 ± 0.7
Experimental*					
A/WSN/33 (H1N1)[a]	-12.1	-9.3	-8.5	-	-
A/Puerto Rico/8/34 (H1N1)[b]	-11.4	-8.8	-8.1	-	-
A/Vietnam/1203/04 (H1N1)[b]	-13.0	-11.2	-8.6	-	-
A/Vietnam/1203/04 (H5N1)[c]	-13.1	-10.5	-9.8	-	-
A/Sydney/5/97 (H3N2)[a]	-12.8	-8.3	-	-8.6	-7.3
A/Wuhan/359/95 (H3N2)[d]	-12.4	-	-	-9.1	-6.2

*$\Delta G_{experiment}$ was calculated from the experimental data using the following references: (a) Abed et al. 2008, (b) Yen et al. 2007, (c) Collins et al. 2008 and (d) Yen et al. 2005.

4.5 CONCLUSIONS

In the present study, multi-MD simulations in conjunction with the LIE method was performed on oseltamivir-NA bound complexes for the four probable NA mutants of influenza A (H1N1-2009): two mutations on the framework residues (N294S and H274Y) and the two others on the direct binding residues (E119V and R292K) of oseltamivir. Reduction in the oseltamivir-enzyme interaction energies, particularly in the ES term, and in the hydrogen bonding were both observed in the two mutated systems with substitution on the direct binding residues, E119V and R292K. Based on the MD/LIE approach, the inhibitory potencies of oseltamivir towards the wild-type and mutants were predicted in accordance with their derived binding free energies (ΔG_{bind}) with: wild-type (-12.3 kcal mol^{-1}) > N294S (-10.4 kcal mol^{-1}) > H274Y (-9.8 kcal mol^{-1}) > E119V (-9.3 kcal mol^{-1}) > R292K (-7.7 kcal mol^{-1}). This means that oseltamivir (which, to date, effectively inhibits the current H1N1-2009 wild-type strain) is less effective in protection and/or treatment of patients with these probable mutants, and especially with the R292K variant. Therefore, surveillance of any mutations in the influenza A (H1N1-2009) needs to be closely watched, and prompt action taken for preparation for the next pandemic of potentially higher oseltamivir resistant H1N1 influenza strains. This calls for the urgent development of new potent anti-influenza agents against both native and mutants forms of H1N1-2009.

CHAPTER V

NA-SUBSTRATE INTERACTIONS

Substrate binding to neuraminidase of influenza viruses: QM/MM MD simulations

Maturos Malaisree[1], Christopher J. Woods[2], Thanyada Rungrotmongkol[1,3], Nadtanet Nantaboot[4], Arthit Vongachariya[1,5], Adrian J. Mulholland[2] and Supot Hannongbua[1]

[1]*Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand.*

[2]*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK.*

[3]*Center of Innovative Nanotechnology, Chulalongkorn University, Bangkok, 10330, Thailand.*

[4]*Department of Chemistry, Faculty of Science, Mahasarakham University, Mahasarakham, 44150, Thailand.*

[5]*Department of Basic R&D, UBE Technical Center (Asia) Limited, Rayong, 21000, Thailand.*

5.1 ABSTRACT

The molecular details of the cleavage mechanism in neuraminidase (NA) were studied using combined QM/MM MD simulations to understand the viral propagation step. To investigate the function of the amino acid residue at position 347 in the binding pocket, different NA subtypes: N1-1918, N1-2005, N1-2009, N2-1967 and N8-1963 bound with human substrate (SA-α-2,6-GAL) were performed. Ring conformation of sialic acid linked with galactose unit was observed to change from chair to twist-boat conformation in all systems except for an avian N8-1963 strain. This conformational change was stabilized via hydrogen bond interactions with the N/Q347 and K431 residues in both the N1 and N2 subtypes. This is not the case for the Y347 residue of avian type in N1-2005 and N8-1963 strains. This indicates that the residue at 347 could be an essential residue to stabilise the conformational change of the natural substrate and could be the source of the avian-human barrier for NA.

5.2 INTRODUCTION

The current outbreak of the novel H1N1-2009 influenza virus has raised global awareness of widespread human-to-human transmission [14, 15]. The first pandemic of 21st century of the new strain has spread rapidly, increasing the number of human deaths to more than 13,000 cases [13]. The combination of gene segments from swine, avian and human influenza viruses in the H1N1-2009 strain has not been detected before [90]. The most interested segment is focused on neuraminidase (NA) which has amino acid residues around the binding pocket that are similar to those in the pandemic Spanish flu H1N1-1918 influenza virus subtype. Understanding of the main interactions between these important amino acid residues and the natural substrate in different NA subtypes could provide useful information for drug development.

NA is a membrane glycoprotein of influenza virus that cleaves the terminal sialic acid and releases the new viruses to infect other cells [17]. Because the cleavage mechanism is an essential step for viral replication, NA has become a potent target for drug development for influenza virus. NA inhibitors have been designed and synthesized based on the transition state analog of sialic acid (SA) [1, 17, 41].

Currently, the potent oral-active agent, oseltamivir, has been widely used against NA function for all subtypes.

Nine subtypes of NA are classified into two groups: group-1 (N1, N4, N5 and N8) and group-2 (N2, N3, N6, N7 and N9) by genetic and structural relationships [21]. Different subtypes of NA have similar structures, which amino acid residues in the active site being highly conserved. Only NA subtypes N1-N2 have been found to infect humans. The ability of influenza virus to infect a host is influenced by the different types of linkage between SA and galactose (GAL) on the host cell surface [18]. Human influenza virus attaches to the host substrate via a SA-α-2,6-GAL linkage, while the avian influenza virus binds preferentially to the SA-α-2,3-GAL linkage [19]. The different specificity of the natural substrates could be the barrier that prevents avian-human viral infections.

Although, all amino acid residues in the active site which NA uses to interact with SA unit are conserved, only the residue in position 347 next to the catalytic pocket keeps changing its nature. Surprisingly, the different types of amino acid at this position do not affect significantly the efficacy of the NA inhibitors [13]. Moreover, the N347 residue in the novel H1N1-2009 strain is the same amino acid type in the pandemic Spanish flu H1N1-1918 strain while the avian H5N1 strain has a Y347 residue at this position [23, 91]. Until now, no experimental data or theoretical studies have revealed the role of this residue in the NA cleavage mechanism.

The mechanism of NA reaction has been proposed based on kinetic isotope effect [92], nuclear magnetic resonance spectroscopy [92] and molecular dynamics (MD) simulations [93-95]. The catalytic pathway for the cleavage mechanism is considered in following steps [92-95]: *(i)* The binding process of sialoside (SA linked with GAL unit) to NA active site involves a conformational change of the sialic acid part from a chair to a twist-boat conformation. This is induced by a steric effect caused be the Y406 and E277 residues at the base of the active site. Consequently, the carboxylate group of sialic acid interacts with the Arg-triad, R118, R292 and R371, via charge-charge interactions. *(ii)* Proton donation and formation of the transition-state intermediate of sialosyl cation are postulated by the activated water molecule by residues D151 and R152. *(iii)* Release of *N*-acetylneuraminic acid (Neu5Ac) is

initially formed by *α*-anomer and then following by mutarotation to *β*-anomer in solution.

In this study, substrate-enzyme interactions between the natural substrate, SA-α-2,6-GAL, and different NA subtypes: N1-1918 (Spanish flu), N1-2005 (avian H5N1 flu isolated from human), N1-2009 (the novel H1N1-2009 flu), N2-1967 (Asian flu) and N8-1963 (avian flu isolated from duck), were studied using QM/MM MD simulations. The aims of this study were to study the first step of the cleavage mechanism in different NA subtypes and to investigate the function of the amino acid residue at position 347 in the binding pocket.

5.3 METHODS

5.3.1 System preparation

Due to the lack of X-ray structure of NA bound with the natural substrate, the initial modelled structures for NAs complexed with SA-α-2,6-GAL were obtained via the following steps. The human SA-2,6-GAL receptor was taken from Protein Data Bank (PDB) code: 1RVT [96]. Superimposition of SA-2,6-GAL with ligand in the active site of NAs was obtained from X-ray structures from PDB codes: 3B7E [91], 2HU4 [23], 2BAT [22] and 2HTR [23] to obtain N1-1918, N1-2005, N2-1967 and N8-1963 systems, respectively. Finally, homology modelling was used to model N1-2009 from the N1-1918 system using the sequence template of the novel influenza virus with the A/California/04/2009(H1N1) strain [97]. Ionization states of amino acids were assigned using the PROPKA program [71]. All missing hydrogen atoms were added using HBUILD module in CHARMM [98]. This was then followed by 2,500 steps of steepest descents (SD) minimisation. Every system was solvated with 25 Å radius of TIP3P [51] water model centered on C2 atom (see Figure 5.1 for atomic position). The solvated water molecules were minimized using 600 steps of SD and 1,000 steps of adopted basis Newton Raphson (ABNR). The solvent sphere was then equilibrated for 25 ps with protein and ligand fixed to relax the system.

5.3.2 QM/MM MD simulations

The QM region contained only SA-2,6-GAL and the overall charge was -1 with 63 atoms. The QM atoms were represented using the semiempirical PM3carb model while the other residues were simulated using the classical CHARMM

forcefield [99]. Each system was divided into the reaction zone (0 - 21 Å) treated by Newtonian dynamics and the buffer regions (22 – 25 Å) treated by Langevin dynamics. A stochastic boundary MD simulation [100] was used to keep water molecules from moving outside of a 25 Å sphere. Amino acid residues in the buffer region were restrained at the minimized positions. Langevin dynamics of protein heavy atoms and water oxygen atoms were used by friction coefficients of 250 and 62 ps^{-1}, respectively [101]. The rest of system that was located more than 26 Å from the center atom was restrained during the simulations. The SHAKE algorithm [54] was applied to constrain the length of all bonds involving hydrogen atoms. A nonbonded cutoff with 13 Å and a dynamics integrator using a 1 fs of time step was used. Before dynamics of each substrate-NA complex was minimized by performing minimization on the whole system with 500 steps of SD, followd by 1,000 steps of conjugated gradient (POWELL). The system was heated to 300 K in three steps with 10 ps intervals (10-100 K, 100-200 K and 200-300 K). Then, the system was simulated in equilibration phase at 300 K and production phase was taken from 300 to 1,000 ps.

5.4 RESULTS AND DISCUSSION

5.4.1 Conformational changes of natural substrate in the first step

In the first step of cleavage mechanism, the natural substrate (SA-α-2,6-GAL) has been proposed to change the main ring conformation at the sialic acid (SA) part via steric repulsion by Y406 and E277 at the base of the active site [92-95]. To study the dynamics behaviour of SA-α-2,6-GAL in the binding process, three torsion angles were measured at carboxylate position (Tor1), main ring of SA (Tor2) and galactose (GAL) conformation (Tor3). The results are shown in Figure 5.1 as defined by sets of four atoms in Table 5.1.

At the beginning, carboxylate group (Tor1 in Figure 5.1(a)) of SA taken from X-ray structure [96] was -67° in the axial position of the pyranose ring. The observed results of Tor1 along QM/MM MD simulations were ~150° for every system except N8-1963 system. This is indicated that the carboxylate group could transform from the axial position to a pseudo-equatorial position as has been proposed by several studies [92-95]. The distorted conformation of the pyranose ring was measured further

by Tor2 in Figure 5.1(b) at the main ring conformation. The calculated torsion angle of ~0° was used to confirm conformational change from boat (-60°) to a twist-boat (~0°) conformation. This unfavourable conformation of SA induced the second step of the cleavage mechanism due to leaving of GAL unit as can be seen in Tor3 in Figure 5.1(c). Tor3 indicates conformation of GAL to SA in SA-α-2,6-GAL structure which O6 atom of GAL links with C2 atom of SA at the equatorial position (~180°) at the beginning. After SA changed its structure to the twist-boat conformation, the GAL unit was pushed to the perpendicular position (~90°) of the pyranose ring, making the suitable conformation of GAL to be cleaved for the second step of mechanism [95]. Although, the conformational changes were detected in the N1 and N2 subtypes (N1-1918, N1-2005, N1-2009 and N2-1967 systems), this was not observed in the avian N8 strain (N8-1963) where only the boat conformation was found. This could lead to the low ability of the avian virus to cleave the glycoside bond of the human receptor.
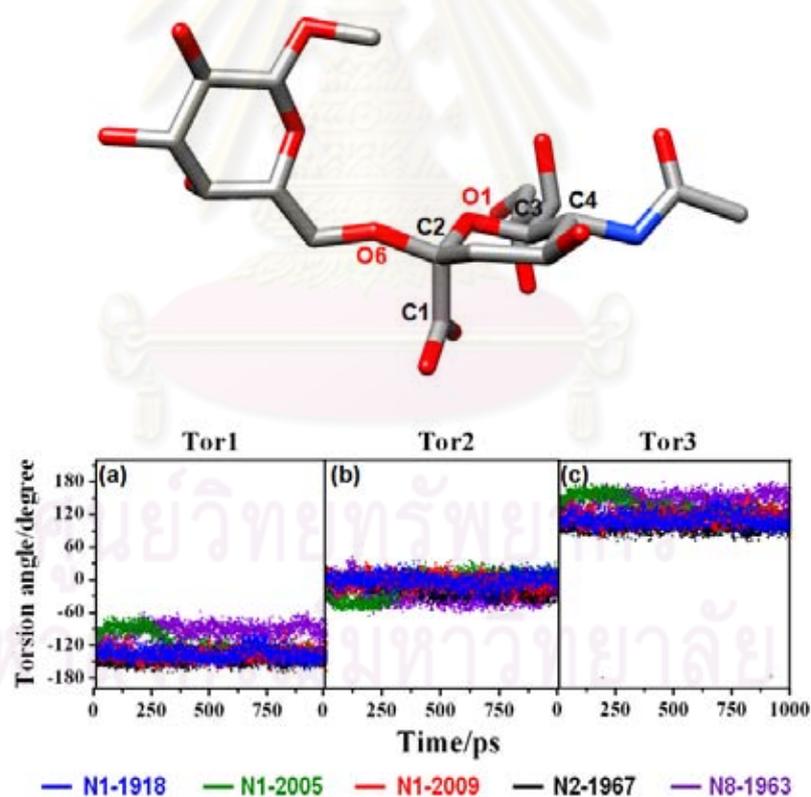
Figure 5.1 Torsion angles (Tor1-3) for each of the five systems as a function of QM/MM MD simulation time, representing the conformational change of the natural substrate in human (SA-2,6-GAL) receptor

Table 5.1 Definition of torsion angles (Tor1-3 shown in Figure 5.1) for human SA-2,6-GAL receptor

Torsion angle	Definition
Tor1	C1-C2-O1-C3
Tor2	C2-O1-C3-C4
Tor3	O6-C2-O1-C3

5.4.2 Substrate binding to NAs

To study the interactions between the natural substrate and NAs, hydrogen bond frequencies were determined based on the following criteria: (*i*) proton donor-acceptor distance ≤ 3.5 Å, and (*ii*) donor-H-acceptor bond angle $\geq 120°$. From hydrogen bond results in Figure 5.2, no significant difference was found in the SA part. The main interactions were conserved with the hydrogen bonds of the carboxylate group of SA and the Arg-triad, R118, R292 and R371. The hydroxyl group interacted with the negative charge residue, E119. For the *N*-Acetyl part, R152 formed hydrogen bonds with the carbonyl group. The last substitution group, the bulky hydrophilic group was found to interact with the E276 and E277 residues in every system.

The most interesting part was observed in the GAL unit. Hydrogen bond interactions were observed only in the N1 and N2 subtypes with the N294, N/Q347, G348 and K431 residues while no such hydrogen bond interactions were found in N8 subtype (see in Figure 5.2). The human viruses, N1-1918, N1-2009 and N2-1967 strains, show the same hydrogen bond pattern between GAL and N/Q347 and K431 residues. This is different pattern in the avian viruses, N1-2005 and N8-1963, in which these hydrogen bond interactions were detected only in the avian N1-2005 virus isolated in human [23] with N294 and G348 instead of Y347 and K/E431. In contrast with N8-1963, an avian virus isolated from duck [23], G294 and G348 could not interact with GAL in the same way with the avian N1-2005. Moreover, the positions 347 and 431 in avian N8-1963 virus are Y and E, respectively, and these were not found to be able to interact with GAL. This could be the reason that the

unstable twist-boat conformation in N8-1963 system could not be formed due to the lack of stabilisation by surrounding residues.
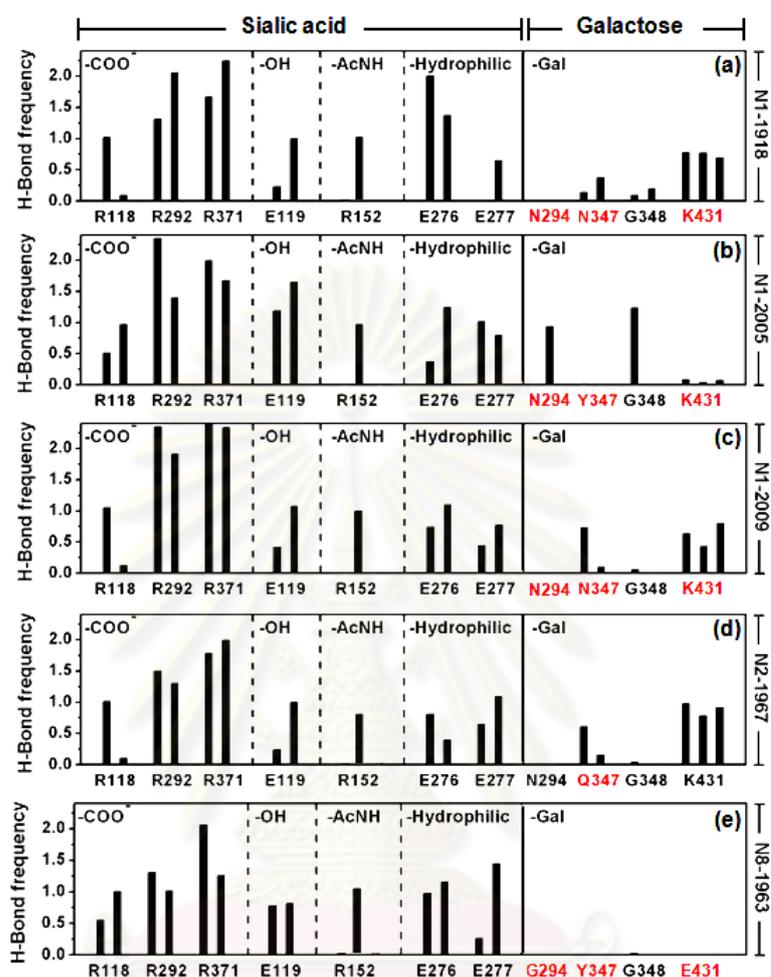
Figure 5.2 Hydrogen bond pattern of SA-2,6-GAL with the surrounding residues in NA active site

5.4.3 Role of residues 347 and 431

Normally, the residue at position 347 in the NA active site does not make any significantly different interactions with SA or inhibitors. However, this residue keeps changing its nature. Interestingly, the novel H1N1-2009 strain has the same N347 with the pandemic Spanish H1N1-1918 strain while the avian H5N1 strain has Y347 residue taking its place at this position. To study the role of residue 347, the distances between N/Q/Y347 and GAL were plotted in Figure 5.3 with the distances from residues K/E431 and Tor3 which was used to indicate the positional change of the GAL part in SA-α-2,6-GAL from a parallel to perpendicular position of the main ring.

At the beginning, the distances between these two residues were far from GAL ~4 - 8 Å in every system. However, when SA changed its conformation and pushed GAL in the perpendicular position (see the red lines at ~90° for Tor3), N/Q347 and K431 were found to come closer to GAL. It can be seen as the blue and black lines in Figure 5.3(a), (c) and (d) for N1-1918, N1-2009 and N2-1967, respectively. This could make hydrogen bonds and stabilise the conformational change of the natural substrate from the chair to the twist-boat conformation. In contrast, Y347 and K/E431 in N1-2005 and N8-1963 could not interact with GAL (Figure 5.3(b) and (e)), especially in the avian N8-1963 virus that has completely lost interaction with these two residues. For N1-2005, even if Y347 could not form a hydrogen bond with GAL (Figure 5.2(e)), the stabilisation of twist-boat conformation could be formed via the next residues in the active site, N294 and G348 (Figure 5.2(b)).
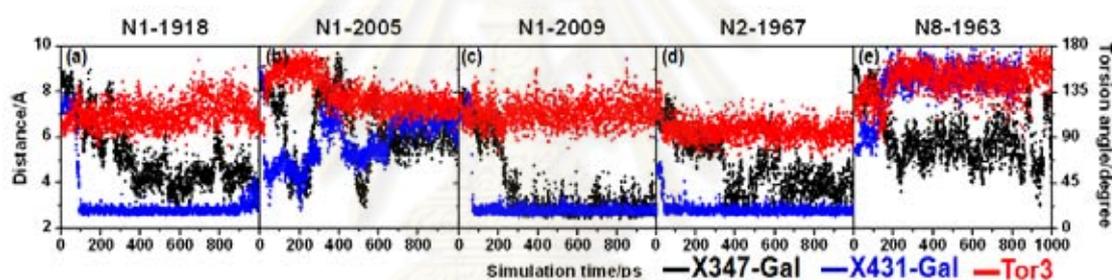
Figure 5.3 Distances of N/Q/Y347 and K/E431 to GAL with conformational change of GAL calculated by Tor3
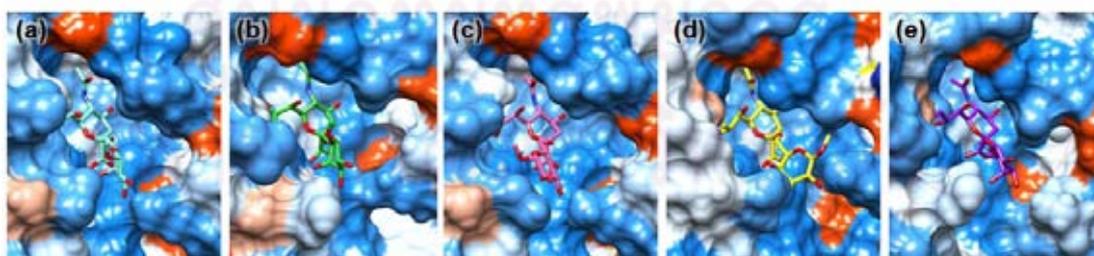
Figure 5.4 Three dimensional structure of (a) N1-1918, (b) N1-2005, (c) N1-2009, (d) N2-1967 and (e) N8-1963 bound with SA-α-2,6-GAL in the first step of the cleavage mechanism.

All together from QM/MM MD simulations, the first step of the cleavage mechanism is performed by the changing of the chair to the twist-boat conformations via the steric effect in the base of the active site. The twist-boat conformation was stabilised by hydrogen bonds from the N/Q347 and the K431 residues in N1 and N2 subtypes (see the three dimensional structure in Figure 5.4). In contrast, the avian virus N8 could not change the conformation of the SA of the natural substrate in human, SA-α-2,6-GAL, that is cleaved in the second step of the cleavage mechanism.

5.5 CONCLUSIONS

The combined QM/MM MD simulations were used to study the natural substrate in human (SA-α-2,6-GAL) bound to different NA subtypes (N1-1918, N1-2005, N1-2009, N2-1967 and N8-1963). The ring conformation of sialic acid linked with the galactose unit was observed to change from a chair to a twist-boat conformation in all systems except for an avian N8-1963 strain. This conformational change is stabilized via hydrogen bond interactions with the N/Q347 and K431 residues in both the N1 and N2 subtypes. This is not the case for the Y347 residue of avian type in N1-2005 and N8-1963 strains. The hydroxyl group of the tyrosine residue at position 347 in avian strains dose not accommodate a twist-boat conformation of the natural substrate. The obtained results indicate that the residues 347 and 431 could be the source of the avian-human barrier in the substrate binding process.

CHAPTER VI

METHOD DEVELOPMENT

A Water-swap reaction coordinate for the calculation of absolute protein-ligand binding free energies

Christopher J. Woods[1], Maturos Malaisree[2], Supot Hannongbua[2] and Adrian J. Mulholland[1]

[1]*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK*

[2]*Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand*
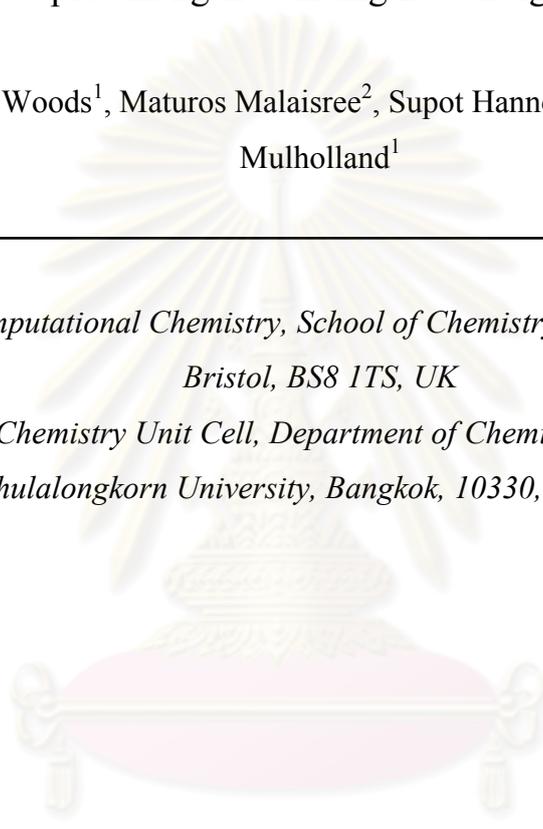
6.1 ABSTRACT

The accurate prediction of absolute protein-ligand binding free energies is one of the grand challenge problems of computational science. Binding free energy measures the strength of binding between a ligand and a protein, and an algorithm that would allow its accurate prediction would be a powerful tool for rational drug design. Here we present the development of a new method that allows for the absolute binding free energy of a protein-ligand complex to be calculated from first principles, using a single simulation. Our method involves the use of a novel reaction coordinate that swaps a ligand bound to a protein, with an equivalent volume of bulk water. This water-swap reaction coordinate is built using an identity constraint, which identifies a cluster of water molecules from bulk water that occupies the same volume as the ligand in the protein active site. A dual topology algorithm is then used to swap the ligand from the active site with the identified water cluster from bulk water. The free energy is then calculated using replica exchange thermodynamic integration (RETI). This returns the free energy change of simultaneously transferring the ligand to bulk water, as an equivalent volume of bulk water is transferred back to the protein active site. This, directly, is the absolute binding free energy. It should be noted that while this reaction coordinate models the binding process directly, an accurate forcefield and sufficient sampling are still required to allow for the binding free energy to be predicted correctly. In this paper we present the details and development of this method, and demonstrate how the potential of mean force (PMF) along the water-swap coordinate can be improved by calibrating the soft-core Coulomb and Lennard-Jones parameters used for the dual topology calculation. The optimal parameters were applied to calculations of protein-ligand binding free energies of a neuraminidase inhibitor (oseltamivir), with these results compared to experiment. These results demonstrate that the water-swap coordinate provides a viable and potentially powerful new route for the prediction of protein-ligand binding free energies.

6.2 INTRODUCTION

A method that could predict the absolute binding free energy of a ligand to a protein would be a powerful tool for rational drug design. Binding free energy measures the strength of binding between a ligand and a protein, and its accurate prediction is one of the grand challenge problems of computational science. However, despite decades of international research, no method exists currently that allows from first principles, the binding free energy of a ligand to a protein to be predicted accurately, and within a reasonable timescale (*i.e.* in less time than the ligand could be synthesized and the binding affinity measured experimentally). The current state of the art in the field are methods based on statistical thermodynamics, which use either free energy perturbation (FEP) [102-105] or thermodynamic integration (TI) [106-108]. These are used to calculate the free energy change of either perturbing one ligand into another to obtain a relative binding free energy [109, 110], or of annihilating a ligand both while bound to the protein, and while in bulk solvent (so-called double-annihilation [111] or double-decoupling [112,113]), used to obtain the absolute binding free energy [114, 115]. Several excellent reviews of binding free energy methods have been written [116-119], many of which detail the difficulties of their application [120-122].

Existing double-decoupling absolute binding free energy methods use two separate reaction coordinates that perform the decoupling of the ligand independently from both protein and from bulk water. To return the correct binding free energy, they require the addition of extra terms that account for the restraints used to limit the accessible volume of the ligand while it is being decoupled [112, 113]. The decoupling free energies must be calculated with a high accuracy. This is because, for large or polar ligands, they are typically of an order of 5 to 20 times the magnitude of the binding affinity, and the binding free energy is calculated as the difference between the free energy to decouple fully the ligand from bulk water and the free energy to fully decouple the ligand from the protein. Methods that reduce errors by avoiding these large decoupling free energies are the subject of active development. For example, Helms and Wade have developed an ingenious reaction coordinate [123] that perturbs a ligand into water. However, the method only works for buried binding sites, as otherwise, as the ligand is turned into water, the appearing waters would be free to diffuse into bulk. Helms and Wade thus limited the application of their method

to the calculation of the absolute binding free energy of camphor to P450cam. Warshel *et. al.* have also proposed a similar route as a means of calculating absolute binding free energies [116]. They proposed a free energy cycle that exchanges the bound ligand with a corresponding volume of bulk water. The cycle uses separate non-physical perturbations that mutate the ligand to water, and a cluster of bulk water to ligand. These perturbations proceed via uncharged forms of the ligand and water [116]. Woo and Roux have also devised a method that avoids calculating large decoupling free energies, by using restraints along a new reaction coordinate that, in effect, pulls the ligand out of the binding site and moves it to bulk water [124]. A similar, but perhaps simpler approach has also been devised by Doudou, Burton and Henchman [125].

In this paper we describe a new reaction coordinate, which we call the water-swap reaction coordinate (WSRC). This reaction coordinate acts, within a single simulation, to swap the ligand bound to the protein, with an equivalent volume of bulk water. The effect is to allow for the calculation of the free energy of transferring the ligand from being bound to the protein, to being solvated by bulk water, while simultaneously an equivalent volume of bulk water is transferred back to the protein. This free energy is, directly, the absolute binding free energy. This is a significant advance on previous methods, as the WSRC avoids the problem of large decoupling free energies, does not require the use of any restraints, and it allows for the binding free energy to be calculated directly from a single simulation along a single reaction coordinate. In the next section we present the details of the implementation of the WSRC. We then test and refine the coordinate in calculations for which the expected free energy change is known exactly. Finally, we present the application of the refined WSRC to the calculation of the absolute binding free energy of the drug, oseltamivir, to influenza neuraminidase.

6.3 METHOD DESCRIPTION

The idea of the water-swap reaction coordinate is based on a method, which we have developed recently, that allows for the identity of water clusters to be constrained during a simulation. This identity constraint (described in detail elsewhere [126, 127]) works by changing how water molecules are labelled during a simulation. Rather than labelling water molecules based on their index in an input coordinate file,

water molecules can instead be labeled by their location in space. This is achieved by placing identity points in space. For example, let us consider a blue and a green identity point (Figure 6.1(a)). These are used to label the blue and green water, with the labels assigned so as to minimize the sum of the distances between the center of the blue water molecule and blue point, and the center of the green water molecule and green point (Figure 6.1(b)). As the simulation progresses, and the water molecules move, this constraint is continually re-evaluated. If another water molecule moves such that, for example, it becomes closer to the green point (Figure 6.1(c)) then the identities of the waters are swapped, and it then becomes the green water molecule (Figure 6.1(d)).
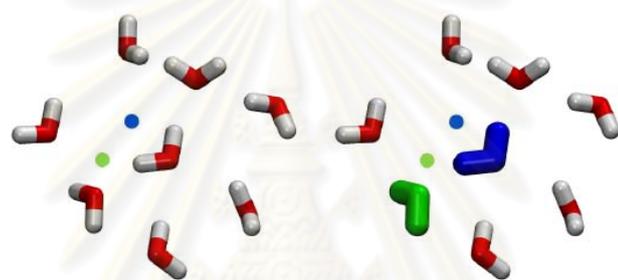
(a)A green and a blue identity point are added to a box of water.

(b)These identify the green and blue waters (the two waters that minimise the total distance between the green water and green point, and the blue water and blue point).
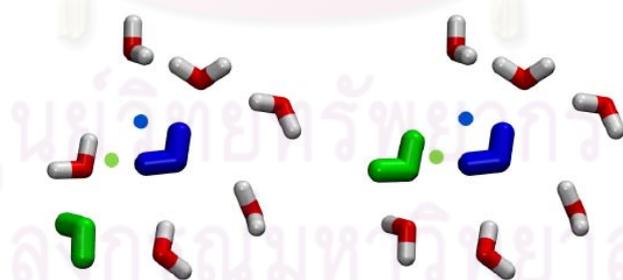
(c)This constraint is re-evaluated as the waters are moved, e.g. if another water moves closer to the green point...

(d)...then it swaps its identity with the green water, and it then becomes the green water.

Figure 6.1 Illustration of the identity constraint algorithm used to constrain the identity of a cluster of water molecules (two water molecules in this case) within a box of bulk water.

In this way, the identity of the cluster is maintained, as water molecules that diffuse away from the cluster are replaced by water molecules that diffuse into it. Note that the identity constraint only affects the labelling of the water molecules. The molecular coordinates are unchanged, no forces or restraints are required, and the thermodynamics of the system is unaffected. The WSRC uses the identity constraint to constrain the identity of a cluster of water molecules that occupy the same volume as the ligand. It achieves this by using two separate periodic boxes (Figure 6.2(a));

1. The bound protein-ligand complex solvated in a periodic box of water. This is called the protein-box.

2. A periodic box of water that is centered on the coordinates of the ligand. This is called the bulk-water-box.

The protein- and bulk-water-boxes are then connected together by attaching them to the same thermostat and barostat (Figure 6.2(b)). This has the effect of putting both boxes into the same heat bath, and ensures that both boxes are in thermal equilibrium with one another (i.e. they both have the same temperature and pressure). While the identity constraint can be used in molecular dynamics simulations [126], to simplify its implementation, here, we have developed the water-swap coordinate to use Monte Carlo [128] (MC) sampling. Connecting the boxes to the same thermostat and barostat is thus achieved easily by using the total energy of both boxes, and total volume of both boxes in the set of MC acceptance tests used during the simulation. With the two boxes connected together, the next step is to add identity points onto a selection of the atoms of the ligand. These identity points move with the atoms of the ligand, and are used to identify a cluster of water molecules in the bulk water box that occupies roughly the same volume as the ligand (Figure 6.2(c)). As the bulk-water-box is centered on the ligand, the identified water cluster has about the same cartesian coordinates as the ligand (in Figure 6.2 the boxes are drawn side-by-side only for clarity. In reality they lie directly on top of each other). The choice of atoms is arbitrary, with the aim being to pick a subset of atoms that allows a cluster to be identified which has roughly the same shape and volume as the ligand.
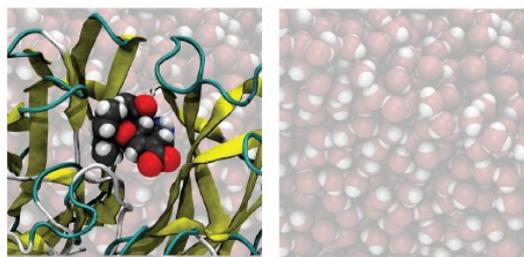
(a)The water-swap coordinate requires two
periodic boxes, a protein-box (left) and a
bulk-water-box (right).

T = 298 K, P = 1 atm
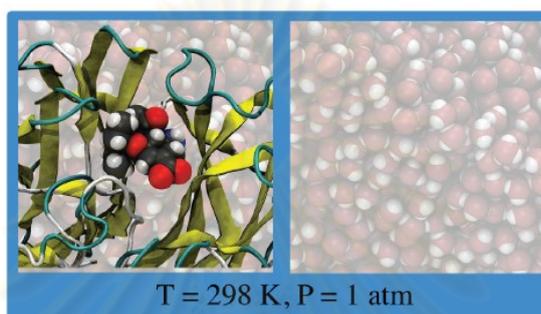
(b)These two independent boxes are connected
to the same thermostat and barostat. This has
the effect of putting both boxes into the same
heat bath, here represented using a blue
rectangle.

T = 298 K, P = 1 atm
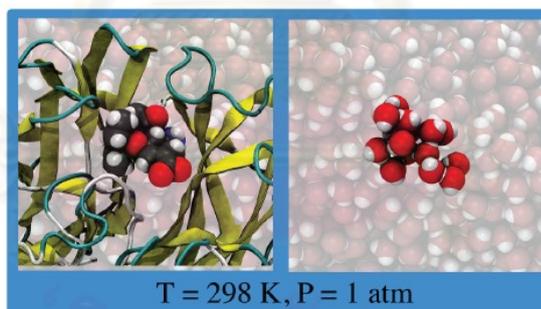
(c)Identity points are added to the atoms of the
ligand such that a cluster of waters in the
bulk-water box can be identified that occupies
roughly the same shape and volume as the
ligand.

Figure 6.2 The procedure necessary to construct the connected protein- and bulk-
water-boxes needed for the water-swap coordinate. Note that these two boxes are both
centered around the ligand, so actually lie directly on top of each other. They are only
shown next to each other here for clarity.

Finally, we now create a dual-topology [104, 109, 129] reaction coordinate, λ,
that switches off the ligand in the protein-box and switches off the water cluster in the
bulk-water-box, while simultaneously switching on the ligand in the bulk-water-box

and switching on the water cluster in the protein-box (Figure 6.3). This is the water-swap reaction coordinate (WSRC), which acts to transfer the ligand from being bound to the protein (at $\lambda = 0$) to the ligand being solvated in bulk water (at $\lambda = 1$), with a corresponding volume of water transferred back from bulk to the protein binding site. The total energy of the system, $H(\lambda)$, is calculated via:

$$H(\lambda) = H_{protein} + H_{bulk} + H_{ligand} + H_{cluster} +$$
$$(1 - \lambda)\,(H_{ligand:protein} + H_{cluster:bulk}) +$$
$$\lambda\,(H_{ligand:bulk} + H_{cluster:protein}) \tag{1}$$

where $H_{protein}$ is the energy of all of the molecules in the protein-box, H_{ligand} is the energy of the ligand, H_{bulk} is the energy of all of the waters in the bulk-water-box, $H_{cluster}$ is the energy of the water molecules identified in the cluster, $H_{ligand:protein}$ is the interaction energy between the ligand and all of the molecules in the protein-box, $H_{ligand:bulk}$ is the interaction energy between the ligand and all of the water molecules in the bulk-water-box, $H_{cluster:protein}$ is the interaction energy between the cluster waters and all of the molecules in the protein-box, and $H_{cluster:bulk}$ is the interaction energy between the cluster waters and all of the water molecules in the bulk-water-box. This energy function allows there to be just one copy of the ligand, which exists simultaneously in both the protein-box and the bulk-water-box. The action of λ can be viewed as "switching off" the ligand in the protein-box, while at the same time, it is being "switched on" in the bulk-water-box (while simultaneously, a water-cluster is switched off in the bulk-water-box and switched on in the protein-box). From the perspective of the ligand, this is equivalent to sampling the ligand while the protein-box is being switched off, and the bulk-water-box is being switched on. Because free energy is a state function, the result of this non-physical perturbation is exactly equivalent to pulling the ligand out of the protein binding site, dragging the ligand through an infinite volume of water, until it is an infinite distance away from the protein, while simultaneously a corresponding volume of bulk water is transported from an infinite distance from the protein, and pushed back into the protein binding site. This is the free energy change associated with the unbinding process of a single ligand from a single protein within an infinite box of water, and thus the difference between $\lambda = 0$ and $\lambda = 1$ calculated from the WSRC can be used directly to obtain the absolute binding free energy.
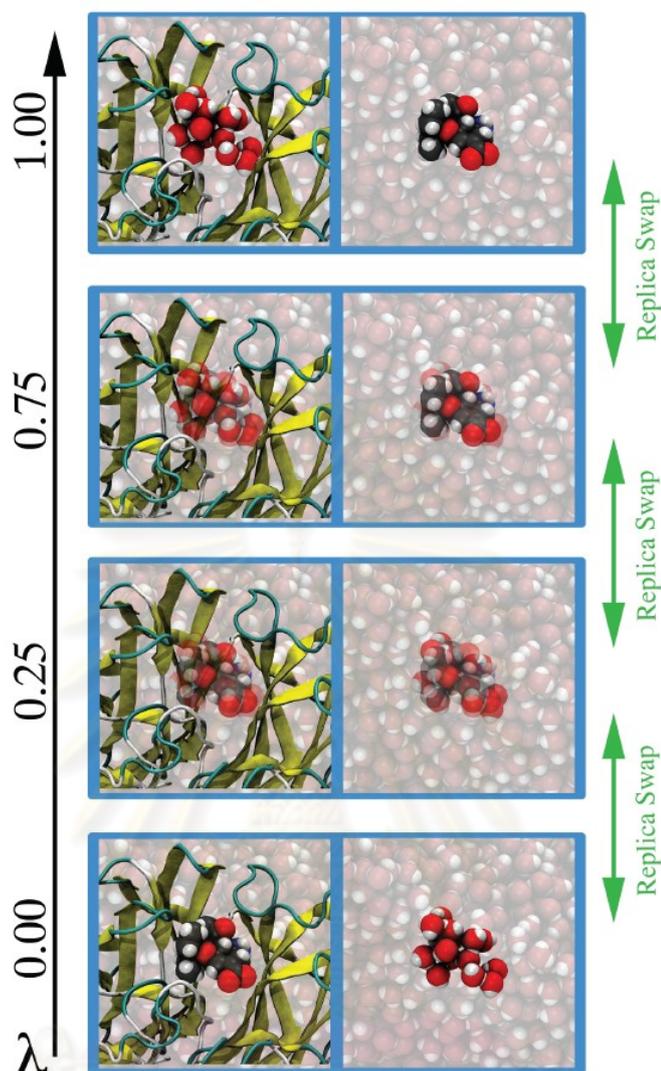
Figure 6.3 Illustration of the water-swap reaction coordinate (λ) used to exchange a ligand bound to a protein with an equivalent volume of bulk water. It acts to "switch off" the ligand in the protein-box (left), "switch off" the water cluster from the bulk-water-box (right), while simultaneously "switching on" the ligand in bulk-water-box and "switching on" the water cluster in the protein-box. The RETI method is used to calculate the free energy along this reaction coordinate. Replicas of the system are distributed across λ (only four replicas are shown here, normally many more are used). The free energy gradient with respect to λ is calculated for each replica. Replica exchange moves are periodically attempted between neighboring replicas, thereby giving each one the freedom to explore the entire water-swap reaction coordinate. At λ = 0 the ligand is bound to the protein, with the water cluster free in bulk (shown at the bottom, with a solid ligand in the protein-box and a solid water cluster in the bulk-water-box). At λ = 1 the ligand is free in bulk water, while the

water cluster is bound to the protein (shown at the top, with a solid ligand in the bulk-water-box and a solid water cluster in the protein-box). Intermediate λ values represent a non-physical combination of the ligand and water cluster existing simultaneously in both the protein- and bulk-water-boxes (shown in the middle with both a transparent ligand and transparent water cluster occupying both boxes).

Many methods exist to calculate free energy changes along reaction coordinates, e.g. FEP [102-105] and TI [106-108] One of the most efficient methods is replica exchange thermodynamic integration (RETI) [130, 131]. This method is simply TI, or finite-difference TI (FDTI) [121, 132, 133], with the addition of replica exchange moves along the free energy reaction coordinate. To use RETI with the water-swap coordinate (λ), we create a set of replicas across λ and use FDTI to calculate the gradient of the free energy with respect to λ at each replica (Figure 6.3). These gradients are integrated to return the potential of mean force (PMF) along the water-swap reaction coordinate. During the simulation, replica exchange moves are performed periodically between neighboring replicas and accepted or rejected using a Monte Carlo test [130]. The effect is to allow each replica to travel fully across the water-swap coordinate, thereby allowing each connected protein-box / bulk-water-box simulation to sample the full range of states between the ligand-bound and ligand-free end-points. This enhances sampling within each replica, and also ensures that the effects of any rare conformations sampled in individual replicas are included in the free energy gradients calculated at each point along the water-swap reaction coordinate. The water-swap coordinate is a dual-topology reaction coordinate that swaps a ligand with an equivalent volume of bulk water. Dual-topology reaction coordinates can be problematic as they allow for the possibility of overlap between the atoms that are being switched on and off, and the atoms of other molecules in the system [121]. The LJ potential becomes highly repulsive as atoms overlap, and both the Coulomb and LJ potentials have singularities at $r_{ij} = 0$ (where r_{ij} is the interatomic distance between atoms i and j). This leads to large energies of overlap, which can significantly increase the random error in the free energy differences or free energy gradients calculated using FEP or TI. This is a particular problem at the end-points of a perturbation, where atoms are fully switched on or off [121]. Soft-core potentials [134, 135] have been developed to overcome this problem. They work by changing the form of the Coulomb and LJ potentials so that, as the molecule being perturbed is

turned off, its atoms become softer, thereby reducing the energy penalty of overlap. Several different soft-core potentials are available. After initial experimentation, we chose and adapted the following soft-core potential [110, 134] for use with the water-swap coordinate,

$$E_{\text{coulomb}} = \frac{(1-\alpha)^n q_i q_j}{4\pi\epsilon_0 \sqrt{\alpha + r_{ij}^2}},$$
(2)

$$E_{\text{LJ}} = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}^{12}}{(\alpha\delta\sigma_{ij} + r_{ij}^2)^6} \right) - \left(\frac{\sigma_{ij}^6}{(\alpha\delta\sigma_{ij} + r_{ij}^2)^3} \right) \right]$$
(3)

where q_i and q_j are the charges on the atoms, σ_{ij} and ε_{ij} are the standard LJ parameters for atoms i and j, and ε_0 is the permittivity of free space. n and δ are the soft-core parameters, which control the shape of the softened potentials, while α is a variable that controls the amount of softening, e.g. at $\alpha = 0$ these equations are equivalent to the original (hard) Coulomb and LJ potentials, while as α is increased, so the potentials become increasingly non-physical and soft. We introduced an extra parameter, α_{scl} to control the maximum degree of softening. For the WSRC, $\alpha = \alpha_{scl}\lambda$ for the interactions between the ligand and the protein-box, and for the interactions between the water cluster and the bulk-water-box, while $\alpha = \alpha_{scl}(1-\lambda)$ for the interactions between the ligand and the bulk-water-box and for the interactions between the water cluster and the protein box.

6.4 VALIDATION AND DEVELOPMENT

Before we could apply the water-swap reaction coordinate to protein-ligand systems, it was important to investigate its behavior in a calculation where we knew the correct answer. We therefore set up a simulation which transferred a ligand from one bulk water simulation box to another. The free energy for this transfer (ΔG) should be zero. Running this calculation therefore allowed for the method to be validated and for the soft-core Coulomb and Lennard-Jones parameters in equations 2 and 3 to be calibrated. Our ideal was to generate a PMF where the curvature was low, any peaks or troughs were small, and the predicted total ΔG was zero. The ligand chosen for this validation was oseltamivir (brand name Tamiflu, Figure 6.4).
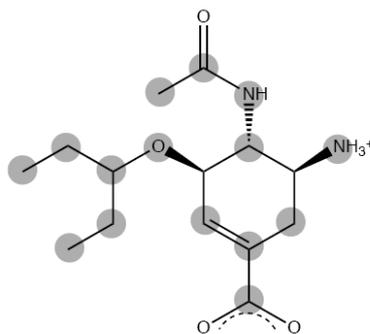
Figure 6.4 The chemical structure of the neuraminidase ligand, oseltamivir. The 15 atoms used as the identity points to identify the 15 water molecules in the cluster are highlighted using grey circles. The identity points are attached to these atoms, and move with these atoms during the simulation.

The coordinates for oseltamivir were obtained from the crystal structure with PDB code: 2HU4 [23]. Amber GAFF [136] parameters were obtained using Antechamber [137]. Ligand charges were generated using the RESP module of AMBER 9 [138], with HF/6-31G* electrostatic potentials calculated using Gaussian03 [48]. The ligand was placed in a periodic water box of 1665 TIP4P [51] water molecules, and equilibrated. A 12 Å electrostatic and non-bonded cutoff was used, with interactions harmonically feathered over the last 0.5 Å [139]. A bulk-water box of 1680 TIP4P water molecules was then connected to the same thermostat and barostat, and, using 15 identity points on the atoms of the ligand (see Figure 6.4), a cluster of 15 water molecules was identified in this bulk-water box (Figure 6.5). The number and position of the identity points were chosen such that the identified water cluster occupied roughly the same shape and volume as the ligand. A water-swap coordinate was then constructed to swap the ligand with this water cluster, with RETI used to calculate the PMF (Figure 6.5). 24 replicas were used across this coordinate (0.001, 0.02, 0.05, then every 0.05 to 0.95, then 0.975, 0.99 and 0.999), with 30 million (M) MC moves with a temperature of 298.15 K and pressure of 1 atm performed on each replica. Preferential sampling [140, 141] was employed to enhance sampling of all of the water molecules, focused around oseltamivir, and replica exchange moves were attempted between neighboring replicas every 100 thousand (k) moves. Soft-core parameters $n = 1$, $\delta = 1$ and $\alpha_{scl} = 1$ were used. The free energy along the WSRC was calculated using FDTI, with a value of $\Delta\lambda$ of 0.001. To estimate and control errors, double-wide sampling [142] was used to generate both a forwards-

difference and backwards-difference estimate of $\Delta G/\Delta \lambda$ with ΔG calculated using the Zwanzig equation [102]. Both a forwards and a backwards difference PMF were generated, and we report here the average of the two. To estimate the error on the calculation, the difference between the forwards and backwards PMFs was monitored, and, in addition, the standard errors on the average free energy gradients for each value of λ were calculated at the 95% confidence level, and then integrated across λ. The size of this error was used, post-hoc, to divide the 30 M MC steps into equilibration and production, by calculating the free energy average and standard error for every 100 k steps moving back from the end of the trajectory, and splitting the trajectory into equilibration and production at the point where this error was minimised. This allowed for automatic post-processing of the results. We visually confirmed that this point corresponded to the end of equilibration by plotting the free energy calculated for blocks of 100 k steps and checking that this value had stabilised. The WSRC was implemented in, and simulations performed using our molecular simulation framework, Sire [143, 144]; the program and all input files necessary to run these simulations are available for download at http://siremol.org/waterswap.

The calculated PMF is shown in Figure 6.6. The difference in free energy between the end-points of the water-swap coordinate was -0.8 kcal mol^{-1}, with standard error 1.5 kcal mol^{-1}. Within error, this is zero, as hoped. However, as is obvious, the PMF is far from ideal. The peak at $\lambda = 0.5$ is greater than 100 kcal mol^{-1}! It is surprising, and indeed encouraging, that the RETI calculation predicts a free energy difference of zero despite such a large total change in free energy across the reaction coordinate. However, this result demonstrates that refinement of the soft-core parameters was required to allow the approach of the WSRC to be applied to calculate absolute protein-ligand free energies. This refinement is described below.
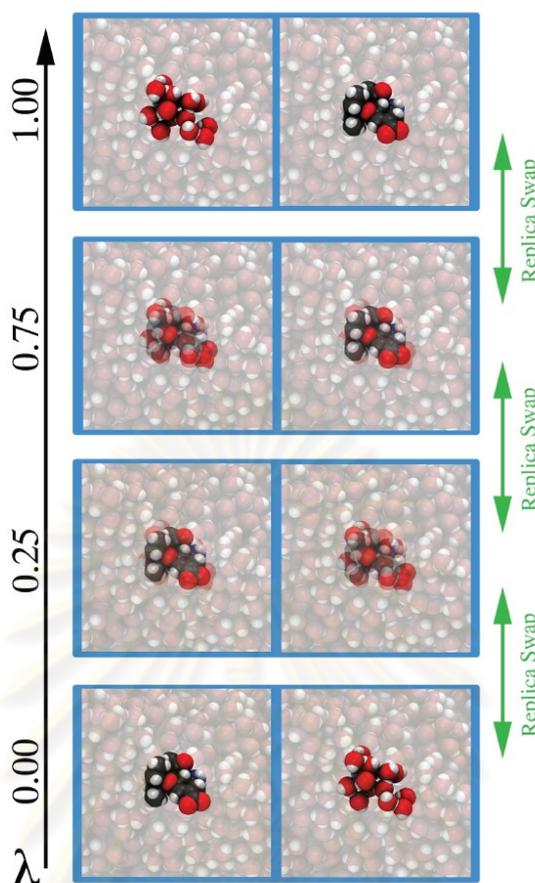
Figure 6.5 The system used to test the water-swap reaction coordinate. The ligand in the left bulk-water-box is swapped with an equivalent volume of water in the right bulk-water-box. The λ coordinate is used to switch off the ligand in the left box, switch on the water cluster in the right box, while simultaneously switching on the ligand in the right box and switching on the water cluster in the left box. The overall free energy change for this process should be zero.
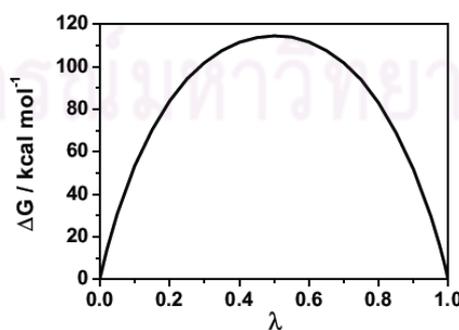
Figure 6.6 The initial PMF along the water-swap reaction coordinate of transferring oseltamivir between two water boxes. This PMF was generated using non-optimized soft-core parameters $n = 1$, $\delta = 1$ and $\alpha_{scl} = 1$.

6.4.1. Soft-core parameter investigation

To discover the source of the large peak, we next performed a systematic investigation of the effect of each soft-core parameter on the shape of the PMF. To speed up the investigation, only 20 M steps of MC sampling was performed for each replica. For the soft-core Coulomb parameters, it can be seen clearly in Figure 6.7(a) that the optimal values of parameter n in equation 2 are either 0 or 0.25. These smoothly reduce the peak down from 110 kcal mol^{-1}, to 35 kcal mol^{-1}. To reduce this further, the Coulomb interaction energy term between cluster water molecules was scaled using an extra function, $S_{cluster}$ (equation 4) to control the interaction between the molecules in the water cluster.

$$S_{cluster} = S - 4(S - 1)(\lambda - 0.5)^2 \qquad (4)$$

This is a simple function that is equal to 1 at $\lambda = 0$ and $\lambda = 1$, and is equal to S at $\lambda = 0.5$. This function was used to strengthen the electrostatic interactions between cluster water molecules at $\lambda = 0.5$ to compensate for the loss of hydrogen bonds to bulk solvent. The effect of modifying the interactions is shown in Figure 6.7(b). It is clear from this graph that a value of 1.5 is optimal, because it brings the PMF down while simultaneously attening it between $\lambda = 0.2$ and $\lambda = 0.8$. The parameter δ in equation 3 is applied to shift the Lennard-Jones potential [110]. The effect of changing δ is shown in Figure 6.7(c). Values of δ of 1.5 to 1.75 can be seen to be most suitable. When the δ parameter is increased to 2.0, the soft-core potential is over-softened, which causes a dip in the PMF at $\lambda = 0.5$. The α_{scl} parameter changes the maximum range of α, which affects the softening of both the Coulomb and LJ terms. The effect of changing α_{scl} is shown in Figure 6.7(d). Values of α_{scl} above one increase the peak of the PMF. Values less than one therefore appear to be ideal. However, α_{scl} is used to control the maximum degree of softening, and therefore values less than one reduce the amount of softening available. This strongly affects the end-points of the PMF (see Figure 6.8). As α_{scl} is reduced, the PMF becomes increasingly discontinuous. The reason for this becomes clear when α_{scl} is set to zero. Using α_{scl} equal to zero disables the soft-core potential, and is equivalent to running the simulation with no softening. Overlap between the atoms of the ligand and cluster waters, with the atoms in the protein- and bulk-water-boxes, leads to large energies, and large changes in the gradients of the free energy with respect to λ. This leads to discontinuities in the PMF at the end-points. It is clear that these discontinuities are

smallest for $\alpha_{scl} = 1$, thus leading to the conclusion that this is the optimal value of this parameter.
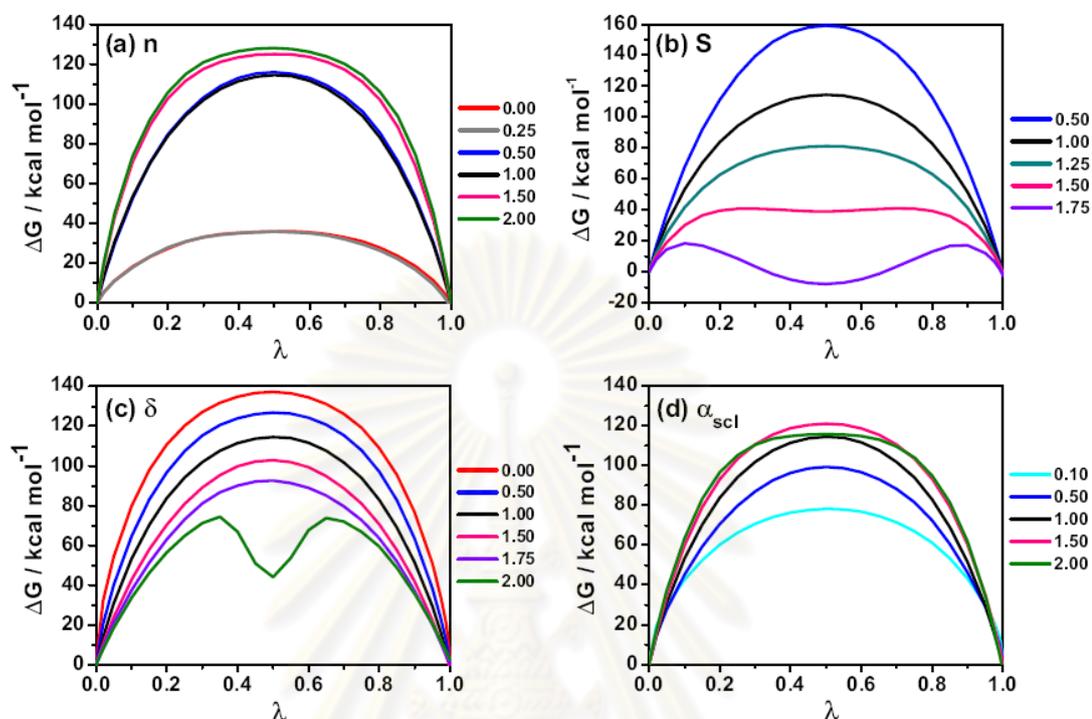
Figure 6.7 The effect on the PMF of transferring oseltamivir between two water boxes, of changing the soft-core parameters n, S, δ and α_{scl} (see equations 2, 3 and 4 for the meaning of these symbols).
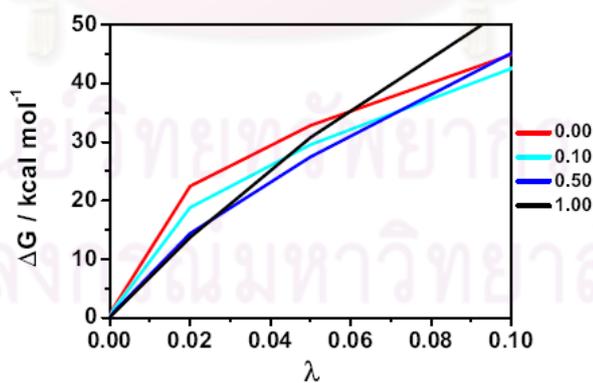
Figure 6.8 The effect on the PMF, near $\lambda = 0$, of decreasing the soft-core parameters α_{scl} from one to zero. A value of $\alpha_{scl} = 0$ turns off softening, and is equivalent to not using a soft-core Coulomb or LJ potential (*i.e.* using just the standard Coulomb and LJ potentials with no softening). An almost identical effect is seen at $\lambda = 1$.

6.4.2. Soft-core parameter optimisation

The next stage in optimising the PMF was to combine and balance the values of the soft-core Coulomb and LJ parameters. As a first step, a Coulomb parameter $n = 0$ was selected, and then used in combination with the other parameters. The results are shown in Figures 6.9(a)-(c). These graphs show that values of S in the range of 1.1 to 1.25, and values of δ in the range of 1.1 to 1.2 were reasonable. This is because these values worked to bring down the peak in the PMF, without the peak inverting and becoming a significant dip at $\lambda = 0.5$. For α_{scl}, values less than one lead to an increase in the peak of the PMF. This is unsurprising, given what was observed in Figure 6.8. Because of this, we decided to set α_{scl} to one for all subsequent simulations.

For the final stage of optimisation, two values of the soft-core Coulomb parameter, n ($n = 0$ and $n = 0.25$), were tested in combination with values of S and δ in the range $S = 1.1 - 1.25$ and $\delta = 1.1 - 1.2$. The resulting PMFs are shown in Figure 6.10. Three combinations of parameters were found to give suitable PMFs, with peaks in range 1 - 10 kcal mol^{-1}, and these were selected for testing in the protein-ligand system. The optimal value sets for the soft-core Coulomb and LJ parameters are summarized in Table 6.1. Within error, all of the optimised parameters returned a free energy change of zero. While changing the soft-core parameters does not affect the free energy change, it does affect the rate of convergence of the calculation. The original, set A and set C parameters resulted in simulations that converged within 5 M steps, while simulations with the set B equilibrated more slowly. Indeed, the set B simulation had not converged after the full 20 M MC moves, and so an extra 20 M steps had to be performed. Equilibration with set B occurred around 25 M steps.
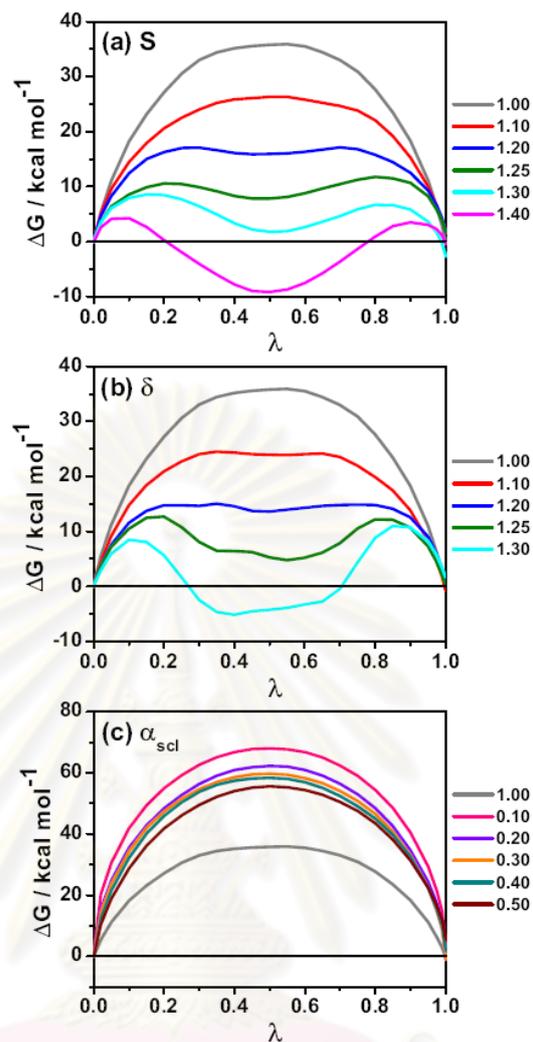
Figure 6.9 The PMFs generated by varying the soft-core parameters (a) S, (b) δ and (c) α_{scl} while keeping the Coulomb soft-core parameter, n, equal to zero.
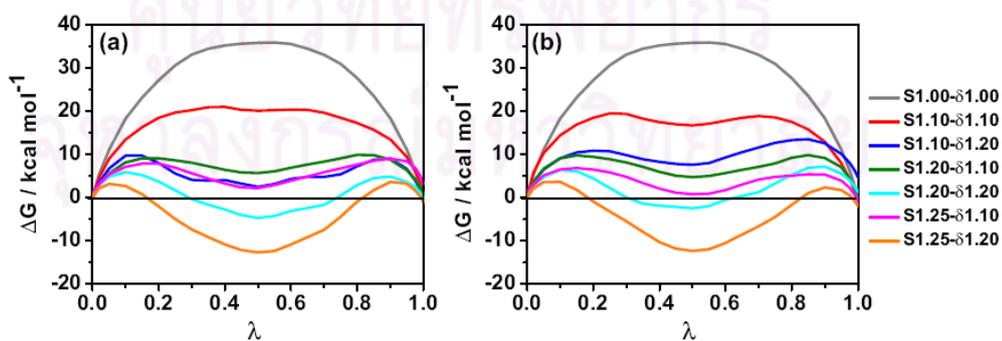
Figure 6.10 The PMFs generated by combining soft-core parameters S from the range 1.1 to 1.25 and δ from the range 1.1 to 1.2, with two values of the Coulomb soft-core parameter (a) $n = 0$ and (b) $n = 0.25$.

Table 6.1 The values of the soft-core parameters from the simulation used to calculate the original PMF (Figure 6.6), and the three sets of soft-core parameters derived to optimise the PMF. The free energy difference between the end-points for each parameter set is shown, complete with two error estimates; the difference between the forwards and backwards free energy, shown in round brackets, and the standard error on the free energy average, at the 95% confidence level, shown in square brackets. Set B was found to equilibrate slowly, so this simulation used 40 M steps per replica, as compared to 20 M steps for sets A and C, and 30 M steps for the original simulation.

Parameter	n	S	δ	α_{scl}	$\Delta G/\text{kcal mol}^{-1}$
Original	1.00	1.00	1.00	1.00	-0.8 (0.0) [1.5]
Set A	0.00	1.10	1.20	1.00	-0.4 (1.9) [3.1]
Set B	0.00	1.25	1.10	1.00	-1.8 (1.5) [1.9]
Set C	0.25	1.25	1.10	1.00	-0.8 (1.4) [1.7]

6.5 APPLICATION TO A PROTEIN SYSTEM

The optimized soft-core Coulomb and Lennard-Jones parameters were used to predict protein-ligand binding free energies of the influenza neuraminidase (NA) inhibitor, oseltamivir (brand name Tamiflu), to NA. NA is a membrane glycoprotein that plays an essential role in influenza virus replication by cleaving the terminal sialic acid from the receptors to release the progeny viruses [17]. Currently, oseltamivir is the only oral-active agent against NA function. Due to the conserved binding site of this protein in every subtype, NA has become a potent target for drug development. We chose this system as a test case because the binding site was relatively rigid and conserved, and computational methods have previously been used successfully on this system [8, 145]. The initial structure of NA was taken from the protein data bank (PDB) code: 2BAT [22]. All missing hydrogen atoms were added using the LEaP module in AMBER 9 [138], with the predicted ionization states calculated using PROPKA [71, 146]. The system was solvated with a TIP3P [51] water box, and then energy minimized using the SANDER module of AMBER 9 [138]. Simulated annealing was carried out with the temperature raised from 10 K to

350 K in seven steps of 20 ps interval. NVT molecular dynamics (MD) equilibration
was performed for 160 ps at 300 K. This was followed by 500 ps of NPT MD
simulation at 300 K and 1 atm pressure. Equilibration was ensured by monitoring
fluctuations in RMSD and total energy. The final structure was used as the starting
point for the Monte Carlo (MC) simulations. The TIP3P water box was converted into
a TIP4P [51] water box to improve the modelling of electrostatic interactions in the
MC simulations [147]. To perform the free energy calculation for oseltamivir and NA,
the oseltamivir coordinates were obtained from the crystal structure with PDB code:
2HU4 [23] and replaced at the NA active site. 24 replicas across the water-swap
coordinate were used, with the same λ values as for the simulations that transferred
oseltamivir between two water boxes. 40 M MC moves were attempted for each
replica, with preferential sampling [140, 141] used to enhance sampling around
oseltamivir, and with replica exchange moves attempted every 100 k steps. A 12 Å
electrostatic and non-bonded cutoff was used, with a harmonic feather applied over
the last 0.5 Å [139]. During the simulation, the backbone of NA was fixed, with only
the side chain motion of the residues sampled. Free energy averages and errors were
calculated in the same way as for the simulations that transferred oseltamivir between
two water boxes. The simulations were performed using Sire [143, 144]; the program
and all input files necessary to run these simulations are available for download at
http://siremol.org/waterswap. PMFs are shown in Figure 6.11 and the calculated
binding free energies are shown in Table 6.2. Note that the binding free energies
($\Delta G_{binding}$) are the negative of the difference in the PMF between $\lambda = 0$ and $\lambda = 1$,
because the WSRC models unbinding. The predicted absolute binding free energies
show reasonable agreement with experimental values, which is encouraging. It should
be noted, however, that agreement with experiment is fortuitous. Exact agreement
with experiment should not be expected as the potential function does not include
important physical effects, such as protein backbone exibility and ligand polarization
[149]. Use of a quantum mechanics (QM) model of the ligand, and calculating a
QM/MM free energy [144] may be one route to including the effect of polarisation.
The results demonstrate that the optimised soft-core parameters give PMFs that are
generally smooth, with no large peaks. Set A is the best of the three, as it does not
contain the dip between $\lambda = 0.3$ and $\lambda = 0.6$. However, all three PMFs could be
improved further by smoothing the large changes in free energy near the end-points. It

is these that are likely to be the source of the large random errors (1.5-2.0 kcal mol^{-1}), and the spread of values from these three simulations. These errors could be reduced by increasing the number of replicas at the end-points, and by increasing the number of MC moves per replica. Overall, these results demonstrate that the WSRC, when combined with balanced soft-core parameters, is a viable, and potentially powerful new route for the calculation of absolute protein-ligand binding free energies.
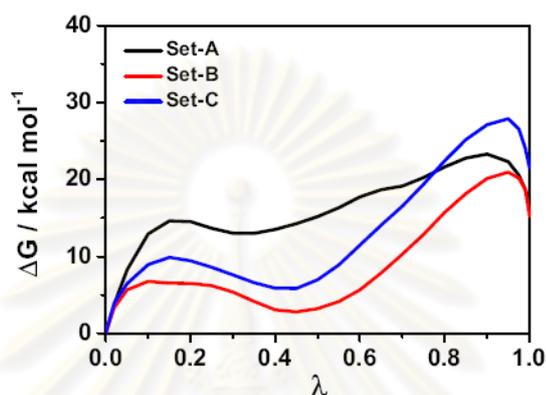
Figure 6.11 The potential of mean force along the water-swap reaction coordinate for the unbinding of oseltamivir from NA. Three PMFs are shown, for the three different sets of optimised soft-core parameters (given in Table 6.1).

Table 6.2 The predicted absolute binding free energies of oseltamivir to neuraminidase as calculated using the three different sets of optimised soft-core parameters. For comparison, reported experimental measures [25] of the binding free energy vary between -12.0 kcal mol^{-1} and -15.0 kcal mol^{-1} (as calculated from $IC_{50}s$ using [7, 73, 148] $\Delta G_{binding}$ = RT ln IC_{50}).

Parameter	$\Delta G_{binding}$ (kcal mol^{-1})
Set A	-17.0 (1.7) [1.6]
Set B	-15.2 (1.4) [2.0]
Set C	-21.5 (1.4) [1.5]

6.6 FUTURE WORK

This paper presents the initial development and optimisation of the water-swap reaction coordinate. Several questions about its implementation still remain to be answered. For example, it is necessary to investigate whether this coordinate would be useful for other protein-ligand systems, and whether or not the optimised soft-core parameters are transferable. It is also necessary to investigate the sensitivity of the calculated free energy to the number of waters swapped with the ligand. We have performed some initial experiments that investigate the sensitivity of the free energy to the number of swapped waters, and these suggest that even 12 or 18 waters could have been swapped in place of oseltamivir. However, a detailed investigation of the affect of both the number and position of the identity points used to identify the swapped water cluster needs to be performed. Further development of the method is also required to allow it to be applied to a wide range of protein-ligand systems. While we believe the WSRC will perform well for proteins with solvent-exposed binding sites, it would not perform well for proteins with buried binding sites. This is because the amount of water to swap in place of the ligand is unclear. One way to overcome this problem would be to supplement the set of MC moves used during the simulation with some Gibbs-ensemble [150, 151] type moves, in which randomly selected water molecules are deleted from one box and added to the other. By adding these moves, both boxes would then be in complete thermal equilibrium, having both the same temperature, pressure and chemical potential. Any inbalance in the amount of water swapped would then be corrected automatically, e.g. if too many waters were being swapped, then the Gibbs-ensemble moves would automatically transfer waters back, so as to balance the chemical potentials of the two boxes. While Gibbs-ensemble water creation and deletion moves would have a low probability of being accepted at ambient temperature, there are algorithms, such as the continuous fractional component move [152] which can be applied to increase the acceptance ratio, and that could be used with the WSRC.

Finally, it should be noted that the development of the WSRC, while highly promising, has not, in itself solved the problem of accurately predicting absolute protein-ligand binding affinities. The binding affinity calculated using the WSRC is actually just the binding affinity for a single binding mode and binding orientation of the ligand, and it corresponds to the binding free energy in an infinitely dilute

solution. This is because the protein- and bulk-water-boxes are effectively windows into a single infinite water box, which contains just a single protein and single ligand. The WSRC has calculated the free energy of moving the ligand from a specific binding mode and orientation in the protein on one side of this infinite water box, to a point that is an infinite distance away, on the other side of this box. It is still an open question to how this relates to the standard state absolute binding free energy [153]. Moreover, the WSRC does little to enhance sampling of the protein-ligand system, will probably perform poorly if the protein undergoes large conformational changes upon ligand binding, and it will not overcome any errors in the parameterisation of the ligand. It will also not solve problems caused by any lack of polarization [149] in the forcefield used for the protein-ligand binding prediction, any deficiencies in the water model to adapt to the different environments in the protein binding site, and it will not correct errors caused by selecting the wrong titration states of the protein residues, or wrong tautomer or binding mode of the ligand. There are many problems that must be overcome before the grand challenge of predicting protein-ligand binding can be considered to be solved. We believe that the WSRC solves one of the main problems, that of defining the optimal reaction coordinate to predict binding. We leave the solution of the other problems as challenges for future work.

6.7 CONCLUSIONS

We have developed a water-swap reaction coordinate (WSRC), which allows for the absolute binding free energy of a protein-ligand complex to be calculated directly from a single simulation. To validate the method, we have performed simulations that swapped a ligand between two bulk water boxes, and then used these simulations to optimize soft-core parameters such that the change in free energy along the reaction coordinate was smooth, with a small total difference in free energy. The optimized parameters were used to predict the protein-ligand binding free energy of NA and oseltamivir. The agreement of the results with experiment is encouraging. The final results show that the WSRC provides a promising new direction in the search for the solution of the challenge of predicting protein-ligand binding.

CHAPTER VII

CONCLUSIONS

7.1 CONCLUSIONS

Several techniques of computational chemistry were used to study structural properties, drug-target interactions and substrate binding of neuraminidase (NA) in influenza virus. The study covers both NA wild-type and mutant strains in H5N1 and H1N1-2009. In addition, the water-swap reaction coordinate (WSRC) which provides a promising route to predict absolute protein-ligand binding free energies, was newly developed.

In drug-target interaction part, the dynamic nature of the three available drugs, oseltamivir, zanamivir and peramivir, complexed with N1 was studied using molecular dynamics (MD) simulations. Structural differences of the substitution groups at the main ring scaffold were observed. The carboxylate and guanidinium groups of peramivir were seen to form many more hydrogen bonds with the surrounding residues in comparison with the other two drugs. Moreover, the D151 residue located in the 150-loop region interacted strongly via the –OH group of peramivir in a similar way to sialic acid. For the bulky substitution group, direct hydrogen bonds were detected only with the hydrophilic side chain of zanamivir. Due to the bulky substitution in oseltamivir and peramivir, the flexibility and steric effects of the hydrophobic group led to rearrangement of the surrounding residues, E276 and R224, to form a hydrophobic pocket for this substitution. While the bulky group of peramivir shows only one preferential conformation, the equivalently ended branches of oseltamivir does not fit well into the active site of N1 due to the rotation of this side chain. This leads to the lower efficacy of oseltamivir against N1 strain with respect to the lowest binding affinity calculated by the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method in comparison with the other drugs.

Consequently, the source of oseltamivir resistance caused by the H274Y mutation was observed to be the reduction of the hydrophobic pocket size. The phenol ring of the mutated Y274 was found to drive the next residue, E276, toward an ethyl moiety of oseltamivir's bulky group. The rotation of E276 side chain decreased the space to accommodate the bulky hydrophobic group of oseltamivir. In addition, lower hydrogen bond interactions between Y274 and the framework residues reduced the

stability of the hydrophobic pocket and subsequently decreased the capacity of the oseltamivir binding to the N1 active site. The H274Y mutation led to a decrease in the $\Delta G_{binding}$ from -14.6 ± 4.3 kcal mol^{-1} to -9.9 ± 6.4 kcal mol^{-1}. Within the large random error in the MM/PBSA calculation, the accurate prediction of binding free energy for the novel pandemic of H1N1-2009 subtype could not be obtained. Therefore, the linear interaction energy (LIE) method, which has higher precision and accuracy, was applied to predict oseltamivir susceptibility against the H1N1-2009 strain. The probable H1N1-2009 mutant strains (R292K, E119V, H274Y and N294S) were modelled. The calculated efficiency of oseltamivir towards the mutants are in the following order: wild-type (-12.8 ± 0.9 kcal mol^{-1}) > N294S (-10.4 ± 0.9 kcal mol^{-1}) > H274Y (-9.8 ± 1.0 kcal mol^{-1}) > E119V (-9.3 ± 0.8 kcal mol^{-1}) > R292K (-7.7 ± 0.7 kcal mol^{-1}). Reduction in oseltamivir-NA interaction energies was observed in terms of lower hydrogen bonds, electrostatic and van der Waals interactions relative to the wild-type H1N1-2009 strain. This implies that oseltamivir is less effective in treatment of patients with these probable mutants, especially with the high level of oseltamivir resistance with the R292K mutant.

To understand the first step of the cleavage mechanism of NA, the natural substrate in humans (SA-α-2,6-GAL) bound to different NA subtypes (N1-1918, N1-2005, N1-2009, N2-1967 and N8-1963) was studied using QM/MM MD simulations. The SA-α-2,6-GAL was found to change ring conformation from chair to twist-boat conformation when it bound to all systems except for an avian N8-1963 strain. This conformational change in both the N1 and N2 subtypes is stabilised by the N/Q347 and K431 residues via hydrogen bond interactions. This is not the case for the Y347 residue of the avian type in the N1-2005 and N8-1963 strains. The obtained results indicate that the residue at 347 position could be an essential residue to facilitate the conformational change of the natural substrate. The human virus has N/Q347 and K431 to bind with the galactose unit of the SA-α-2,6-GAL, making the suitable conformation of sialoside to be cleaved in the second step of the cleavage mechanism. This is in contrast for the avian virus where Y347 and E431 were not found to form hydrogen bond to the natural substrate in humans. Designing new inhibitors to interact with these residues could provide a new route to discover highly effective inhibitors that can fit better into the NA active site in both the wild-type and mutant strains.

To achieve a faster way to discover the new NA inhibitors, a new method to accurately predict the strength of binding between a ligand and a protein was developed using a single simulation. The water-swap reaction coordinate (WSRC) method was found to provide a promising new route to predict absolute protein-ligand binding affinities within a reasonable timescale. The main idea of the novel binding method is the use of the identity constraint method to identify a cluster of water molecules that would occupy the same volume as the ligand in the protein active site. A dual topology reaction coordinate (λ) is then applied to swap the ligand of the protein ligand complex, with the water cluster from bulk water using replica exchange thermodynamic integration (RETI). The free energy change of swapping the water cluster with the ligand is, directly, the absolute binding free energy. To improve the method, the soft-core Coulomb and Lennard-Jones parameters were optimised and applied for prediction of NA-oseltamivir binding free energies, with these results compared to experiment.

7.2 RESEARCH LIMITATIONS

Although, computational chemistry is a powerful tool to provide information at molecular level that would be useful for drug development, there are still problems due to the assumptions made in the calculations using current methods.

1. **System preparation:** due to the lack of X-ray structure for the initial structure in MD simulations, therefore, molecular modelling was commonly used to set up the protein-ligand complex systems. This would affect the reliability of the predicted results. To improve the quality of the research in a future study, multiple simulations are required to test the modelled systems.

2. **Running simulations:** because of limitated computer facilities, time-consuming simulations could not be performed to obtain the dynamical properties of the protein-ligand complexes over longer time periods. Moreover, the use of short residue-based cutoffs could affect in the results, longer non-bonded cutoffs may be required.

3. **Method assumptions:** this is a limitation of every current available program, including the WSRC method in Sire, caused by the lack of protein-ligand sampling. To calculate the binding free energy accurately, this methods need to be improved to cover all important protein-ligand conformation.

7.3 SUGGESTIONS FOR FUTURE WORK

Neuraminidase (NA) is still an important target for drug development against influenza virus. Although, NA inhibitors have been widely used for clinical treatment, the current available commercially agents are still problems due to drug resistance with several mutation strains. In this dissertation, the obtained results could provide the molecular-level knowledge for drug-target interactions in both wild-type and mutant strains for N1 subtype, H5N1 and H1N1-2009 influenza viruses. However, several questions about the NA protein still remain to be answered. Moreover, the new water-swap reaction coordinate (WSRC) method needs to be developed in order to get higher accurate prediction of absolute protein-ligand binding free energies. To improve the research with NA, there are many interesting projects to provide the other answers for discovering and designing the new potent NA inhibitors.

- Due to the current speed of computer, the extensive simulations for classical MD with longer time period or QM/MM MD with higher accurate method could be performed to achieve more accurate results, especially dynamic properties of NA system.

- The new interesting analogs of NA inhibitor are in development. This includes A-315675 [28], R-125489 [29], glycomonomers [30], Phospha-Oseltamivir (Tamiphosphor) monoester [31] and Sialoglycoconjugates analog [32]. Running MD simulations to investigate molecular properties of these new potent inhibitors properties could provide useful information to speed up the development process.

- There are many open questions in the other mutation residues in different NA subtypes. For instances, why NA mutant strains were found with subtype specific, H274Y occur only in N1 subtype while R292K and E119V take place in N2 subtype, and R152K and D198N in influenza B virus [3, 33, 37].

- In extensive from chapter V of this work which the first step of cleavage mechanism was studied using QM/MM MD simulations, it is also interesting to complete the process by studying the second step. In this case, the trajectories from this research can be then used as starting structure.

- Application of the WSRC method to calculate the binding free energy of different inhibitors, different NA strains as well as different protein-ligand systems would be challenges for future works.

REFERENCES

[1] von Itzstein, M. The war against influenza: discovery and development of sialidase inhibitors. Nat. Rev. Drug. Discov. 6 (2007) : 967-974.

[2] McKimm-Breschkin, J. L., Selleck, P. W., Usman, T, B., Johnson, M. A. Reduced sensitivity of influenza A (H5N1) to oseltamivir. Emerg. Infect. Dis. 13 (2007) : 1354-1357.

[3] Reece, P. A. Neuraminidase inhibitor resistance in influenza viruses. J. Med. Virol. 79 (2007) : 1577–1586.

[4] Kiso, M., Mitamur, K., Sakai-Tagawa, Y., Shiraishi, K., Kawakami, C., Kimura, K., Hayden, F. G., Sugaya, N., Kawaoka, Y. Resistant influenza A viruses in children treated with oseltamivir: descriptive study. Lancet 364 (2004) : 759–765.

[5] Le, Q. M., Kiso, M., Someya, K., Horimoto, T., Takada, A., Goto, H., Suzuki, T., Suzuki, Y., Kawaoka, Y. Avian flu: isolation of drug-resistant H5N1 virus. Nature 437 (2005) : 1108.

[6] Malaisree, M., Rungrotmongkol, T., Decha, P., Intharathep, P., Aruksakunwong, O., Hannongbua, S. Understanding of known drug-target interactions in the catalytic pocket of neuraminidase subtype N1. Proteins 71 (2008) : 1908-1918.

[7] Malaisree, M., Rungrotmongkol, T., Nunthaboot, N., Aruksakunwong, O., Intharathep, P., Decha, P., Sompornpisut, P., Hannongbua, S. Source of oseltamivir resistance in avian influenza H5N1 virus with the H274Y mutation. Amino Acids 37 (2009) : 725–732.

[8] Rungrotmongkol, T., Malaisree, M., Nunthaboot, N., Sompornpisut, P., Hannongbua, S. Molecular prediction of oseltamivir efficiency against probable influenza A (H1N1-2009) mutants: molecular modelling approach. Amino Acid 39 (2010) : 393-398.

[9] Malaisree, M., Woods, C. J., Rungrotmongkol, T., Nantaboot, N., vongachariya, A., Mulholland, A. J., Hannongbua, S. Substrate binding to neuraminidase of influenza viruses: QM/MM MD simulations. (2010) *In preparation.*

[10] Woods, C. J., Malaisree, M., Hannongbua, S., Mulholland. A. J. A water-swap reaction coordinate for the calculation of absolute protein-ligand binding free energies. (2010) *Submitted*.

[11] Avian influenza. Available from: http://www.who.int/csr/disease/avian_influenza/en [2010, September].

[12] de Jong, M. D., Hien. T. T. Avian influenza A (H5N1). J. Clin. Virol. 35 (2006) : 2-13.

[13] Pandemic (H1N1) 2009. Available from: http://www.who.int/csr/disease/ swineflu/en [2010, September].

[14] Operario, D. J., Moser, M. J., St George, K. Highly sensitive and quantitative detection of the H274Y oseltamivir resistance mutation in seasonal A/H1N1 influenza. J. Clin. Microbiol. (2010) *In press.*

[15] Baz, M., Abed, Y., Simon, P., Hamelin, M. E., Boivin, G. Effect of the neuraminidase mutation H274Y conferring resistance to oseltamivir on the replicative capacity and virulence of old and recent human influenza A(H1N1) viruses. J. Infect. Dis. 201 (2010) : 740-745.

[16] Horimoto, T., Kawaoka, Y. Pandemic threat posed by avian influenza a viruses. Clin. Microbiol. Rev. 14 (2001) : 129-149.

[17] De Clercq, E. D. Antiviral agents active against influenza A virus. Nature 5 (2006) : 1015-1025.

[18] Matrosovich, M. N., Gambaryan, A. S., Teneberg, S., Piskarev, V. E., Yamnikova, S. S., Lvov, D. K., Robertson, J. S., Karlsson, K. A. Avian influenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. Virology 23 (1997) : 224–234.

[19] Chandrasekaran, A., Srinivasan, A., Raman, R., Viswanathan, K., Raguram, S., Tumpey, T. M., Sasisekharan, V., Sasisekharan, R. Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. Nat. Biotechnol. 26 (2008) : 107–113.

[20] McKimm-Breschkin, J. L. Neuraminidase inhibitors for the treatment and prevention of influenza. Expert. Opin. Pharmacother. 3 (2002) : 103-112.

[21] Thompson, J. D., Higgins, D. G., Gibson, T. J. Improved sensitivity of profile searches through the use of sequence weights and gap excision. Comput. Appl. Biosci. 10 (1994) : 19-29.

[22] Varghese, J. N., McKimm-Breschkin, J. L., Caldwell, J. B., Kortt, A. A., Colman, P. M. The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. Proteins 14 (1992) : 327-332.

[23] Russell, R. J., Haire, L. F., Stevens, D. J., Collins, P. J., Lin, Y. P., Blackburn, G.M., Hay, A.J., Gamblin, S. J., Skehel, J. J. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature 44 (2006) : 45-49.

[24] Colman, P. M., Varghese, J. N., Laver, W. G. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature 303 (1983) : 41–44.

[25] Bantia, S., Parker, C. D., Ananth, S. L., Horn, L. L., Andrie, K., Chand, P., Kotian, P. L., Dehghani, A., Babu, Y. S. Comparison of the anti-influenza virus activity of RWJ-270201 with those of oseltamivir and zanamivir. Antimicrob. Agents Chemother. 45 (2001) : 1162-1167.

[26] Gubareva, L. V., Webster, R. G., Hayden, F. G. Comparison of the activities of zanamivir, oseltamivir, and RWJ-270201 against clinical isolates of influenza virus and neuraminidase inhibitor-resistant variants. Antimicrob. Agents Chemother. 45 (2001) : 3403-3408.

[27] Influenza (Flu) Antiviral drugs and related information. Available from: http://www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/ucm100228.htm [2010, September].

[28] Abed, Y., Nehmé, B., Baza, M., Boivin, G. Activity of the neuraminidase inhibitor A-315675 against oseltamivir-resistant influenza neuraminidases of N1 and N2 subtypes. Antiviral Res. 77 (2008) : 163-166.

[29] Koyama, K., Takahashi, M., Oitate, M., Nakai, N., Takakusa, H., Miura, S., Okazaki, O. CS-8958, a prodrug of the novel neuraminidase inhibitor R-125489, demonstrates a favorable long-retention profile in the mouse respiratory tract. Antimicrob. Agents Chemother. 53 (2009) : 4845-4851.

[30] Suzuki, K., Sakamoto, J., Koyama, T., Yingsakmongkon, S., Suzuki, Y., Hatano, K., Terunuma, D., Matsuoka, K. Synthesis of sialic acid derivatives having a C=C double bond substituted at the C-5 position and their glycopolymers. Bioorg. Med. Chem. Lett. 19 (2009) : 5105-5108.

[31] Carbain, B., Collins, P. J., Callum, L., Martin, S. R., Hay, A. J., McCauley, J., Streicher. H. Efficient synthesis of highly active phospha-isosteres of the influenza neuraminidase inhibitor oseltamivir. Chem. Med. Chem. 4 (2009) : 335-337.

[32] Carbain, B., Martin, S. R., Collins, P. J., Hitchcock, P. B., Streicher, H. Galactose-conjugates of the oseltamivir pharmacophore—new tools for the characterization of influenza virus neuraminidases. Org. Biomol. Chem. 7 (2009) : 2570-2575.

[33] Ferraris, O., Lina, B. Mutations of neuraminidase implicated in neuraminidase inhibitors resistance. J. Clin. Virol. 41 (2008) : 13-19.

[34] Yen, H. L., Ilyushina, N. A., Salomon, R., Hoffmann, E., Webster, R. G., Govorkova, E. A. Neuraminidase inhibitor-resistant recombinant A/Vietnam/1203/04 (H5N1) influenza viruses retain their replication efficiency and pathogenicity in vitro and in vivo. J. Virol. 81 (2007) : 12418–12426.

[35] Wang, M. Z., Tai, C. Y., Mendel, D. B. Mechanism by which mutations at His274 alter sensitivity of influenza A virus N1 neuraminidase to oseltamivir carboxylate and zanamivir. Antimicrob. Agents Chemother. 46 (2002) : 3809–3816.

[36] Baz, M., Abed, Y., McDonald, J., Boivin, G. Characterization of multidrug resistant influenza A/H3N2 viruses shed during 1 year by an immunocompromised child. Clin. Infect. Dis. 43 (2006) : 1555–1561.

[37] McKimm-Breschkin, J. L., Sahasrabudhe, A., Blick, T., McDonald, M. Mechanisms of resistance of influenza virus to neuraminidase inhibitors. Int. Congr. Ser. 1219 (2001) : 855-861.

[38] de Jong, M. D., Tran, T. T., Truong, H. K., Vo, M. H., Smith, G. J., Nguyen, V. C., Bach, V. C., Phan, T. Q., Do, Q. H., Guan, Y., Peiris, J. S., Tran, T. H., Farrar, J. Oseltamivir resistance during treatment of influenza A (H5N1) infection. N. Engl. J. Med. 353 (2005) : 2667–2672.

[39] Ives, J. A., Carr, J. A., Mendel, D. B., Tai, C. Y., Lambkin, R., Kelly, L., Oxford, J. S., Hayden, F. G., Roberts, N. A. The H274Y mutation in the influenza A/H1N1 neuraminidase active site following oseltamivir phosphate

treatment leave virus severely compromised both in vitro and in vivo. <u>Antiviral Res.</u> 55 (2002) : 307–317.

[40] Moscona, A. Oseltamivir resistance-disabling our influenza defenses. <u>N. Engl. J. Med.</u> 353 (2005) : 2633–2636.

[41] Moscona, A. Neuraminidase inhibitors for influenza. <u>N. Engl. J. Med.</u> 353 (2005) : 1363-1373.

[42] Webster, R. G., Govorkova, E. A. H5N1 Influenza — Continuing Evolution and Spread. <u>N. Engl. J. Med.</u> 355 (2006) : 2174-2177.

[43] Bantia, S., Arnold, C. S., Parker, C. D., Upshaw, R., Chand, P. Anti-influenza virus activity of peramivir in mice with single intramuscular injection. <u>Antiviral Res.</u> 69 (2006) : 39-45.

[44] Chand. P., Bantia, S., Kotian, P., El-Kattan, Y., Lin, T., Babu, Y. Comparison of the anti-influenza virus activity of cyclopentane derivatives with oseltamivir and zanamivir in vivo. <u>Bioor. Med. Chem.</u> 13 (2005) : 4071-4077.

[45] Gubareva, L. V., Kaiser, L., Matrosovich, M. N., Soo-Hoo, Y., Hayden, F. G. Selection of influenza virus mutants in experimentally infected volunteers treated with oseltamivir. <u>J. Infect. Dis.</u> 183 (2001) : 523-531.

[46] McKimm-Breschkin, J. L. Resistance of influenza viruses to neuraminidase inhibitors - a review. <u>Antiviral Res.</u> 47 (2000) : 1-17.

[47] Case, D. A., and others. <u>AMBER 7</u>, University of California, San Francisco. 2002.

[48] Frisch, M. J., and others. <u>Gaussian 03</u>, Gaussian, Inc., Wallingford CT. 2004.

[49] Cornell, W. D., Cieplak, P., Bayly, C. I., Kollmann, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. <u>J. Am. Chem. Soc.</u> 115 (1993) : 9620-9631.

[50] Wang, J. M., Wang, W., Kollman, P. A. Antechamber: an accessory software package for molecular mechanical calculations. <u>J. Am. Chem. Soc.</u> 222 (2001) U403-U403 (Abstr).

[51] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. Comparison of simple potential functions for simulating liquid water. <u>J. Chem. Phy.</u> 79 (1983) : 926-935.

[52] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., Kollman, P. A. A second

generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J. Am. Chem. Soc. 117 (1995) : 5179 - 5197.

[53] Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F. V., DiNola, A., Haak, J. R. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81 (1984) : 3684-3690.

[54] 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. J. Comput. Phys. 23 (1997) : 327-341.

[55] York, D. M., Darden, T. A., Pedersen, L. G. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. J. Chem. Phys. 99 (1993) : 8345-8348.

[56] Sitkoff, D., Sharp, K. A., Honig, B. Accurate calculation of hydration free-energies using macroscopic solvent models. J. Phys. Chem. 98 (1994) : 1978-1988.

[57] Rocchia, W., Alexov, E., Honig, B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. J. Phys. Chem. B 105 (2001) : 6507-6514.

[58] Lee, M. R., Duan, Y., Kollman, P. A. Use of MM-PB/SA in estimating the free energies of proteins: application to native, intermediates, and unfolded villin headpiece. J. Mol. Biol. 313 (2001) : 417-430.

[59] Babu, Y. S., Chand, P., Bantia, S., Kotian, P., Dehghani, A., El-Kattan, Y., Lin, T. H., Hutchison, T. L., Elliott, A. J., Parker, C. D., Ananth, S. L., Horn, L. L., Laver, G. W., Montgomery, J. A. BCX-1812 (RWJ-270201): discovery of a novel, highly potent, orally active, and selective influenza neuraminidase inhibitor through structure-based drug design. J. Med. Chem. 43 (2000) : 3482-3486.

[60] Bonnet, P., Bryce, R. A. Molecular dynamics and free energy analysis of neuraminidase-ligand interactions. Protein Sci. 13 (2004) : 946-957.

[61] Masukawa, K. M., Kollman, P. A., Kuntz, I. D. Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. J. Med. Chem. 46 (2003) : 5628-5637.

[62] Varghese, J. N., Epa, V. C., Colman, P. M. Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. Protein Sci. 4 (1995) : 1081-1087.

[63] Aruksakunwong, O., Malaisree, M., Decha, P., Sompornpisut, P., Parasuk, V., Pianwanit, S., Hannongbua, S. On the lower susceptibility of oseltamivir to in.uenza neuraminidase subtype N1 than those in N2 and N9. Biophys. J. 92 (2007) : 798-807.

[64] Todd, M. J., Luque, L., Velazquez-Campoy, A., Freire, E. Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. Biochemistry 39 (2000) : 11876-11883.

[65] Govorkova, E. A., Leneva, I. A., Goloubeva, O. G., Bush, K., Webster, R. G. Comparison of efficacies of RWJ-270201, zanamivir, and oseltamivir against H5N1, H9N2, and other avian influenza viruses. Antimicrob. Agents Chemother. 45 (2001) : 2723-2732.

[66] Boivin, G., Goyette, N. Susceptibility of recent Canadian influenza A and B virus isolates to different neuraminidase inhibitors. Antiviral Res. 54 (2002) : 143-147.

[67] McSharry, J. J., McDonough, A. C., Olson, B. A., Drusano, G. L. Phenotypic drug susceptibility assay for influenza virus neuraminidase inhibitors. Clin. Diagn. Lab. Immunol. 11 (2004) : 21-28.

[68] Mishin, V. P., Hayden, F. G., Gubareva, L.V. Susceptibilities of antiviral-resistant influenza viruses to novel neuraminidase inhibitors. Antimicrob. Agents Chemother. 49 (2005) : 4515-4520.

[69] Baz, M., Abed, Y., Boivin, G. Characterization of drug-resistant recombinant influenza A/H1N1 viruses selected in vitro with peramivir and zanamivir. Antiviral Res. 74 (2006) : 159-162.

[70] Hurt, A. C., Selleck, P., Komadina, N., Shaw, R., Brown, L., Barr, I. G. Susceptibility of highly pathogenic A(H5N1) avian influenza viruses to the neuraminidase inhibitors and adamantanes. Antiviral Res. 35 (2006) : 345-354.

[71] Li, H., Robertson, A. D., Jensen, J. H. Very fast empirical prediction and interpretation of protein pKa values. Proteins 61 (2005) : 704-721.

[72] Srinivasan, J., Cheatham, T. E. III., Cieplak, P., Kollman, P. A., Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. J. Am. Chem. Soc. 120 (1998) : 9401-9409.

[73] Wang, J., Morin, P., Wang, W., Kollman, P. A. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J. Am. Chem. Soc. 123 (2001) : 5221-5230.

[74] Gilson, M. K., Sharp, K. A., Honig, B. H. Calculating the Electrostatic Potential of Molecules in Solution - Method and Error Assessment. J. Comput. Chem. 9 (1988) : 327-335.

[75] Prompers, J. J., Brüschweiler, R. Dynamic and structural analysis of isotropically distributed molecular ensembles. Proteins 46 (2002) : 177-189.

[76] Saarela, J. T., Tuppurainen, K., Peräkylä, M., Santa, H., Laatikainen, R. Correlative motions and memory effects in molecular dynamics simulations of molecules: Principal components and rescaled range analysis suggest that the motions of native BPTI are more correlated than those of its mutants. Biophys. Chem. 95 (2002) : 49-57.

[77] Zhou, Z., Madrid, M., Evanseck, J. D., Madura, J . D. Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. J. Am. Chem. Soc. 127 (2005) : 17253-17260.

[78] Ferrari, A. M., Degliesposti, G., Sgobba, M., Rastelli, G. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. Bioor. Med. Chem. 15 (2007) : 7865–7877.

[79] Sak, K., Karelson, M., Järv, J. Modeling of the amino acid side chain effects on peptide conformation. Bioorg. Chem. 27 (1999) : 434-442.

[80] Collins, P. J., Haire, L. F., Lin, Y. P., Liu, J., Russell, R. J., Walker, P. A., Skehel, J. J., Martin, S. R., Hay, A. J., Gamblin, S. J. Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants. Nature 453 (2008) : 1258-1261.

[81] Rungrotmongkol, T., Intharathep, P., Malaisree, M., Nunthaboot, N., Kaiyawet, N., Sompornpisut, P., Payungporn, S., Poovorawan, Y., Hannongbua, S.

Susceptibility of antiviral drugs against 2009 influenza A (H1N1) virus. Biochem. Biophys. Res. Commun. 385 (2009) : 390-394.

[82] Rungrotmongkol, T., Udommaneethanakit, T., Malaisree, M., Nunthaboot, N., Intharathep, P., Sompornpisut, P., Hannongbua, S. How does each substituent functional group of oseltamivir lose its activity against virulent H5N1 influenza mutants? Biophys. Chem. 145 (2009) : 29-36.

[83] Case, D. A., and others. AMBER 10, University of California, San Francisco. 2008.

[84] King, G., Warshel, A. A surface constrained all-atom solvent model for effective simulations of polar solutions. J. Chem. Phys. 91 (1989) : 3647-3661.

[85] Marelius, J., Kolmodin, K., Feierberg, I., Åqvist, J. Q: A molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. J. Mol. Graph. Model. 16 (1998) : 213-225.

[86] Åqvist, J., Medina, C., Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. Protein. Eng. 7 (1994) : 385-391.

[87] Hansson, T., Marelius, J., Åqvist, J. J. Ligand binding affinity prediction by linear interaction energy methods. J. Comput. Aided. Mol. Des. 12 (1998) : 27-35.

[88] Wall, I. D., Leach, A. R., Salt, D. W., Ford, M. G., Essex, J. W. Binding constants of neuraminidase inhibitors: an investigation of the linear interaction energy method. J. Med. Chem. 42 (1999) : 5142-5152.

[89] Chachra, R., Rizzo, R. C. Origins of resistance conferred by the R292K neuraminidase mutation via molecular dynamics and free energy calculations. J. Chem. Theory. Comput. 4 (2008) : 1526-1540.

[90] Emedicine. Available from: http://emedicine.medscape.com/article/1673658-overview [2009, May].

[91] Xu, X., Zhu, X., Dwek, R. A., Stevens, J., Wilson, I. A. Structural characterization of the 1918 influenza virus H1N1 neuraminidase. J. Virol. 82 (2008) : 10493-10501.

[92] Chong, A. K., Pegg, M. S., Taylor, N. R., von Itzstein, M. Evidence for a sialosyl cation transition-state complex in the reaction of sialidase from influenza virus. Eur. J. Biochem. 207 (1992) : 335-343.

[93] Tiralongo, J., Pegg, M. S., von Itzstein, M. Effect of substrate aglycon on enzyme mechanism in the reaction of sialidase from influenza virus. <u>FEBS Lett.</u> 372 (1995) : 148-150.

[94] Taylor, N. R., von Itzstein, M. Molecular modeling studies on ligand binding to sialidase from influenza virus and the mechanism of catalysis. <u>J. Med. Chem.</u> 37 (1994) : 616-624.

[95] Thomas, A., Jourand, D., Bret, C., Amara, P., Fieldm, M. J. Is there a covalent intermediate in the viral neuraminidase reaction? a hybrid potential free-energy study. <u>J. Am. Chem. Soc.</u> 121 (1999) : 9693–9702.

[96] Gamblin, S. J., Haire, L. F., Russell, R. J., Stevens, D. J., Xiao, B., Ha, Y., Vasisht, N., Steinhauer, D. A., Daniels, R. S., Elliot, A., Wiley, D. C., Skehel, J. J. The structure and receptor-binding properties of the 1918 influenza hemagglutinin. <u>Science</u> 303 (2004) : 1838–1842.

[97] CDC, Antiviral Drugs and H1N1 Flu (Swine Flu). Available from: <u>http://www.cdc.gov/h1n1flu/antiviral.htm</u> [2009, May].

[98] 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. <u>J. Comput. Chem.</u> 4 (1983) : 187–217.

[99] MacKerell, A. D. and others. All-atom empirical potential for molecular modeling and dynamics studies of proteins. <u>J. Phys. Chem. B</u>. 102 (1998) : 3586–3616.

[100] Brunger, A. T., Brooks, C. L. III., Karplus, M. Active site dynamics of ribonuclease. <u>Proc. Natl. Acad. Sci. USA</u>. 82 (1985) : 8458–8462.

[101] Brooks, C. L., Karplus, M. Solvent effects on protein motion and protein effects on solvent motion. Dynamics of the active site region of lysozyme. <u>J. Mol. Biol.</u> 208 (1989) : 159–181.

[102] Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar Gases. <u>J. Chem. Phys.</u> 22 (1954) : 1420-1426.

[103] Kollman, P. A. Free energy calculations: Applications to chemical and biochemical phenomena. <u>Chem. Rev.</u> 93 (1993) : 2395-2417.

[104] Jorgensen, W. L., Ravimohan, C. J. Monte Carlo simulation of differences in free energies of hydration. <u>J. Chem. Phys.</u> 83 (1985) : 3050-3055.

[105] Shurki, A., Warshel, A. Structure/function correlations of proteins using MM, QM/MM, and related approaches: methods, concepts, pitfalls, and current progress. Adv. Protein Chem. 66 (2003) : 249-313.

[106] Pearlman, D. A., Charifson, P. S. Are free energy calculations useful in practice? a comparison with rapid scoring functions for the p38 MAP kinase protein system J. Med. Chem. 44 (2001) : 3417-3423.

[107] Oostenbrink, B. C., Pitera, J. W., Lipzig, M. M. H., Meerman, J. H. N., van Gunsteren, W. F. Simulations of the estrogen receptor ligand-binding domain: affinity of natural ligands and xenoestrogens. J. Med. Chem. 43 (2000) : 4596-4605.

[108] Barril, X., Orozco, M., Luque, F. J. Predicting relative binding free energies of tacrine−huperzine a hybrids as inhibitors of acetylcholinesterase. J. Med. Chem. 42 (1999) : 5110-5119.

[109] Tembe, B. L., McCammon, J. A. Ligand-receptor interactions. Comput. Chem. 8 (1984) : 281-283.

[110] Michel, J., Verdonk, M. L., Essex, J. W. Protein−ligand complexes: computation of the relative free energy of different scaffolds and binding modes. J. Chem. Theory Comput. 3 (2007) : 1645-1655.

[111] Jorgensen, W. L., Buckner, J. K., Boudon, S., Tirado-Rives, J., Efficient computation of absolute free energies of binding by computer simulations, application to the methane dimer in water. J. Chem. Phys. 89 (1988) : 3742-3746.

[112] Hamelburg, D., McCammon, J. A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: double-decoupling method. J. Am. Chem. Soc. 126 (2004) : 7683-7689.

[113] Hermans J., Wang, L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. application to a complex of benzene and mutant T4 lysozyme. J. Am. Chem. Soc. 119 (1997) : 2707-2714.

[114] Singh, S. B., Kollman, P. A. calculating the absolute free energy of association of netropsin and DNA. J. Am. Chem. Soc. 121 (1999) : 3267-3271.

[115] Barillari, C., Taylor, J., Viner, R., Essex, J. W. Classification of water molecules in protein binding sites. J. Am. Chem. Soc. 129 (2007) : 2577-2587.

[116] Kato, M., Braun-Sand, S., Warshel, A. Computational and structural approaches to drug discovery: ligand-protein interactions. Royal Society of Chemistry : (2008)

[117] Reynolds, C. A., King, P. M., Richards, W. G. Free energy calculations in molecular biophysics. Mol. Phys. 76 (1992) : 251-275.

[118] Michel, J., Foloppe, N., Essex, J. W. Rigorous free energy calculations in structure-based drug design. Mol. Inf. (2010) *In press*.

[119] Jorgensen, W. L. The many roles of computation in drug discovery. Science 303 (2004) :1813-1818.

[120] Singh, N., Warshel, A. Absolute binding free energy calculations: on the accuracy of computational scoring of protein–ligand interactions. Proteins 78 (2010) : 1705-1723.

[121] Michel, J., Essex, J. W. Prediction of protein-ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. J. Comput. Aided Mol. Des. 24 (2010) : 639-658.

[122] Aleksandrov, A., Thompson, D., Simonson, T. Alchemical free energy simulations for biological complexes: powerful but temperamental. J. Mol. Recognit. 23 (2010) : 117-127.

[123] Helms, V., Wade, R. C. Computational alchemy to calculate absolute protein−ligand binding free energy. J. Am. Chem. Soc. 120 (1998) : 2710-2713.

[124] Woo, H. J., Roux, B. Chemical theory and computation special feature: calculation of absolute protein–ligand binding free energy from computer simulations. Proc. Natl. Acad. Sci. U.S.A. 102 (2005) : 6825-6830.

[125] Doudou, S., Burton, N. A., Henchman, R. H., Standard free energy of binding from a one-dimensional potential of mean force. J. Chem. Theory Comput. 5 (2009) : 909-918.

[126] Tyka, M. D., Sessions, R. B., Clarke, A. R. Absolute free-energy calculations of liquids using a harmonic reference state. J. Phys. Chem. B 111 (2007) : 9571-9580.

[127] Woods, C. J., Tyka, M. D., Sessions, R. B., Mulholland, A. J. A method to allow the identity of a cluster of solvent molecules in bulk to be maintained during a simulation. (2010) *Submitted*.

[128] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E. Equation of state calculations by fast computing machines. J. Chem. Phys. 21 (1953) : 1087-1092.

[129] Pearlman, D. A. A comparison of alternative approaches to free energy calculations. J. Phys. Chem. 98 (1994) : 1487-1493.

[130] Woods, C. J., King, M. A., Essex, J. W. The development of replica-exchange-based free-energy methods. J. Phys. Chem. B 107 (2003) : 13703-13710.

[131] Woods, C. J., King, M. A., Essex, J. W. Enhanced configurational sampling in binding free-energy calculations. J. Phys. Chem. B 107 (2003) : 13711-13728.

[132] Mezei, M. The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water. J. Chem. Phys. 86 (1987) : 7084-7088.

[133] Guimaraes, C. R. W., Alencastro, R. B. Evaluating the relative free energy of hydration of new thrombin inhibitor candidates using the finite difference thermodynamic integration (FDTI) method. Int. J. Quantum. Chem. 85 (2001) : 713-726.

[134] Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R., van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chem. Phys. Lett. 222 (1994) : 529-539.

[135] Zacharias, M., Straatsma, T. P., McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. J. Chem. Phys. 100 (1994) : 9025-9031.

[136] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. Case, D. A. Development and testing of a general amber force field. J. Comput. Chem. 25 (2004) : 1157-1174.

[137] Wang, J., Wang, W., Kollman, P. A., Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model. 25 (2006) : 247-260.

[138] Case, D. A., and others. AMBER 9, University of California, San Francisco. 2006.

[139] Jorgensen, W. L., Maxwell, D. S., Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. 118 (1996) : 11225-11236.

[140] Owicki, J. C., Scheraga, H. A. Preferential sampling near solutes in Monte Carlo calculations on dilute solutions. Chem. Phys. Lett. 47 (1977) : 600-602.

[141] Jorgensen, W. L. Theoretical studies of medium effects on conformational equilibria. J. Phys. Chem. 87 (1983) : 5304-5314.

[142] Jorgensen, W. L., Ravimohan, C. Monte Carlo simulation of differences in free energies of hydration. J. Chem. Phys. 83 (1985) : 3050-3054.

[143] Woods, C. J. Sire: a complete molecular simulation framework (2010), in development. Avaliable from: http://siremol.org [2010, September].

[144] Woods, C. J., Manby, F. R., Mulholland, A. J. An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. J. Chem. Phys. 128 (2008) : 0141091-0141098.

[145] Michel, J., Verdonk, M. L., Essex, J. W. Protein-ligand binding affinity predictions by implicit solvent simulations: a tool for lead optimization? J. Med. Chem. 49 (2006) : 7427-7439.

[146] 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.

[147] Shaw, K. E., Woods, C. J., Mulholland, A. J. Compatibility of quantum chemical methods and empirical (MM) water models in quantum mechanics/molecular mechanics liquid water simulations. J. Phys. Chem. Lett. 1 (2010) : 219-233.

[148] Rizzo, R. C., Udier-Blagović, M., Wang, D. P., Watkins, E. K., Smith, M. B., Smith, R. H. Jr., Tirado-Rives, J., Jorgensen, W. L. Prediction of activity for

nonnucleoside inhibitors with HIV-1 reverse transcriptase based on monte carlo simulations. J. Med. Chem. 45 (2002) : 2970-2987.

[149] Hensen, C., Hermann, J. C., Nam, K. H., Ma, S. H., Gao, J. L., Oltje, H. D. H. A combined QM/MM approach to protein−ligand interactions: polarization effects of the HIV-1 protease on selected high affinity inhibitors. J. Med. Chem. 47 (2004) : 6673-6680.

[150] Panagiotopoulos, A. Z. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. Mol. Phys. 61 (1987) : 813-826.

[151] Panagiotopoulos, A. Z., Quirke, N., Stapleton, M., Tildesley, D. Phase equilibria by simulation in the Gibbs ensemble: alternative derivation, generalization and application to mixture and membrane equilibria. J. Mol. Phys. 63 (1988) : 527-545.

[152] Shi, W., Maginn, E. J. Improvement in molecule exchange efficiency in Gibbs ensemble Monte Carlo: development and implementation of the continuous fractional component move. J. Comput. Chem. 29 (2008) : 2520-2530.

[153] General, I. J. A note on the standard state's binding free energy. J. Chem. Theory Comput. 6 (2010) : 2520-2524.

APPENDIX

A. CONFERENCES (POSTER PRESENTATION)

1. **The 32th congress on science and technology Thailand**, 10-12 October 2006 at Queen Sirikit National Convention Center, Thailand

2. **The Third Asian-Pacific Conference of Theoretical and Computational Chemistry**, 22-25 September 2007 at Beijing, China

3. **The 33th congress on science and technology Thailand**, 18-24 October 2007 at Walailuk University, Thailand

4. **Bangkok International Conference on Avian Influenza 2008: Integration from Knowledge to Control**, 23-25 January 2008 at the Dusit Thani Bangkok, Thailand

5. **Pure and Applied Chemistry International Conference 2008**, 30 January – 1 February 2008 at Sofitel Centara Grand Bangkok, Bangkok, Thailand

6. **Computational Molecular Science 2008**, 22-25 June 2008 at the Royal Agricultural College, Cirencester, UK

7. **The Third Collaborative Computational Project for Biomolecular Simulation (CCPB)**, 6-8 January 2009 at the Yorkshire Museum and Gardens, York, UK

8. **Computational Molecular Science 2010**, 27-30 June 2010 at the Royal Agricultural College, Cirencester, UK

B. CONFERENCES (ORAL PRESENTATION)

1. **Global Concerns, Recent Outbreak and Molecular Insight into Avian Influenza H5N1, The TRF Senior Scholarship, The Thailand Research Fund**, 18 April 2008 at Asia Hotel, Bangkok, Thailand

2. **The Science Forum 2010,** 11-12 March 2010 at Chulalongkorn University, Bangkok, Thailand

3. **Eleventh RGJ-Ph.D. Congress (RGJ-Ph.D. Congress XI),** 1-3 April, 2010 at Jomthien Palm Beach and Resort, Pattaya, Thailand

4. **The 240th American Chemical Society National Meeting & Exposition,** 22–26 August 2010 at Boston, USA

C. RESEARCH VISITING

1. **April, 2008 - March, 2009** Centre of Computational Chemistry, School of Chemistry, University of Bristol, UK, with Prof. Dr. Adrian J. Mulholland and Dr. Christopher J. Woods

2. **May, 2009 – December, 2009** International Centre for Science and High Technology (ICS-UNIDO), Trieste, Italy, with Prof. Dr. Stanislav Miertus

3. **April, 2010 – July, 2010** Centre of Computational Chemistry, School of Chemistry, University of Bristol, UK, with Prof. Dr. Adrian J. Mulholland and Dr. Christopher J. Woods

D. SCHOLARSHIP AND AWARDS

2003   Top ten award in Department of Chemistry, Faculty of Science, Naresuan University, Thailand

2004   Top ten award in Department of Chemistry, Faculty of Science, Naresuan University, Thailand

2005   The Royal Golden Jubilee Ph.D. Program, Thailand Research Fund

2008   Outstanding poster presentation award at Pure and Applied Chemistry International Conference, Thailand

2010   The best oral presentation award at The Science Forum 2010, Thailand

2010   Outstanding oral presentation award at RGJ-Ph.D. Congress XI

E. PUBLICATIONS

1) O. Aruksakunwong, **M. Malaisree**, P. Decha, P. Sompornpisut, V. Parasuk, S. Pianwanit, and S. Hannongbua. On the lower susceptibility of oseltamivir to influenza neuraminidase subtype N1 than those in N2 and N9. Biophys. J. 92 (2007) : 789-807.

2) **M. Malaisree**, T. Rungrotmongkol, P. Decha, P. Intharathep, O. Aruksakunwong and S. Hannongbua. Understanding of known drug-target interactions in the catalytic pocket of neuraminidase subtype N1. Proteins 71 (2008) : 1908-1918.

3) P. Decha, T. Rungrotmongkol, P. Intharathep, **M. Malaisree**, O. Aruksakunwong, C. Laohpongspaisan, V. Parasuk, P. Sompornpisut, S. Pianwanit, S. Kokpol and S. Hannongbua. Source of high pathogenicity of an avian influenza virus H5N1: Why H5 is better cleaved by furin. Biophys. J. 95 (2008) : 1-7.

4) P. Intharathep, C. Laohpongspaisan, T. Rungrotmongkol, A. Loisruangsin, **M. Malaisree**, P. Decha, O. Aruksakunwong, K. Chuenpennit, N. Kaiyawet, P. Sompornpisut, S. Pianwanit, and S. Hannongbua. How amantadine and rimantadine inhibit proton transport in the M2 protein channel. J. mol. Graph. Model. 27 (2008) :342-348.

5) T. Rungrotmongkol, P. Decha, **M. Malaisree**, P. Sompornpisut, and S. Hannongbua. Comment on ''Cleavage mechanism of the H5N1 hemagglutinin by trypsin and furin''. Amino Acids 35 (2008) :511-512.

6) V. Nukoolkarn, V. S. Lee, **M. Malaisree**, O. Aruksakunwong and S. Hannongbua. Molecular dynamic simulations analysis of ritronavir and lopinavir as SARS-CoV 3CL^{pro} inhibitors. J. Theor. Biol. 254 (2008) : 861-867.

7) **M. Malaisree**, T. Rungrotmongkol, N. Nunthaboot, O. Aruksakunwong, P. Intharathep, P. Decha, P. Sompornpisut and S. Hannongbua. source of oseltamivir resistance in avian influenza H5N1 virus with the H274Y mutation. Amino Acids 37 (2009) : 725-732.

8) T. Rungrotmongkol , **M. Malaisree**, T. Udommaneethanakit , S. Hannongbua. Comment on "Another look at the molecular mechanism of the resistance of H5N1 influenza A virus neuraminidase (NA) to oseltamivir (OTV)". Biophys. Chem. 141 (2009) : 131-132.

9) T. Rungrotmongkol, P. Decha, P. Sompornpisut, **M. Malaisree**, P. Intharathep, N. Nunthaboot, T. Udommaneethanakit, O. Aruksakunwong, and S. Hannongbua. Combined QM/MM mechanistic study of the acylation process in furin complexed with the H5N1 avian influenza virus hemagglutinin's cleavage site. Proteins 76 (2009) : 62-71.

10) C. Laohpongspaisan, T. Rungrotmongkol, P. Intharathep, **M. Malaisree**, P. Decha, O. Aruksakunwong, P. Sompornpisut and S. Hannongbua. Why amantadine loses its function in influenza M2 mutants: MD simulations. J. Chem. Inf. Model. 49 (2009) : 847-852.

11) T. Rungrotmongkol, P. Intharathep, **M. Malaisree**, N. Nunthaboot, N. Kaiyawet, P. Sompornpisut, C. Payungporn, Y. Poovorawan and S. Hannongbua. Susceptibility of Antiviral Drugs Against 2009 Influenza A (H1N1) Virus. Biochem. Biophys. Res. Commun. 385 (2009) : 390-394.

12) T. Rungrotmongkol, T. Udommaneethanakit, **M. Malaisree**, N. Nunthaboot, P. Intharathep, P. Sompornpisut and S. Hannongbua. How does each substituent functional group of oseltamivir lose its activity against virulent H5N1 mutants? <u>Biophys. Chem.</u> 145 (2009) : 29-36.

13) T. Rungrotmongkol, **M. Malaisree**, N. Nunthaboot, P. Sompornpisut, and S. Hannongbua. Molecular prediction of oseltamivir efficiency against probable influenza A (H1N1-2009) mutants: Molecular modelling approach. <u>Amino Acids</u> 39 (2010) : 393-398.

14) N. Nunthaboot, T. Rungrotmongkol, **M. Malaisree**, P. Decha, N. Kaiyawet, P. Intharathep, P. Sompornpisut, Y. Poovorawan and S. Hannongbua. Molecular insights of human receptor binding to 2009 H1N1 influenza A hemagglutinin. <u>Monatsh. Chem.</u> 141 (2010) : 801-807.

15) N. Nunthaboot, T. Rungrotmongkol, **M. Malaisree**, N. Kaiyawet, P. Decha, P. Sompornpisut, Y. Poovorawan, S. Hannongbua. Evolution of human receptor binding affinity of H1N1 hemagglutinins from 1918 to 2009 pandemic influenza A virus. <u>J. Chem. Inf. Model.</u> 50 (2010) : 1410-1417.

16) C. J. Woods, **M. Malaisree**, S. Hannongbua, A. J. Mulholland. A water-swap reaction coordinate for the calculation of absolute protein-ligand binding free energies. (2010) ***Submitted***.

17) T. Rungrotmongkol, N. Nunthaboot, **M. Malaisree**, N. Kaiyawet, P. Intharathep, A. Meepraset 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. <u>J. mol. Graph. Model.</u> (2010) ***Accepted***.

18) T. Rungrotmongkol, **M. Malaisree**, V. Frecer, S. Hannongbua and S. Miertus. Inhibitory efficiency of oseltamivir phosphonate against oseltamivir-resistant influenza neuraminidases of H5N1 and novel pandemic H1N1 viruses: A MD/LIE approach. (2010) ***In preparation.***

19) **M. Malaisree**, C. J. Woods, T. Rungrotmongkol, N. Nantaboot, A. Vongachariya, A. J. Mulholland and S. Hannongbua. Substrate Binding to Neuraminidase of Influenza Viruses: QM/MM MD Simulations. (2010) ***In preparation.***

VITAE

PERSONAL DATA

Name   Maturos Malaisree

Nationality Thai

Date of Birth 5 October 1982

E-mail   maturos_ros@hotmail.com

EDUCATION

2001-2005   B.Sc. in Chemistry (1st honor)

             Department of Chemistry, Faculty of Science, Naresuan University,

             Phitsanulok, Thailand

2005-2010   Ph.D. in Physical Chemistry

             Department of Chemistry, Faculty of Sciences, Chulalongkorn

             University, Bangkok, Thailand

PUBLICATIONS IN THIS DESSERTATION

1) M. Malaisree, T. Rungrotmongkol, P. Decha, P. Intharathep, O. Aruksakunwong and S. Hannongbua. Understanding of known drug-target interactions in the catalytic pocket of neuraminidase subtype N1. Proteins 71 (2008) : 1908-1918.

2) M. Malaisree, T. Rungrotmongkol, N. Nunthaboot, O. Aruksakunwong, P. Intharathep, P. Decha, P. Sompornpisut and S. Hannongbua. Source of Oseltamivir Resistance in Avian Influenza H5N1 Virus with the H274Y Mutation. Amino Acids 37 (2009) : 725-732.

3) T. Rungrotmongkol, M. Malaisree, N. Nunthaboot, P. Sompornpisut, and S. Hannongbua. Molecular prediction of oseltamivir efficiency against probable influenza A (H1N1-2009) mutants: Molecular modelling approach. Amino Acid 39 (2010) : 393-398.

4) M. Malaisree, C. J. Woods, T. Rungrotmongkol, N. Nantaboot, A. Vongachariya, A. J. Mulholland and S. Hannongbua. Substrate Binding to Neuraminidase of Influenza Viruses: QM/MM MD Simulations. (2010) *In preparation.*

5) C. J. Woods, M. Malaisree, S. Hannongbua, A. J. Mulholland. A water-swap reaction coordinate for the calculation of absolute protein-ligand binding free energies. (2010) *Submitted*.