

การยืมเนื้อหาของabstractและโอเชลทามิเวียร์ในนิรรมินิเตสของไวรัสไข้หวัดใหญ่ด้วยการจำลอง
พลวัตเชิงโมเลกุล



นางจิราพร พานิชย์

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

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

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาเคมี ภาควิชาเคมี

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2559

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

SUBSTRATE AND OSELTAMIVIR BINDING INTO INFLUENZA VIRUS
NEURAMINIDASE BY MOLECULAR DYNAMICS SIMULATIONS

Mrs. Jiraphorn Phanich



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 2016
Copyright of Chulalongkorn University

Thesis Title SUBSTRATE AND OSELTAMIVIR BINDING
INTO INFLUENZA VIRUS
NEURAMINIDASE BY MOLECULAR
DYNAMICS SIMULATIONS

By Mrs. Jiraphorn Phanich

Field of Study Chemistry

Thesis Advisor Professor Dr.Supot Hannongbua

Thesis Co-Advisor Assistant Professor Dr.Thanyada
Rungrotmongkol

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

..... Dean of the Faculty of Science
(Associate Professor Dr.Polkit Sangvanich)

THESIS COMMITTEE

..... Chairman
(Associate Professor Dr.Vudhichai Parasuk)

..... Thesis Advisor
(Professor Dr.Supot Hannongbua)

..... Thesis Co-Advisor
(Assistant Professor Dr.Thanyada Rungrotmongkol)

..... Examiner
(Associate Professor Dr.Pornthep Sompornpisut)

..... Examiner
(Assistant Professor Dr.Somsak Pianwanit)

..... External Examiner
(Assistant Professor Arunee Thitithanyanont)

..... External Examiner
(Dr.Saree Phongphanphanee)

5572848623 : MAJOR CHEMISTRY

KEYWORDS: INFLUENZA VIRUS / NEURAMINIDASE / SUBSTRATE SPECIFICITY / MD SIMULATIONS / BINDING FREE ENERGY

JIRAPHORN PHANICH: SUBSTRATE AND OSELTAMIVIR BINDING INTO INFLUENZA VIRUS NEURAMINIDASE BY MOLECULAR DYNAMICS SIMULATIONS. ADVISOR: PROF. SUPOT HANNONGBUA, CO-ADVISOR: ASST. PROF. THANYADA RUNGROTMONGKOL, 74 pp.

Neuraminidase (NA), a glycoprotein on influenza virus membrane, attaches to the terminal sialic acid (SIA) then cleaves the glycosidic linkage from the subsequent sugar of host cell receptor to release new virions. Types of receptor based on glycosidic linkage are SIA-2,3-GAL-1,4-NAG (3SL) and SIA-2,6-GAL-1,4-NAG (6SL). NA has been an important target for developing anti-influenza agents; however, the substrate binding and specificity of NA are not well understood. In addition, the widespread use of NA inhibitor leads to the mutation of gene. For instance, the pandemic of novel H7N9 virus has a R292K substitution of NA conferring oseltamivir resistance. Therefore, this work aims to understand the substrate binding into NAs from H1N1, H2N2 and H3N2 human viruses and H5N1, H7N9 avian viruses as well as to elucidate insight into oseltamivir resistance of H7N9 substitution using MD simulations and binding free energy predictions. QM/MM-GBSA binding free energies predicted that NAs of human viruses show strong binding to 6SL, whereas avian NAs have preference for 3SL. Although SIA is the main part of glycan for binding into NA, the difference of binding patterns is caused by the diversity of 6SL revealing the rotation of torsion angles between GAL and NAG. The prediction of oseltamivir susceptibilities of H7N9 NA using MM/PB(GB)SA and MM/3D-RISM is in good agreement with the experimental data. The primary source of oseltamivir resistance is due to the decrease of drug-target interaction. The shorter side chain of lysine at mutated residue results in the loss of hydrogen bond forming with carboxylate group and increasing of water molecules.

Department: Chemistry Student's Signature

Field of Study: Chemistry Advisor's Signature

Academic Year: 2016 Co-Advisor's Signature

ACKNOWLEDGEMENTS

I would like to thank my advisors, Prof. Dr. Supot Hannongbua and Asst. Prof. Dr. Thanyada Rungrotmongkol for their support throughout my graduate study. This research could not be done without the suggestions and guidance from my advisors. I deeply appreciate their encouragement, enthusiasm and freedom given me which enabled me to grow as a computational chemistry scientist. In addition, I especially thank Assoc. Prof. Dr. Vudhichai Parasuk for his kindness and useful suggestions in our group meetings.

Family are the power of my life. I would like to give all gratitude to my parents who have always supported me to accomplish my greatest mission. I am thankful for my husband, Phaisat and beloved daughter, Marisa for their love and inspiration to keep me going as well as my brother, Kittikorn.

The Science Achievement Scholarship of Thailand (SAST) is gratefully acknowledged for financial support throughout my graduate study. The research is supported by the Rachadaphiseksomphot Endowment Fund Part of the Strengthen Chulalongkorn University Researcher's Project and the 90th of Chulalongkorn University. The computing facilities were supported by the computational chemistry unit cell (CCUC), Chulalongkorn University and College of Life Science, Ritsumeikan University. I have had a good opportunity to collaborate with the RISM team headed by Prof. Dr. Fumio Hirata and his nice colleagues: Dr. Daniel Sindhikara, Dr. Saree Phongphanphanee, Assoc. Prof. Dr. Norio Yoshida and Asst. Prof. Dr. Masatake Sugita. I honestly thank them for their scientific discussions on the theory of liquid and for their computer facilities.

I thank Dr. Arthitaya Meeprasert and Dr. Kiattisak Luksanangam for always fruitful discussions and suggestions about research, Dr. Andrew King for English proof reading. All CCUC members are acknowledged for all kinds of help. Finally, I also thank all others who somehow in some ways have supported me during my Ph.D. study.

CONTENTS

	Page
THAI ABSTRACT	iv
ENGLISH ABSTRACT.....	v
ACKNOWLEDGEMENTS	vi
CONTENTS.....	vii
LIST OF FIGURES	1
LIST OF TABLES	3
LIST OF ABBREVIATIONS.....	4
CHAPTER I INTRODUCTION.....	6
1.1 Research rationales	6
1.2 Influenza neuraminidase	7
1.3 Mutation conferring receptor recognition.....	12
1.4 Neuraminidase inhibitors and resistance mutation	13
1.5 Literature reviews	15
1.5.1 Substrate specificity of IV NA	15
1.5.2 Reduced susceptibility of NAI conferring NA mutation.....	16
1.6 Scope of this work	17
CHAPTER II THEORETICAL BACKGROUND.....	18
2.1 Computational simulation methods	18
2.2 Molecular mechanics potential function.....	18
2.2.1 Bonded interactions	19
2.2.2 Non-bonded interactions.....	20
2.3 Solvation treatment	21
2.3.1 Explicit solvation	22
2.3.2 Continuum model	22
2.3.3 3D-RISM approach.....	27
2.4 Statistical mechanics	30
2.5 MD simulation	31
2.6 Binding free energy calculation	33

	Page
CHAPTER III CALCULATION DETAILS	36
3.1 MD simulation of substrate binding to NAs	36
3.1.1 Setup of substrate/NA complexes	36
3.1.2 MD simulations of substrate/NA complexes	37
3.1.3 QM/MM-GBSA calculation of substrate specificity	38
3.2 MD simulation of H7N9 NAs and oseltamivir complex	39
3.2.1 H7N9/Oseltamivir complex preparation	39
3.2.2 MD simulation and protocol	39
3.2.3 Binding free energy calculations	40
CHAPTER IV RESULTS AND DISCUSSION.....	42
4.1 MD simulation of substrate binding to NAs	42
4.1.1 Substrate specificity of influenza NAs	42
4.1.2 Dynamics of neuraminidase	43
4.1.3 Glycan topology.....	45
4.1.4 Substrate binding pattern	49
4.2 MD simulation of H7N9 NAs and oseltamivir complex	52
4.2.1 Stability of complexes	52
4.2.2 Binding free energy calculation of oseltamivir susceptibility	53
4.2.3 Change of drug binding interaction	55
4.2.4 Change of solvation conferring R292K substitution	57
CHAPTER V CONCLUSIONS	60
REFERENCES	62
VITA.....	74

LIST OF FIGURES

	Page
Figure 1 Two-dimensional structures of trisaccharide substrate (a) 3SL and (b) 6SL.....	7
Figure 2 Hemagglutinin and neuraminidase glycoproteins on cell membrane of influenza virus that interact with terminal sialic acid of glycoprotein of host cell.....	8
Figure 3 Distribution of 3SL and 6SL receptors found in respiratory tract of human, ferret, swine and chicken ¹⁵	9
Figure 4 Proposed cleavage mechanism of influenza virus neuraminidase.....	10
Figure 5 Highly charged binding pocket of influenza NA.....	10
Figure 6 Two-dimensional chemical structures of NAI including zanamivir, oseltamivir, peramivir and laninamivir.....	13
Figure 7 Structure of oseltamivir binding in the H7N9 NA active site (a) wild-type and (b) R292K substitution which results in multidrug resistance	14
Figure 8 Illustration of Periodic boundary condition (http://isaacs.sourceforge.net/phys/pbc.html)	21
Figure 9 Illustration of different sites (3-, 4-, 5- and 6-sites) of water model (https://en.wikipedia.org/wiki/Water_model).....	22
Figure 10 Presentation of neck region (shaded area) and illustration of its geometry defining by the atoms radius R_1 , R_2 , its separated distance d and radius of solvent molecule R_W ⁸⁰	26
Figure 11 Scheme of 3D-RISM calculation.....	29
Figure 12 Scheme of MD simulation.....	32
Figure 13 Important loops of NA and 150-cavity formation observed by the distance between the residues 149 and 431	44
Figure 14 Distribution plot of distance between the alpha carbon of the residues 149 and 431 of (a) 3SL complexes (b) 6SL complexes.....	45
Figure 15 (Upper) Illustration of topological angle (θ) and glycosidic dihedral angles (τ) of (a) 3SL (b) 6SL substrates, (Lower) the angle's distribution of substrates in complex with different NAs as well as the representative conformations at the highest probability for (c) 3SL and (d) 6SL complexes.....	46

Figure 16 Plot of torsion angle's distribution of (a) 3SL and (b) 6SL glycans within binding pocket of different NAs, including H1N1, H2N2, H3N2, H5N1 and H7N9	48
Figure 17 Percentage of intermolecular hydrogen bonds between 3SL/6SL trisaccharide and NAs residues of avian and human IV	49
Figure 18 Substrate binding pattern of 6SL (green) and 3SL (grey) in the NA binding pocket with the hydrogen bond formations (dash line) demonstrated from the representative snapshot of each system	50
Figure 19 Difference in per-residue decomposition energy between 3SL (yellow) and 6SL (green) in complex with (a) H5N1 and (b) H7N9 NAs, whereas the positive and negative values were displayed as spectrum from red (3 kcal/mol), white (0 kcal/mol) and blue (-3 kcal/mol) surfaces	51
Figure 20 RMSD calculations in term of the backbone atoms of the NA protein (upper) and the oseltamivir ligand atoms (lower) versus the simulation time.....	53
Figure 21 Percentage occupation of the hydrogen bond interactions between oseltamivir function groups and the H7N9 NA binding residues for (a) wild-type and (b) R292K strains. The hydrogen bonds (black dashed line) of oseltamivir-NA taken from the last snapshot are also shown for (c) wild-type and (d) R292K systems.....	56
Figure 22 Distribution of the torsion angles of the pentyl group of oseltamivir calculated while in complexation with the binding pockets	57
Figure 23 (Upper) 3D-RISM function of O-water atoms with $g(r) > 5$ as cyan contour shown within 5 Å of oseltamivir in complex with (a) wild-type and (b) R292K NAs taken from the representative stable structure, (Lower) (c) 3D-RISM RDF of water molecules around heteroatoms of drug averaged from 100 snapshots of production phase.....	58

LIST OF TABLES

	Page
Table 1 Sequence alignment and structural superimposition of emerging N1, N2 and N9 NAs.....	11
Table 2 Comparison of catalytic site and frame work site residues between H1N1, H5N1, H2N2, H3N2 and H7N9 NAs	11
Table 3 Drug susceptibilities of plaque-purified H7N9 viruses and fold resistance of R292K substitution compare to those of wild-type	14
Table 4 Summary of various GB models in AMBER program	25
Table 5 All simulated substrate/NA complexes	37
Table 6 QM/MM-GBSA binding free energies of 6SL and 3SL substrates binding towards different influenza NAs (in kcal/mol)	42
Table 7 Binding free energy of oseltamivir toward wild-type and R292K NAs using MM/3D-RISM and MM/PB(GB)SA methods (in kcal/mol)	54

LIST OF ABBREVIATIONS

3D	= Three dimension
3SL	= SIA-2,3-GAL-1,4-NAG
6SL	= SIA-2,6-GAL-1,4-NAG
Ala (A)	= Alanine
Arg (R)	= Arginine
Asn (N)	= Asparagine
Asp (D)	= Aspartic acid
CG	= Conjugated gradient
Cys (C)	= Cysteine
CYX	= Cysteine involved in disulfide bond
DFTB	= Density functional tight-binding
Ele	= Electrostatics
GAL	= Galactose
GB	= generalized Born
GB ^{HCT}	= generalized Born model developed by Hawkins, Cramer and Truhlar
GB ^{Neck}	= generalized Born model developed by Mongan et al.
GB ^{OBC}	= generalized Born model developed by Onufriev, Bashford and Case
Gln (Q)	= Glutamine
Glu (E)	= Glutamic acid
Gly (G)	= Glycine
HA	= Hemagglutinin
His (H)	= Histidine
HNC	= Hypernetted-chain closure relation
IAV	= Influenza A virus
IC ₅₀	= Median inhibition concentration reduced the effect by 50%
Ile (I)	= Isoleucine
IV	= Influenza virus
KH	= Kovalenko-Hirata closure relation
Leu (L)	= Leucine
Lys (K)	= Lysine
MD	= Molecular dynamics
MM	= Molecular mechanics
MM/3D-RISM	= Molecular mechanic/ Three-dimensional reference interaction site model
MM/GBSA	= Molecular mechanic/generalized Born surface area
MM/PB(GB)SA	= Molecular mechanic/ Poisson-Boltzmann or generalized Born surface area
MM/PBSA	= Molecular mechanic/Poisson-Boltzmann surface area
NA	= Neuraminidase
NAG	= N-acetyl glucosamine

NAI	= Neuraminidase inhibitors
NPT	= Isobaric-isothermal ensemble
NVE	= Microcanonical ensemble
NVT	= Canonical ensemble
PB	= Poisson-Boltzmann
PBC	= Periodic boundary condition
PDB	= Protein Data Bank
PME	= Particle mesh Ewald
Pro (P)	= Proline
PSE- <i>n</i>	= Partial series expansion of order- <i>n</i>
PY	= Percuc-Yevick
QM	= Quantum mechanics
QM/MM	= Quantum mechanics/molecular mechanics
QM/MM-GBSA	= Quantum mechanics/molecular mechanics generalized Born surface area
RDF	= Radial distribution function
REMD	= Replica exchange Molecular dynamics
RISM	= Reference interaction site model
RMSD	= Root mean square displacement
SASA	= Solvent accessible surface area
SCC-DFTB	= Self-consistent charge density functional tight-binding
SD	= Steepest descent
Ser (S)	= Serine
SIA	= Sialic acid
Thr (T)	= Threonine
Trp (W)	= Tryptophan
Tyr (Y)	= Tyrosine
vdW	= van der Waals
WHO	= World Health Organization

CHAPTER I

INTRODUCTION

1.1 Research rationales

Neuraminidase (NA), one of glycoprotein on influenza virus (IV) surface membrane, has been focusing as effective target for designing and developing the anti-influenza agents. NA acts as enzyme to cleave off the glycosidic linkage between terminal sialic acid (SIA) and the next sugar of galactose (GAL) releasing new virions to infect the other cells. The glycosidic linkage is classified into two types of SIA-2,3-GAL-1,4-NAG (3SL or avian receptor) and SIA-2,6-GAL-1,4-NAG (6SL or human receptor) depicted in Figure 1. Previous studies found that NA can cleave both glycosidic linkages but they prefer for the 3SL. NA cleavage activity is directly related to the substrate binding. However, the substrate binding to influenza virus NA and its specificity are not well understood. In addition, the widespread uses of NA inhibitors (NAI) lead to the substitution of NA gene conferring drug resistance. For instance, some strains of currently emerging H7N9 IV have a R292K mutation in NA conferring high-level resistance to frequently used oseltamivir. Therefore, the aims of this research are (i) to elucidate the substrate binding and specificity towards emerged NAs of human IV from H1N1, H2N2 and H3N2 and avian IV from H5N1, H7N9 and (ii) to understand the source of oseltamivir resistance toward R292K mutation of H7N9 NA using of molecular dynamics (MD) simulations and binding free energy calculations. This atomistic information could clarify the key catalytic residues and the specificity function of influenza NA, which is the fundamental knowledge for understanding influenza replication and designing the potent inhibitor.

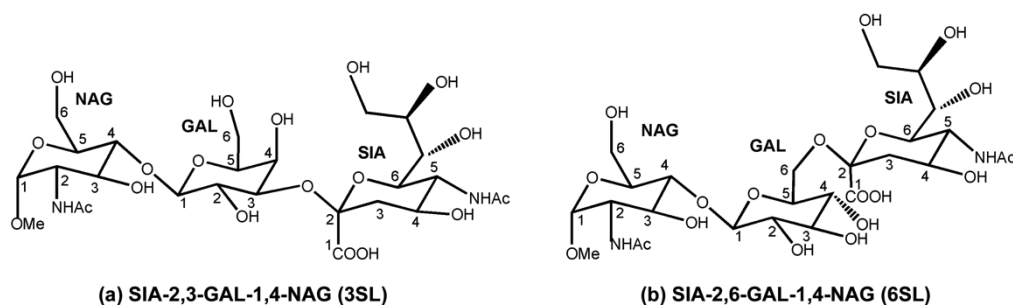


Figure 1 Two-dimensional structures of trisaccharide substrate (a) 3SL and (b) 6SL

1.2 Influenza neuraminidase

Influenza is an infectious disease in avian and mammals caused by IV resulting in sporadic zoonotic infections, seasonal epidemics and pandemics. The currently circulating IVs that cause human infection are A and B types. The seasonal epidemics has associated with an antigenic drift of influenza A and B viruses that develops the mutations within gene to escape from pre-existing human immunity and are responsible for a quarter to a half of million deaths in each year.¹ Whilst the pandemics is caused by an antigenic shift, a re-assortment among different species of influenza A viruses (IAV) leading to a novel subtype, allowing human infection from zoonotic origin and efficient human-to-human transmission.² There are several emergences of human pandemic influenza; in 1918 the Spanish flu of H1N1 caused by avian H1N1 viruses killed over 50 million people, followed by the pandemics of H2N2 in 1957 and H3N2 in 1968. In 2004 the avian H5N1 caused by re-assortment of mainly avian strains resulted in zoonotic infections and followed by human infection.³ The pandemics of IAV H1N1 2009 was originated from the genetic translocation of three different species: human H3N2, avian H1N1 and swine H1N1.⁴ Since March 2013, China has been faced with epidemics fifth times of human infection from novel avian H7N9 caused by the multiple re-assortment of at least four avian IAV and avian H9N2.⁵⁻⁷

IV has two important glycoproteins on their surface: hemagglutinin (HA) and neuraminidase (NA) revealing virus replication (Figure 2). IAV are categorized into different subtypes based on a combination of these two glycoproteins.⁸ Both

glycoproteins recognize the oligosaccharides of host cell; however, their functions are contradictory. HA has a receptor-binding function to initiate into the host cell, whereas NA has a receptor-destroying function to cleave the terminally bound SIA of glycans on host cell surface and viral envelop in order to spread of new virions and also prevent their aggregation on host cell surface.^{9, 10}

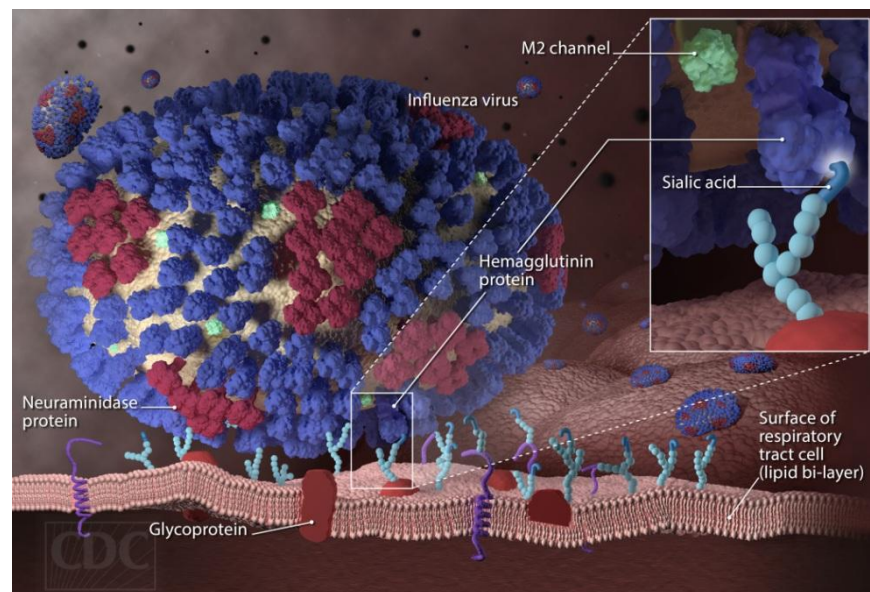


Figure 2 Hemagglutinin and neuraminidase glycoproteins on cell membrane of influenza virus that interact with terminal sialic acid of glycoprotein of host cell

The oligosaccharide substrate is classified into two types based on glycosidic linkage: SIA-2,3-GAL-1,4-NAG (3SL) and SIA-2,6-GAL-1,4-NAG (6SL). The 6SL is mainly found in human epithelial cell surfaces, called as human receptor, whereas the human lower respiratory tract and avian epithelial cell surfaces are predominant by the 3SL, referred as avian receptor.^{11, 12} The respiratory tracts of swine have both 3SL and 6SL, which can be infected by human and avian IV as intermediate host,^{13, 14} while the ferret's respiratory tract is used to represent a model of human respiratory tract. The distribution of receptors in respiratory tract for different hosts is shown in Figure 3.

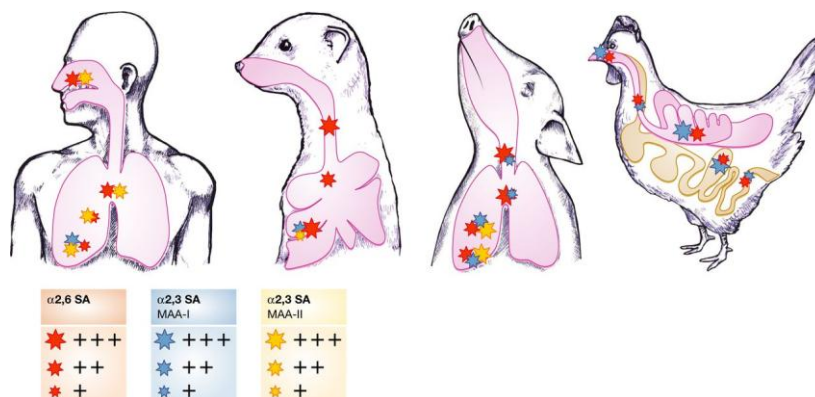


Figure 3 Distribution of 3SL and 6SL receptors found in respiratory tract of human, ferret, swine and chicken ¹⁵

The substrate specificity of HA from different host origins is well investigated. HA from human and avian IV preferentially recognize to 6SL and 3SL, while HA from swine viruses can bind both linkages.^{13, 16-20} The different preference is a barrier for interspecies transmission. A change of receptor preference could be the source of a new pandemic.²¹ In contrast, the substrate specificity of NA is less elaborated and remains to be elucidated. Experimentally, NA can cleave both types of linkage but shows preference for cleaving 3SL rather than 6SL; however, NAs isolated from human viruses show the increased substrate specificity of 6SL compare to avian isolated NA.^{10, 22-27} Hence, many researchers believe that the HA binding affinity and the NA enzymatic activity of SIA-linked oligosaccharides could be balanced for the efficient replication and transmission.^{2, 28, 29}

NA has been the currently successful anti-influenza target for influenza treatment. NA inhibitors (NAI) were designed based on the transition-state analogue of NA cleavage mechanism.³⁰⁻³² NA acts as enzyme to cleave off the glycosidic linkage between terminal sialic acid and the adjacent sugar to release new virions infected other host cells. The NA cleavage mechanism is introduced from the substrate binding into the its pocket involving a structural change of SIA ring to active conformation, followed by the cleaving off glycosidic bond between terminal SIA and adjacent galactose with a consequence of hydrolysis reaction of sialoyl cation intermediate (rate determining step), and finally the mutarotation from α -anomer to β -anomer of cleaved sialic acid in the releasing process, illustrated in Figure 4.³³⁻³⁷ In

the absence of NA, the IV cannot complete the viral life cycle and releasing of progeny virions by aggregating on host cell surface.

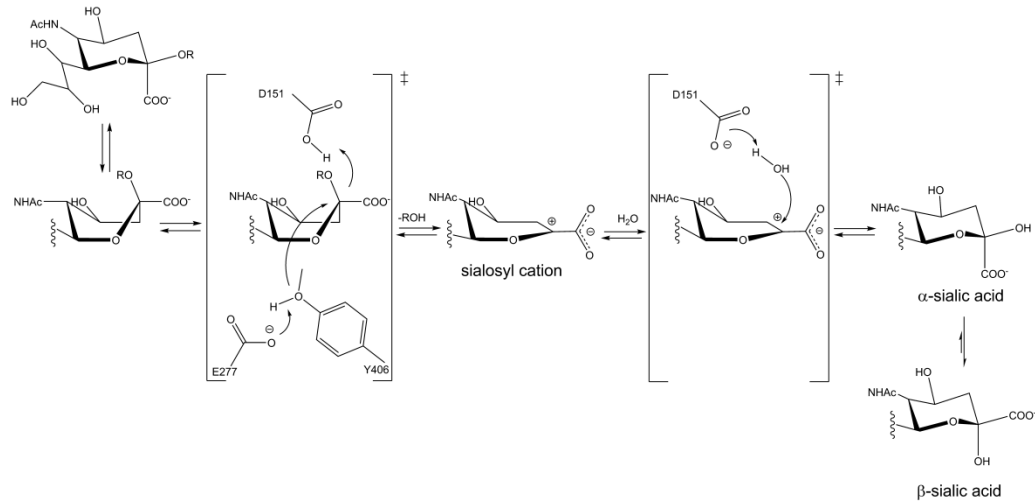


Figure 4 Proposed cleavage mechanism of influenza virus neuraminidase

IAV NAs can be classified into two distinct groups based on conformational differences: group-1 (N1, N4, N5 and N8 subtypes) and group-2 (N2, N3, N6, N7 and N9),³⁸ while N10 and N11 have recently found in bats.³⁹ NAs contain many charged amino acids such as R118, E119, D151, R152, E276, E277, R224, R292, R371 for binding to negatively charged SIA showed in Figure 5.

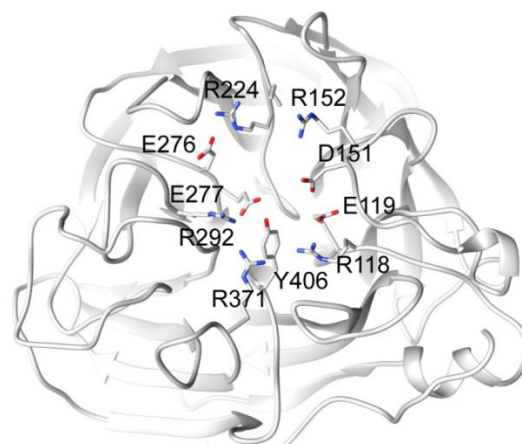


Figure 5 Highly charged binding pocket of influenza NA

The important emerging IAVs, which lead to several outbreaks in human, are H1N1, H2N2, H3N2, H5N1, and H7N9. Although the sequence identities among N1,

N2 and N9 NAs are less than 50% (Table 1), the amino acid residues embedded in the active site are mostly conserved (Table 2). The difference within binding site has mainly been found at residue 347, which involved with galactose unit of substrate and a conserved Ca^{2+} ion. Using sequence analysis, Y347 is highly conserved in N1 NAs of avian viruses, in particular for H5N1 subtype, whilst N347 is mainly observed in that circulating in human populations (*e.g.* pandemic H1N1). This might trigger the difference in the molecular recognition regarding to both types of substrate. Unfortunately, the systematic experimental and computational studies on NA specificity from those pandemic IV have never been carried out. Hence, the *in silico* investigation on the substrate binding and specificity of the emerging IV NAs from human (pandemic H1N1, H2N2 and H3N2) and avian (H5N1 and H7N9) strains was the aim of this work.

Table 1 Sequence alignment and structural superimposition of emerging N1, N2 and N9 NAs

%identity	H1N1	H5N1	H2N2	H3N2	H7N9
H1N1(3TI6.pdb)	-	0.288 Å*	0.752 Å	0.783 Å	0.763 Å
H5N1(2HU4.pdb)	91.2%	-	0.780 Å	0.790 Å	0.782 Å
H2N2(3TIC.pdb)	45.7%	46.7%	-	0.355 Å	0.748 Å
H3N2(4GZW.pdb)	45.3%	45.4%	85.3%	-	0.784 Å
H7N9(4MWQ.pdb)	47.7%	48.2%	49.0%	46.4	-

Table 2 Comparison of catalytic site and frame work site residues between H1N1, H5N1, H2N2, H3N2 and H7N9 NAs

N2 numbering	118	119	151	152	156	178	179	198	222	224	227	274	276	277	292	294	347	371	406
H1N1 (3TI5.pdb)	R	E	D	R	R	W	S	D	I	R	E	H	E	E	R	N	N	R	Y
H5N1 (2HU4.pdb)	R	E	D	R	R	W	S	D	I	R	E	H	E	E	R	N	Y	R	Y
H2N2 (3TIC.pdb)	R	E	D	R	R	W	S	D	I	R	E	H	E	E	R	N	Q	R	Y
H3N2 (4GZW.pdb)	R	E	D	R	R	W	S	D	I	R	E	H	E	E	R	N	H	R	Y
H7N9 (4MWR.pdb)	R	E	D	R	R	W	S	N	I	R	E	H	E	E	R	N	N	R	Y

1.3 Mutation conferring receptor recognition

The receptor specificity of HA is also responsible for determining the host recognition and being a species barrier of transmission. Avian IV needs ability to bind with human receptor and loses affinity with avian receptor in order to transmit from human to human. However, a few changes in amino acid of HA can cause a switch of receptor recognition. The human viruses of H1N1 in 1918 and 2009, H2N2 in 1957, and H3N2 in 1968 strains were derived from the avian viruses. The double E190D and G225D mutations were found in H1 from 1918 and 2009 H1N1 strains, while the single Q226L and G228S mutations were required for H2 and H3 HAs.^{11, 18, 20, 40} The introduction of Q226L and G228S mutations into H5N1 HA have ability to bind with 6SL;⁴¹ however, this variant was not found in H5N1 viruses isolated from human. In contrast, the natural substitutions of Q226L and G186V were found in the recent H7N9 HA, which is able to bind with both receptors compared to the avian receptor preference of its wild-type, indicating the change of receptor-binding properties.^{42, 43} Meanwhile, the HA from novel H6N1 subtype avian IV associated with the recombinant D190V and G228S substitutions still has preference toward 3SL.⁴⁴

Interestingly, the mutations associated the change in receptor-binding specificity could be identified to discover the possibility of transmission in mammals. For instance HA, the single mutations of N182K, Q192R, and N193K in H5N1 enhanced the recognition of 6SL.⁴⁵ The previous studies have confirmed that the introduction of site-directed mutagenesis in H5N1 HA was conducted then passed in ferrets, which represent a human receptor, to become transmissible in mammals.⁴⁶⁻⁴⁸ The amino acid mutations that enable to preferentially bind with human receptor are N158D, N224K, Q226L, and T381I isolated from Vietnam and H110Y, T160A, Q226L, and G228S from Indonesian strain.

For NA, some sequences isolated from human H3N2 contained the D151G mutation, which has ability to bind with red blood cells.⁴⁹ This mutation of NA was crystallized as the N2 NA complex with sialylated glycan.⁵⁰ The G147R substitution isolated from H1N1 and H5N1 led to enhance the binding affinity with their substrates, approximately 2-fold compared with the wild-type.⁵¹ In addition, the crystal structure of H7N9 NA has a secondary binding, which enhance the receptor

binding for both linkages. However, this N9 NA poorly cleaves the human receptor resulting in an inefficiency releasing after the viral replication.⁵² We expect that the crucial residues obtained from substrate/NA interactions could be further studied by mutagenesis for discovering the host jump ability of human-to-human transmission.

1.4 Neuraminidase inhibitors and resistance mutation

Currently, the commercially available NAI are zanamivir, oseltamivir, peramivir and laninamivir (Figure 6). However, the widespread treatment of NAI leads to the mutation of virus. For instance the novel H7N9 IAV caused the epidemic in China, the sequence analysis of NA have been shown to contain the R292K substitution after oseltamivir treatment such as in A/Shanghai/1/2013.^{6, 53} The arginine residue 292 (R292) is one of the arginine triad members (R118, R292 and R371 in N2 numbering) located in the NA active site and directly interacts with the carboxylate group of SIA or with other drugs that bind in a similar fashion to SIA (Figure 7).

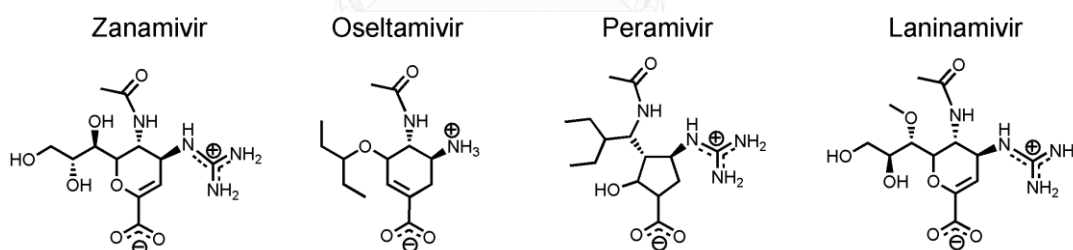


Figure 6 Two-dimensional chemical structures of NAI including zanamivir, oseltamivir, peramivir and laninamivir

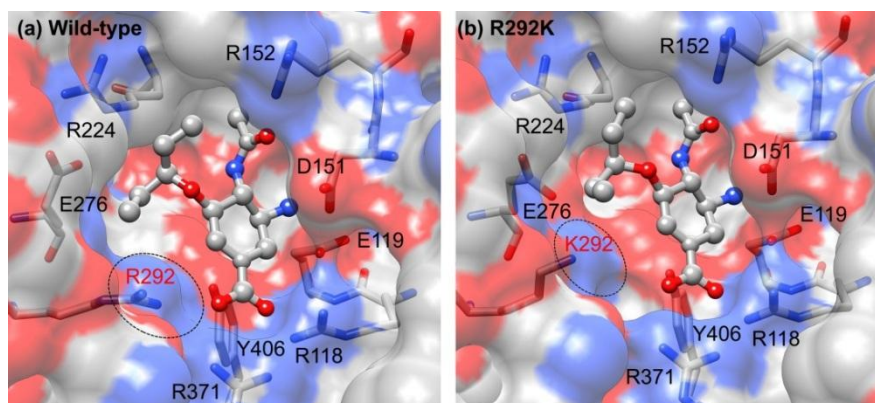


Figure 7 Structure of oseltamivir binding in the H7N9 NA active site (a) wild-type and (b) R292K substitution which results in multidrug resistance

In general, the R292K substitution, which causes the resistance of multi NAI, is mostly found in the N2 and N9 NA subtypes.^{54, 55} The novel H7N9 containing R292K mutation can be treated by NAI; however, the protein- and virus-based assays clearly indicated drug resistance to oseltamivir and peramivir and a mild resistance to zanamivir and laninamivir, as listed in Table 3.⁵⁶⁻⁵⁹

Table 3 Drug susceptibilities of plaque-purified H7N9 viruses and fold resistance of R292K substitution compare to those of wild-type

NAI	NA change	<i>IC</i> ₅₀ (nM)			
		Yen et al.*	Sleeman et al.**	Wu et al.***	Hai et al.***
Oseltamivir	Wild-type	1.0	0.3	1.1	1.9
	R292K	> 100.0	4,987.0	149,467.0	8,620.0
	Fold resistance	> 100.0	16623.3	135879.1	4536.8
Zanamivir	Wild-type	1.8	0.5	1.3	3.8
	R292K	52.0	23.5	58.5	40.1
	Fold resistance	28.9	47.0	45.0	10.6
Peramivir	Wild-type	-	0.1	0.4	0.3
	R292K	-	101.9	165.0	191.0
	Fold resistance	-	1019.0	412.5	636.7
Laninamivir	Wild-type	-	0.4	2.0	-
	R292K	-	15.4	53.4	-
	Fold resistance	-	38.5	26.7	-

*Plaque-purified virus ** Recombinant (100% of K292) *** Reverse genetic

Of particular importance is oseltamivir as it is used frequently in clinical treatment. Hence, this work will also clarify the role of R292K mutation in H7N9 toward reduced susceptibility of oseltamivir compared with its wild-type strain.

1.5 Literature reviews

1.5.1 Substrate specificity of IV NA

The NA activities toward 3SL and 6SL receptors have been characterized by several methods such as BODIPY-labeled substrate, glycan array and library screening format to differentiate the substrate binding.^{23, 50, 60, 61} Although NA can cleave both types of linkages; the difference in NA cleavage activity from different hosts was characterized. The avian and early human N2 NAs showed higher activity on 3SL than 6SL; however, an increase of 6SL specificity to match with their HA was detected in the later human N2.²³ Similarly, H1N1 NAs from different origins exhibit the preference ratio between 3SL and 6SL is of 50:1, 20:1 and 5:1 for avian, swine and human, respectively.²⁶ Although H5N1 NAs from avian IV show high specific with 3SL, the 3SL/6SL ratios of substrate preference between avian and human origins are not significantly different.¹⁰ These suggest that NA function could be match to its HA for optimal viral infection.

The binding of substrate into NA pocket is the first step of cleavage mechanism. The previous study of NMR spectroscopy confirmed that the cleavage efficiency of different N1 strains (H5N1 and H1N1 from and 2009) is dependent to the substrate binding affinity.²⁷ This suggests that the strength of binding affinity directly affects the cleavage activity of NA. Hence, the high binding affinity of substrate within NA binding pocket might be resulted to the efficiency NA cleavage activity.

MD simulations of avian H5N1 NA in complex with 6SL and 3SL provided that the 6SL is more flexible than 3SL regarding the different binding modes.⁶² The simulations of N1 NAs from avian H5N1 and human H1N1 demonstrate distinguish substrate selectivity by a critical hydrogen bonding with different residues: Y347 of avian H5N1 and Q430 of human H1N1 for recognition of 3SL and 6SL,

respectively.⁶³ However, these previous studies were carried out using docked substrate/NA complexes and mainly focused on the N1 subtype without considering the binding free energy prediction. In 2012, the crystal structures of substrate in complex with D151G mutation of human H3N2 NA were firstly crystalized.⁵⁰ Therefore, this work aims to investigate the substrate binding and specificity of emerged NAs (pandemic H1N1, H2N2, H3N2, H5N1 and H7N9) based on the crystal structures by performing MD simulation and binding free energy calculation of QM/MM-GBSA method.

1.5.2 Reduced susceptibility of NAI conferring NA mutation

In case of R292K mutation in H7N9 NA, MD simulations of drug-H7N9 NA complexes have previously been carried out by others based on homology modeling and their predicting binding affinities.^{64, 65} Later, the crystal structures of H7N9 NA in complex with NAI have been available for both wild-type and mutant strains.⁵⁸ In addition, the previous simulations do not clarify the effect of R292K NA mutation in term of solvation.

In general, molecular mechanics combined with Poisson-Boltzmann or generalized Born electrostatics and surface area nonpolar solvation, MM/PB(GB)SA approaches, has been popular for predicting the ligand binding affinity, however, these methods are based on implicit solvent model, which cannot provide the solvation structure. Instead, three-dimension reference interaction site model (3D-RISM) complemented with Kovalenko-Hirata (KH) closure is a molecular theory of solvation operating the explicit water model around solute using statistical mechanics calculation, which has been deeply described elsewhere.⁶⁶⁻⁷⁰ Subsequently, the combination of molecular mechanics and 3D-RISM, MM/3D-RISM method, is reliable useful to discovery the phenomena of solvation. The oseltamivir binding affinity toward wild-type and R292K mutation will be predicted and compared by different methods of MM/3D-RISM and MM/PB(GB)SA.

1.6 Scope of this work

(i) to investigate *in silico* the substrate binding specificity to the emerging NAs of H1N1, H2N2 and H3N2 human IV and H5N1 and H7N9 avian IV by means of MD simulation and QM/MM-GBSA binding free energy calculation.

(ii) to understand the oseltamivir binding to the novel H7N9 NA and provide insight into the role of the R292K mutation using MD simulations and three different binding free energy calculations of MM/PB(GB)SA and MM/3D-RISM.



CHAPTER II THEORETICAL BACKGROUND

2.1 Computational simulation methods

The simulation approaches are commonly used in computational chemistry to calculate the properties of studied molecules in a variety of manners by solving chemical problems through the cooperating of theoretical chemistry methods and efficient computer programs. A modeled structure directly relevant to the energy, must be sufficient to represent the investigated calculations for both accuracy and time-consuming. There are three main types of models, which represent the interested energy: molecular mechanics (MM), quantum mechanics (QM), and hybrid quantum mechanical/molecular mechanical (QM/MM) methods. The MM considers a molecule as a charged sphere connected together with springs that allows to observe the conformational dynamics. On the other hand, the QM describes a molecule as a wave function by using the Schrödinger equation, which allows to predict a change of electronic structure such as chemical reaction (bond breaking and bond forming), electron transfer, charge densities *etc.* However, the QM approach is incompatible for a large system according to the time consuming. To study a chemical reaction in biological system, the combined QM/MM method is popularly used, where a whole system is divided into two zones; the region of chemical interest is treated with QM and the rest of complex is calculated using MM. The important thing in computational chemistry is the reliable model, hence the sufficiently detailed model should be considered for investigating the interested problem. In this research, we mainly explore the ligand binding pattern and binding affinity toward the neuraminidase protein (nearly 400 residues). Therefore, our model system is entirely based on MM, which is deeply described in the next section.

2.2 Molecular mechanics potential function

Molecular mechanics (MM) model considers a molecule as balls connected by spring. Each atom is represented as charge spheres of a given radius. The

conformational changes of molecule are sourced by stretching of bond, bending of angle, rotating of dihedral angles. The MM potential energy (U) of system is the sum of the bonded (U_{bonded}) and non-bonded ($U_{non-bonded}$) interactions presented as a function. The simplest MM potential function is the sum of the bond stretching (U_{bonds}), bond bending (U_{angles}), torsion angle ($U_{dihedrals}$), and the non-bonded interaction of electrostatic (U_{ele}) and van der Waals (U_{vdW}) terms (eq.1-2).

$$U = U_{bonded} + U_{non-bonded} \quad (1)$$

$$U = \left(U_{bonds} + U_{angles} + U_{dihedrals} \right)_{bonded} + \left(U_{ele} + U_{vdW} \right)_{non-bonded} \quad (2)$$

The calculation of potential energy can be carried out using force field parameters, equilibrium parameters, and constant values, which are specific to the atom types. The force field parameters were derived from both experimental data and *ab initio* calculations. There are many force fields that can be used to represent biomolecules such as AMBER, CHARMM, OPLS, GROMOS *etc.* Using the force field parameters, the coordination of the initial structure is an input for computing the potential energy function. It is worth to note that the obtained value is the relative energy, not exact energy of system.

2.2.1 Bonded interactions

In term of bonded interactions, the geometry change of molecule was considered through harmonic oscillator functions including bond, angle, and dihedral angle. All pair atoms are computed the bond stretching (eq.3); where k_b is a bond parameter and r_0 is an equilibrium bond distance.

$$U_{bonds} = \sum_{bonds} k_b (r - r_0)^2 \quad (3)$$

Similar to the bond stretching, the bond bending is described by the change in angle between three atoms (eq.4); where k_θ is an angle parameters and θ_0 is an equilibrium angle

$$U_{angles} = \sum_{angles} k_\theta (\theta - \theta_0)^2 \quad (4)$$

The rotating of dihedral angles relevant with four atoms is assigned as periodic function from 0 to 360 degree (eq.5). The parameters of dihedral function are rotational barrier height (V_n), periodicity of rotation (n), dihedral angles in equilibrium (ϕ) and in radians (τ). The 1-5 intramolecular interactions are assigned as non-bonded interactions.

$$U_{dihedrals} = \sum_{dihedrals} \frac{V_n}{2} (1 + \cos [n \tau - \phi]) \quad (5)$$

2.2.2 Non-bonded interactions

By considering a molecule as point charges connected by springs, the partial atomic charge (q) of atom is introduced. The electrostatic interaction of all pairs between i and j atoms is derived from the Coulombic potential function; where ϵ is a dielectric constant and r_{ij} is a distance of each pair of atoms (eq.6).

$$U_{ele} = \frac{1}{4 \pi \epsilon} \sum_{i < j}^{atoms} \frac{q_i q_j}{r_{ij}} \quad (6)$$

On the other hand, the dispersion force of nonpolar interactions is approximated by using the Lennard-Jones potential function (eq.7). The Lennard-Jones potential assigns all pairs of i and j atoms separated by a distance (r_{ij}), while the ϵ and σ parameters are Lennard-Jones well depth and collision diameter depended on the specific type of pair atoms. The repulsive term is denoted by $1/r^{12}$, while the attractive term is presented by $1/r^6$.

$$U_{vdW} = \sum_{i < j}^{atoms} 4 \epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \quad (7)$$

However, the calculation of the non-bonded interaction is the most computational consuming by considering over all atom pairs with N^2 , where N is the number of atom. There are several approaches to reduce this bottleneck calculation. The treatments of non-bonded interaction employed in this work are **(i) cutoff distance**: this technique calculates interactions within the cutoff radius and neglects the rest interactions as zero based on the assumption that the interaction

potential energy will be converged to zero at a specific distance, (ii) **Periodic Boundary Condition (PBC)**: a system in a regular-shaped box is treated as infinite, which is surrounded with its identical systems in all spatial directions to avoid the problem of the boundary effects. For instance, a particle, which goes out from the simulation box, will be replaced by a particle in the neighboring box as depicted in Figure 8. There are various shapes of PBC box such as cubic, orthorhombic, parallelepiped, truncated octahedral, rhombic dodecahedral *etc.*, (iii) **Particle Mesh Ewald (PME)**: the PME is taken into account for computing the electrostatic interaction using the Fourier transform technique, which is faster than the Ewald summation.⁷¹

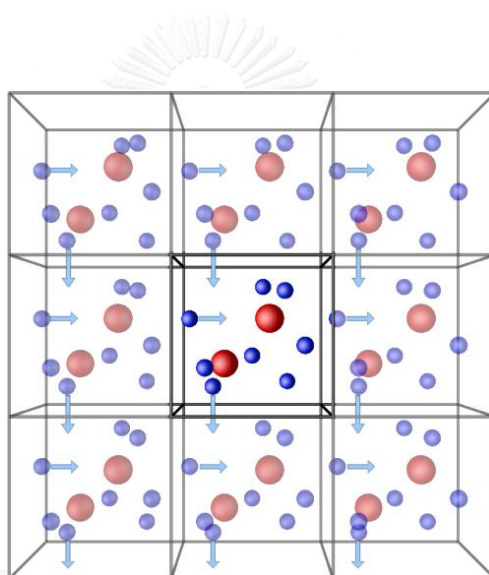


Figure 8 Illustration of Periodic boundary condition
(<http://isaacs.sourceforge.net/phys/psc.html>)

2.3 Solvation treatment

Most of biological reactions occur in aqueous solution, thus, the solvent in particular water plays an important role in the processes. In enzymatic reaction, water in binding pocket is essential for cleavage mechanism. In molecular simulation of the biological system, the aqueous environment cannot be ignored. There are three ways to represent the solvent: explicit solvation, continuum solvation and 3D-RISM approaches.

2.3.1 Explicit solvation

The explicit solvation represents the three-dimensional (3D) structure of water molecules around the solute system. The presence of water molecules can provide physical insight into the molecular structure such as hydrogen bonding formation, steric effect; however, the computational cost is also dependent on the amount of water molecules. Based on the MM model, the water molecule consists of a point charge held together by a spring. The charged site can be denoted at atom and lone pair electrons to exhibit the non-bonded interactions of the electrostatic and dispersion forces by the previous equations resulting in the different shapes of water models such as 3-site, 4-site, 5-site and, 6-site, depicted in Figure 9. The geometric parameters of water model including O-H distance, H-O-H angle, charge of each atom and Lennard-Jones parameters are widely available. The 3-site water models having three interaction points over three atoms of a molecule achieve an efficiency calculation which is widely used in MD simulations such as TIP3P⁷², SPC⁷³ and SPC/E⁷⁴.

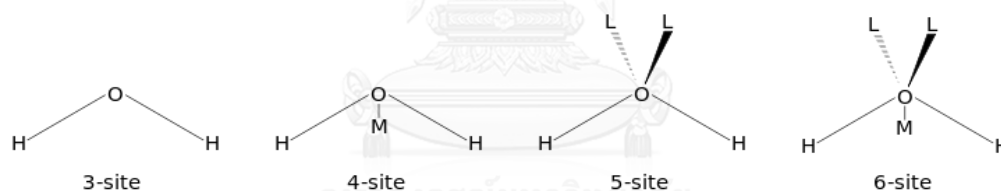


Figure 9 Illustration of different sites (3-, 4-, 5- and 6-sites) of water model (https://en.wikipedia.org/wiki/Water_model)

2.3.2 Continuum model

The continuum model treats the solvation as an electrostatic interaction of dielectric media and a solute molecule using a specified dielectric constant to represent the average properties of the real solvent, which is unique for each solvent. The total solvation free energy is the combination between the non-polar ($\Delta G_{solv}^{non-polar}$) and electrostatic (ΔG_{solv}^{ele}) contributions (eq.8).

$$\Delta G_{solv} = \Delta G_{solv}^{non-polar} + \Delta G_{solv}^{ele} \quad (8)$$

The implicit solvation considers that the continuous medium is disturbed by a cavity formation of a solute molecule. The favorable dispersion interaction and the unfavorable cavity formation between solute structure and solvent environment are generally assigned as the non-polar solvation energy. The $\Delta G_{solv}^{non-polar}$ contribution can be approximated by the linear function of the solvent accessible surface area (SASA) approach (eq.9).⁷⁵ The SASA is determined by using an exact radius of solvent as a probe rolling on the van der Waals surface of a solute molecule, whilst the ΔG_{solv}^{ele} can be calculated using the two well-known approaches: Poisson-Boltzmann equation and generalized Born model, which are deeply described as below.

$$\Delta G_{solv}^{non-polar} = \Delta G_{cavity} + \Delta G_{dispersion} \approx \gamma SASA + \beta \quad (9)$$

where γ and β are surface tension and offset values.

The SASA is determined by using an exact radius of solvent as a probe rolling on the van der Waals surface of solute molecule. Whilst the ΔG_{solv}^{ele} can be calculated using two well-known approaches: Poisson-Boltzmann equation and generalized Born model which deeply described below.

Poisson-Boltzmann (PB) model

For Poisson's equation, the electrostatic potential, $\varphi(r)$, is computed as a function of charge distribution, $\rho(r)$, and position-dependent dielectric constant, $\varepsilon(r)$, (eq.10-11).

$$\nabla \varepsilon(r) \nabla \varphi(r) = -4\pi\rho(r) \quad (10)$$

$$\nabla^2 \varphi(r) = \frac{-4\pi\rho(r)}{\varepsilon} \quad (11)$$

However, this equation is valid under the absence of mobile ions. For the existing of electrolyte in solvation, the Boltzmann distribution is employed to describe the distribution of mobile ions inside the potential field leading to Poisson-Boltzmann (PB) equation. The linear PB equation is simplified for applying in biomolecular

simulations (eq.12), where q_i and n_i is the atomic charge of electrolyte and density of each ion, k_B is the Boltzmann constant and T is temperature.

$$\nabla \varepsilon(r) \nabla \varphi(r) = -4\pi \left[\rho(r) + \sum_i q_i n_i \exp\left(-\frac{q_i \varphi(r)}{k_B T}\right) \right] \quad (12)$$

For linear PB equation, the ΔG_{solv}^{ele} is the difference of the electrostatic potentials between solution φ_{sol} and vacuum φ_{vac} phases (eq.13) at every time the conformation of molecule changes.

$$\Delta G_{solv}^{ele} = \frac{1}{2} \int \rho(r) (\varphi_{sol} - \varphi_{gas}) dr \quad (13)$$

Generalized Born (GB) model

The generalized Born (GB) model is a spherical cavity model approximation, which is more simplicity and computational efficiency compared to the PB calculation. On the basis of GB model, each atom in a molecule is represented as a sphere of radius R_i and a charge q_i having lower dielectric constant than environment. In AMBER program, the interior atom uniformly defined by a dielectric constant of 1 is surrounded by the high dielectric constant of solvent ($\varepsilon = 80$ of water). The GB model calculates the ΔG_{solv}^{ele} (eq.14)^{76,77}

$$\Delta G_{solv}^{ele} = -\frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{GB}(r_{ij}, R_i, R_j)} \left(1 - \frac{\exp(-\kappa f_{GB})}{\varepsilon} \right) \quad (14)$$

where r_{ij} is the distance between atoms i and j , R_i is the effective Born radii, f_{GB} is a smooth function interpolating the atomic radii and distance between pair of atoms (eq.15).

$$f_{GB} = \left[r_{ij}^2 + R_i R_j \exp\left(\frac{-r_{ij}^2}{4 R_i R_j}\right) \right]^{1/2} \quad (15)$$

The effect of the salt concentration can be evaluated by the Debye-Hückel screening parameter, κ . For a single ion, the atom radius is equal to its van der Waals radius, ρ_i . In pure water ($\kappa = 0$), the ΔG_{solv}^{ele} of a single ion is written (eq.16).

$$\Delta G_{solv}^{ele} = -\frac{q_i^2}{2\rho_i} \left(1 - \frac{1}{\epsilon} \right) \quad (16)$$

The effective radii is approximated equal to the distance from atom to the molecular surface. Hence, the conformational change of molecule directly affects the effective radii leading to a problem of calculation. In the case of atom is surrounded by the others, its electrostatics will be screened according to the low dielectric constant of environment, instead of high dielectric of solvent, called as de-screening. Hence, the effective radii of each atom can be determined from the degree of de-screening. The different atomic de-screening leads to many GB models. In AMBER, the available GB models assigned by Igb value as summarized in Table 4.

Table 4 Summary of various GB models in AMBER program

GB model	GB ^{HCT}	GB ^{OBC}	GB ^{OBC2}	GB ^{Neck}	GB ^{Neck2}
Igb value	1	2	5	7	8
Calculation of effective Born radii	Pairwise approximated integration	Re-scaling of radii		Adding neck regions to HCT model	
Reference	Hawkins et al. ⁷⁸	Onufriev et al. ⁷⁹		Mogan et al. ⁸⁰	

GB^{HCT}: The 3D integral is carried out over the van der Waals spheres of solute atoms resulting in the solute volume, instead of molecular surface in PB calculation. It should be noted that this method give underestimate the effective radii for buried atoms. The assumption of the pairwise de-screening approach can be expressed (eq.17)

$$R_i^{-1} = \tilde{\rho}_i^{-1} - I \quad (17)$$

where

$$I = \frac{1}{4\pi} \int \theta(|\vec{r}| - \tilde{\rho}_i) \frac{1}{r^4} d^3\vec{r} \quad (18)$$

$\tilde{\rho}_i = \rho_i - 0.09 \text{ \AA}$ and θ is the step-function which excluded the volume of atom i from the integration.

GB^{OBC}: The effective Born radii are re-scaled to account for the small space between van der Waals spheres of protein atoms that can be filled with water molecules. For this model, the rescaling function was included to above equation (eq.19), where α, β, γ are adjustable parameters.

$$R_i^{-1} = \tilde{\rho}_i^{-1} - \rho_i^{-1} \tanh(\alpha I \tilde{\rho}_i - \beta (I \tilde{\rho}_i)^2 + \gamma (I \tilde{\rho}_i)^3) \quad (19)$$

There are two sets of $\{\alpha, \beta, \gamma\}$: 0.8, 0.0, 2.91 for GB^{OBC} and 1.0, 0.8, 4.85 for GB^{OBC2}. However, the rescaling function of OBC model only affects the sufficiently buried atoms.

GB^{Neck}: This model is to eliminate the interstitial region of high dielectric smaller than a solvent molecule called neck region (depicted in Figure 10) by adding a molecular volume correction based on geometric to the underestimate approximation of GB^{HCT}.

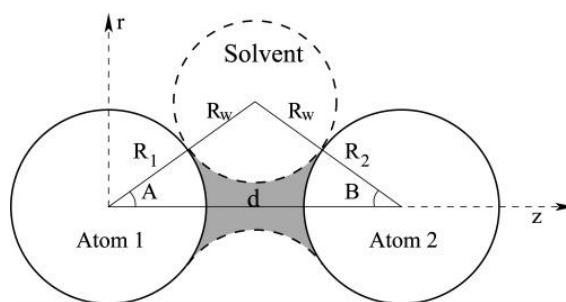


Figure 10 Presentation of neck region (shaded area) and illustration of its geometry defining by the atoms radius R_1, R_2 , its separated distance d and radius of solvent molecule R_w ⁸⁰

The correction region can be approximated by a sum of the neck regions based on pairwise of atoms. There are two steps of scaling calculation. First, each neck integral is evaluated (eq. 20)

$$neck_integral(d) = \frac{m_0}{1 + (d - d_0)^2 + 0.3(d - d_0)^6} \quad (20)$$

where d_0 is position and m_0 is the maximum value determined from the optimization over the neck region. Afterward, the calculation of effective radii is carried out using eq.18 of OBC model. The different parameters of $\{\alpha, \beta, \gamma\}$ lead to two options of neck model: GB^{Neck} and GB^{Neck2} . However, this model is not recommended for nucleic acids systems.

2.3.3 3D-RISM approach

Simulation of explicit solvent can illustrate the exist interaction of solvent and solute; however, this explicit sampling would be long enough (maybe until microsecond) to converge the flux of solvent throughout the biomolecular interior. Whilst the continuum model of PB and GB does not provide the molecular view of solvent and cannot exactly define the dielectric constant of solute molecule. Interestingly, the attractive solvent model is three-dimensional reference interaction site model (3D-RISM). The 3D-RISM is a statistical mechanics theory of liquids calculating the equilibrium distribution of solvent around solute molecule,⁸¹ which derived from the principle of the molecular Ornstein-Zernike (MOZ) equation.

In generally, the structure of liquid can be determined by using radial distribution function (RDF), $g(\bar{r})$ which is the probability of finding a particle at the position (\bar{r}) from the interested particle compared to that of an uncorrelated particle . Based on MOZ equation, the total correlation function $h(\bar{r}_{12}, \Omega_1, \Omega_2)$ can be decomposed into direct correlation function $c(\bar{r}_{12}, \Omega_1, \Omega_2)$ and indirect correlation function (eq.21)

$$h(\bar{r}_{12}, \Omega_1, \Omega_2) = c(\bar{r}_{12}, \Omega_1, \Omega_2) + \rho \int c(\bar{r}_{13}, \Omega_1, \Omega_3) h(\bar{r}_{32}, \Omega_3, \Omega_2) d\bar{r}_3 d\Omega_3 \quad (21)$$

where \bar{r}_{12} is the distance, Ω_1 and Ω_2 are the angular orientation of particles 1 and 2. The definition of total correlation function is $h(\bar{r}_{12}, \Omega_1, \Omega_2) = g(\bar{r}_{12}, \Omega_1, \Omega_2) - 1$. The MOZ equation can be written in terms of an infinite series in which the direct correlation function expanded in the power of ρ (eq.22).

$$h(\bar{r}_{12}, \Omega_1, \Omega_2) = c(\bar{r}_{12}, \Omega_1, \Omega_2) + \rho \int c(\bar{r}_{13}, \Omega_1, \Omega_3) c(\bar{r}_{32}, \Omega_3, \Omega_2) d\bar{r}_3 d\Omega_3 + \rho^2 \int c(\bar{r}_{13}, \Omega_1, \Omega_3) c(\bar{r}_{34}, \Omega_3, \Omega_4) c(\bar{r}_{42}, \Omega_4, \Omega_2) d\bar{r}_3 d\Omega_3 d\bar{r}_4 d\Omega_4 + \dots \quad (22)$$

In case of homogenous and isotropic, the molecule does not depend on its orientation, the equation can be written in real space, r (eq.23).

$$h(r) = c(r) + \rho \int c(|r - r'|) h(r') dr' \quad (23)$$

Using Fourier transform, this equation is simplified in reciprocal space (k -space) (eq.24).

$$\hat{h}(k) = \frac{\hat{c}(k)}{1 - \rho \hat{c}(k)} \quad (24)$$

However, the solution of MOZ equation cannot achieve since containing two unknown functions. Hence, the closure relations have been proposed to related two functions such as Hypernetted-chain (HNC), Percuc-Yevick (PY), Kovalenko and Hirata (KH) and Partial series expansion of order- n (PSE- n) etc. The general closure relation to the MOZ equation is (eq.25)

$$g(r_{12}, \Omega_1, \Omega_2) = \exp\left[-\frac{u(\bar{r})}{k_B T} + h(r_{12}, \Omega_1, \Omega_2) - c(r_{12}, \Omega_1, \Omega_2) + b(r_{12}, \Omega_1, \Omega_2)\right] \quad (25)$$

where $u(\bar{r})$ is interaction potential between two particles in a system, k_B is Boltzmann constant at temperature T and $b(r_{12}, \Omega_1, \Omega_2)$ is bridge function.

To apply with polyatomic molecules, the reference interaction-site model (RISM) has been proposed by Chandler and Andersen through averaging the angular coordinates of molecules and fixing the distance between a pair of interaction sites.⁸² The six-coordinates of solvent are generalized to one dimension called as 1D-RISM. The distribution function of solvent around solute molecule can be calculated by three-dimensional RISM (3D-RISM) by taking statistical average the angular coordinate of solvent. Similar to MOZ equation, the 1D-RISM and 3D-RISM equations need the closure relation to solve until self-consistency.

The procedure of 3D-RISM calculation is depicted as Figure 11. First, the site-site correlation functions between solvent molecules based on specified atomic interaction potential of water model was calculated as 1D-RISM. Then, the correlation functions, or susceptibilities, are utilized for solvent-solute interaction using 3D-RISM equations coupled with closure relations. This couple is iterated until reaching the convergence to obtain the equilibrium solvation properties including the 3D solvent distribution function $g(\vec{r})$, the total correlation function $h(\vec{r})$, the direct correlation function $c(\vec{r})$ and other solvent properties such as solvation free energy.

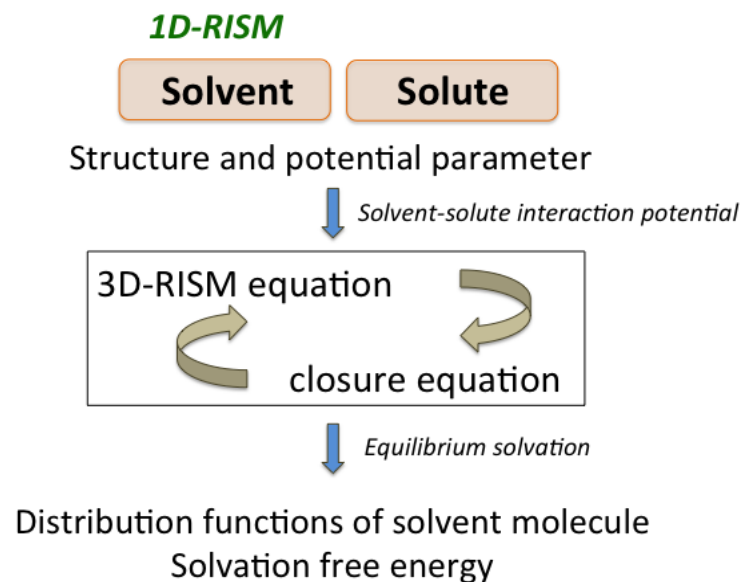


Figure 11 Scheme of 3D-RISM calculation

In this work, the KH closure was adopted to 1D-RISM and 3D-RISM calculations. The 3D-RISM equation with KH closure is shown as the following criteria (eq.26)

$$\begin{aligned}
 g(\vec{r}) &= \exp(-d(\vec{r})) & \text{for } d(\vec{r}) \leq 0 \\
 g(\vec{r}) &= 1 + d(\vec{r}) & \text{for } d(\vec{r}) > 0
 \end{aligned}
 \tag{26}$$

where $d(\vec{r}) = -\frac{u(\vec{r})}{k_B T} + h(\vec{r}) - c(\vec{r})$. Then, the solvation free energy can be computed

from the distribution function (eq. 27).

$$\Delta \mu_{KH} = \rho^v k_B T \sum_{\gamma} \int d\vec{r} \left[\frac{1}{2} h_{\gamma}^{uv}(\vec{r})^2 \Theta(-h_{\gamma}^{uv}(\vec{r})) - c_{\gamma}^{uv}(r) - \frac{1}{2} h_{\gamma}^{uv}(\vec{r}) c_{\gamma}^{uv}(\vec{r}) \right] \quad (27)$$

where u and v denote for solute and solvent molecule, γ labels atom types of model and Θ is the Heaviside step function.

2.4 Statistical mechanics

Statistical mechanics act as a bridge linked between thermodynamic of bulk properties and molecular structure or properties of individual atoms. A fundamental of statistical mechanics is Boltzmann distribution calculating the possibility p_i of a particular energy level E_i at a given temperature T in the state i (eq.28)

$$p_i = \frac{\exp(-E_i / k_B T)}{Q} \quad (28)$$

where k_B is the Boltzmann constant and Q is a partition function obtained from all possible configurations (eq.29).

$$Q = \sum \exp(-E_j / k_B T) \quad (29)$$

The partition function contains thermodynamics properties of system such as internal energy, entropy, heat capacity and so on. Once obtaining the partition function, the bulk properties of system can be predicted.

In simulation, Boltzmann distribution was operated for collecting all possible configurations called ensemble in a specific manner. The specific constraints and properties are applied to define the types of ensembles: canonical ensemble (NVT), grand canonical ensemble (μ VT), microcanonical ensemble (NVE) and isobaric-isothermal ensemble (NPT). For instance, canonical ensemble (NVT) fixes the total number of atoms (N), volume of system and temperature (T).

Thus, the important thermodynamics quantities in computational chemistry such as free energy and entropy can be computed through the partition function. The free energy of a system provides the kinds of change: spontaneous or non-spontaneous. The calculation of absolute binding free energy can be directly related to its partition function Q (eq.30)

$$G = -k_B T \ln Q \quad (30)$$

where G is Gibbs free energy for NPT ensemble.

Using Boltzmann, the entropy (S) of a system can be calculated from the expression (eq. 31)

$$S = k_B \ln W \quad (31)$$

where w is the number of different ways that molecules in a system can be arranged to give the same total energy. The entropy is zero when there is only one way for obtaining the same energy ($w = 1$). In contrast, the entropy is greater than zero when there are many ways to achieve a given total energy. To relate with partition function (q), the entropy can be written as (eq.32)

$$S = \frac{U - U(0)}{T} + Nk_B \ln q \quad (32)$$

where U and $U(0)$ are the internal energies at temperature of T and 0 K, N represents the number of molecules.

2.5 MD simulation

Once the initial structure and force field are available, the collection of structures can be generated using sampling methods. However, a simulation system is the microscopic world compared to an experiment, which is a macroscopic world. To represent the macroscopic properties or experimental observation, the ensemble average is equal to the time average of a property of a simulation system based on ergodic hypothesis.

MD simulation is one of the sampling methods, which are a great tool to study insight into structural dynamics and conformational change of biomolecules as well as protein-ligand interactions.⁸³⁻⁸⁹ The scheme of MD simulation is depicted in Figure 12.

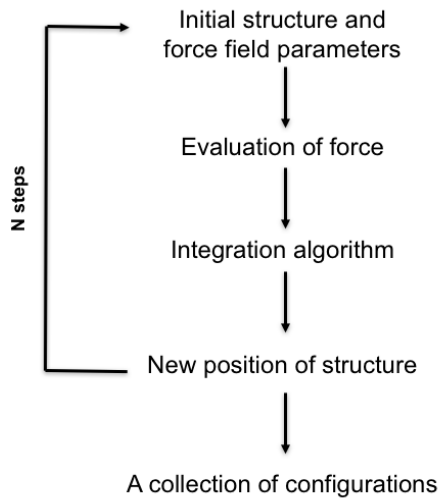


Figure 12 Scheme of MD simulation

The starting configuration of atomic coordinates as well as force field parameters are the raw materials for performing the MD simulation. The coordinate of bio-macromolecule generally obtains from the X-ray or NMR spectroscopy crystal structure or homology modeling technique. The initial velocity of atoms is generated using the Maxwell-Boltzmann distribution. MD simulation employs the Newton equations of motion to find the new position (\vec{r}_i) of atom i (eq.33)

$$\vec{F}_i = -\nabla_i U = m_i \frac{d\vec{v}_i}{dt} = m_i \frac{d^2\vec{r}_i}{dt^2} \quad (33)$$

where U represents the potential energy of molecule and \vec{F}_i , \vec{v}_i and \vec{r}_i are force velocity and position of each atom.

MD simulation generates the structures as the time-evolution yielding a trajectory of system. Hence, the integration algorithm is needed to predict the movement via a time step (δt). The integration algorithms are Verlet, Leapfrog and Beeman etc. The Leapfrog algorithm was used throughout this work as described below. The velocities are determined using a half of time step, $\frac{\delta t}{2}$ to improve the accuracy of simulation (eq.34-35).

$$v\left(t + \frac{\delta t}{2}\right) = v(t) + \left(\frac{dv}{dt}\right)\left(\frac{\delta t}{2}\right) + \frac{1}{2}\left(\frac{d^2v}{dt^2}\right)\left(\frac{\delta t}{2}\right)^2 \quad (34)$$

$$v\left(t - \frac{\delta t}{2}\right) = v(t) - \left(\frac{dv}{dt}\right)\left(\frac{\delta t}{2}\right) + \frac{1}{2}\left(\frac{d^2v}{dt^2}\right)\left(\frac{\delta t}{2}\right)^2 \quad (35)$$

After subtraction the above equations, the leapfrog equation in term of velocity and position are obtained in the similar way (eq.36-37).

$$v\left(t + \frac{\delta t}{2}\right) = v\left(t - \frac{\delta t}{2}\right) + a(t)\delta t = v\left(t - \frac{\delta t}{2}\right) + \left(\frac{F}{m}\right)\delta t \quad (36)$$

$$r(t + \delta t) = r(t) + v\left(t + \frac{\delta t}{2}\right)\delta t \quad (37)$$

2.6 Binding free energy calculation

Binding free energy calculation is a great tool in computational simulation to predict the ligand binding affinity into a target. The more negative value means the stronger binding interaction with a ligand. As mentioned above, the binding free energy can be directly evaluated from a partition function; however, it is very difficult to calculate the absolute binding free energy of a large system such as protein or membrane. Although the relative binding free energy free energy perturbation, thermodynamics integration promise the sufficient quantity of binding energy, but these approaches poorly converge and need high computational consuming. To avoid this problem, the change in energy upon binding is approximated in the implicit solvation. The binding free energy is determined by the difference between complex form and its free forms (eq.38) called end-state free energy calculations.

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (38)$$

The free energy of each term is obtained from two main contributions: enthalpy (H) and entropy (S) (eq.39).

$$G = H - TS \quad (39)$$

H is sum of the ligand binding interactions in vacuum or gas-phase including van der Waals, electrostatics as well as internal energies, and the solvation free energy

(ΔG_{solv}) (eq.40). While ΔG_{solv} can be decomposed into the polar (ΔG_{solv}^{polar}) and non-polar ($\Delta G_{solv}^{non-polar}$) contributions (eq.41).

$$H = (E_{vdW} + E_{ele} + E_{int}) + \Delta G_{solv} \quad (40)$$

$$\Delta G_{solv} = \Delta G_{solv}^{polar} + \Delta G_{solv}^{non-polar} \quad (41)$$

ΔG_{solv}^{polar} can be calculated using several implicit solvent models: PB, GB and 3D-RISM. The conventional methods of MM combined with PB or GB electrostatics solvation and linear function of SASA are MM/PB(GB)SA methods which have been successful applied to various ligand binding affinities.^{85, 87, 90-92} These approaches compute $\Delta G_{solv}^{non-polar}$ as the linear function of SASA equation. In contrast, the combination of MM and 3D-RISM solvation energy method (MM/3D-RISM) determines the total ΔG_{solv} , then calculates the $\Delta G_{solv}^{non-polar}$ by removing all the charges from the solute molecule. The ΔG_{solv}^{polar} is obtained from the subtraction of those terms. The 3D-RISM calculation is able to predict not only solvation thermodynamics, but also provides the solvation structure in the form of 3D distribution function around the solute molecule.⁶⁶⁻⁷⁰ Then, the MM/3D-RISM method can predict the ligand bindings, which is in good agreement with the experimental data.^{70, 88, 93}

The conformational entropy ($T\Delta S$) upon the ligand binding is useful to correct the quantity of binding free energy including translational, rotational and vibrational entropies. The translational and rotational entropies can be calculated using standard statistical mechanics formulas,^{94, 95} while the vibrational entropy is estimated through the vibrational frequencies on the basis of ideal-gas rigid-rotor harmonic-oscillator approximation. Using a normal mode analysis, the conformation of solute is minimized to a local minimum, then the Hessian matrix is constructed and diagonalized to achieve the vibrational frequencies.⁹⁶

In addition, the H term (in eq.39) is not computed by MM level but also performed by QM/MM approach. The QM/MM system is partitioned into QM and MM regions. In sander module of AMBER program, the semi-empirical methods and density functional tight-binding (DFTB) approach are available as the QM implements.⁹⁷ The QM/MM calculation has been coupled with implicit solvent of

GBSA method developed by Pellegrini and Field.² The introduction of this couple to the binding free energy calculation is referred as QM/MM-GBSA method, which successfully predicts the ligand binding affinity as well as substrate specificity.⁹⁸⁻¹⁰³ Hence, the principle of the QM/MM-GBSA calculation is similar to MM/GBSA approach by using QM/MM method to compute the H energy. In this work, the self-consistent charge DFTB method (SCC-DFTB) is adopted as QM/MM approach.

The SCC-DFTB is an approximation of density functional theory (DFT), which estimates Hamiltonian in term of the electron density, instead of wave function-based calculation. In SCC-DFTB, the DFT energy is expanded to second-order in terms of reference density ρ_0 and charge density fluctuations ρ_1 (eq.42)⁷³

$$E[\rho] = \sum_i^{\text{valence orbitals}} n_i \langle \psi_i | \hat{H}[\rho_0] | \psi_i \rangle + \sum_i^{\text{core orbitals}} n_i \langle \psi_i | \hat{H}[\rho_0] | \psi_i \rangle + E_{xc}[\rho_0] - \frac{1}{2} \int_{R^3} \rho_0 V_H[\rho_0] - \int_{R^3} \rho_0 V_{xc}[\rho_0] + E_{\text{nuclei}} + \frac{1}{2} \int_{R^3} \rho_1 V_H[\rho_1] + \frac{1}{2} \iint_{R^3} \left. \frac{\delta^2 E_{xc}}{\delta \rho^2} \right|_{\rho_0} \rho_1^2 + o(3) \quad (42)$$

where \hat{H} is the matrix element of the Hamiltonian and ψ_i is the Kohn-Sham molecular orbitals, E_{xc} is the exchange-correlation functional, E_{nuclei} is the nuclei-nuclei interaction, V_{xc} corresponds to the exchange-correlation potential energy and V_H is the Coulomb potential of electron density interaction.

Then SCC-DFTB is derived from those terms using some approximations, which can be written as (eq.43)

$$E^{\text{SCC-DFTB}} = \sum_i^{\text{valence orbitals}} n_i \varepsilon_i + \frac{1}{2} \sum_{A \neq B}^{\text{atoms}} E_{\text{rep}}^{AB} + \frac{1}{2} \sum_{A, B}^{\text{atoms}} \gamma_{AB} \Delta q_A \Delta q_B \quad (43)$$

where n_i and ε_i are the occupation and the orbital energy of Kohn-Sham eigenstate, E_{rep}^{AB} is a pairwise repulsive interaction between atoms A and B , γ_{AB} is the distance-dependent function of charge-charge interaction, Δq_A and Δq_B represent the charge of each atom obtained from the Mulliken population analysis.

CHAPTER III CALCULATION DETAILS

3.1 MD simulation of substrate binding to NAs

3.1.1 Setup of substrate/NA complexes

The structure preparation and MD simulations were performed by AMBER 12 package.¹⁰⁴ The crystal structures of emerged IV NAs of H5N1, H7N9, H1N1, H2N2 and H3N2 were obtained from Protein Data Bank (PDB) codes: 2HU4³⁸, 4MWR⁵⁸, 3TI5, 3TIC¹⁰⁵ and 4GZW/4GZX⁵⁰, while their ligand were removed. Since the avian H5N1 and H7N9 IV have found to interact with avian receptor of 3SL and human receptor 6SL, while the human IV can only bind with 6SL, thus, the avian NAs were prepared in complex with both linkages as summarized in Table 5. By superimposition, the substrates from the co-crystal structures of human H3N2 NA-D151G mutant (4GZW.pdb for trisaccharide 3SL and 4GZX.pdb for disaccharide 6SL) were adopted as the ligand structures for preparation of the substrate complexes with the other influenza NAs. To prepare the trisaccharide 6SL, the GlcNAc molecule was extracted from the trisaccharide 6SL binding to H7N9 HA (4BSB.pdb)¹⁰⁶ and then was minimized by 2,000 steps of steepest descent (SD) then followed by 3,000 steps of conjugated gradient (CG) using SANDER module in AMBER to remove the bad contact with NA. Both trisaccharide substrates were terminated with a methoxy group. For H3N2 NA D151G variant, the residue 151 was mutated to aspartate using the rotamer library of Dunbrack¹⁰⁷ in Chimera program.¹⁰⁸

Table 5 All simulated substrate/NA complexes

NA subtypes	PDB code	Sequence at 345-348	6SL	3SL
H5N1	2HU4 (Russel 2006)	GAYG	✓	✓
H7N9	4MWR (Wu 2013)	NNNN	✓	✓
H1N1	3TI5 (Vavricka 2011)	GANG	✓	-
H2N2	3TIC (Vavricka 2011)	GTQG	✓	-
H3N2	4GZW, 4GZX (Zhu 2013)	GGHG	✓	-

The protonation state of titratable amino acid was considered at pH 7 by PROPKA and manual verification. Disulfide bonds in protein were assigned as CYX notation. All hydrogen atoms and missing atoms were added using LEaP module in AMBER. The calcium ions in crystal structure of NA were kept, hence, the counter ions were neutralized the complex. Each complex of substrate was firstly solvated with TIP3P water model using placevent algorithm¹⁰⁹ to replace the crystal water molecules around the complex, then entirely immerge into the cubic box of 90*90*90 Å³ through LEaP. To optimize the complex prior to simulation, the energy minimization of hydrogen atoms, followed by water molecules and entire system was performed in ordering using SD (1,000 steps) and CG (2,000 steps) methods by forcing the rest of minimized structures.

3.1.2 MD simulations of substrate/NA complexes

The simulation was carried out with pmemd.CUDA module. The parameter of substrate is GLYCAM06,¹¹⁰ while the force field parameter of NA protein is ff12SB. A cutoff of 10 Å was set for non-bonded interactions and particle mesh Ewald (PME) method¹¹¹ was applied for long-range electrostatics to reduce the computer

consuming. SHAKE algorithm was used to fix the bond length of all relevant hydrogen atoms¹¹² as well as periodic boundary condition was applied to avoid edge effect during simulation. There were three steps of simulations. First, the simulated system was heat up from 10 to 300 K with NVT ensemble for 100 ps. Temperature was controlled by a Langevin thermostat with a collision frequency of 2.0 ps⁻¹. Subsequently, the complex was equilibrated at target temperature with NVT for 100 ps followed by NPT ensemble of 1atm and 300 K until 1ns through position-restrained simulations of gradually decreased force constants from 10.0 to 2.5 kcalmol⁻¹Å⁻¹ to the substrate and its binding residues. Finally, the unrestraint NPT simulation was performed with 2 fs of time step till 20 ns. The production phase was determined by considering of root mean square displacement (RMSD) in respect to initial structure along simulation time for further analyzed and interpret the substrate binding and specificity toward different NAs.

3.1.3 QM/MM-GBSA calculation of substrate specificity

The binding affinity of substrate toward NA was determined by QM/MM-GBSA method using the MMPBSA.py program⁹⁶ in AMBER. The calculations were performed over 100 snapshots extracted from the production phase of simulation. In this research, the conformation of substrate was entirely characterized by QM calculation with SCC-DFTB level, while NA protein was determined by MM level. The ΔG_{solv}^{polar} is obtained from modified GB model developed by Onufriev, Bashford and Case¹¹³, whereas the $\Delta G_{solv}^{non-polar}$ with a probe radius of 1.4 Å and the values of γ and β are 0.0072 kcal/mol·Å² and 0.0 kcal/mol.

Finally, the $T\Delta S$ was approximated over the 25 snapshots obtained from the production phase in order to reduce the computational demanding. Using normal mode analysis, each conformation was minimized using a maximum of 1,000 steps until energy gradient is less than 0.01 kcal mol⁻¹ Å⁻¹ similar to our previous work of NA.¹¹⁴

3.2 MD simulation of H7N9 NAs and oseltamivir complex

3.2.1 H7N9/Oseltamivir complex preparation

Oseltamivir binds to both wild-type and R292K mutation H7N9 NA. Complexes were prepared from crystal structure of the unbound NA PDB structures, with codes of 4MWJ and 4MWL respectively, with superimposition to their oseltamivir structure of 4MWQ and 4MWW respectively.⁵⁸ These systems were then prepared for calculation and carried out using the AMBER 14 software package.¹¹⁵ The partial charges and parameters of oseltamivir was obtained from previous work,¹¹⁶ and the force fields of the protein were set to the ff12SB model of Amber 14. The ionized states of all side chains of lysine (K), arginine (R), aspartate (D), glutamate (E) and histidine (H) were configured to be at pH 7.0 using PROPKA3.1.¹¹⁷ All disulfide bonds between pairs of cysteine (C) residues were defined to stabilize the protein structure. The LEaP module added the missing hydrogen atoms in the oseltamivir-NA complex. A crystal calcium ion was kept, while water molecules were deleted. The placevent algorithm based on the 3D-RISM distribution function was employed to firstly solvate TIP3P water model around solute molecule.¹⁰⁹ The TIP3P waters were then filled with a minimum distance of 10 Å from the solute to the edge of the simulation box (88 x 90 x 81 Å³) and two Cl⁻ ions were added to neutralize the system.

3.2.2 MD simulation and protocol

To optimize the initial structure before performing MD simulations, the complexes were minimized in the following order, hydrogen atoms, water molecules and then the entire system implementing a 1,000 steps using the steepest descent algorithm followed by 2,000 steps with the conjugated gradient algorithm. The energy minimization process and MD simulations were performed using the pmemd.CUDA module of AMBER. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm.¹¹⁸ To reduce the computational time of non-bonded interactions, a cutoff distance of 10 Å was used and the particle mesh Ewald method¹¹⁹ was employed to calculate the long-range electrostatic interaction.

The simulation employed the periodic boundary condition using Langevin dynamics with a collision frequency of 2 ps^{-1} and a simulation time step of 2 fs. The system was heated up from 10 K to 300 K for 100 ps using canonical ensemble (NVT) in the thermalization step. Subsequently, the equilibration step was firstly conducted for 100 ps of NVT followed by NPT ensemble at 1 atm and 300 K until 30 ns. Finally, the system was further simulated for 50 ns. The production phase of the root mean square displacement (RMSD) analysis was assigned from 30 to 50 ns.

3.2.3 Binding free energy calculations

In this work, three end-state free energy methods: MM/PB(GB)SA and MM/3D-RISM, are carried out to predict the oseltamivir susceptibilities toward both wild-type and mutant H7N9 NA. The calculations were averaged over 100 snapshots extracted from the production phase using the MMPBSA.py program.⁹⁶ MM/PBSA computes $\Delta G_{\text{solv}}^{\text{polar}}$ by solving the linear PB equation using a pbsa implementation in *sander*.¹⁰⁴ The grid spacing was set to 0.125 \AA and a probe radius of water molecule is 1.4 \AA . Whilst, the $\Delta G_{\text{solv}}^{\text{polar}}$ of MM/GBSA is calculated from GB model of Onufriev et al.⁷⁹ However, both methods similarly define the dielectric constant for solute and solvent with 1 and 80, respectively. The $\Delta G_{\text{solv}}^{\text{non-polar}}$ was obtained by solvent accessible surface area (SASA) using a probe radius of 1.4 \AA and a surface tension constant (γ) of $0.0072 \text{ kcal mol}^{-1} \text{ \AA}^2$.

On the other hand, MM/3D-RISM approach determines the equilibrium distribution of solvent around solute and calculates the solvation free energy. This method computes the entire solvation free energy then obtained the nonpolar solvation energy by using none charge of an identical solute to decompose solvation energy into polar and nonpolar terms. The calculation of 3D-RISM solvent distribution is similar to the previous study of influenza B NA.⁸⁸ The explicit water of TIP3P model was employed to prepare solvent susceptibility function using the convergence criteria of 0.025 \AA grid size and $1e-12$ tolerance. The 3D-RISM complemented with Kovalenko-Hirata closure (3D-RISM-KH) was computed with a grid size of 0.5 \AA and a tolerance of $1e-5$.¹²⁰

To reduce the high computer consuming, the structural entropy calculation is estimated over the 25 snapshots of stable MD structures by normal mode analysis in the same protocol with previous work.



CHAPTER IV

RESULTS AND DISCUSSION

4.1 MD simulation of substrate binding to NAs

4.1.1 Substrate specificity of influenza NAs

The NA cleavage activity is directly related to the strength of substrate binding affinity. In order to elucidate the substrate specificity toward different influenza NAs, QM/MM-GBSA binding free energy calculation was carried out over 100 snapshots extracted from the production phase of simulation. Each trisaccharide substrate was considered as quantum mechanics (QM) region computing by DFT method with SCC-DFTB level of theory, while the rest atoms were treated by molecular mechanics (MM) level. The binding free energies of substrate into NA (ΔG_{bind}) and their components are summarized in Table 6.

Table 6 QM/MM-GBSA binding free energies of 6SL and 3SL substrates binding towards different influenza NAs (in kcal/mol)

Energetics	Simulated system					
	Substrate	H1N1	H2N2	H3N2	H5N1	H7N9
ΔE_{vdw}	6SL	-36.02 ± 4.71	-38.41 ± 3.43	-40.37 ± 3.71	-26.66 ± 3.34	-29.84 ± 3.66
	3SL				-43.50 ± 4.10	-33.95 ± 3.47
ΔG_{qm}	6SL	-123.42 ± 14.11	-100.94 ± 12.41	-65.15 ± 14.04	-75.84 ± 11.48	-118.66 ± 11.29
	3SL				-96.92 ± 13.12	-131.07 ± 11.79
ΔG_{solv}	6SL	116.98 ± 10.84	96.14 ± 9.48	66.74 ± 10.54	72.29 ± 9.81	123.37 ± 9.20
	3SL				97.44 ± 10.13	131.26 ± 9.20
ΔH	6SL	-42.46 ± 5.90	-43.20 ± 4.39	-38.78 ± 5.44	-30.22 ± 3.83	-25.14 ± 4.54
	3SL				-42.98 ± 6.50	-33.76 ± 4.23
TAS	6SL	-30.51 ± 7.95	-27.53 ± 14.39	-26.49 ± 6.89	-26.62 ± 9.42	-25.20 ± 7.64
	3SL				-31.82 ± 10.92	-28.27 ± 7.53
ΔG_{bind}	6SL	-11.95	-15.67	-12.29	-3.60	0.06
	3SL				-11.16	-5.49

Table 6 shows that the 6SL substrate provides much stronger binding to H1N1, H2N2 and H3N2 NAs of human IV with QM/MM-GBSA energies of -11.95, -15.67 and -12.29 kcal/mol, respectively, compared to its avian H5N1 and H7N9 complexes (-3.60 and 0.06 kcal/mol). Interestingly, an increase in binding strength was found in the H5N1 and H7N9 NAs in complex with 3SL receptor (-11.16 and -5.49 kcal/mol). This result evidently suggests that avian NAs is preferential to bind with 3SL receptor, while human NAs show a strong binding affinity with 6SL receptor corresponding to their HA proteins. Such low binding efficiency of avian NAs with 6SL indicates the unfavorable binding, which might be the barrier of avian-to-human transmission of avian IV.

To consider the energy components, a highly negative value of ΔG_{qm} contributes as the main driving force for the binding of substrate containing charged SIA terminus ($-1e$) into highly charged NA pocket. The negative contribution of total binding interaction energy, a summation of ΔG_{qm} and ΔE_{vdw} , is generally subtle by the large positive value of solvation free energy (ΔG_{sol}) referred as dehydration free energy, the loss of solute-solvent interactions during ligand binding into NA.^{88, 114} A dramatic decrease in magnitude of binding free energy of H5N1 and H7N9 NAs in complex with 6SL compared to 3SL by 7.56 and 5.55 kcal/mol, respectively, is affected by a reduction in ΔH and $T\Delta S$ terms. Such lower entropy contribution of the two 3SL/NAs (by c.a. 5 kcal/mol) implies a more stable complex.

4.1.2 Dynamics of neuraminidase

NA contains the four important loops including 150-loop (residue 147-152), 270-loop (267-276), 380-loop (380-390), and 430-loop (429-433) shown in Figure 13.¹²¹⁻¹²³ However, only the 150- and 430-loops have an important role to interact with a substrate as well as drug.^{124, 125} The previous simulation found that the couple motions of these loops results in a closed or open cavity of NA binding pocket.¹²⁶ The opening and closing of NA cavity are required to fit a substrate into the active site of NA. The inherent flexibility of 150- and 430-loops could have a role for receptor recognition. In this work, the formation of NA 150-cavity was explored by

determining of the distance between the residue 149 in the 150-loop and the residue 431 in the 430-loop, depicted in Figure 13.

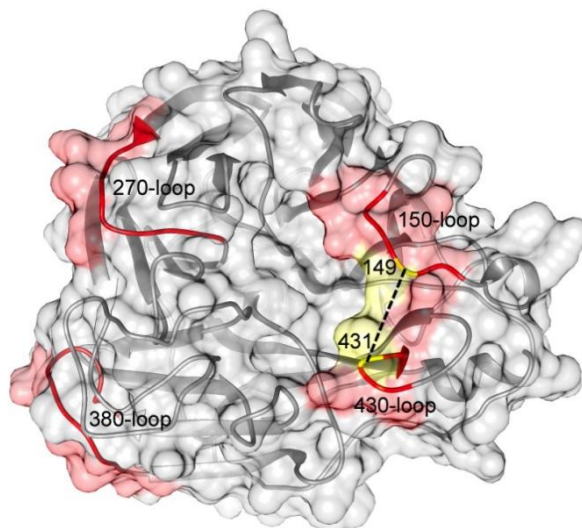


Figure 13 Important loops of NA and 150-cavity formation observed by the distance between the residues 149 and 431

From previous calculation of Amaro et al., the greater than ~ 15 Å of residues 149-431 distance was proposed as an open 150-cavity.¹²⁶ In Figure 14, the width between 150- and 430-loops was determined as probability for all studied systems. For 3SL complexes, the H5N1 NA has a closed form of 150-cavity with ~ 11 Å, while the H7N9 NA extends the cavity to open conformation with ~ 16 Å (Figure 14a). This results in the much stronger binding of 3SL with H5N1 (-11.16 kcal/mol) compared to H7N9 (-5.49 kcal/mol). However, both avian NAs in complexed with 6SL show a wide-open 150-cavity as can be seen from the wide distribution and a distance shifting to 18 Å of H5N1 (Figure 14b), indicating the unfavorable binding of avian NA with human receptor. In contrast, all human NAs with 6SL bound show the closed form of cavity with a broad distribution in the range of 7 - 12 Å (Figure 14b). This result confirms that the formation of 150-cavity plays an essential role for substrate binding, which is consistent to the prediction of binding free energy in Table 6.

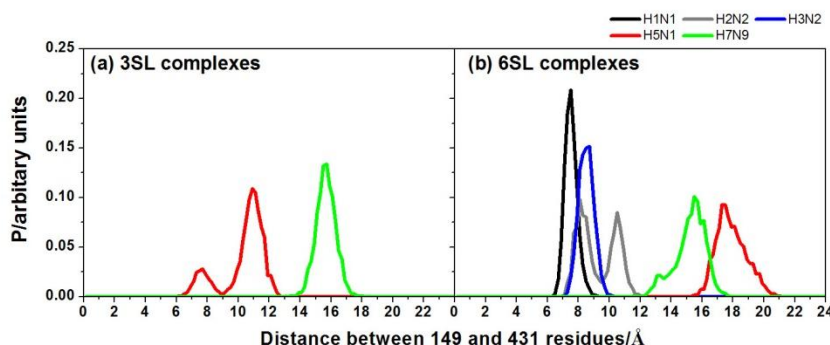


Figure 14 Distribution plot of distance between the alpha carbon of the residues 149 and 431 of (a) 3SL complexes (b) 6SL complexes

Although H1N1 and H5N1 NAs show high sequence identity with 91%, the dynamics of the 150-cavities are different. The H1N1 NA presents a deficient 150-cavity for binding with human receptor, while H5N1 NA in complex with substrate can be either closed or open conformations. Generally, the N1 NA (except for pandemic H1N1) adopts in an open 150-cavity due to the loss of salt bridge interaction between D147 and H150 forming closed conformation in N2 NA.¹⁰⁵ The H1N1 NA does not contain 150-cavity; however, the previous MD simulation showed that its 150-cavity exhibits an open form.¹²⁶ In addition, the difference in the 150-loop of H1N1 (I149) and H5N1 (V149) does not affect the 150-loop due to both of them are able to adopt an open conformation. Therefore, the difference in 150-loop conformation of human H1N1 and avian H5N1 NAs is related to the substrate specificity.

4.1.3 Glycan topology

The global shape of substrate is generally defined using topological angle between the C2, C1 and C1 atoms of terminal SIA and the subsequent GAL and NAG, θ depicted in Figure 15a,b. The different glycosidic linkages of substrate result in the distinct conformations which play a key role for HA specificity.¹²⁷ The 3SL has $\theta > 110^\circ$ adopting a cone-like shape, whilst the 6SL has $\theta < 110^\circ$ referring an umbrella-like topology. Thus, in this present study the topological angle of trisaccharide substrate was determined and plotted in Figure 15c,d as the distribution over the

production phase to characterize the substrate's conformation in complex with different NAs.

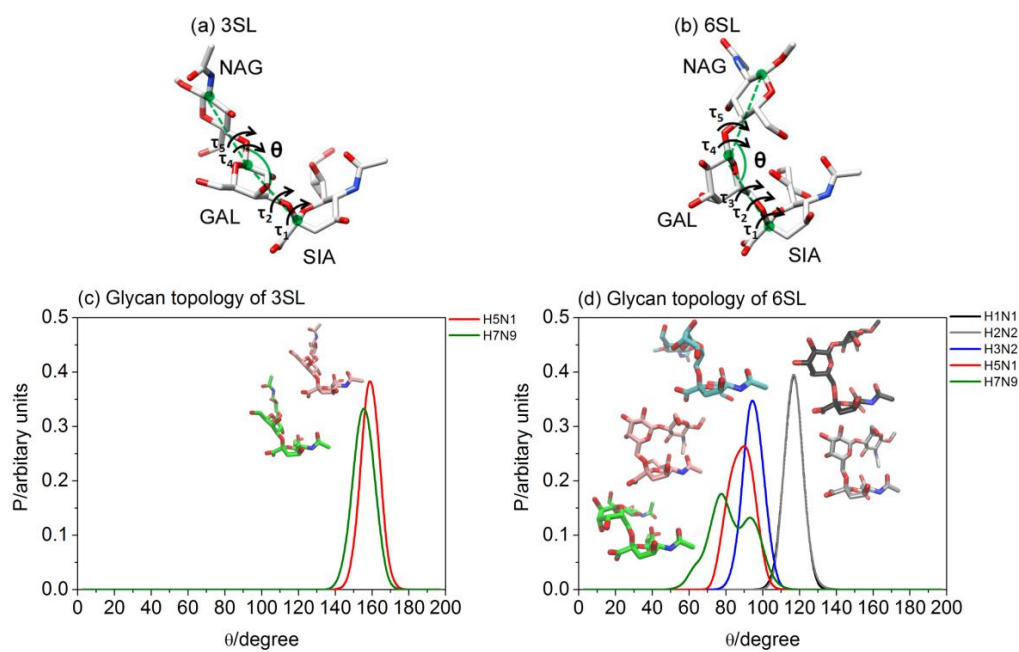


Figure 15 (Upper) Illustration of topological angle (θ) and glycosidic dihedral angles (τ) of (a) 3SL (b) 6SL substrates, (Lower) the angle's distribution of substrates in complex with different NAs as well as the representative conformations at the highest probability for (c) 3SL and (d) 6SL complexes

In Figure 15, the topological angle showed that the 3SL and 6SL trisaccharide substrates have distinct conformations within NA binding pocket. The 3SL in complex with avian H5N1 and H7N9 NAs shows a sharp peak of θ distribution at 157° and 155° , respectively, conferring a cone-like topology (Figure 15c) stably occupied in avian NAs. The topological angle value of NA-bound glycans is consistent to their free form, which has angle of $\sim 160^\circ$ from our replica exchange MD (REMD) simulation and 152° from the previous MD simulation.¹²⁸ In contrast, the bound 6SL show either sharp or board θ distribution differently placed in the range of 60° to 130° toward all five NAs indicating the various shapes of human receptor (Figure 15d). Among three human NAs, the 6SL shows an identical distribution of sharp peak at 117° in H1N1 and H2N2 NAs (black and grey, Figure 15d) suggesting

the open-umbrella topology, while the θ peak is shifted to 92° in H3N2 NA due to a rotation of NAG unit after ~ 5 ns through the fifth repeated simulations. On the other hand, the rather broad distribution of 6SL in avian H5N1 and H7N9 NAs centered at 87° and 77° . Their lower topological angle results in the change of 6SL glycan shapes as evidenced in pink and green structures in Figure 15d. This suggests that NA-bound 6SL is much more flexible in respect to 3SL which agrees to the previous simulations of H5N1 NA in complex with both glycans.⁶² The high flexibility of 6SL in avian NAs, in particular H7N9 maybe a source of lower 6SL binding affinities (described in Table 6). To compare with the free 6SL pentasaccharide, the θ distribution is mostly of $\sim 120^\circ$ from REMD, $\sim 100^\circ$ from MD¹²⁸ and $\sim 90^\circ$ from NMR-based MD,¹²⁹ indicating the significant change of this glycan to bind with NA protein. However, the orientation of SIA-GAL-NAG trisaccharide likely affects to the glycan topology and plays a role for identifying the glycan change rather than topological angle parameter alone. Sasaki *et al.*¹²⁸ proposed that a difference in glycan topology relates to the change in glycosidic torsion angles. To seek the source of structural change in NA-bound glycans, the distribution of glycosidic torsion angles was determined and plotted in Figure 16. The τ_1 - τ_3 dihedral angles are the linkages between SIA and GAL, while the τ_4 - τ_5 dihedral angles are GAL-NAG bridges, illustrated in Figure 15.

Figure 16a shows that all four dihedral angle's distributions of 3SL binding to avian NAs are almost similar consistent to the considerably topological angle θ . The τ_1 and τ_2 angles are predominantly at -40° and -100° , respectively, which agree to the dynamic structures of the unbound 3SL.¹²⁸ For GAL-NAG linkages, their τ_4 angles present a sharp peak at 170° , while the τ_5 is substantially distributed around 120° . The situation is rather different for the NA-bound 6SLs. Only τ_1 shows the practically similar pattern of distribution over the range of -100° to -50° toward NAs, whilst the significant variances were evidenced at the other torsion angles in particular the broadest distribution of τ_5 as given in Figure 16b. Taken altogether, the obtained data suggests that the 3SL does not require conformational change to fit within the avian NAs binding pocket.

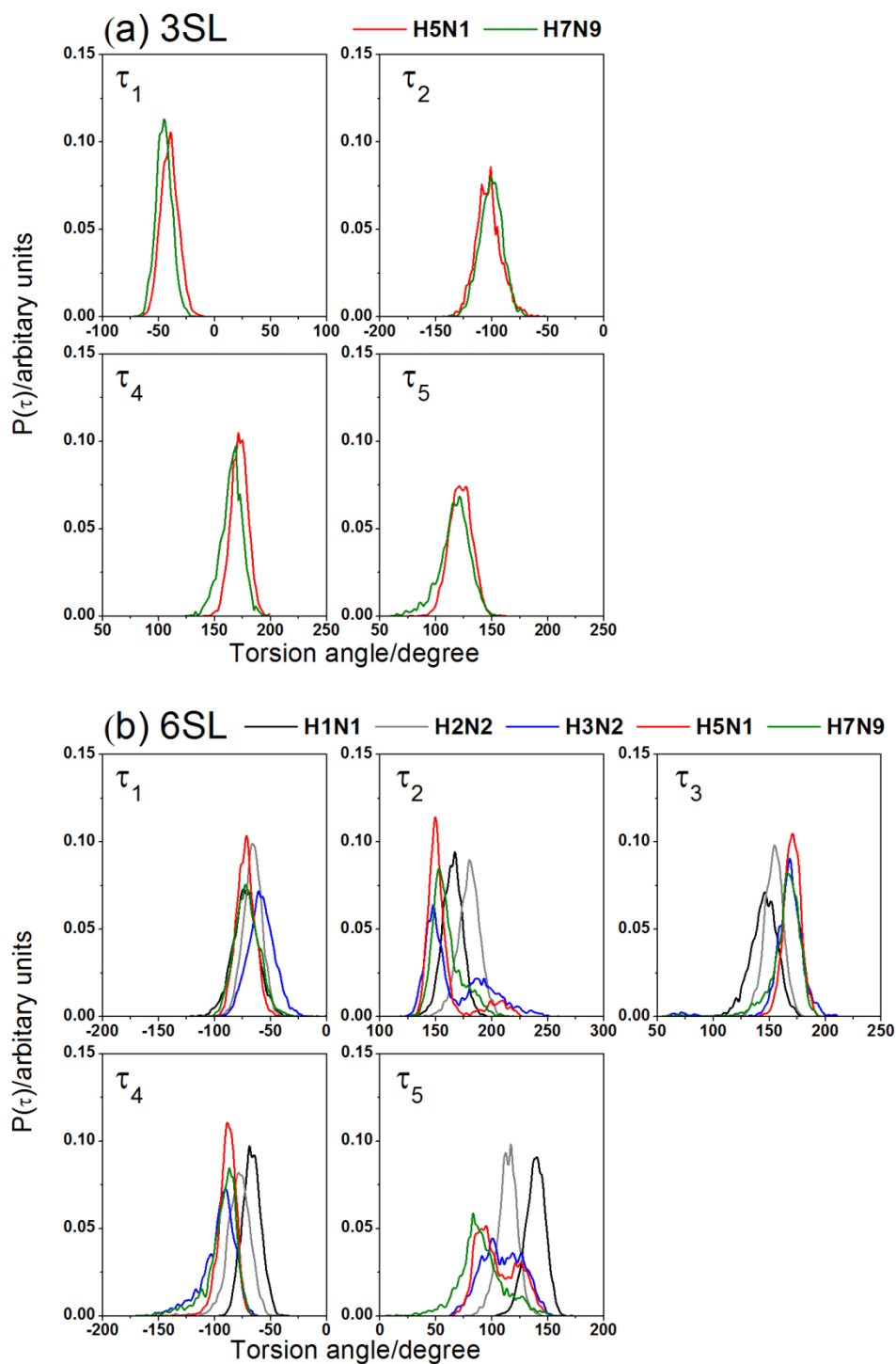


Figure 16 Plot of torsion angle's distribution of (a) 3SL and (b) 6SL glycans within binding pocket of different NAs, including H1N1, H2N2, H3N2, H5N1 and H7N9

4.1.4 Substrate binding pattern

In order to discover the binding pattern of substrates into the human and avian NAs, the occupation of hydrogen bond (H-bond) pairs between substrate and contact residues were determined using the criteria of distance between proton donor and acceptor atoms of $\leq 3.5 \text{ \AA}$ and the angle of those atoms linked by proton of $> 120^\circ$. The results are depicted and compared in Figure 17.

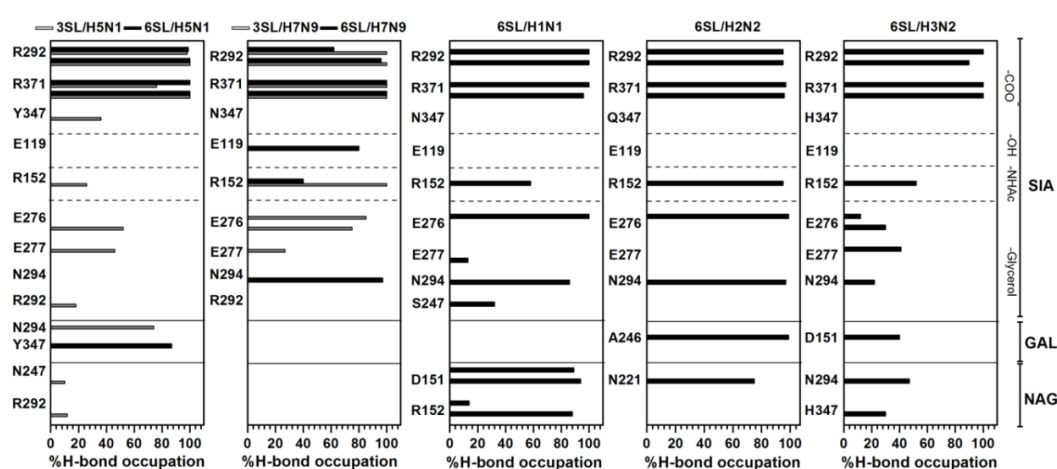


Figure 17 Percentage of intermolecular hydrogen bonds between 3SL/6SL trisaccharide and NAs residues of avian and human IV

Either 3SL or 6SL binding toward all different NAs in Figure 17 shows that the terminally bound SIA plays a major role in forming the intermolecular hydrogen bonds with many charged residues at the NA active site (R292, R371, R152, E276 and E277 as illustrated in Figure 18) better than the subsequent glycan units, GAL and NAG. Among all NAs with 6SL bound, the substrate-protein interactions in human NAs are relatively comparable and significantly stronger than those in avian NAs. Note that the 6SL in umbrella-like topology could intensely bind with human H1N1 and H2N2 NAs through 9 highly formed hydrogen bonds with D151, R152, N221, A246, E276, R292, N294 and R371 (Figures 17 and 18a-b) as supported by a greater binding energy in these two complexes ($\Delta H = -42.46$ and -43.20 kcal/mol, Table 6). A reduction in binding susceptibility of 6SL in human H3N2 (4 strong

hydrogen bonds with R292 and R371; $\Delta H = -38.78$ kcal/mol) is probably due to its twisted conformation with NAG pointing outward the binding site (Figure 18c). High conformational flexibility of such twisted 6SL in avian H5N1 and H7N9 NAs as mentioned earlier in Figure 15b results in even less hydrogen bond formation and energy stabilization ($\Delta H = -30.22$ and -25.14 kcal/mol). Instead, the avian IV could prefer to accommodate the cone-like topology of 3SL with a presence of more intermolecular hydrogen bonds and attractive interaction by ~ 10 kcal/mol than the 6SL in twisted form.

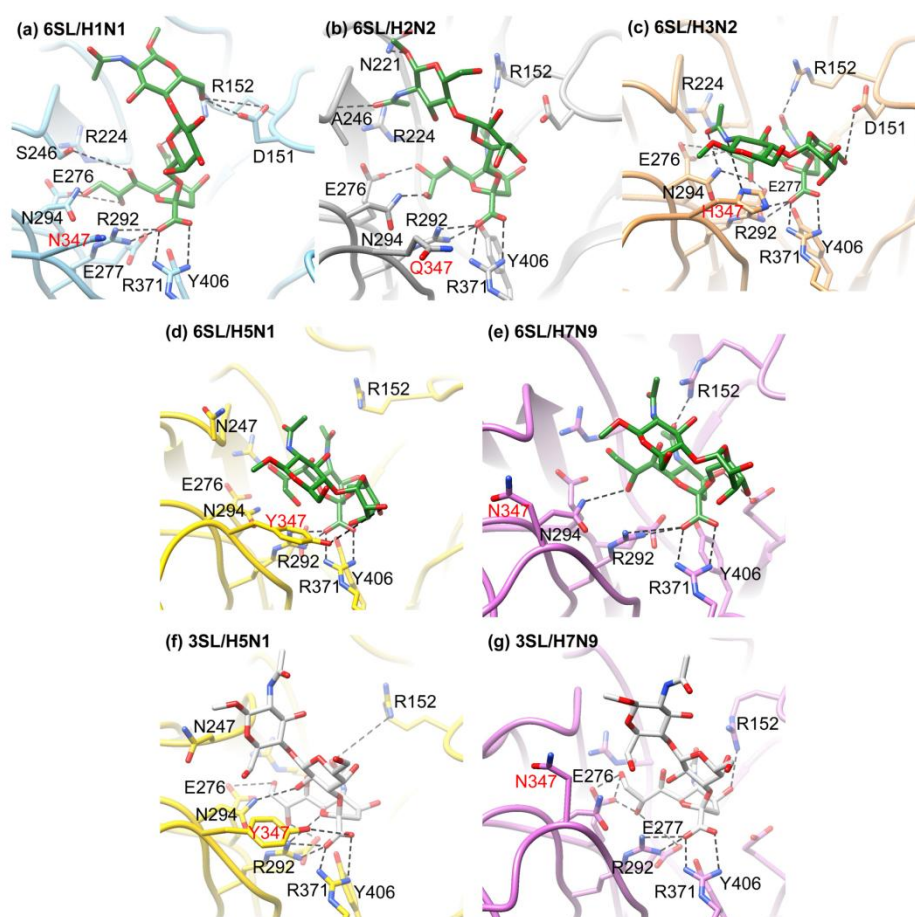


Figure 18 Substrate binding pattern of 6SL (green) and 3SL (grey) in the NA binding pocket with the hydrogen bond formations (dash line) demonstrated from the representative snapshot of each system

To distinctly clarify the preference of avian NAs, the difference in per-residue decomposition free energy between 3SL and 6SL ($\delta\Delta G_{\text{residue}}$) on the basis of MM/GBSA method was given in Figure 19. The negative value indicates a more favorable binding with 3SL, whereas the positive value reveals the preference with 6SL of NA. A much more negative fingerprint firmly shows that both H5N1 and H7N1 have significant preference with 3SL avian receptor over 6SL human receptor.

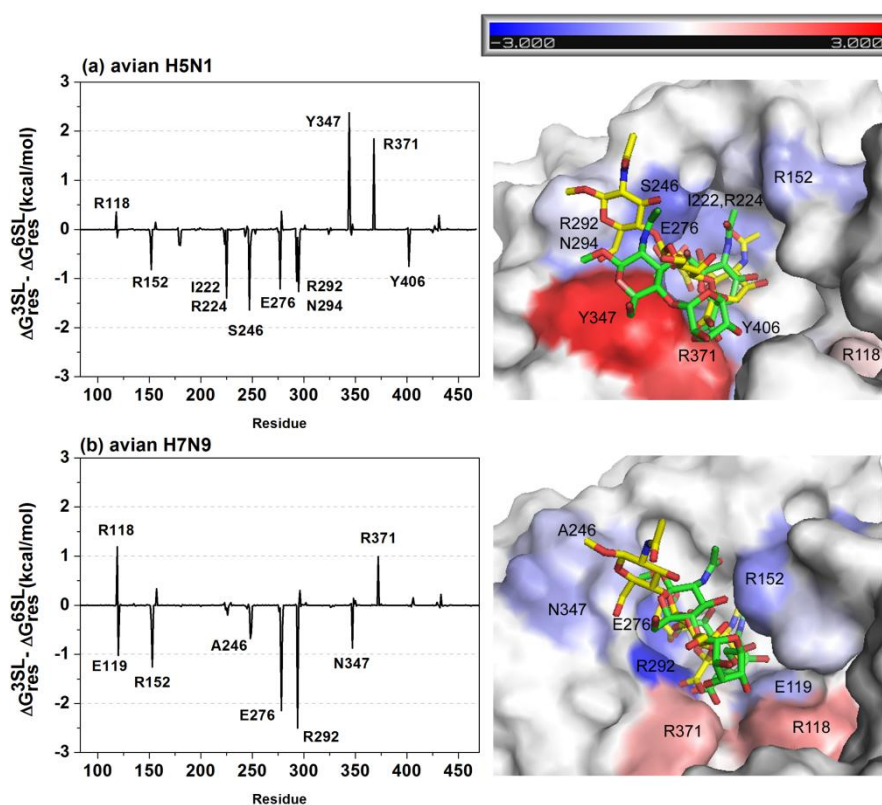


Figure 19 Difference in per-residue decomposition energy between 3SL (yellow) and 6SL (green) in complex with (a) H5N1 and (b) H7N9 NAs, whereas the positive and negative values were displayed as spectrum from red (3 kcal/mol), white (0 kcal/mol) and blue (-3 kcal/mol) surfaces

For substrate/H5N1 complexes (Figure 19a), the eight residues positioned in the NA binding pocket R152, I222, R224, S246, E276, R292, N294 and Y406 conferring the negative value of $\delta\Delta G_{\text{residue}}$ of < -0.5 kcal/mol better stabilize the 3SL avian receptor; however, the Y347 (2.38 kcal/mol) and R371 (1.84 kcal/mol) favorably interact with the GAL unit and the SIA carboxylate group of the 6SL

human receptor, respectively (Figure 17). Similar to H5N1, except for R118 and R371, H7N9 NA has a more contribution with 3SL binding through only six residues E119, R152, A246, E276, R292 and N347.

4.2 MD simulation of H7N9 NAs and oseltamivir complex

4.2.1 Stability of complexes

The stabilities of each complex were evaluated using root-mean square displacement (RMSD) to explore the maintained fluctuation of the MD structure compared to their initial structure. In Figure 20, the RMSD was calculated in term of backbone atoms, which belong to the NA protein and the ligand atoms of oseltamivir versus simulation time. For backbone RMSD, both wild-type and R292K complexes show increasing RMSD from ~0.5 to 1.0 Å during the first 10 ns and after which it oscillates at this value until 30 ns. Finally, the fluctuation stabilizes after 30 ns. With regards to the ligand RMSD of oseltamivir the values varied more ranging from 0.5 to 2.0 Å for the entire simulation. It can be clearly seen that oseltamivir structure of R292K system has significantly higher fluctuations compared to the wild-type system. As a result, the MD structures from 30-50 ns were adopted as the production phase for further analysis.

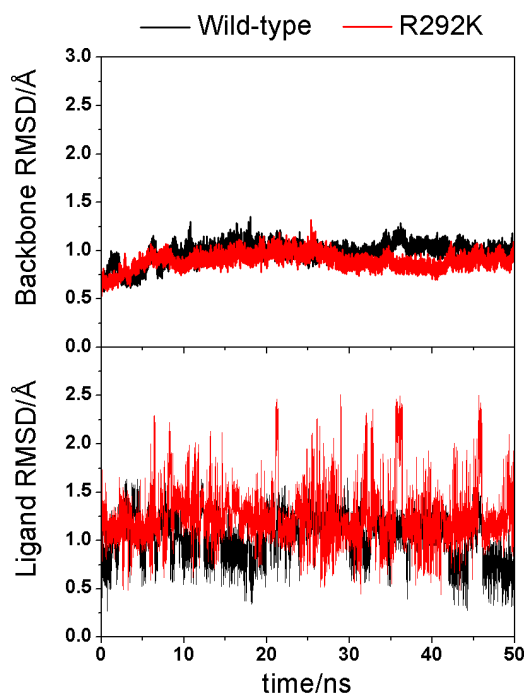


Figure 20 RMSD calculations in term of the backbone atoms of the NA protein (upper) and the oseltamivir ligand atoms (lower) versus the simulation time

4.2.2 Binding free energy calculation of oseltamivir susceptibility

The aspects of this study are to understand the oseltamivir efficiency towards the H7N9 NAs and to explore the source of reduce susceptibility concerning the R292K mutation. The MM/PB(GB)SA and MM/3D-RISM binding free energies and their components are shown in Table 7 and are also compared to the experimental values.

Table 7 Binding free energy of oseltamivir toward wild-type and R292K NAs using MM/3D-RISM and MM/PB(GB)SA methods (in kcal/mol)

Energy	Wild-type			R292K		
	MM/3D-RISM	MM/GBSA	MM/PBSA	MM/3D-RISM	MM/GBSA	MM/PBSA
ΔE_{ele}^{MM}		-180.4±7.1			-166.6±1.2	
ΔE_{vdw}^{MM}		-26.1±0.8			-21.0±0.4	
ΔE^{MM}		-206.5±6.4			-187.6±1.5	
$T\Delta S^{MM}$		-18.1±5.6			-17.7±6.0	
ΔG_{bind}^{MM}		-188.4			-169.9	
$\Delta\Delta G_{ele}^{solv}$	168.5±5.0	165.2±9.1	178.7±10.8	152.1±2.1	154.9±8.3	170.4±8.9
$\Delta\Delta G_{cav}^{solv}$	9.6±0.8	-5.3±0.2	-5.0±0.1	14.2±0.6	-4.8±0.2	-5.0±0.1
$\Delta\Delta G_{bind}^{solv}$	178.1±5.7	159.9±9.0	173.7±10.8	166.3±2.1	150.1±8.2	165.4±8.9
ΔG_{bind}	-10.3	-28.5	-14.7	-3.6	-19.8	-4.5
$\Delta\Delta G_{bind}$		-		6.7	8.7	10.2
ΔG^{IC50}		-13.1 and -12.2			-7.2 and -5.2	

In the wild-type complex, MM/PBSA and MM/3D-RISM methods qualitatively predicts the binding free energy of -14.7 and -10.3 kcal/mol, respectively, which are consistent with the experimental values of -12 to -13 kcal/mol,⁵⁶⁻⁵⁹ while MM/GBSA approach overestimates the value with a prediction of -28.5 kcal/mol. The total binding free energy (ΔG_{bind}) of these methods is a subtle balance between the negative value of the binding interaction (ΔG_{MM}) and the positive value of the solvation free energy ($\Delta\Delta G_{solv}$) called the dehydration penalty which involves the loss of solvation around the protein after ligand binding and is clarified in more detail in a previous study.⁸⁸ The electrostatic contribution plays a major role for oseltamivir binding as can be seen from the large free energy values of the electrostatic interaction ΔE_{ele}^{MM} and the polar solvation free energy $\Delta\Delta G_{solv}^{polar}$. This is due to the zwitter ion nature of oseltamivir accommodating the highly charged binding pocket of NA.

The reduction of binding free energy in the R292K mutated system was calculated for MM/GBSA (-19.8 kcal/mol), MM/PBSA (-4.5 kcal/mol) and MM/3D-RISM (-3.6

kcal/mol) to be 8.7, 10.2 and 6.7 kcal/mol, respectively, compared to those of wild-type (-5.2 or -7.2 kcal/mol). Similar to the wild-type complex, the MM/GBSA method overestimates the prediction of the binding free energy compared to the rest approaches. The source of reduced oseltamivir susceptibility conferred by the R292K substitution is a reduction in binding interaction (negative value) and solvation free energy (positive energy). The molecular mechanics energy shows a loss of drug interaction by 18.5 kcal/mol, and a decrease in dehydration penalty resulting in a change of solvation energy by 9.8 kcal/mol for MM/GBSA, 8.3 kcal/mol for MM/PBSA and 11.8 kcal/mol for MM/3D-RISM. In the following sections we will further elucidate the effect of the R292K substitution towards oseltamivir binding.

4.2.3 Change of drug binding interaction

The favorable electrostatic interaction of oseltamivir binding is obtained from strong hydrogen bond (H-bond) formations between polar moieties of the drug (including $-\text{COO}^-$, $-\text{NHAc}$ and $-\text{NH}_3^+$) and the highly charged binding pocket of NA. In addition, the nonpolar side chain of the pentyl group gives van der Waals contribution stabilizing the hydrophobic pocket formed by the E276 and R224 residues. The H-bond formation is illustrated as a percentage occupation, where occupation is defined as having maximum distance of 3.5 Å between hydrogen donor (D) and acceptor (A) together with the minimum bond angle of 120 degree between three atoms (D...H...A) depicted in Figure 21 for the wild-type and the R292K mutant NAs.

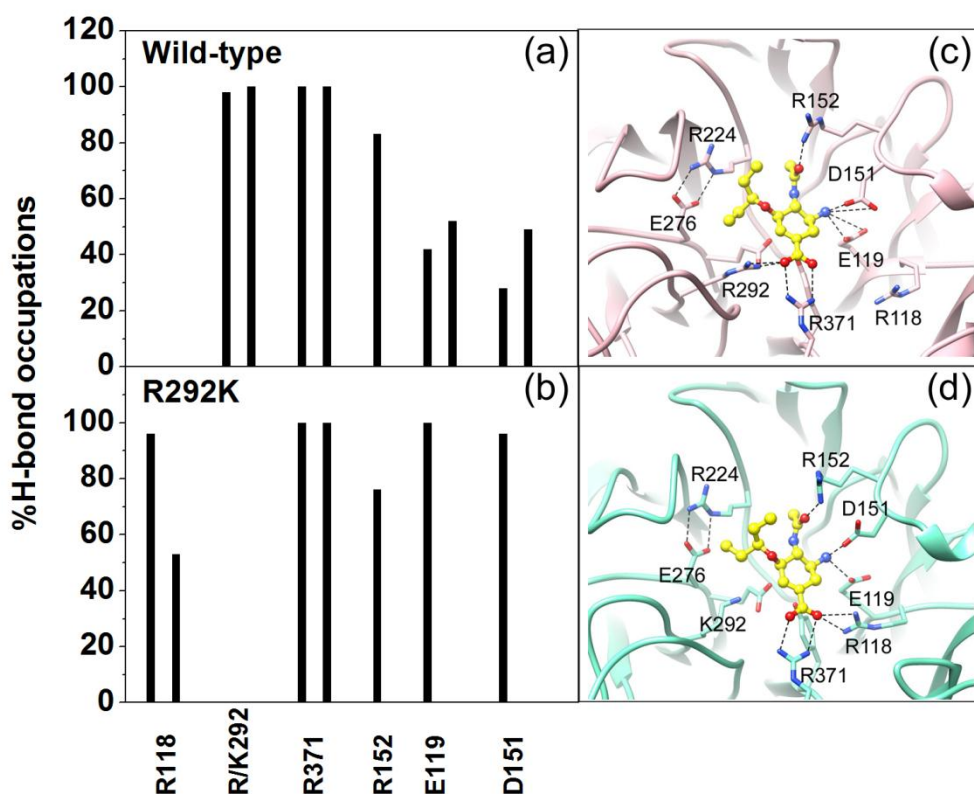


Figure 21 Percentage occupation of the hydrogen bond interactions between oseltamivir function groups and the H7N9 NA binding residues for (a) wild-type and (b) R292K strains. The hydrogen bonds (black dashed line) of oseltamivir-NA taken from the last snapshot are also shown for (c) wild-type and (d) R292K systems.

In Figure 21, oseltamivir forms H-bond interaction with residues of the arginine triad (R118, R292 and R371, according to N2 numbering), R152, E119 and D151 as is generally found to be the case in other NA subtypes.^{85, 88, 116, 130} In the R292K complex, there is a complete loss of the carboxylate-K292 interaction as previously found in the H1N1 subtype,¹³⁰ however, this is compensated for by the R118 interaction (Figure 21b). A smaller side chain of lysine (K) compared to arginine (R) at position 292 cannot form the necessary H-bond with the carboxylate group (distance of around 5 Å), leading to the larger cavity and shifting toward the R118 residue of ligand (Figure 21d). It is worth noting that the conformation of this K292 mutated residue does not perturb the formation of hydrophobic pocket made by E276 and R244 residues as found in the crystal structure.^{58, 65}

The conformational change of the pentyl group was evaluated by the distribution of torsion angles (Figure 22). Both wild-type and R292K systems have different angular distributions, particularly for τ_z . The torsion distribution of τ_z in the mutant system dramatically changes from the normal ~ 60 degrees of the wild-type, while τ_x and τ_y show the same pattern but different distributions. This suggests that the drug in the R292K complex is oscillating. This oscillation can clearly be seen from the high fluctuation of ligand RMSD during simulation (in Figure 20). The twist of pentyl group was found to disrupt the hydrophobic binding in a previous study.⁶⁴

To summarize, the unsuitable accommodation of oseltamivir in R292K NA results in a decrease of drug binding interaction by 13.8 kcal/mol for electrostatic interaction and 5.1 kcal/mol for van der Waals interaction compared to the wild-type complex.

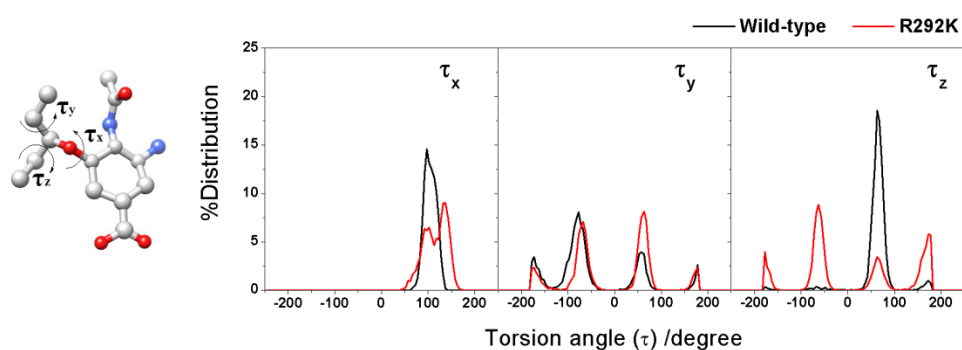


Figure 22 Distribution of the torsion angles of the pentyl group of oseltamivir calculated while in complexation with the binding pockets

4.2.4 Change of solvation conferring R292K substitution

The latter source of the high-level oseltamivir resistance toward R292K substitution is due to the decrease of the positive contribution of the solvation free energy resulting in a smaller dehydration penalty compared to the wild-type as evidenced in Table 7. Since 3D-RISM method can depict the distribution of solvent around solute molecule, this method was further adopted to analyze the change of

solvation in term of equilibrium distribution around complex and radial distribution function (RDF) of oseltamivir heteroatoms, depicted in Figure 23.

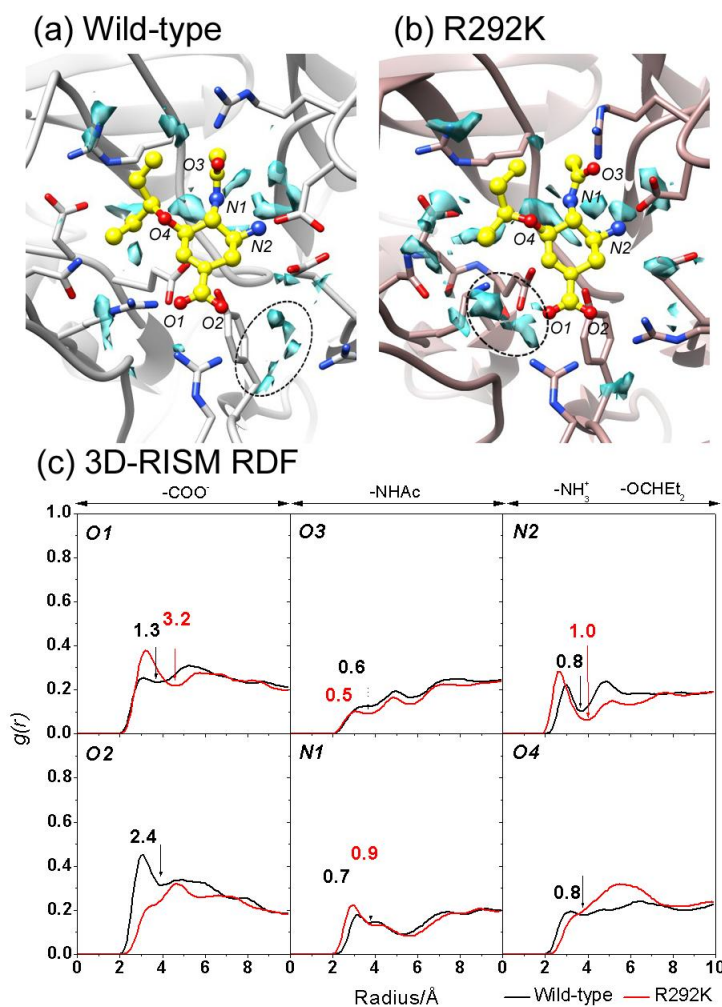


Figure 23 (Upper) 3D-RISM function of O-water atoms with $g(r) > 5$ as cyan contour shown within 5 Å of oseltamivir in complex with (a) wild-type and (b) R292K NAs taken from the representative stable structure, (Lower) (c) 3D-RISM RDF of water molecules around heteroatoms of drug averaged from 100 snapshots of production phase

As mentioned in previous sections, the shorter side chain of K292 causes poor binding with the drug carboxylate group and a larger cavity. This results in greater accessibility for water molecules towards the NA binding pocket as evident by the RDF around oseltamivir heteroatoms. In Figure 23a, the first peak of water

distribution around oseltamivir heteroatoms occurs at 3.5 Å and is mostly sharp, which suggests that the accessible water molecules strongly interact with oseltamivir. The nonzero value at the terminal of the first peak as found in the O1, O2 and N1 graphs indicate a high degree of transfer of water molecules in this shell. Switching of water molecules around the oseltamivir molecule was found to occur between the O1 and O2 atoms at the oseltamivir carboxylate group (see illustration in Figure 23c,d), while the rest of the heteroatoms are conserved for both wild-type and R292K systems. Additionally, two water molecules were found to bind around the K292 residue, compared to the one molecule of the wild-type (Figure 23b). This mutated residue was strongly solvated by three water molecules as can be seen by the approximately close to zero values at the first minimum, which is in good agreement with the crystal structures of H7N9 and H11N9.^{58, 131} Thus, the R292K mutation of NA can affect the accessibility of water in the binding pocket, which causes a decreased dehydration penalty.

CHAPTER V

CONCLUSIONS

The predicted binding free energies by QM/MM-GBSA method suggests that the avian H5N1 and H7N9 NAs prefer to bind with an avian receptor of 3SL rather than 6SL human receptor, while the human H1N1, H2N2 and H3N2 NAs show a strong binding affinity towards 6SL, which corresponds to HA protein. Since NA cleavage activity is directly related to the substrate binding, this could clarify the functional balance between HA and NA glycoproteins on influenza virus membrane for viral replication. The difference in receptor binding of influenza NAs maybe relate to the specific pathogenicity of influenza virus. The flexibility of 150- and 430-loops has a role for receptor recognition. Both glycans adopt distinct topologies for binding into different NAs. The 3SL has a cone-like topology in complex with avian NAs, while the 6SL can display open-umbrella and twisted conformations toward different NAs conferring the high flexibility of glycosidic torsion angles between GAL and NAG molecules. In comparison between the glycan conformations in unbound and bound states, a lower conformational diversity upon binding to the targeted NAs implied that the glycan could adopt its conformation to be fitted well in the binding pocket. Although the terminal SIA unit strongly interacted with NA residues the binding site through many hydrogen bond formations, the distinct glycan topologies of substrate can affect the different binding patterns, in particular the interactions at GAL and NAG units. This suggests that the glycan topology has an important role for NA binding which is found in HA recognition.

The oseltamivir binding towards the novel H7N9 influenza NA and the high resistance R292K substitution was investigated using MD simulations and MM/PB(GB)SA and MM/3D-RISM binding free energy calculations. All methods were able to predict the reduced susceptibility of oseltamivir toward the R292K strain which is in good agreement with the energy converted from the IC_{50} data, although predicted binding free energy of MM/GBSA method is overestimate compared to MM/PBSA and MM/3D-RISM. A structural entropy calculation was necessary to improve the qualitative of the predicted binding free energy. Similar to the previous study of influenza B NA, the main factor for oseltamivir binding toward H7N9 NA is

a subtle balance of the large negative value of binding interaction and the large positive contribution of solvation conferring a smaller dehydration penalty during drug binding. The major interaction of oseltamivir binding is the electrostatic contribution, as is generally found for all NA subtypes, caused by the opposite charges of oseltamivir and NA binding residues. A dramatic decrease in binding free energy was found in the R292K system compared to the wild-type. The role of R292K mutation in H7N9 NA is to decrease the binding interaction as was shown by the high fluctuation of drug RMSD in the mutant pocket and the reduction of the dehydration penalty resulting in an increased accessibility for water molecules around the K292 mutated residue and the carboxylate group of drug as evidenced by 3D-RISM calculation. The results from this work suggest that a hydrophilic bulky group as found in zanamivir and laninamivir, instead of the pentyl group of oseltamivir, promises to be a potent NAI by maintaining a strong interaction toward residue 292 of both arginine and lysine amino acids. This work also shows that MM/3D-RISM binding free energy calculation has the ability to predict the oseltamivir susceptibility providing both solvation free energy and solvation structure.



REFERENCES

1. WHO Influenza (Seasonal). <http://www.who.int/mediacentre/factsheets/fs211/en/index.html> (September 10, 2010),
2. Gaymard, A.; Le Briand, N.; Frobert, E.; Lina, B.; Escuret, V., Functional balance between neuraminidase and haemagglutinin in influenza viruses. *Clinical Microbiology and Infection* **2016**, *22*, 975-983.
3. Lipatov, A. S.; Govorkova, E. A.; Webby, R. J.; Ozaki, H.; Peiris, M.; Guan, Y.; Poon, L.; Webster, R. G., Influenza: Emergence and Control. *Journal of Virology* **2004**, *78*, 8951-8959.
4. Al-Muharrmi, Z., Understanding the Influenza A H1N1 2009 Pandemic. *Sultan Qaboos University Medical Journal* **2010**, *10*, 187-195.
5. M. Fry; Erin Burns; Susan Trock; Suizan Zhou; Jacqueline M. Katz; Daniel B. Jernigan, A. D. I. Y. J. J. C. T. D. D. E. W. T. M. U. K. R. M. G. T. L. G. A., Increase in Human Infections with Avian Influenza A(H7N9) Virus During the Fifth Epidemic — China, October 2016–February 2017. *Morbidity and Mortality Weekly Report* **2017**, *66*, 254-255.
6. Gao, R.; Cao, B.; Hu, Y.; Feng, Z.; Wang, D.; Hu, W.; Chen, J.; Jie, Z.; Qiu, H.; Xu, K.; Xu, X.; Lu, H.; Zhu, W.; Gao, Z.; Xiang, N.; Shen, Y.; He, Z.; Gu, Y.; Zhang, Z.; Yang, Y.; Zhao, X.; Zhou, L.; Li, X.; Zou, S.; Zhang, Y.; Li, X.; Yang, L.; Guo, J.; Dong, J.; Li, Q.; Dong, L.; Zhu, Y.; Bai, T.; Wang, S.; Hao, P.; Yang, W.; Zhang, Y.; Han, J.; Yu, H.; Li, D.; Gao, G. F.; Wu, G.; Wang, Y.; Yuan, Z.; Shu, Y., Human infection with a novel avian-origin influenza A (H7N9) virus. *New England Journal of Medicine* **2013**, *368*, 1888-1897.
7. Liu, D.; Shi, W.; Shi, Y.; Wang, D.; Xiao, H.; Li, W.; Bi, Y.; Wu, Y.; Li, X.; Yan, J.; Liu, W.; Zhao, G.; Yang, W.; Wang, Y.; Ma, J.; Shu, Y.; Lei, F.; Gao, G. F., Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: phylogenetic, structural, and coalescent analyses. *The Lancet* **2013**, *381*, 1926-1932.
8. Hay, A. J.; Gregory, V.; Douglas, A. R.; Lin, Y. P., The evolution of human influenza viruses. *Philosophical Transactions of the Royal Society of London. Series B* **2001**, *356*, 1861-1870.
9. Wagner, R.; Matrosovich, M.; Klenk, H. D., Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Reviews in Medical Virology* **2002**, *12*, 159 - 166.
10. Onsirirakul, N.; Nakakita, S.-i.; Boonarkart, C.; Kongchanagul, A.; Suptawiwat, O.; Puthavathana, P.; Chaichuen, K.; Kittiniyom, K.; Suzuki, Y.; Auewarakul, P., Substrate specificity of avian influenza H5N1 neuraminidase. *World Journal of Virology* **2014**, *3*, 30-36.
11. Matrosovich, M.; Tuzikov, A.; Bovin, N.; Gambaryan, A.; Klimov, A.; Castrucci, M. R.; Donatelli, I.; Kawaoka, Y., Early Alterations of the Receptor-Binding Properties of H1, H2, and H3 Avian Influenza Virus Hemagglutinins after Their Introduction into Mammals. *Journal of Virology* **2000**, *74*, 8502-8512.
12. Shinya, K.; Ebina, M.; Yamada, S.; Ono, M.; Kasai, N.; Kawaoka, Y., Avian Flu: Influenza Virus Receptors in the Human Airway. *Nature* **2006**, *440*, 435 - 436.

13. Ito, T.; Couceiro, J. N. S. S.; Kelm, S.; Baum, L. G.; Krauss, S.; Castrucci, M. R.; Donatelli, I.; Kida, H.; Paulson, J. C.; Webster, R. G.; Kawaoka, Y., Molecular Basis for the Generation in Pigs of Influenza A Viruses with Pandemic Potential. *Journal of Virology* **1998**, *72*, 7367-7373.
14. Bateman, A. C.; Karamanska, R.; Busch, M. G.; Dell, A.; Olsen, C. W.; Haslam, S. M., Glycan Analysis and Influenza A Virus Infection of Primary Swine Respiratory Epithelial Cells: THE IMPORTANCE OF NeuAca2-6 GLYCANS. *The Journal of Biological Chemistry* **2010**, *285*, 34016-34026.
15. de Graaf, M.; Fouchier, R. A. M., Role of receptor binding specificity in influenza A virus transmission and pathogenesis. *The EMBO Journal* **2014**, *33*, 823.
16. Rogers, G. N.; Paulson, J. C., Receptor Determinants of Human and Animal Influenza Virus Isolates: Differences in Receptor Specificity of the H3 Hemagglutinin Based on Species of Origin. *Virology* **1983**, *127*, 361-373.
17. Rogers, G. N.; D'Souza, B. L., Receptor Binding Properties of Human and Animal H1 Influenza Virus Isolates. *Virology* **1989**, *173*, 317-322.
18. Connor, R. J.; Kawaoka, Y.; Webster, R. G.; Paulson, J. C., Receptor Specificity in Human, Avian, and Equine H2 and H3 Influenza Virus Isolates. *Virology* **1994**, *205*, 17-23.
19. Gambaryan, A. S.; Tuzikov, A. B.; Piskarev, V. E.; Yamnikova, S. S.; Lvov, D. K.; Robertson, J. S.; Bovin, N. V.; Matrosovich, M. N., Specification of Receptor-Binding Phenotypes of Influenza Virus Isolates from Different Hosts Using Synthetic Sialylglycopolymers: Non-Egg-Adapted Human H1 and H3 Influenza A and Influenza B Viruses Share a Common High Binding Affinity for 6'-Sialyl(N-acetyl)lactosamine. *Virology* **1997**, *232*, 345-350.
20. 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* **1997**, *233*, 224-234.
21. Gamblin, S. J.; Skehel, J. J., Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins. *Journal of Biological Chemistry* **2010**, *285*, 28403-28409.
22. Baum, L. G.; Paulson, J. C., The N2 Neuraminidase of Human Influenza Virus Has Acquired a Substrate Specificity Complementary to the Hemagglutinin Receptor Specificity. *Virology* **1991**, *180*, 10-15.
23. Kobasa, D.; Kodihalli, S.; Luo, M.; Castrucci, M. R.; Donatelli, I.; Suzuki, Y.; Suzuki, T.; Kawaoka, Y., Amino Acid Residues Contributing to the Substrate Specificity of the Influenza A Virus Neuraminidase. *Journal of Virology* **1999**, *73*, 6743-6751.
24. Kobasa, D.; Wells, K.; Kawaoka, Y., Amino Acids Responsible for the Absolute Sialidase Activity of the Influenza A Virus Neuraminidase: Relationship to Growth in the Duck Intestine. *Journal of Virology* **2001**, *75*, 11773-11780.
25. França de Barros Jr, J.; Sales Alviano, D.; da Silva, M. H.; Dutra Wigg, M.; Sales Alviano, C.; Schauer, R.; dos Santos Silva Couceiro, J. N., Characterization of Sialidase from an Influenza A (H3N2) Virus Strain: Kinetic Parameters and Substrate Specificity. *Intervirology* **2003**, *46*, 199-206.

26. Mochalova, L.; Kurova, V.; Shtyrya, Y.; Korchagina, E.; Gambaryan, A.; Belyanchikov, I.; Bovin, N., Oligosaccharide Specificity of Influenza H1N1 Virus Neuraminidases. *Arch Virol* **2007**, *152*, 2047-2057.
27. Garcia, J.-M.; Lai, J. C. C.; Haselhorst, T.; Choy, K. T.; Yen, H.-L.; Peiris, J. S. M.; von Itzstein, M.; Nicholls, J. M., Investigation of the Binding and Ceavage Characteristics of N1 Neuraminidases from Avian, Seasonal, and Pandemic Influenza Viruses using Saturation Transfer Difference Nuclear Magnetic Resonance. *Influenza Other Respir. Viruses* **2014**, *8*, 235-242.
28. Yen, H.-L.; Liang, C.-H.; Wu, C.-Y.; Forrest, H. L.; Ferguson, A.; Choy, K.-T.; Jones, J.; Wong, D. D.-Y.; Cheung, P. P.-H.; Hsu, C.-H.; Li, O. T.; Yuen, K. M.; Chan, R. W. Y.; Poon, L. L. M.; Chan, M. C. W.; Nicholls, J. M.; Krauss, S.; Wong, C.-H.; Guan, Y.; Webster, R. G.; Webby, R. J.; Peiris, M., Hemagglutinin–neuraminidase balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in ferrets. *Proceedings of the National Academy of Sciences* **2011**, *108*, 14264-14269.
29. Mitnaul, L. J.; Matrosovich, M. N.; Castrucci, M. R.; Tuzikov, A. B.; Bovin, N. V.; Kobasa, D.; Kawaoka, Y., Balanced Hemagglutinin and Neuraminidase Activities Are Critical for Efficient Replication of Influenza A Virus. *Journal of Virology* **2000**, *74*, 6015-6020.
30. Stoll, V.; Stewart, K. D.; Maring, C. J.; Muchmore, S.; Giranda, V.; Gu, Y.-g. Y.; Wang, G.; Chen, Y.; Sun, M.; Zhao, C.; Kennedy, A. L.; Madigan, D. L.; Xu, Y.; Saldivar, A.; Kati, W.; Laver, G.; Sowin, T.; Sham, H. L.; Greer, J.; Kempf, D., Influenza Neuraminidase Inhibitors: Structure-Based Design of a Novel Inhibitor Series. *Biochemistry* **2003**, *42*, 718-727.
31. von Itzstein, M., The war against influenza: discovery and development of sialidase inhibitors. *Nature Reviews Drug Discovery* **2007**, *6*, 967-974.
32. Vavricka, C. J.; Liu, Y.; Kiyota, H.; Sriwilaijaroen, N.; Qi, J.; Tanaka, K.; Wu, Y.; Li, Q.; Li, Y.; Yan, J.; Suzuki, Y.; Gao, G. F., Influenza neuraminidase operates via a nucleophilic mechanism and can be targeted by covalent inhibitors. *Nature Communications* **2013**, *4*, 1491.
33. Chong, A. K. J.; 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. *European Journal of Biochemistry* **1992**, *207*, 335-343.
34. Taylor, N. R.; von Itzstein, M., Molecular Modeling Studies on Ligand Binding to Sialidase from Influenza Virus and the Mechanism of Catalysis. *Journal of Medicinal Chemistry* **1994**, *37*, 616-624.
35. Tiralongo, J.; Pegg, M. S.; von Itzstein, M., Effect of Substrate Aglycon on Enzyme Mechanism in the Reaction of Sialidase from Influenza Virus. *FEBS Letters* **1995**, *372*, 148-150.
36. Thomas, A.; Jourand, D.; Bret, C.; Amara, P.; Field, M. J., Is There a Covalent Intermediate in the Viral Neuraminidase Reaction? A Hybrid Potential Free-Energy Study. *Journal of the American Chemical Society* **1999**, *121*, 9693-9702.
37. Kati, W. M.; Montgomery, D.; Maring, C.; Stoll, V. S.; Giranda, V.; Chen, X.; Laver, W. G.; Kohlbrenner, W.; Norbeck, D. W., Novel α - and β -Amino Acid Inhibitors of Influenza Virus Neuraminidase. *Antimicrobial Agents and Chemotherapy* **2001**, *45*, 2563-2570.

38. 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* **2006**, 443, 45-49.
39. Tong, S.; Zhu, X.; Li, Y.; Shi, M.; Zhang, J.; Bourgeois, M.; Yang, H.; Chen, X.; Recuenco, S.; Gomez, J.; Chen, L.-M.; Johnson, A.; Tao, Y.; Dreyfus, C.; Yu, W.; McBride, R.; Carney, P. J.; Gilbert, A. T.; Chang, J.; Guo, Z.; Davis, C. T.; Paulson, J. C.; Stevens, J.; Rupprecht, C. E.; Holmes, E. C.; Wilson, I. A.; Donis, R. O., New World Bats Harbor Diverse Influenza A Viruses. *PLoS Pathog.* **2013**, 9, e1003657.
40. Tumpey, T. M.; Maines, T. R.; Van Hoven, N.; Glaser, L.; Solórzano, A.; Pappas, C.; Cox, N. J.; Swayne, D. E.; Palese, P.; Katz, J. M.; García-Sastre, A., A Two-Amino Acid Change in the Hemagglutinin of the 1918 Influenza Virus Abolishes Transmission. *Science* **2007**, 315, 655-659.
41. Stevens, J.; Blixt, O.; Tumpey, T. M.; Taubenberger, J. K.; Paulson, J. C.; Wilson, I. A., Structure and Receptor Specificity of the Hemagglutinin from an H5N1 Influenza Virus. *Science* **2006**, 312, 404-410.
42. Shi, Y.; Zhang, W.; Wang, F.; Qi, J.; Wu, Y.; Song, H.; Gao, F.; Bi, Y.; Zhang, Y.; Fan, Z.; Qin, C.; Sun, H.; Liu, J.; Haywood, J.; Liu, W.; Gong, W.; Wang, D.; Shu, Y.; Wang, Y.; Yan, J.; Gao, G. F., Structures and receptor binding of hemagglutinins from human-infecting H7N9 influenza viruses. *Science* **2013**, 342, 243-247.
43. Dortmans, J. C. F. M.; Dekkers, J.; Wickramasinghe, I. N. A.; Verheije, M. H.; Rottier, P. J. M.; van Kuppeveld, F. J. M.; de Vries, E.; de Haan, C. A. M., Adaptation of novel H7N9 influenza A virus to human receptors. *Scientific Reports* **2013**, 3, 3058.
44. Tzarum, N.; de Vries, R. P.; Zhu, X.; Yu, W.; McBride, R.; Paulson, J. C.; Wilson, I. A., Structure and receptor binding of the hemagglutinin from a human H6N1 influenza virus. *Cell host & microbe* **2015**, 17, 369-376.
45. Yamada, S.; Suzuki, Y.; Suzuki, T.; Le, M. Q.; Nidom, C. A.; Sakai-Tagawa, Y.; Muramoto, Y.; Ito, M.; Kiso, M.; Horimoto, T.; Shinya, K.; Sawada, T.; Kiso, M.; Usui, T.; Murata, T.; Lin, Y.; Hay, A.; Haire, L. F.; Stevens, D. J.; Russell, R. J.; Gamblin, S. J.; Skehel, J. J.; Kawaoka, Y., Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* **2006**, 444, 378-382.
46. Herfst, S.; Schrauwen, E. J. A.; Linster, M.; Chutinimitkul, S.; de Wit, E.; Munster, V. J.; Sorrell, E. M.; Bestebroer, T. M.; Burke, D. F.; Smith, D. J.; Rimmelzwaan, G. F.; Osterhaus, A. D. M. E.; Fouchier, R. A. M., Airborne Transmission of Influenza A/H5N1 Virus Between Ferrets. *Science (New York, N.Y.)* **2012**, 336, 1534-1541.
47. Imai, M.; Watanabe, T.; Hatta, M.; Das, S. C.; Ozawa, M.; Shinya, K.; Zhong, G.; Hanson, A.; Katsura, H.; Watanabe, S.; Li, C.; Kawakami, E.; Yamada, S.; Kiso, M.; Suzuki, Y.; Maher, E. A.; Neumann, G.; Kawaoka, Y., Experimental adaptation of an influenza H5 haemagglutinin (HA) confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* **2012**, 486, 420-428.
48. Chen, L.-M.; Blixt, O.; Stevens, J.; Lipatov, A. S.; Davis, C. T.; Collins, B. E.; Cox, N. J.; Paulson, J. C.; Donis, R. O., In vitro evolution of H5N1 avian influenza virus toward human-type receptor specificity. *Virology* **2012**, 422, 105-113.
49. Lin, Y. P.; Gregory, V.; Collins, P.; Kloess, J.; Wharton, S.; Cattle, N.; Lackenby, A.; Daniels, R.; Hay, A., Neuraminidase Receptor Binding Variants of Human Influenza

- A(H3N2) Viruses Resulting from Substitution of Aspartic Acid 151 in the Catalytic Site: a Role in Virus Attachment? *Journal of Virology* **2010**, 84, 6769-6781.
50. Zhu, X.; McBride, R.; Nycholat, C. M.; Yu, W.; Paulson, J. C.; Wilson, I. A., Influenza Virus Neuraminidases with Reduced Enzymatic Activity That Avidly Bind Sialic Acid Receptors. *Journal of Virology* **2012**, 86, 13371-13383.
51. Hooper, K. A.; Bloom, J. D., A mutant influenza virus that uses an N1 neuraminidase as the receptor-binding protein. *Journal of Virology* **2013**.
52. Benton, D. J.; Wharton, S. A.; Martin, S. R.; McCauley, J. W., Role of Neuraminidase in Influenza A(H7N9) Virus Receptor Binding. *Journal of Virology* **2017**, 91.
53. Hu, Y.; Lu, S.; Song, Z.; Wang, W.; Hao, P.; Li, J.; Zhang, X.; Yen, H.-L.; Shi, B.; Li, T.; Guan, W.; Xu, L.; Liu, Y.; Wang, S.; Zhang, X.; Tian, D.; Zhu, Z.; He, J.; Huang, K.; Chen, H.; Zheng, L.; Li, X.; Ping, J.; Kang, B.; Xi, X.; Zha, L.; Li, Y.; Zhang, Z.; Peiris, M.; Yuan, Z., Association between adverse clinical outcome in human disease caused by novel influenza A H7N9 virus and sustained viral shedding and emergence of antiviral resistance. *The Lancet* **2013**, 381, 2273-2279.
54. McKimm-Breschkin, J. L., Resistance of influenza viruses to neuraminidase inhibitors — a review. *Antiviral Research* **2000**, 47, 1-17.
55. Yen, H.-L.; Hoffmann, E.; Taylor, G.; Scholtissek, C.; Monto, A. S.; Webster, R. G.; Govorkova, E. A., Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. *Journal of Virology* **2006**, 80, 8787-8795.
56. Yen, H. L.; McKimm-Breschkin, J. L.; Choy, K. T.; Wong, D. D.; Cheung, P. P.; Zhou, J.; Ng, I. H.; Zhu, H.; Webby, R. J.; Guan, Y.; Webster, R. G.; Peiris, J. S., Resistance to neuraminidase inhibitors conferred by an R292K mutation in a human influenza virus H7N9 isolate can be masked by a mixed R/K viral population. *MBio* **2013**, 4.
57. Sleeman, K.; Guo, Z.; Barnes, J.; Shaw, M.; Stevens, J.; Gubareva, L. V., R292K substitution and drug susceptibility of influenza A(H7N9) viruses. *Emerging Infectious Diseases* **2013**, 19, 1521-1524.
58. Wu, Y.; Bi, Y.; Vavricka, C. J.; Sun, X.; Zhang, Y.; Gao, F.; Zhao, M.; Xiao, H.; Qin, C.; He, J.; Liu, W.; Yan, J.; Qi, J.; Gao, G. F., Characterization of two distinct neuraminidases from avian-origin human-infecting H7N9 influenza viruses. *Cell Research* **2013**, 23, 1347-1355.
59. Hai, R.; Schmolke, M.; Leyva-Grado, V. H.; Thangavel, R. R.; Margine, I.; Jaffe, E. L.; Krammer, F.; Solorzano, A.; Garcia-Sastre, A.; Palese, P.; Bouvier, N. M., Influenza A(H7N9) virus gains neuraminidase inhibitor resistance without loss of in vivo virulence or transmissibility. *Nature communications* **2013**, 4, 2854.
60. Li, Y.; Cao, H.; Dao, N.; Luo, Z.; Yu, H.; Chen, Y.; Xing, Z.; Baumgarth, N.; Cardona, C.; Chen, X., High-throughput neuraminidase substrate specificity study of human and avian influenza A viruses. *Virology* **2011**, 415, 12-19.
61. Xu, R.; Zhu, X.; McBride, R.; Nycholat, C. M.; Yu, W.; Paulson, J. C.; Wilson, I. A., Functional Balance of the Hemagglutinin and Neuraminidase Activities Accompanies the Emergence of the 2009 H1N1 Influenza Pandemic. *Journal of Virology* **2012**, 86, 9221-9232.

62. Raab, M.; Tvaroška, I., The Binding Properties of the H5N1 Influenza Virus Neuraminidase as Inferred from Molecular Modeling. *J. Mol. Model.* **2011**, *17*, 1445-1456.
63. Jongkon, N.; Sangma, C., Receptor Recognition Mechanism of Human Influenza A H1N1 (1918), Avian Influenza A H5N1 (2004), and Pandemic H1N1 (2009) Neuraminidase. *J. Mol. Model.* **2012**, *18*, 285-293.
64. Woods, C. J.; Malaisree, M.; Long, B.; McIntosh-Smith, S.; Mulholland, A. J., Computational assay of H7N9 influenza neuraminidase reveals R292K mutation reduces drug binding affinity. *Scientific reports* **2013**, *3*, 3561.
65. Tran-To Su, C.; Ouyang, X.; Zheng, J.; Kwoh, C.-K., Structural analysis of the novel influenza A (H7N9) viral neuraminidase interactions with current approved neuraminidase inhibitors oseltamivir, zanamivir, and peramivir in the presence of mutation R289K. *BMC Bioinformatics* **2013**, *14*, S7.
66. Kovalenko, A. Three-dimensional Rism Theory for Molecular Liquids and Solid-Liquid Interfaces. In *Molecular Theory of Solvation*, Hirata, F., Ed.; Springer-Kluwer: Dordrecht, 2003, pp 169-275.
67. Phongphanphanee, S.; Yoshida, N.; Hirata, F., On the Proton Exclusion of Aquaporins: A Statistical Mechanics Study. *Journal of the American Chemical Society* **2008**, *130*, 1540-1541.
68. Phongphanphanee, S.; Rungrotmongkol, T.; Yoshida, N.; Hannongbua, S.; Hirata, F., Proton Transport through the Influenza A M2 Channel: Three-Dimensional Reference Interaction Site Model Study *Journal of the American Chemical Society* **2010**, *132*, 9782-9788.
69. Yoshida, N.; Imai, T.; Phongphanphanee, S.; Kovalenko, A.; Hirata, F., Molecular Recognition in Biomolecules Studied by Statistical-Mechanical Integral-Equation Theory of Liquids. *Journal of Physical Chemistry B* **2008**, *113*, 873-886.
70. Genheden, S.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Ryde, U., An MM/3D-RISM Approach for Ligand Binding Affinities. *J Phys Chem B* **2010**, *114*, 8505-8516.
71. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N [center-dot] log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089-10092.
72. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* **1983**, *79*, 926-935.
73. Xu, R.; de Vries, R. P.; Zhu, X.; Nycholat, C. M.; McBride, R.; Yu, W.; Paulson, J. C.; Wilson, I. A., Preferential Recognition of Avian-Like Receptors in Human Influenza A H7N9 Viruses. *Science* **2013**, *342*, 1230.
74. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The missing term in effective pair potentials. *The Journal of Physical Chemistry* **1987**, *91*, 6269-6271.
75. Lee, B.; Richards, F. M., The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology* **1971**, *55*, 379-IN4.
76. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T., Semianalytical treatment of solvation for molecular mechanics and dynamics. *Journal of the American Chemical Society* **1990**, *112*, 6127-6129.

77. Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A., Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theoretical Chemistry Accounts* **1999**, 101, 426-434.
78. Hawkins, G. D. C., Christopher J.; Truhlar, Donald G., Pairwise solute descreening of solute charges from a dielectric medium. *Chemical Physics Letters* **1995**, 246, 122-129.
79. Onufriev, A.; Bashford, D.; Case, D. A., Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics* **2004**, 55, 383-394.
80. Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A., Generalized Born Model with a Simple, Robust Molecular Volume Correction. *Journal of Chemical Theory and Computation* **2007**, 3, 156-169.
81. Hirata, F., *Molecular Theory of Solvation*. Springer-Kluwer: Dordrecht, 2003.
82. Chandler, D.; Andersen, H. C., Optimized Cluster Expansions for Classical Fluids. II. Theory of Molecular Liquids. *The Journal of Chemical Physics* **1972**, 57, 1930-1937.
83. Decha, P.; Rungrotmongkol, T.; Intharathep, P.; Malaisree, M.; Aruksakunwong, O.; Laohpongspaisan, C.; Parasuk, V.; Sompornpisut, P.; Pianwanit, S.; Kokpol, S.; Hannongbua, S., Source of high pathogenicity of an avian influenza virus H5N1: Why H5 is better cleaved by furin. *Biophysical Journal* **2008**, 95, 128-134.
84. Nunthaboot, N.; Rungrotmongkol, T.; Malaisree, M.; Kaiyawet, N.; Decha, P.; Sompornpisut, P.; Poovorawan, Y.; Hannongbua, S., Evolution of human receptor binding affinity of H1N1 hemagglutinins from 1918 to 2009 pandemic influenza A virus. *Journal of Chemical Information and Modeling* **2010**, 50, 1410-1417.
85. Kar, P.; Knecht, V., Mutation-induced loop opening and energetics for binding of tamiflu to influenza N8 neuraminidase. *Journal of Physical Chemistry B* **2012**, 116, 6137-49.
86. Chen, J.; Wang, J.; Zhu, W.; Li, G., A computational analysis of binding modes and conformation changes of MDM2 induced by p53 and inhibitor bindings. *J Comput Aided Mol Des* **2013**, 27, 965-974.
87. Meeprasert, A.; Hannongbua, S.; Rungrotmongkol, T., Key binding and susceptibility of NS3/4A serine protease inhibitors against hepatitis C virus. *Journal of Chemical Information and Modeling* **2014**, 54, 1208-1217.
88. Phanich, J.; Rungrotmongkol, T.; Sindhikara, D.; Phongphanphanee, S.; Yoshida, N.; Hirata, F.; Kungwan, N.; Hannongbua, S., A 3D-RISM/RISM study of the oseltamivir binding efficiency with the wild-type and resistance-associated mutant forms of the viral influenza B neuraminidase. *Protein Science* **2016**, 25, 147-158.
89. Schaduangrat, N.; Phanich, J.; Rungrotmongkol, T.; Lerdsamran, H.; Puthavathana, P.; Ubol, S., The significance of naturally occurring neuraminidase quasispecies of H5N1 avian influenza virus on resistance to oseltamivir: a point of concern. *Journal of General Virology* **2016**, 97, 1311-1323.
90. Rungrotmongkol, T.; Nunthaboot, N.; Malaisree, M.; Kaiyawet, N.; Yotmanee, P.; Meeprasert, A.; Hannongbua, S., Molecular insight into the specific binding of ADP-ribose to the nsP3 macro domains of chikungunya and venezuelan equine

encephalitis viruses: Molecular dynamics simulations and free energy calculations. *Journal of Molecular Graphics and Modelling* **2010**, 29, 347-353.

91. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling* **2011**, 51, 69-82.
92. Pan, P.; Li, L.; Li, Y.; Li, D.; Hou, T., Insights into susceptibility of antiviral drugs against the E119G mutant of 2009 influenza A (H1N1) neuraminidase by molecular dynamics simulations and free energy calculations. *Antiviral Research* **2013**, 100, 356-364.
93. Yesudas, J. P.; Blinov, N.; Dew, S. K.; Kovalenko, A., Calculation of Binding Free Energy of Short Double Stranded Oligonucleotides Using MM/3D-RISM-KH Approach. *Journal of Molecular Liquids* **2015**, 201, 68-76.
94. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research* **2000**, 33, 889-897.
95. Jensen, F., *Introduction to computational chemistry*. John Wiley & Sons: England, 2006.
96. Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation* **2012**, 8, 3314-21.
97. Case, D. A.; Babin, V.; Berryman, J.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A., In; University of California: 2014.
98. Gräter, F.; Schwarzl, S. M.; Dejaegere, A.; Fischer, S.; Smith, J. C., Protein/Ligand Binding Free Energies Calculated with Quantum Mechanics/Molecular Mechanics. *The Journal of Physical Chemistry B* **2005**, 109, 10474-10483.
99. Söderhjelm, P.; Aquilante, F.; Ryde, U., Calculation of Protein–Ligand Interaction Energies by a Fragmentation Approach Combining High-Level Quantum Chemistry with Classical Many-Body Effects. *The Journal of Physical Chemistry B* **2009**, 113, 11085-11094.
100. Ibrahim, M. A. A., Performance Assessment of Semiempirical Molecular Orbital Methods in Describing Halogen Bonding: Quantum Mechanical and Quantum Mechanical/Molecular Mechanical-Molecular Dynamics Study. *Journal of Chemical Information and Modeling* **2011**, 51, 2549-2559.
101. Wichapong, K.; Rohe, A.; Platzer, C.; Slynko, I.; Erdmann, F.; Schmidt, M.; Sippl, W., Application of Docking and QM/MM-GBSA Rescoring to Screen for Novel

- Myt1 Kinase Inhibitors. *Journal of Chemical Information and Modeling* **2014**, 54, 881-893.
102. Chen, J.; Wang, J.; Zhang, Q.; Chen, K.; Zhu, W., Probing Origin of Binding Difference of inhibitors to MDM2 and MDMX by Polarizable Molecular Dynamics Simulation and QM/MM-GBSA Calculation. *Scientific Reports* **2015**, 5, 17421.
103. Chen, J.; Wang, J.; Zhang, Q.; Chen, K.; Zhu, W., A comparative study of trypsin specificity based on QM/MM molecular dynamics simulation and QM/MM GBSA calculation. *Journal of Biomolecular Structure and Dynamics* **2015**, 33, 2606-2618.
104. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A., In; University of California: San Francisco, 2012.
105. Vavricka, C. J.; Li, Q.; Wu, Y.; Qi, J.; Wang, M.; Liu, Y.; Gao, F.; Liu, J.; Feng, E.; He, J.; Wang, J.; Liu, H.; Jiang, H.; Gao, G. F., Structural and Functional Analysis of Laninamivir and its Octanoate Prodrug Reveals Group Specific Mechanisms for Influenza NA Inhibition. *PLoS Pathog.* **2011**, 7, e1002249.
106. Xiong, X.; Martin, S. R.; Haire, L. F.; Wharton, S. A.; Daniels, R. S.; Bennett, M. S.; McCauley, J. W.; Collins, P. J.; Walker, P. A.; Skehel, J. J.; Gamblin, S. J., Receptor binding by an H7N9 influenza virus from humans. *Nature* **2013**, 499, 496-499.
107. Dunbrack, R., Rotamer Libraries in the 21st Century. *Current Opinion in Structural Biology* **2002**, 12, 431-440.
108. Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T., UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004**, 25, 1605-1612.
109. Sindhikara, D. J.; Yoshida, N.; Hirata, F., Placevent: an algorithm for prediction of explicit solvent atom distribution-application to HIV-1 protease and F-ATP synthase. *Journal of Computational Chemistry* **2012**, 33, 1536-43.
110. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J., GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *Journal of computational chemistry* **2008**, 29, 622-655.
111. York, D. M.; Darden, T. A.; Pedersen, L. G., The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *Journal of Chemical Physics* **1993**, 99, 8345-8348.
112. 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.* **1977**, 23, 327-341.
113. Onufriev, A.; Bashford, D.; Case, D. A., Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics* **2004**, 55, 383-394.

114. Phanich, J.; Rungrotmongkol, T.; Kungwan, N.; Hannongbua, S., Role of R292K mutation in influenza H7N9 neuraminidase toward oseltamivir susceptibility: MD and MM/PB(GB)SA study. *J Comput Aided Mol Des* **2016**, *30*, 917-926.
115. D.A. Case, V. B., J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman AMBER 14. 2014.
116. 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: Structure, Function, and Bioinformatics* **2008**, *71*, 1908-1918.
117. Olsson, M. H.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation* **2011**, *7*, 525-37.
118. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23*, 327-341.
119. 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. *The Journal of Chemical Physics* **1993**, *99*, 8345-8348.
120. Kovalenko, A.; Hirata, F., Three-dimensional Density Profiles of Water in Contact with a Solute of Arbitrary Shape: a RISM Approach. *Chemical Physics Letters* **1998**, *290*, 237-244.
121. Li, Q.; Qi, J.; Wu, Y.; Kiyota, H.; Tanaka, K.; Sahara, Y.; Ohrui, H.; Suzuki, Y.; Vavricka, C. J.; Gao, G. F., Functional and Structural Analysis of Influenza Virus Neuraminidase N3 Offers Further Insight into the Mechanisms of Oseltamivir Resistance. *Journal of Virology* **2013**, *87*, 10016-10024.
122. Air, G. M., Influenza neuraminidase. *Influenza and Other Respiratory Viruses* **2012**, *6*, 245-256.
123. Y.A. Shtyrya, L. V., Mochalova, N.V. Bovin, Influenza virus neuraminidase: Structure and function. *Acta Naturae* **2009**, 26-32.
124. Amaro, R. E.; Minh, D. D. L.; Cheng, L. S.; Lindstrom, W. M.; Olson, A. J.; Lin, J. H.; Li, W. W.; McCammon, J. A., Remarkable loop flexibility in avian influenza N1 and its implications for antiviral drug design. *Journal of the American Chemical Society* **2007**, *129*, 7764-+.
125. Han, N.; Mu, Y., Plasticity of 150-loop in influenza neuraminidase explored by Hamiltonian replica exchange molecular dynamics simulations. *PloS one* **2013**, *8*, e60995.
126. Amaro, R. E.; Swift, R. V.; Votapka, L.; Li, W. W.; Walker, R. C.; Bush, R. M., Mechanism of 150-cavity formation in influenza neuraminidase. *Nature Communications* **2011**, *2*, 388.

127. 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. *Nature Biotechnology* **2008**, *26*, 107-113.
128. Xu, D.; Newhouse, E. I.; Amaro, R. E.; Pao, H. C.; Cheng, L. S.; Markwick, P. R. L.; McCammon, J. A.; Li, W. W.; Arzberger, P. W., Distinct Glycan Topology for Avian and Human Sialopentasaccharide Receptor Analogues upon Binding Different Hemagglutinins: A Molecular Dynamics Perspective. *Journal of Molecular Biology* **2009**, *387*, 465-491.
129. Sasaki, G. L.; Elli, S.; Rudd, T. R.; Macchi, E.; Yates, E. A.; Naggi, A.; Shriver, Z.; Raman, R.; Sasisekharan, R.; Torri, G.; Guerrini, M., Human ($\alpha 2 \rightarrow 6$) and Avian ($\alpha 2 \rightarrow 3$) Sialylated Receptors of Influenza A Virus Show Distinct Conformations and Dynamics in Solution. *Biochemistry* **2013**, *52*, 7217-7230.
130. Rungrotmongkol, T.; Malaisree, M.; Nunthaboot, N.; Sompornpisut, P.; Hannongbua, S., Molecular prediction of oseltamivir efficiency against probable influenza A (H1N1-2009) mutants: molecular modeling approach. *Amino Acids* **2010**, *39*, 393-398.
131. Varghese, J. N.; Smith, P. W.; Sollis, S. L.; Blick, T. J.; Sahasrabudhe, A.; McKimm-Breschkin, J. L.; Colman, P. M., Drug design against a shifting target: a structural basis for resistance to inhibitors in a variant of influenza virus neuraminidase. *Structure*. **1998**, *6*, 735-746.

APPENDIX



VITA

NAME Mrs. Jiraphorn Phanich

DATE OF BIRTH 18 May, 1985

EDUCATIONS

2003 – 2007 B.Sc. (2nd Class Honor, Chemistry), Prince of Songkla University, Thailand

2009 – 2012 M.Sc. (Physical chemistry), Chulalongkorn University, Thailand

2012 – 2017 Ph.D. (Physical chemistry), Chulalongkorn University, Thailand

PUBLICATIONS

1 Phanich, J.; Rungrotmongkol, T.; Sindhikara, D.; Phongphanphanee, S.; Yoshida, N.; Hirata, F.; Kungwan, N.; Hannongbua, S. A 3D-RISM/RISM study of the oseltamivir binding efficiency with the wild-type and resistance-associated mutant forms of the viral influenza B neuraminidase. *Protein Sci* 25, 147-158 (2016).

2 Schaduangrat, N.; Phanich, J.; Rungrotmongkol, T.; Lerdsamran, H.; Puthavathana, P.; Ubol, S., The significance of naturally occurring neuraminidase quasispecies of H5N1 avian influenza virus on resistance to oseltamivir: a point of concern. *J Gen Virol* 97, 1311-1323 (2016).

3 Phanich, J., Rungrotmongkol, T., Kungwan, N. & Hannongbua, S. Role of R292K mutation in influenza H7N9 neuraminidase toward oseltamivir susceptibility: MD and MM/PB(GB)SA study. *J Comput Aided Mol Des* 30, 917-926 (2016).

4 Phanich, J., Siraphob, T., Kungwan, N. Rungrotmongkol, T. & Hannongbua, S. Glycan binding and specificity of viral influenza neuraminidases by classical molecular dynamics and replica exchange molecular dynamics simulations. (In preparation)