

**INCORPORATING MOLECULAR DYNAMICS SIMULATIONS INTO DRUG  
DESIGN TARGETING INFLUENZA N1 NEURAMINIDASES**

by

Ly Le

A dissertation submitted to the faculty of  
The University of Utah  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

May 2011

Copyright © Ly Le 2011

All Rights Reserved

# The University of Utah Graduate School

## STATEMENT OF DISSERTATION APPROVAL

The dissertation of Ly Le

has been approved by the following supervisory committee members:

Thanh N. Truong, Chair May 4<sup>th</sup> 2010  
Date Approved

Gregory A. Voth, Member May 4<sup>th</sup> 2010  
Date Approved

Jack Simons, Member May 4<sup>th</sup> 2010  
Date Approved

Peter F. Flynn, Member May 4<sup>th</sup> 2010  
Date Approved

Thomas Cheatham, Member May 4<sup>th</sup> 2010  
Date Approved

and by Henry S. White, Chair of

the Department of Chemistry

and by Charles A. Wight, Dean of The Graduate School.

## **ABSTRACT**

The FDA antiviral drug, Tamiflu (oseltamivir) is the front-line antiviral drug for the fight against 2003 avian flu (H5N1) as well as, more recently, for the 2009 swine flu (H1N1pdm). The drug functions as a neuraminidase inhibitor that prevents the release of new virions. Unfortunately, there is emerging evidence that the neuraminidase mutations H274Y and N294S render oseltamivir ineffective against the H5N1 virus. Of greater concern is the growing likelihood of the emergence of similar oseltamivir-resistant strains of H1N1pdm. It is therefore critical to understand the mechanisms for mutation-induced drug resistance in the H5N1 and H1N1pdm flu viruses in order to develop new and effective therapies. As molecular dynamics (MD) simulations have become an important tool for the study of biological systems, this dissertation aims to employ MD simulations for computer-aided rational drug design. Specifically, different MD simulation techniques were utilized in the investigation of oseltamivir-resistant mechanisms of H5N1/H1N1pdm and for the development of new antiviral drugs.

Chapter 1 is simply a general introduction to the whole thesis. Chapter 2 presents top-hits for H1N1pdm neuraminidase identified by virtual screening using ensemble-based docking technique, which incorporates protein flexibility into molecular docking.

Next in Chapter 3, progress in the development of two related methodologies for calculation of solvation free energy, one called the Coupled Reference Interaction Site model-hyper-netted chain/molecular dynamics (RISM/MD) approach, and another called Molecular Mechanics Poisson–Boltzmann linear response approximation and surface area contributions (MMPB-LRA-SA), are presented. The methods are expected to be applicable to the lead refinement process since they provide more reliable results than the continuum model but are less computationally expensive than conventional MD methods. In Chapter 4, we discuss our observations based on drug-protein endpoint interactions on how the mutations H274Y and N294S induce oseltamivir resistance in neuraminidase N1 subtypes. However, since the two mutations are non-active-site, endpoint interactions alone cannot fully account for the drug resistance. In Chapter 5, we present our finding of the drug binding pathway through electrostatic surface potential and steered MD simulation. The results reveal a novel oseltamivir-resistant mechanism in which the mutations rupture the drug binding funnel, in conjunction with the findings reported in Chapter 4. Our study not only assists understanding of oseltamivir-resistance in neuraminidase N1 subtypes, but also bears several important consequences for the intelligent design of new inhibitors that can overcome the established resistance strain.

## CONTENTS

ABSTRACT .....	iii
LIST OF FIGURES .....	vii
LIST OF TABLES .....	ix
ACKNOWLEDGMENTS .....	x
1. GENERAL INTRODUCTION .....	1
2. PROMISING ANTIVIRAL DRUGS FOR SWINE FLU H1N1PDM REVEALED BY VIRTUAL SCREENING USING ENSEMBLE-BASED DOCKING .....	6
2.1 Receptor .....	6
2.1.1 Molecular model of H1N1dpm neuraminidase .....	6
2.1.2 MD simulation of ionized oseltamivir-neuraminidase complex .....	9
2.1.3 RMSD clustering to extract receptor ensembles from an all-atom MD simulation .....	10
2.2 Ligands .....	12
2.3 Molecular docking .....	13
2.4. Results and discussion .....	15
2.4.1 Comparing binding modes of oseltamivir from molecular docking and MD simulation .....	15
2.4.2. Ranking of the top hit compounds .....	18
2.4.3. Hydrogen bond analysis of top binding compounds.....	30
2.5. Conclusion .....	34
3. NEW APPROACHES FOR SOLVATION FREE ENERGY CALCULATION .....	36
3.1 Improving the performance of the coupled reference interaction site model – hyper-netted chain (RISM-HNC) / MD method for free energy of solvation .....	36
3.1.1 Introduction of RISM-HNC/MD .....	36
3.1.2 Using the hard-sphere (HS) free energy as the reference and a linear response approximation (LRA) to improve the performance of the RISM-HNC/MD method.....	37
3.1.3 Methodology assessment .....	38
3.1.4 Conclusion and outlook .....	41
3.2 Explicitly-solvated ligand contribution to continuum solvation models for binding free energies: Selectivity of Theophylline binding to an RNA aptamer .....	42
3.2.1 Introduction of MM-PB/LRA-SA.....	42

3.2.2 Using MM-PB/LRA-SA to study selectivity of Theophylline binding to an RNA aptamer .....	44
3.2.3 Conclusion and outlook .....	48
<b>4. MOLECULAR MODELING OF MUTATION-INDUCED OSELTAMIVIR DRUG RESISTANCE FOR INFLUENZA N1 NEURAMINIDASES .....</b>	<b>50</b>
4.1 Literature review of oseltamivir-resistance and our proposed study .....	50
4.2 Computational details .....	51
4.3 Results and discussion .....	53
4.4 Conclusion and outlook .....	61
<b>5. AN ELECTROSTATIC FUNNEL DIRECTS BINDING OF OSELTAMIVIR TO INFLUENZA N1 SUBTYPE NEURAMINDASES .....</b>	<b>64</b>
5.1 Characteristics of the electrostatic surface potential of N1 neuraminidases shed light for further study on drug binding and drug resistance .....	64
5.2 Computational details .....	66
5.3 Results .....	68
5.4 Discussion and future studies .....	76
5.5 Supplementary information .....	78
APPENDIX .....	80
REFERENCES .....	83

## LIST OF FIGURES

1.1 Oseltamivir functions as a neuraminidase inhibitor, preventing the release of new swine flu H1N1pdm viruses.....	2
2.1 Discrete optimized protein energy (DOPE) per-residue of homology and molecular model of H1N1pdm neuraminidase in comparison with that of template avian H5N1 neuraminidase. ....	7
2.2 Schematic representations of drug-bound simulation systems. ....	8
2.3 The 13 representative ensembles resulting from clustering analysis, accounting for 96.2% of the configuration space, are ordered by the corresponding simulation time ....	10
2.4 Overlapping structures of the four most populated clusters in Figure 2.3 .....	11
2.5 Oseltamivir docked in SA cavity of H1N1pdm neuraminidase as compared to that from MD simulation, as shown in black with no hydrogen. ....	16
2.6 Binding energies of 33 compounds are listed with their NSC ID.....	19
2.7 Binding energy spectra vs. simulation time of zanamivir, oseltamivir and peramivir. ....	30
2.8 Peramivir (A) and zanamivir (B) docked in the H1N1pdm neuraminidase structure.	32
2.9 Docked structure of top-hits against swine flu neuraminidase .....	34
3.1 Chemical structures of theophylline (A) and derivatives, 1-methylxanthine (B), 3-methylxanthine (C), xanthine (D) and caffeine (E). ....	45
3.2 Conformation of theophylline-bound RNA aptamer after MD simulation, with triples labeled by residue and distinguished by separate colors, visualized using the Pymol viewing program. <sup>56</sup> Grey balls represent Mg <sup>2+</sup> ions. ....	46
3.3 Relative binding free energies calculated by (A) MM-PBSA and (B) MM-PB/LRASA	



for the 5 theophylline analogues compared with experimental data, showing linear fits of data with R <sup>2</sup> values. ....	47
4.1 Root mean squared deviation (RMSD) of WT and mutant avian H5N1 and swine H1N1pdm neuraminidases across six 20 ns production simulations .....	54
4.2 Network and occupancy of hydrogen bonds stabilizing oseltamivir in the sialic acid binding pocket of wild type and drug-resistant mutant avian H5N1 neuraminidases .....	55
4.3 Network and occupancy of hydrogen bonds stabilizing oseltamivir in the sialic acid binding pocket of wild type and drug-resistant mutant avian H1N1pdm neuraminidases	56
4.4. The solvent accessible surface area of oseltamivir's pentyl group (PG-SASA) in H5N1 and H1N1pdm wild type and mutant simulations .....	59
4.5 Electrostatic surface potential of avian H5N1 and swine H1N1pdm neuraminidases in oseltamivir-bound simulations, revealing a positively charged pathway into the binding pocket .....	62
5.1 Electrostatic surface potential of the SA binding pocket of H1N1pdm and oseltamivir .....	65
5.2 Distances between hydrogen bond acceptor-donor pairs between oseltamivir and active site amino acids vs. simulation time in simSMD1 .....	69
5.3 Forced unbinding of oseltamivir from H5N1 neuraminidase .....	70
5.4 Oseltamivir in “flipped” position (position of oseltamivir at 7.5ns in simSMD1) in comparison with its stable equilibrium position, shown with the electrostatic surface potential of the protein .....	71
5.5 Escape of oseltamivir from H5N1 neuraminidase in equilibrium simulations from simFEQ1 and simFEQ4 .....	74
5.6 Escape and rebinding of oseltamivir through the electrostatic binding funnel in H5N1 neuraminidase .....	75

## LIST OF TABLES

2.1 RMSD of the binding structures of oseltamivir from docking results in the 13 receptor ensembles compared to those extracted from the 20ns MD simulation. ....	17
2.2 Docking results for 33 compounds ranked by harmonic mean of binding energy. The first column is the final rank and also the compound ID.....	20
2.3 The frequency of hydrogen bond (in percent) for binding site residues with known drugs and top binding compounds utilizing the hydrogen bond distance and angle cutoffs of 3.5Å and 45°, respectively. ....	31
3.1. Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/ MD compared to experimental values, FEP values, and stand-alone RISM-HNC results of Sato et al. ....	39
3.2 Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/ MD compared to experimental values, and those of Wan et al. ....	40
3.3 Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/ MD with the GAFF force field and HF/3-21G charges, compared to experimental values. ....	41
3.4 Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/MC with the OPLS force field and HF/6-31G(d,p) charges, compared to experimental values and results of TP.....	42
3.5 Calculated binding free energies of theophylline and analogues, compared between different methods and with experimental values.....	48
4.1 Summary of simulations. The “ensemble” column lists the variables held constant during the simulations; N, p and T correspond to the number of atoms, pressure, and temperature, respectively.....	61
5.1 Summary of simulations. The “Structure” column lists the type of neuraminidase, associated mutations, and drug modeled in each system. ....	66
5.2 Summary of FEQ1-10 simulations starting from “flipped” position of oseltamivir taken from simSMD1 at 7.5 ns. ....	72

## ACKNOWLEDGMENTS

First, I am very grateful to my advisor Prof. Thanh N. Truong for his tremendous support during my PhD study. Second, I would especially like to thank Prof. Klaus Schulten for giving me an invaluable opportunity to collaborate with his research group where I had a chance to immerse myself to an exceptionally professional and challenging environment and learn a great deal from that.

Many thanks to all of my past instructors at the University of Utah who made significant contributions to my training. I would like to especially thank Dr. Holly Freedman for explaining to me insights into different methods for free energy of solvation and teaching me how to run an MD simulation for the very first time, Dr. Zingfa Xiao for teaching me homology modeling and docking, Prof. Valeria Molinero for teaching me coarse-grained modeling. At the University of Illinois at Urbana Champaign, I am greatly indebted to Dr. Eric Lee for teaching me SMD simulations and also working closely with me on the “swine flu” project. I would also like to thank Dr. Saam Jan for teaching me drug parameterization.

I am sincerely grateful to my committee members: to Prof. Greg Voth for helping me to put my research plan into perspective, to Prof. Tom Cheatham for teaching me molecular modeling and allowing me to attend his group meeting to gain more experience in the subject, and to Prof. Jack Simons for teaching me how to design QM/MM experiments for my proposal defense.

I also appreciate the help I have received from many other faculty and staff, especially Prof. Michael Morse and Ms. Jo Hoovey from the Department of Chemistry, Dr. Martin Cuma from the Center for High Performance Computing.

I would like to thank my old teacher, Prof. Le Trong Tin, for supporting my dream of becoming as good a teacher as he is. Also, I would like to thank Prof. Ernest Davidson for his support of my PhD study pursuit in the US through the VEF.

Finally, I want to thank my brothers and parents for the unconditional love they have for me, and all of my friends in Utah and UIUC, especially Hieu Le, whose friendship and companionship have made my PhD experience so enjoyable and unforgettable.

This research was supported by grant(s) from The National Science Foundation (grant number 0326027) and the National Institutes of Health P41-RR05969. The computing resources were provided by the University of Utah's Center for High Performance Computing and an allocation at the Texas Advanced Computing Center via Large Resources Allocation Committee grant MCA93S028. I also thank the Vietnam Education Foundation for their financial aid in the first two years of my PhD program.

Most of Chapter 2 presents “ Identified top-hits for H1N1pdm neuraminidase by using ensemble-based docking,” which originally appeared in *PLoS Curr Influenza*, RRN1030 (2009). Chapter 3 describes the prospects and recent progress of the Coupled RISM/MD approach for thermochemistry of solvation and its validations in some organic compound published in *J. Phys. Chem. B.* **112**, 2340 (2008), and a method to improve the accuracy of the MM-PBSA method by adding an explicitly solvated ligand contribution to continuum solvation models for binding free energies and an application to study the

selectivity of theophylline binding to an RNA aptamer appeared in *J. Phys. Chem. B*, **114**, 2227 (2010). All of Chapter 4 originally appeared as “Thermodynamics of Oseltamivir binding to neuraminidase N1 subtype” in *PLoS Curr Influenza*, RRN1015 (2009), and in a manuscript titled “Molecular modeling of mutation-induced Tamiflu drug resistance for swine influenza H1N1pdm and avian H5N1 influenza N1 neuraminidases“ and submitted to *J. Mol. Med.*

Chapter 5 has been submitted and is under revision for publication in *PLoS Computational Biology* as “An electrostatic funnel directs binding of oseltamivir to Influenza N1 neuraminidases.” Works reported in Chapter 4 and 5 were done in collaboration with Eric H. Lee and other members of the Schulten group.

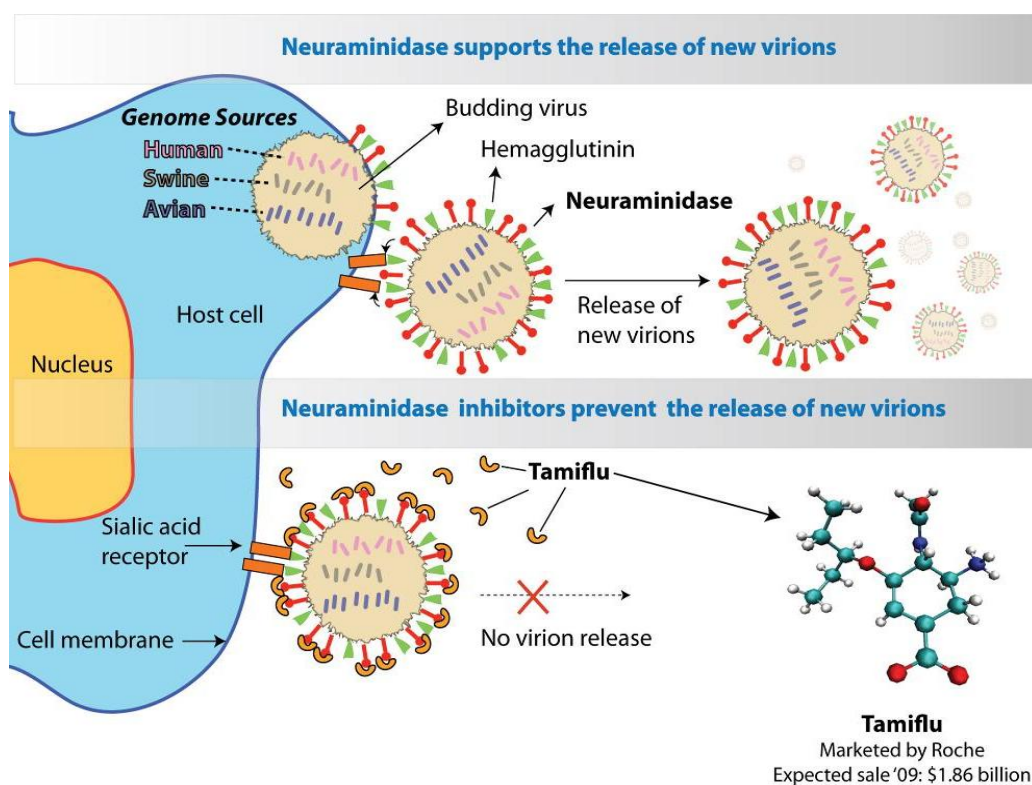
## CHAPTER 1

### GENERAL INTRODUCTION

Nowadays, information needed for computer-aided drug design can be accessed easily. There are more than 64,000 coordinate entries of macromolecules revealed by X-ray, NMR, and Electron Microscopy techniques available in the Brookhaven Protein Data Bank<sup>1</sup>. Many new target proteins whose 3D structures have not been solved experimentally can still be easily predicted by homology modeling with adequate reliability. Similarly, thousands of compounds including their drug-like properties are provided in free chemical databases such as the NCI Diversity set,<sup>2</sup> ZINC,<sup>3</sup> and the Drug Bank<sup>4</sup>. Therefore, it is crucial to develop suitable tools to take advantages of these promising resources for rational drug design. MD simulation has proven itself as a complement to lab experiments to fill in the gaps in our knowledge between 3D structures of biological targets and their potential inhibitors.<sup>5</sup>

In the fight against influenza A viruses, which have killed millions of human being throughout history, it is important to utilize MD simulations to speed up the rational design of antiviral drugs. Due to their ability to mutate quickly, the viruses can easily survive people's immune systems and resist current drugs. Drug resistance<sup>5</sup> became a big concern to public health in 2005 when several patients in Vietnam died in

spite of being given oseltamivir.<sup>6</sup> Currently, there are two pathogenic viruses, the 2003 avian flu (H5N1)<sup>7</sup> virus and recently the 2009 swine flu (H1N1pdm)<sup>8</sup> virus, which share the same subtype 1 (N1) of neuraminidase. Neuraminidase, a glycoprotein component on a flu virus' surface that cleaves the alpha-ketosidic linkage of sialic acid (SA) located on human cells to release new virions that then infect nearby healthy cells, is the most important target for antiviral drug design.<sup>9</sup> Oseltamivir functions as a neuraminidase inhibitor that binds competitively to the SA binding site of the neuraminidase. The presence of oseltamivir prevent the release of new virions, as shown in Figure 1.1. In this dissertation, different MD simulation techniques<sup>10, 11</sup> are applied to investigation of the oseltamivir-resistant mechanism and to rational drug design targeting neuraminidase N1 subtypes.



**Figure 1.1** Oseltamivir functions as a neuraminidase inhibitor, preventing the release of new swine flu H1N1pdm viruses.

In the lead discovery stage, thousands of compounds from free chemical databases are docked against the rigid structures of receptors. Such screening processes are fast but are associated with a high number of false positives and false negatives. Flexible docking has improved the outcome of this process, but to a very limited degree. A recent method that significantly increases the accuracy of the lead identification process is called ensemble-based docking.<sup>12</sup> This method combines the docking algorithm with the dynamic structure of proteins taken from MD simulations as well as statistical analysis of binding energy. Chapter 2 presents top-hits for H1N1pdm neuraminidase identified by virtual screening using the ensemble-based docking technique.<sup>13</sup> The molecular model of H1N1pdm used in the docking experiment was built from the crystal structure of avian H5N1 and the amino acid sequence of swine flu H1N1pdm from GenBank.<sup>14</sup>

The next step in drug development after screening is how to distinguish among equally well bound hits. In lead optimization, accurate calculation of ligand binding affinity is crucial. In most computational methods solvation components and intramolecular components are computed separately and summed to yield the total free energy. A typical mammalian cell consists of more than 85% water, and so most biochemical processes modeled must be treated within an aqueous environment and solvation free energies need to be taken into account. Commonly used solvation free energy methodologies include the computationally inexpensive continuum solvation models, which unfortunately do not always lead to reliable results; and molecular simulation methods, such as thermodynamic integration and thermodynamic perturbation, representing the opposite extreme, being quite accurate but complex and time-consuming.



In Chapter 3, the new methodologies for calculation of solvation free energy called coupled reference interaction site model-hyper-netted chain (RISM-HNC)/MD<sup>15</sup> and molecular mechanics Poisson–Boltzmann/ linear response approximation and surface area contributions (MM-PB/LRA-SA)<sup>16</sup> developed by Freedman et al. are presented. The authors aim to develop methods that yield more reliable results than continuum models but are less computationally expensive than conventional MD methods. The methods have obtained very positive preliminary results and are still under development for application in drug design.

Another important aspect of rational drug design is to understand all drug-protein interactions at an atomistic level and how these interactions are ruptured by certain mutations, leading to drug resistance. Thorough understanding of drug resistant mechanisms is crucial for the design of new agents effective against current drug-resistant strains. To understand how mutations H274Y and N294S induce oseltamivir resistance in neuraminidase N1 subtype, analysis of the drug binding was studied thoroughly by equilibrium MD simulations. We have observed important hydrogen bond and hydrophobic interactions between the protein and drug that can not be seen in static crystal structures. The results discussed in Chapter 4 identify important potential sources of oseltamivir resistance, mainly the loss of hydrogen bonds with R152 and hydrophobic interaction of oseltamivir's bulky pentyl group in the case of mutant H274Y. This leads to our suggestion to replace the pentyl group by a more hydrophilic one by adding a few hydroxyl groups onto the drug. Since the two mutations are non-active-site, endpoint interactions alone cannot fully explain the drug resistance mechanism. Sources of drug resistance in mutant N294S remains unclear. However, the electrostatic surface potential

calculations from equilibrium (EQ) simulations reveal both residues 274 and 294 located on a distinct negatively charged column of residues, that can lead to another drug resistance mechanism.

In Chapter 5, we discuss our finding of the drug binding pathway through electrostatic surface potential and steered MD (SMD) simulation.<sup>17</sup> The findings were then validated by 10 subsequenced EQ simulations (EQ simulations with starting structure taken from SMD simulation). More importantly, mutation H274Y and N294S were found to be located on or adjacent to the binding funnel. The results reveal a novel oseltamivir-resistant mechanism, in which the mutations rupture the drug binding funnel, complementary to what is reported in Chapter 4. The findings in Chapter 4 are important for the rational design of a drug with good binding kinetics. It poses the question of whether similar binding funnels exist in other types of neuraminidases. The mutations affect binding of oseltamivir but certainly not its natural substrate SA. Another FDA antiviral drug, zanamivir (Relenza)<sup>18</sup> might enter the binding pocket in a similar way as SA does, leading to insensitivity to these two mutations.

Our study has provided insights into the oseltamivir-resistant mechanism and suggestions for further studies involving other commercial drugs, other subtypes of neuraminidase as well as intelligent design of new antiviral drugs that work for established drug-resistant variants. As oseltamivir resistance is emerging, the development of new neuraminidase inhibitors and other drug classes (e.g. M2 channel blockers) are crucial in the fight against influenza A virus.

## **CHAPTER 2**

### **PROMISING ANTIVIRAL DRUGS FOR SWINE FLU H1N1PDM REVEALED BY VIRTUAL SCREENING USING ENSEMBLE-BASED DOCKING**

Here we describe how top-binding compounds to the swine flu H1N1pdm were identified by the ensemble-based docking technique. This work is in collaboration with Hung Nguyen. My contribution was preparing ligands and the receptor for docking, repeating some docking procedures for validation, making some figures and writing part of the manuscript.

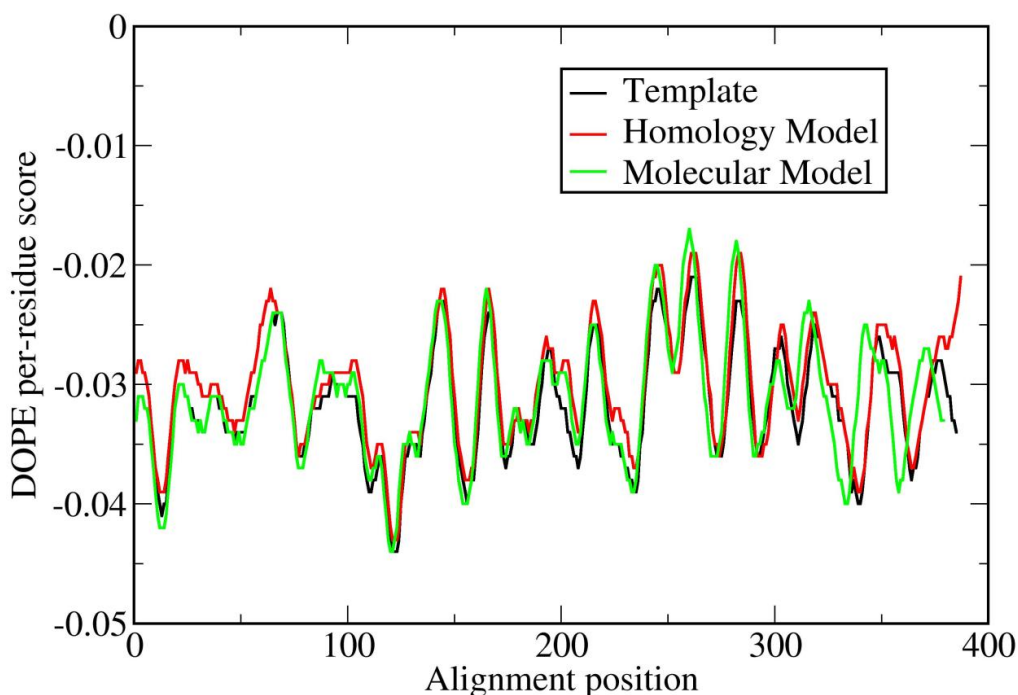
#### **2.1 Receptor**

##### **2.1.1 Molecular model of H1N1pdm neuraminidase**

The atomic structure of H1N1pdm neuraminidase is not known. However, the fact that the H1N1pdm strain responds positively to oseltamivir which is effective against H5N1, suggests similarities to H5N1 neuraminidase. The amino acid sequence of H1N1pdm neuraminidase was obtained from Genbank Locus ID CY041156 and that of H5N1 neuraminidase from the Protein Data Bank (PDB entry 2HU4).<sup>19</sup> Sequence alignment performed using Multiseq in VMD<sup>20</sup> shows that H1N1pdm has a higher percentage sequence identity (91.47 %) with H5N1 excluding the transmembrane and unstructured linker regions, than it has with historic Spanish H1N1 or the H1N1

counterparts in the Americas. At the drug binding pocket, the notable difference between H1N1pdm and H5N1 neuraminidase is the replacement of Y347 by N347.<sup>14</sup>

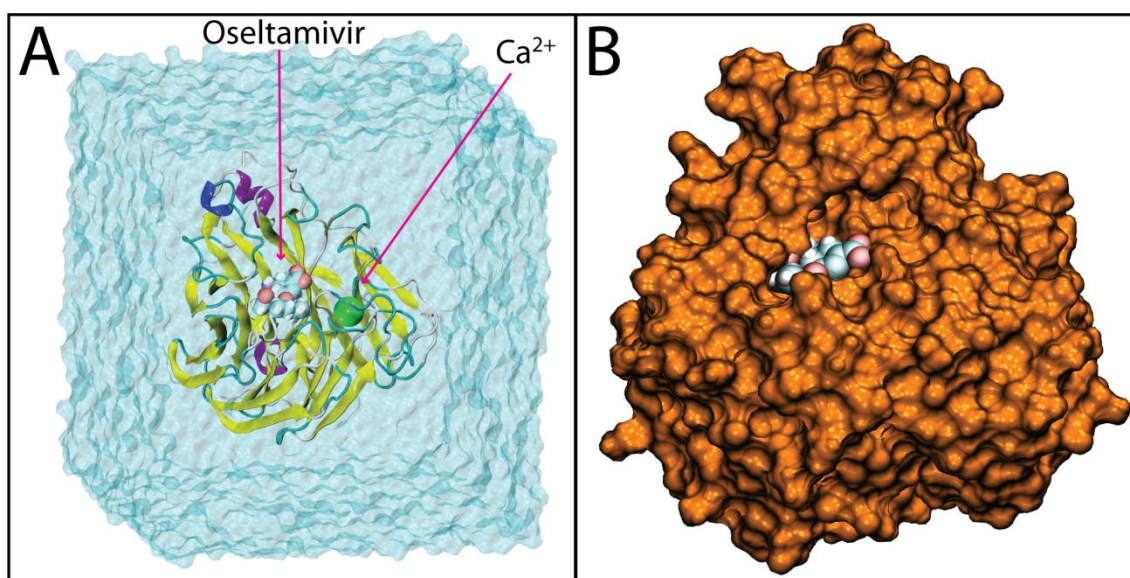
H5N1 neuraminidase was used as a template to build a homology model of H1N1pdm neuraminidase by Modeller9v7<sup>21</sup> and a molecular model of H1N1pdm by mutating the corresponding residues in H5N1 to match the wild type H1N1pdm using a home-made script in VMD. Both the homology and molecular model of swine H1N1pdm have also revealed high similarity to template, in discrete optimized protein energy (DOPE) per-residue score, shown in Figure 2.1. As DOPE is the most important factor to validate the quality of built 3D models,<sup>22</sup> the results indicate that the 3D structure of both the homology and molecular models are highly similar to the 3D model of the template.



**Figure 2.1** Discrete optimized protein energy (DOPE) per-residue of homology and molecular model of H1N1pdm neuraminidase in comparison with that of template avian H5N1 neuraminidase.

In order to easily retain the coordinates of oseltamivir, the crystallographically resolved water molecules, and a structurally relevant calcium ion near the native binding site for SA in the template, the molecular model was used for further study. The protein complexes were then solvated in a TIP3P<sup>23</sup> water box and ionized by NaCl (0.152M) to mimic physiological conditions. The solvated H1N1pdm system with bound oseltamivir is shown in Figure 2.2A, with schematic views of the buried drug in the SA binding site shown in Figure 2.2B.

Simulation parameters for oseltamivir were developed within the CHARMM force field<sup>24</sup> parameter scheme, to complement the CHARMM31 force field for proteins



**Figure 2.2** Schematic representations of drug-bound simulation systems. Shown here as a representative example of simulated systems is H1N1pdm bound to oseltamivir. In A), the simulation system is shown in the full explicit solvation box with oseltamivir and the active-site calcium ion labeled. In B), oseltamivir is shown buried in the SA binding pocket of H1N1pdm rendered in surface view.

with the CMAP correction.<sup>25</sup> Parameters for ligands were prepared using Paratool<sup>26</sup> in VMD. Structure optimization and frequency calculations were performed at the level of HF/6-31G\* by Gaussian03<sup>27</sup> and subsequently imported into Paratool. Atom types and charges already described in an existing CHARMM force field were assigned using those existing parameters; however, the atomic charges for the atoms in oseltamivir's six member ring were newly parameterized by dividing the ring into several small fragments, and atomic charges for the six-member ring atoms recalculated based on the total charge of each fragment. Fragments not explicitly defined in the CHARMM force field were modeled using analogs in CHARMM force field extensions which fit closely to the fragment being parameterized. The dihedral angle potentials, which render the respective torsions highly rigid due to electron delocalizations, were generated from the quantum mechanical calculations.

### 2.1.2 MD simulation of ionized oseltamivir-neuraminidase complex

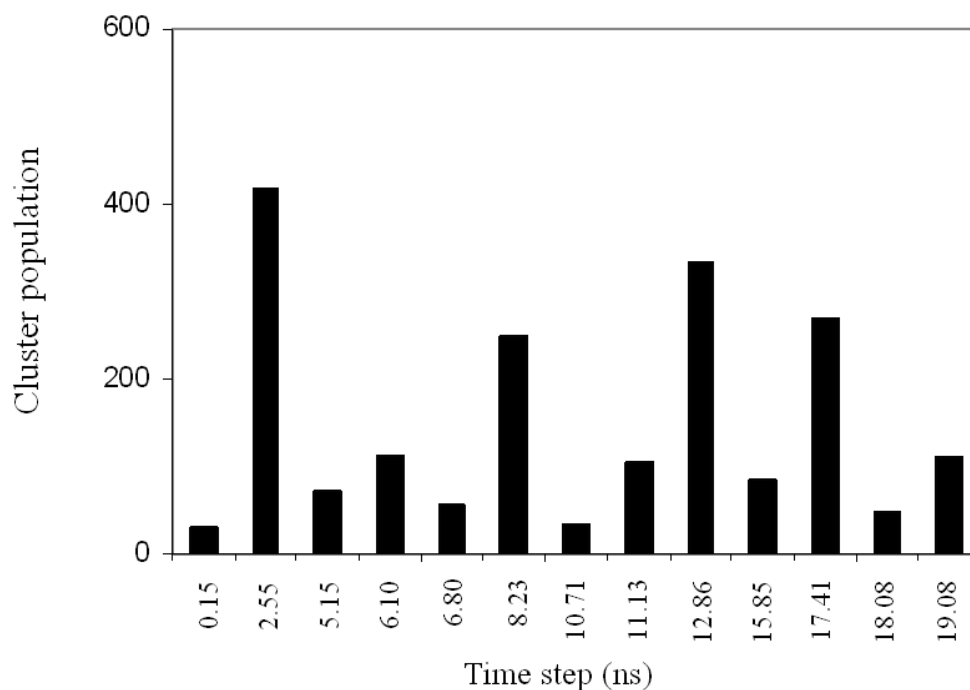
All simulations were performed using NAMD 2.7<sup>20</sup> and the CHARMM31 force field with the CMAP correction. The ionized systems were minimized for 10,000 integration steps and equilibrated for 20 ns with a 1 fs time step. Following this, a 20 ns unconstrained equilibration production run was performed for subsequent trajectory analysis, with frames stored after each picosecond (every 1000 time steps). Constant temperature ( $T = 300$  K) was enforced using Langevin dynamics with a damping coefficient of  $1 \text{ ps}^{-1}$ . Constant pressure ( $p = 1 \text{ atm}$ ) was enforced using the Nosé-Hoover Langevin piston method. Van der Waals interaction cutoff distances were set at  $12 \text{ \AA}$  (smooth switching function beginning at  $10 \text{ \AA}$ ) and long-range electrostatic forces were

computed using the particle-mesh Ewald (PME) summation method. The trajectories from our 20ns production run of EQ simulation of H1N1pdm neuraminidase with oseltamivir bound, were clustered as shown in Figure 2.3, and then used for ensemble-based docking.

### 2.1.3 RMSD clustering to extract receptor ensembles from an all-atom

#### MD simulation

The holo system with oseltamivir removed was used for RMSD clustering and the docking experiments. Clustering analyses were performed on 20 ns MD trajectories using the `g_cluster` tool in the Gromacs package.<sup>28</sup> In brief, snapshots at every 10 ps over



**Figure 2.3** 13 representative ensembles resulting from clustering analysis, accounting for 96.2% of the configuration space, are ordered by the corresponding simulation time. The four largest cluster ensembles account for over 66% of the configuration space from the 20 ns MD simulation.

the 20 ns simulation were recorded. 2000 resulting structures were superimposed, using all  $C_{\alpha}$  atoms to remove possible rotational and translational movements of the whole system. We have visually verified that the binding-site residues of avian H5N1 neuraminidase, which cover the active site and are responsible for interaction with putative inhibitors, can also be used for the swine flu H1N1pdm neuraminidase. The four most populated clusters is shown in Figure 2.4.

The RMSD clustering analysis was performed on this subset (117-119, 133-138, 146-152, 156, 179, 180, 196-200, 223-228, 243-247, 277, 278, 293, 295, 344-347, 368, 401, 402, and 426-441) using all-atoms (including side chains and hydrogen atoms) with



**Figure 2.4** Overlapping structures of the four most populated clusters. The most dominant cluster is colored in orange, the second one in silver, the third one in blue and the fourth one in cyan.



the cutoff of 1.5 Å. A total of 13 representative clusters were obtained, which account for 96.2% of the configuration space from the 20 ns MD trajectories. These 13 clusters were used in the docking experiments.

## 2.2 Ligands

The high percentage sequence identity (91.47 %) of swine H1N1pdm compared to avian H5N1 neuraminidase, and similarities in 3D structures (active-site 150-loop and SA binding site residues) explain why oseltamivir, which has been developed for the H5N1 virus, is also effective against the swine flu virus H1N1pdm. These facts also suggest that top-binding ligands of avian H5N1 are likely to have high affinity to the swine flu H1N1pdm neuraminidase. Therefore the 27 top hits for H5N1 neuraminidase, suggested by Cheng et al.<sup>7</sup> from performing full virtual screening on H5N1 neuraminidase using the diversity set from the National Cancer Institute (NCI) small molecule database, were used in our study.

In this study, 32 ligands were used as a starting point for searching for promising antiviral drugs against the H1N1pdm neuraminidase, using the ensemble-based docking approach. The ligand set includes the 27 top-hit compounds from Cheng et al. and six additional compounds for reference purpose, namely: SA, N-acetyl neuraminic acid, aka, NANA or Neu5Ac), DANA (2,3-didehydro-2-deoxy-N-acetyl neuraminic acid), oseltamivir, zanamivir, peramivir (clinical trial phase 3 drug candidate) and SK (shikimic acid, widely used as a chiral building block in the synthesis of oseltamivir). Ligand structures were obtained from the NCI and optimized by the MOPAC module which was integrated into the Avisto cyberinfrastructure.

### 2.3 Molecular docking

In the docking experiments, the 13 most representative configurations were used as receptors. AutoDockTools 1.5.2 was used to add polar hydrogens, assign Gasteiger charges<sup>11</sup> and create grid binding boxes. The volume of each grid box was 72 x 72 x 72, with the default 0.375 Å spacing. The binding box was positioned to encompass all three possible binding sites, namely the SA, 150 and 430 cavities. AutoGrid version 4.2.1 was used to calculate the binding affinities using the following atom types: A (aromatic carbon), C, N, NA (hydrogen bond accepting N), OA (hydrogen bond accepting O), P, S, SA (hydrogen bond accepting S), Cl, HD (polar hydrogen) and e (electrostatics).

AutoDockTools version 1.5.2 was also used to merge nonpolar hydrogens, add Gasteiger charges and visually set up rotatable bonds for each ligand via AutoTors. The Lamarckian genetic algorithm was used to do the docking experiments using AutoDock 4.2.1.<sup>16</sup> Docking parameters were chosen to reproduce structures of 13 corresponding oseltamivir–neuraminidase complexes in the MD simulation. Other parameters are as follows: trials of 100 dockings, population size of 200, random starting position and conformation, translation step range of 2.0 Å, rotation step range of 50 degrees, maximum number of generations of 27000, elitism of 1, mutation rate of 2%, crossover rate of 80%, local search rate of 6%, 8 million energy evaluations, unbound model was “same as bound”, and docked conformations were clustered with the tolerance of 2.0 Å RMSD.

Docking results were sorted by the lowest binding energy of the most populated cluster in cases of convergence. In the case of no dominant cluster, docking results were visually analyzed using VMD to choose the best binding pose.

Hydrogen bond analysis utilized a distance and angle cutoff of 3.5 Å and 45 degrees, respectively. The appearance frequency (AF) of important hydrogen bonds is calculated as follows:

$$AF = \frac{\sum_{i=1}^{13} n_i h_i}{\sum_{i=1}^{13} n_i} \quad (2.1)$$

where  $i$  is the index number of each ensemble;  $n_i$  the size of each ensemble; and  $h_i = 1$  if hydrogen bond exists, and 0 otherwise.

After the dockings, statistical calculations were performed to obtain the final binding energies for compound ranking, using the arithmetic mean (AM) and harmonic mean (HM) binding energies as defined in the previous study.

The arithmetic means were calculated directly from the binding energies

$$AM = \frac{\sum n E}{\sum_{i=1}^{13} n_i} \quad (2.2)$$

where  $E_i$  is the binding energy of each ensemble with the standard deviation

$$SD = \sqrt{\frac{\sum_{i=1}^{13} n_i (E_i - AM)^2}{\sum_{i=1}^{13} n_i}} \quad (2.3)$$

The harmonic means were calculated by first converting the binding energies into inhibition constant  $K_i$

$$K_i = e^{\frac{1000E_i}{RT}} \quad (2.4)$$

where R is the Boltzmann constant and T is the temperature (298.15K)

The harmonic means  $\overline{K}_i$  were calculated using:

$$\overline{K}_i = \frac{\sum_{i=1}^{13} n_i}{\sum_{i=1}^{13} \frac{n_i}{K_i(i)}} \quad (2.5)$$

then the calculated  $\overline{K}_i$  were converted back to HM binding energies.

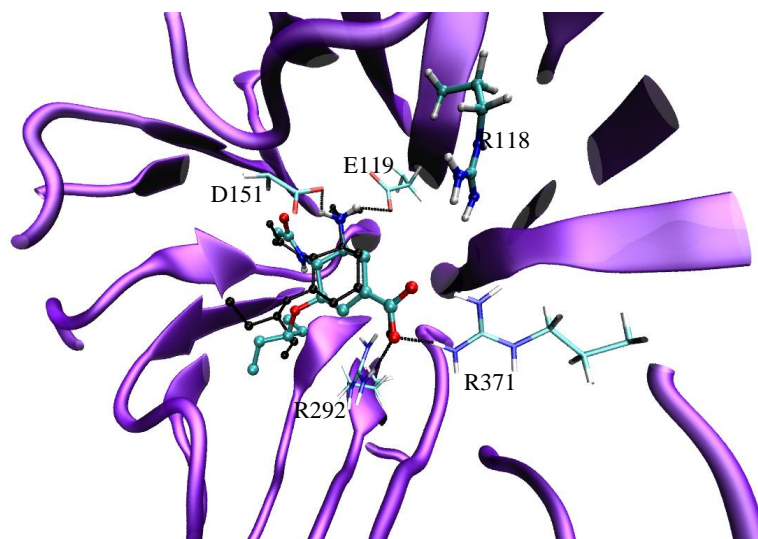
## 2.4. Results and discussion

### 2.4.1 Comparing binding modes of oseltamivir from molecular docking and MD simulation

First, we compared the binding modes of oseltamivir to the H1N1pdm neuraminidase resulting from the present docking experiment, to those from the previous full all-atom MD simulation. By applying the ensemble-based docking approach

suggested by McCammon and co-workers,<sup>29-31</sup> docking results are in good agreement with those from the MD simulation for oseltamivir. As shown in Figure 2.5, the two average structures are nearly superimposed except for the pentyl group. Also shown are residues, which frequently interact with oseltamivir via hydrogen bonds: E119, D151, R292, and R371. One main reason that its binding free energy is less than that in the H5N1 case, as discussed below, is the loss of contact with R118.

As shown in Table 2.1, the maximum all-atom RMSD between the two docking and MD ligand binding poses is less than 2 Å, and the average RMSD is of only 1.5 Å. The difference between the binding poses lies mainly in the orientation of the pentyl ether group, which interacts with the receptor primarily via hydrophobic forces. When excluding this group from the analysis, the average RMSD drops to below 1.1 Å.



**Figure 2.5** Oseltamivir docked in SA cavity of H1N1pdm neuraminidase as compared to that from MD simulation, as shown in black with no hydrogen.

**Table 2.1** RMSD of the binding structures of oseltamivir from docking results in the 13 receptor ensembles compared to those extracted from the 20ns MD simulation. Also listed are the percent of population of the receptor ensemble, the percent contribution of the 20 ns MD simulation, and the percent population of the largest docked oseltamivir cluster.

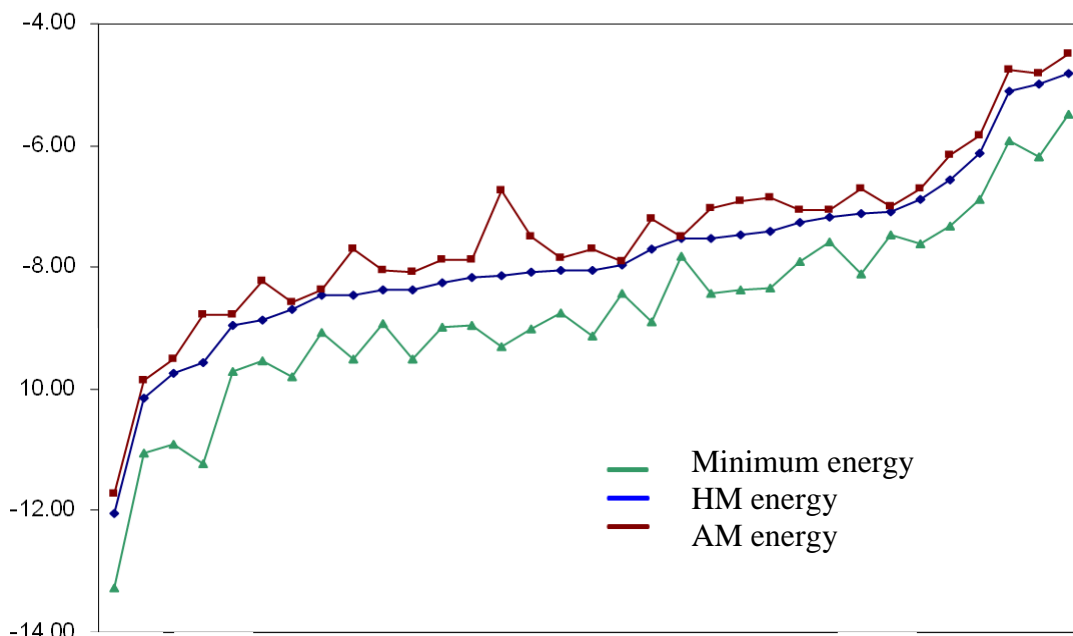
Ensemble	Percent population	All-atom RMSD	Excluding the pentyl group	Percent population of
1	22.65	1.21	1.21	96
2	18.70	1.41	0.63	81
3	10.20	1.74	1.17	73
4	9.45	1.95	1.37	65
5	8.40	1.74	1.04	83
6	7.00	1.20	1.17	88
7	5.20	1.80	1.20	66
8	3.20	1.11	1.15	89
9	3.10	1.13	0.56	72
10	2.30	1.77	0.90	69
11	2.25	1.81	1.17	61
12	1.90	1.55	1.54	74
13	1.50	1.16	1.13	58
Mean	Non-weighted	1.51	1.10	
	Weighted	1.48	1.06	

This comparison validates the accuracy of the ensemble-based docking approach. Docking results over 13 receptor ensembles also show consistent statistics with the most populated cluster having the lowest binding energy and the largest percent population of the ligand cluster is found in the ensemble with largest percent population in the receptor space (see ensemble 1 as compared to ensemble 13).

#### 2.4.2. Ranking of the top binding compounds

The ranking is made according to the harmonic mean of binding free energy instead of the arithmetic mean, since it gives more weight to the more populated clusters. The better performance of the harmonic mean of binding free energy, shown in Figure 2.6, is also seen in previous studies.<sup>32, 33</sup>

First, differences in the present ranking results for these top hits against H1N1pdm from those reported in the previous study for the avian H5N1 neuraminidase were observed. As shown in Table 2.2, some compounds rise (**2**, **5** and **6** as compared to the ranks of 16, 20 and 14, respectively for H5N1) while others move down (**29**, **19** and **25** as compared to 6, 3 and 9, respectively) in the ranking list. Oseltamivir retains its relatively high rank (with new/old rank of 7/4) while the changes in the rank for paramivir and zanamivir are substantial (peramivir jumps from 26 to **8**, right after oseltamivir, and zanamivir falls from 5 to **26**). Still, there are compounds such as **1**, **3**, **10** and **13** whose ranks are relatively unchanged.



**Figure 2.6** Binding energies of 33 compounds are listed with their NSC ID. Minimum energy is the lowest binding free energy docking into all 13 receptor ensembles. HM and AM energy (kcal/mol) are the harmonic mean and arithmetic mean binding free energy, respectively.

There are two possible reasons for the observed differences in the ranking order for the top hits calculated here for H1N1pdm as compared to those in the study for H5N1. One is the differences in the hydrogen-bond networks in the H1N1pdm and H5N1 systems, as discussed in more detail below, and another may be due to the improvement in the free energy scoring function in AutoDock 4.2.1.<sup>34</sup> It is important to point out that the current approach may misidentify candidates in the NCI diversity set that should be in the top candidate list. Since the binding cavities of H1N1pdm and H5N1 are very similar, the number of such candidates should be rather small. However, to identify such candidates, a complete screening of the diversity set should be done.



**Table 2.2** Docking results for 33 compounds ranked by harmonic mean of binding energy. The first column is the final rank and also the compound ID. The predicted  $K_i$  calculated according to the harmonic mean binding free energies are also shown, as well as the arithmetic mean binding free energies and their corresponding standard deviations.

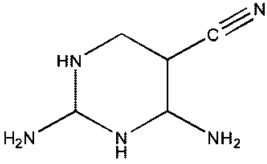
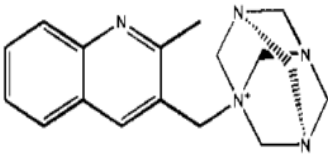
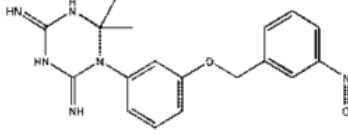
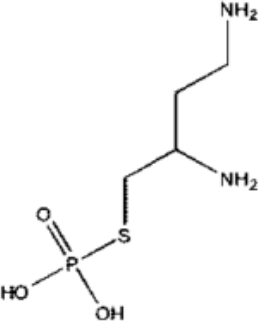
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted $K_i$ ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
1	2	211332	-12.05	0.001	-11.73	0.61	
2	16	141562	-10.14	0.037	-9.87	0.63	
3	1	109836	-9.74	0.072	-9.50	0.47	
4	8	350191	-9.56	0.099	-8.77	0.68	

Table 2.2 continued

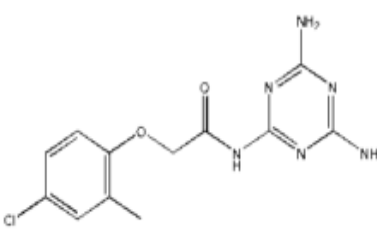
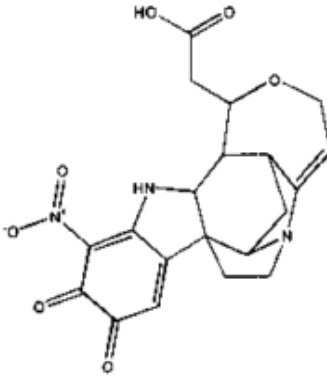
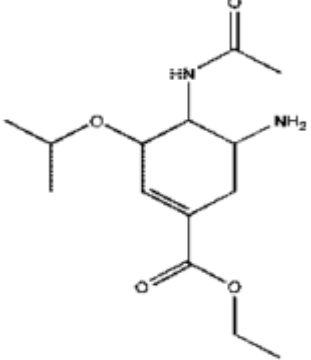
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
5	20	164640	-8.95	0.274	-8.78	0.42	
6	14	5069	-8.85	0.326	-8.23	1.11	
7	4	Oseltamivir	-8.69	0.429	-8.57	0.35	

Table 2.2 continued

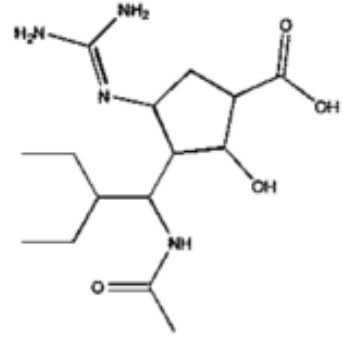
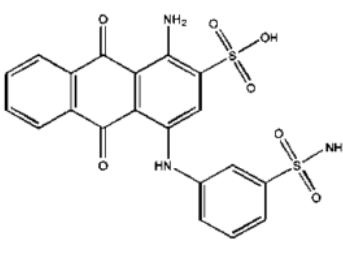
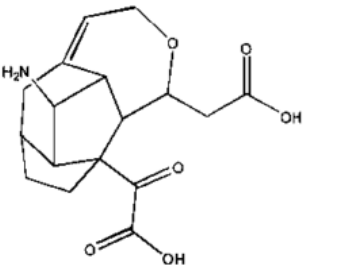
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
8	26	Peramivir	-8.47	0.615	-8.38	0.36	 The chemical structure of Peramivir is a complex molecule featuring a central five-membered ring. This ring is substituted with a guanidino group (NH-C(=NH)-NH2), a hydroxyl group (-OH), and a carboxylic acid group (-COOH). Additionally, it is linked to a side chain containing two ethyl groups and a methylamino group (-NH-C(=O)-CH3).
9	17	117079	-8.46	0.633	-7.70	0.98	 The chemical structure of compound 117079 consists of a central benzene ring fused to a naphthalene-like system. It features an amino group (-NH2) and a sulfonic acid group (-SO3H) on the benzene ring, and an amine group (-NH-) connected to another benzene ring that has a sulfonamide group (-SO2NH2) attached.
10	10	131612	-8.37	0.731	-8.05	0.65	 The chemical structure of compound 131612 is a complex polycyclic system, possibly a bicyclic cage structure. It features a primary amine group (-NH2) and a carboxylic acid group (-COOH) attached to the framework.

Table 2.2 continued

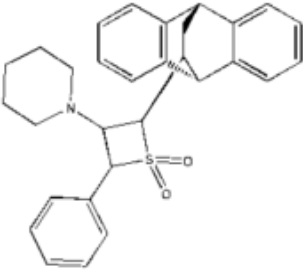
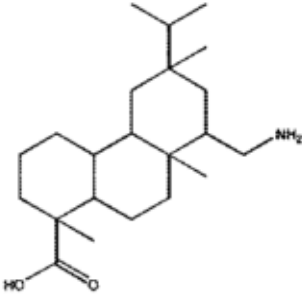
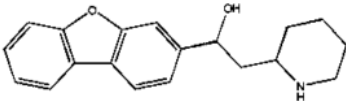
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
11	21	135371	-8.36	0.745	-8.08	0.55	
12	27	70194	-8.26	0.882	-7.89	0.88	
13	13	90630	-8.17	1.023	-7.89	0.52	

Table 2.2 continued

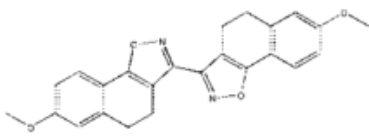
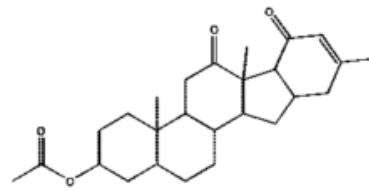
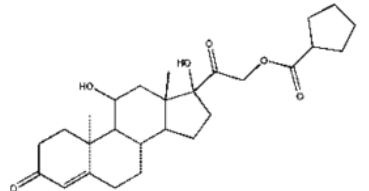
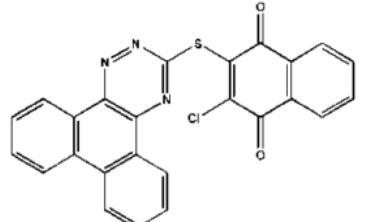
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
14	30	371688	-8.15	1.067	-6.75	1.15	
15	25	59620	-8.07	1.220	-7.50	0.82	
16	7	17245	-8.07	1.223	-7.85	0.54	
17	15	327705	-8.04	1.273	-7.69	0.57	

Table 2.2 continued

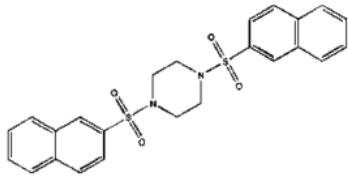
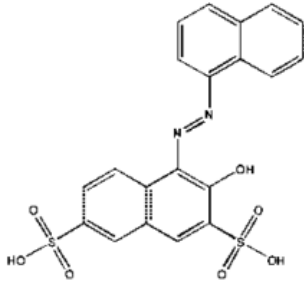
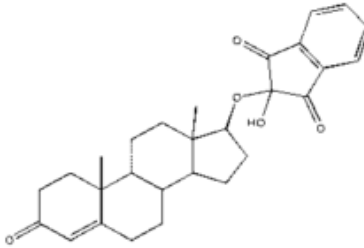
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
18	11	37245	-7.96	1.454	-7.92	0.24	
19	3	45583	-7.69	2.292	-7.21	0.75	
20	19	72254	-7.52	3.092	-7.49	0.19	

Table 2.2 continued

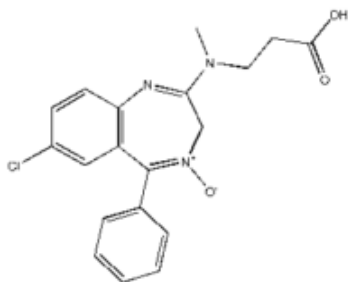
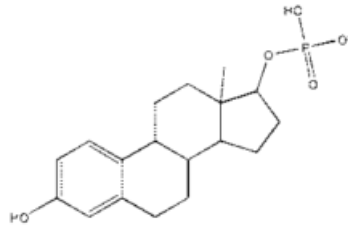
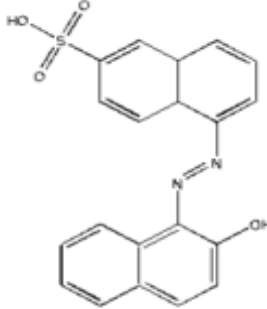
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
21	23	46080	-7.51	3.110	-7.04	0.69	
22	28	18312	-7.46	3.376	-6.91	0.79	
23	22	45576	-7.40	3.770	-6.85	0.74	

Table 2.2 continued

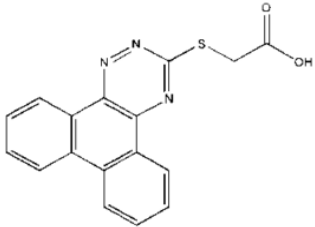
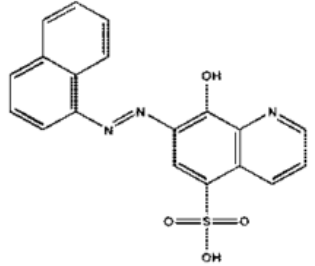
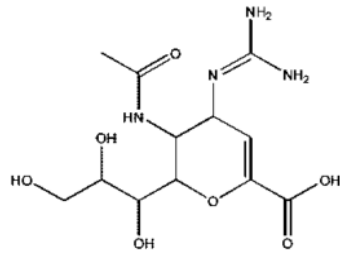
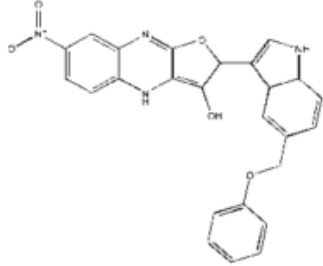
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
24	18	327704	-7.27	4.700	-7.05	0.53	
25	9	148354	-7.19	5.381	-7.07	0.39	
26	5	Zanamivir	-7.12	6.056	-6.72	0.78	
27	24	372294	-7.08	6.407	-7.00	0.34	



Table 2.2 continued

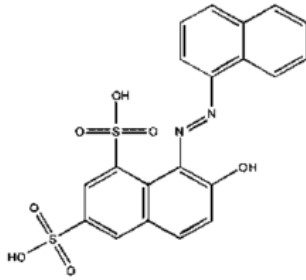
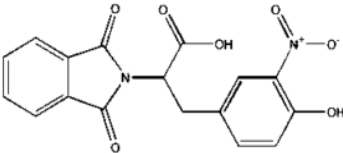
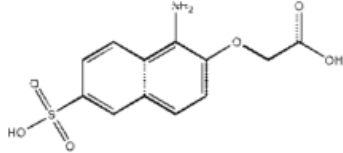
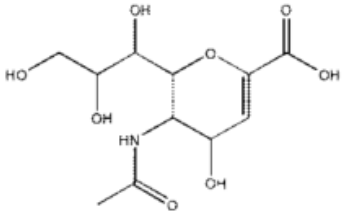
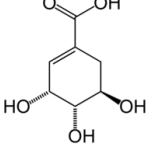
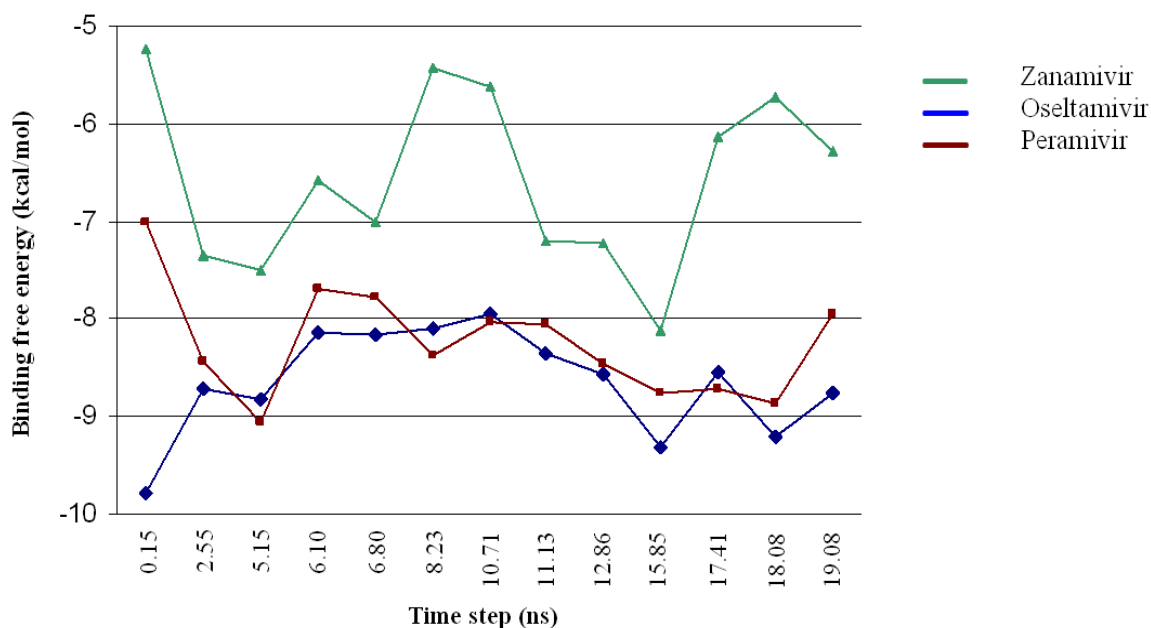
Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
28	12	45582	-6.88	8.980	-6.71	0.43	
29	6	106920	-6.57	15.338	-6.16	0.66	
30	32	16163	-6.12	32.802	-5.84	0.55	
31	31	DANA	-5.10	181.88 0	-4.75	0.65	
32	N/A	SK	-4.99	221.56 0	-4.81	0.35	

Table 2.2 continued

Rank / Comp. ID	Rank in H5N1	NSC	HM mean energy (kcal/mol)	Predicted Ki ( $\mu\text{M}$ )	AM mean energy (kcal/mol)	Standard deviation (kcal/mol)	Structure
33	29	SA	-4.82	292.03 0	-4.50	0.66	

Figure 2.7 shows the binding energy spectra of the three known bird flu drugs, zanamivir, oseltamivir and peramivir along the MD simulation time of H1N1pdm. Specifically it is the plot of the binding energies of these ligands with each of the H1N1pdm cluster ensembles. The results show zanamivir as consistently having a smaller binding energy compared to the other two drugs. A particularly interesting result here is the prediction of zanamivir to be much less effective against H1N1pdm than peramivir, whereas the previous study predicted that it is more effective for treating H5N1. The present calculated result confirms the observed effectiveness of oseltamivir for treating both H1N1pdm and H5N1. This confirmation not only provides another validation of the accuracy of the ensemble-based docking approach used in this study but also demonstrates the capability of the virtual screening technique in general.



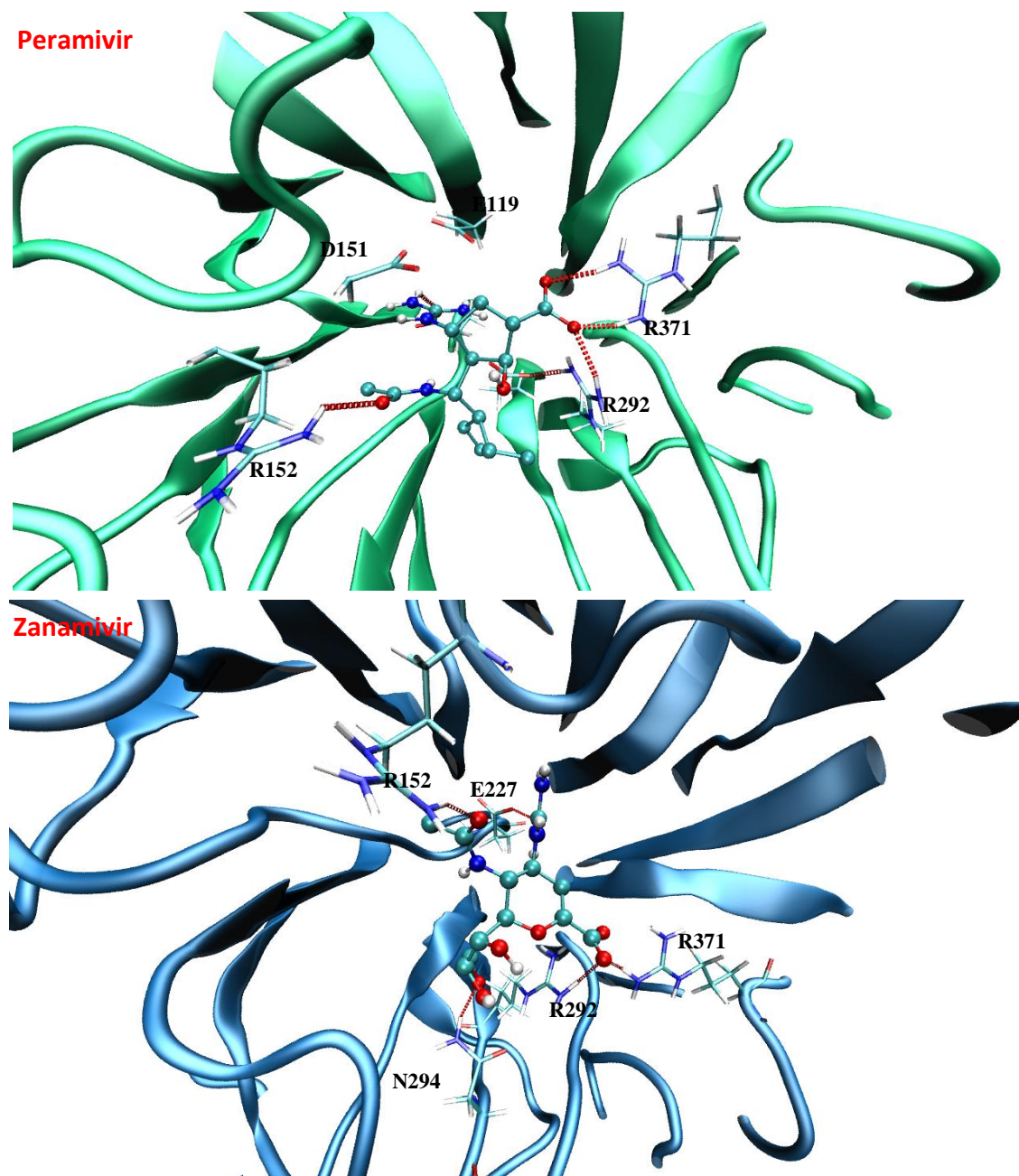
**Figure 2.7** Binding energy spectra vs. simulation time of zanamivir, oseltamivir and peramivir.

### 2.4.3. Hydrogen bond analysis of top binding compounds

Consistent with results from the recent MD simulation, docking results also show that oseltamivir forms strong hydrogen bonds with E119, D151, R292 and R371. In particular, the amino ( $\text{NH}_3^+$ ) group of oseltamivir forms strong hydrogen-bonds with the carboxyl ( $\text{COO}^-$ ) groups of E119 and D151, whereas its carboxyl group forms hydrogen bonds with the guanidinium ( $\text{NHC}(=\text{NH}_2^+)\text{NH}_2$ ) groups of R292 and R371. These hydrogen bonds are highly conserved in all 13 receptor ensembles. The Guanidinium group of R152 and the carboxyl groups of R227 and R277 sometimes form hydrogen bonds with the amide group of oseltamivir (Table 2.3, Figure 2.8).

**Table 2.3** The frequency of hydrogen bond (in percent) for binding site residues with known drugs and top binding compounds utilizing the hydrogen bond distance and angle cutoffs of 3.5Å and 45°, respectively.

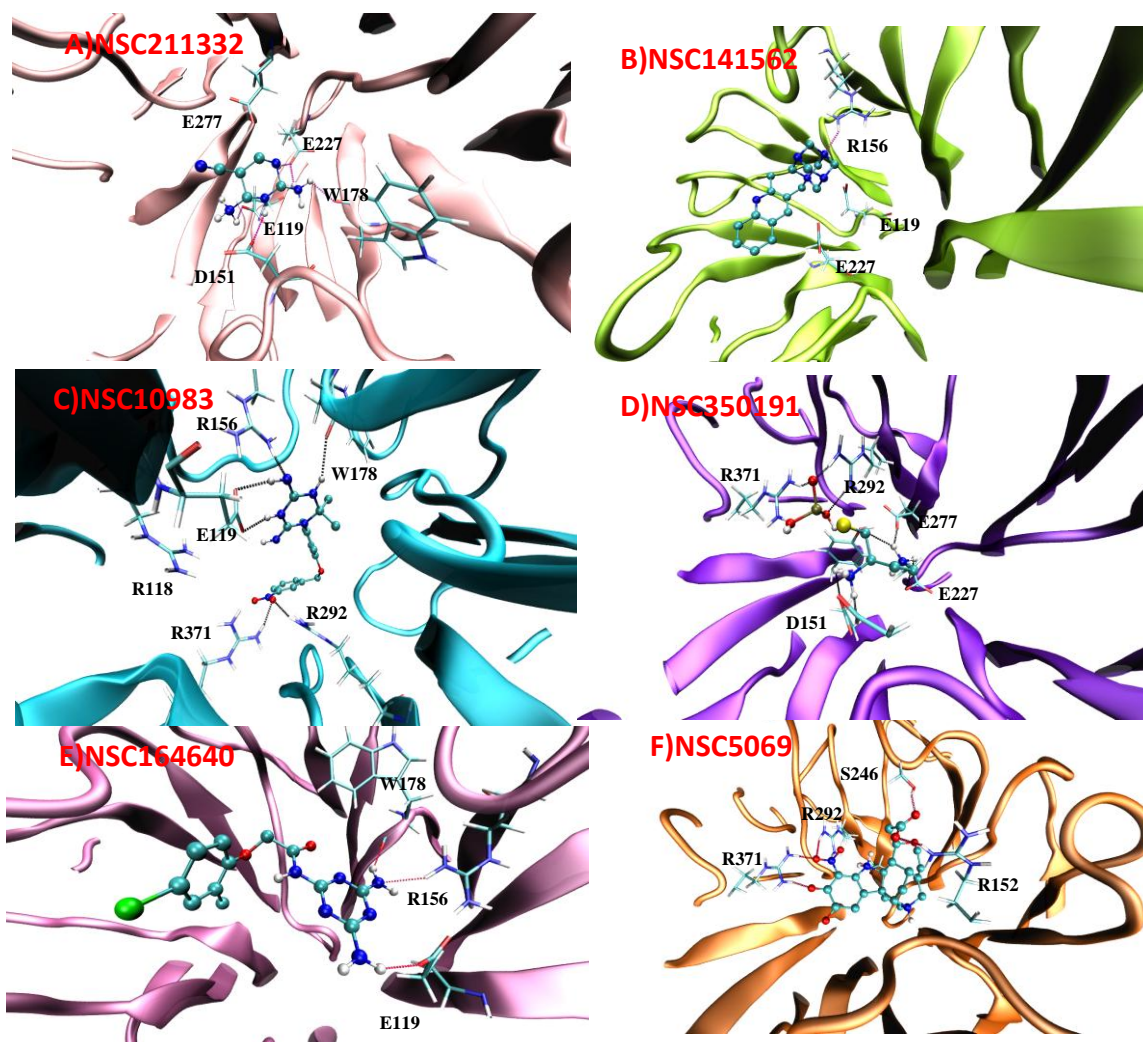
Residues	1	2	3	4	5	6	Oseltamivir	Peramivir	Zanamivir
E119	100	0	97	100	22	14	100	100	41
D151	97	0	14	100	0	0	98	98	47
R152	0	0	0	0	6	60	42	19	63
R156	0	100	94	0	71	8	0	0	0
W178	49	0	33	6	40	0	0	22	10
E227	98	0	0	94	16	0	13	100	100
E277	51	0	0	96	3	46	27	100	74
R292	0	0	97	100	3	72	100	100	100
R371	0	0	97	100	0	76	100	100	100
Y406	31	0	0	70	3	4	0	23	0



**Figure 2.8** Peramivir (A) and zanamivir (B) docked in the H1N1pdm neuraminidase structure.

Peramivir and zanamivir, like oseltamivir, also form highly conserved hydrogen bonds with the guanidinium groups of R292 and R371 through their carboxyl groups (Table 2.3, Figure 2.9). Peramivir's guanidinium group also has strong contacts via hydrogen bonding and electrostatic interaction with carboxyls of E119, D151, E227 and E277; however, zanamivir does not. Its guanidinium still makes hydrogen bonds with these residues, but less frequently as compared to oseltamivir and peramivir. This partially explains its smaller binding free energy (less negative) compared to the swine flu H1N1pdm neuraminidase. Another reason may be due to the receptors used in the docking procedure having been taken from the oseltamivir-bound holo structures, which possibly do not have good induced-fit effects towards zanamivir.

Compound **1**, which has a pyrimidine carbonitrile framework, binds tightly in the SA cavity in all 13 receptor ensembles regardless of their flexibility. This compound not only forms hydrogen bonds with E119 and E227 reported in a recent study on avian H5N1 neuraminidase,<sup>7</sup> but also forms strong ones with D151, and weak ones with W178, E277 and Y406 (Table 2.3, Figure 2.9A). This can explain its extremely low binding energy and thus its high ranking. Compound **3** (Figure 2.9C), on the other hand, binds less effectively to H1N1pdm than H5N1 neuraminidase due to the loss of interactions between its nitro (NO<sub>2</sub>) and guanidinium of R118. The guanidinium groups of the three residues (R118, R292 and R371) in the swine flu H1N1pdm neuraminidase are not in close proximity like those in H5N1. Therefore, several compounds including **2**, oseltamivir, peramivir and zanamivir lose part of their binding affinity since in H5N1 they usually form stronger hydrogen bonds with this trio. The special case is **2**, which has a tetraazatricyclo ring, and therefore it can form hydrogen bonds only with the



**Figure 2.9** Docked structure of top-hits against swine flu neuraminidase (shown in CPK). A: NSC211332 (1); B: NSC141562 (2); C: NSC109836 (3); D: NSC350191 (4); E: NSC164640 (5); F: NSC5069 (6). Some residues predicted to be important are also shown in sticks.

guanidinium group of R156. Its binding energy is mainly due to electrostatic interactions with surrounding negatively charged carboxyl groups.

## 2.5. Conclusion

Since the H1N1pdm neuraminidase primary structure is 91% similar to that of H5N1 and the active sites of both proteins are similar, this study used the top hit compounds for H5N1 as the starting point of the virtual screening process to find promising antiviral drugs for H1N1pdm using the ensemble-based docking technique. All 33 compounds, including the top binding compounds to avian H5N1 neuraminidase from the NCI diversity set and known drugs, were docked to ensembles of 13 H1N1pdm configurations that represent more than 96% of the configuration space covered by a 20 ns MD simulation. The results obtained reveal that 6 compounds, specifically NSC211332, NSC141562, NSC109836, NSC350191, NSC1644640, and NSC5069, have higher binding affinity to the H1N1pdm neuraminidase protein than oseltamivir does. The results also confirm the observed effectiveness of oseltamivir against both the swine H1N1pdm and H5N1 flu. Detailed analysis on the hydrogen bond networks between top binding candidates with the swine H1N1pdm neuraminidase protein reveals that the mutations H274Y and N294S do not have any direct interactions with these compounds, and thus suggest the possibility of using the present top hit compounds for further computational and experimental studies to design new antiviral drugs against swine H1N1pdm flu virus and its variants. Furthermore, a more complete virtual screening using the full NCI diversity set or larger sets from the ZINC database should be done to identify drug candidates that may have been missed in this study.



## CHAPTER 3

### NEW APPROACHES FOR SOLVATION FREE

### ENERGY CALCULATION

In Chapter 3, recent progress in development of the two new methodologies for solvation free energy called coupled RISM-HNC/MD and MM-PB/LRA-SA developed by Freedman et al., are presented. My contributions to the work here were developing and implementing a parallel method (using MPI – Message Passing Interface) to calculate radial distribution functions used as input for the two methods and running some MD simulations for assessment.

#### **3.1 Improving the performance of the coupled reference interaction site model – hyper-netted chain (RISM-HNC) / MD method for free energy of solvation**

##### 3.1.1 Introduction of RISM-HNC/MD

RISM-HNC/MD<sup>35, 36</sup> is an approach to calculate solvation free energy by solving the RISM equation as a function of radial distribution functions (RDF) taken from MD simulations in combination with the equation HNC as a closure relation. The idea of combining the RISM and HNC equations to solve for RDFs and then using these and the RISM equation to determine solvation free energy has been done.<sup>37-40</sup> However, by using

more accurate RDFs from MD simulations, the authors argue that RISM-HNC/MD will produce more accurate results as well as solve the convergence problems appearing if solving RISM for RDFs.

3.1.2 Using the hard-sphere (HS) free energy as the reference and a linear response approximation (LRA) to improve the performance of the RISM-HNC/MD method

In the prior studies by Freedman *et al.*, the coupled RISM-HNC/MD method has produced reasonable relative solvation free energies for small systems. Specifically, the method has been employed successfully for studies of the potential of mean force curve of  $\text{SN}_2 \text{Cl}^- + \text{CH}_3\text{Cl}$  reaction in water,<sup>41</sup> the conformational analysis of the alanine dipeptide<sup>42</sup> and tautomeric equilibria of 2-Hydroxypyridine/2-Oxopyridine and of cytosine in water<sup>43</sup> However, errors have appeared when it has been applied to larger-sized systems. A hard-sphere reference free energy and a linear response have been introduced to attempt to eliminate the contribution to the solvation free energy from the unknown bridge function, which otherwise introduces a sizeable error in proportion to the solute's size.<sup>15</sup> This approach aims to reduce the size of the matrices in the RISM equation to a more manageable scale.

Using the hard-sphere reference combined with the coupled RISM-HNC/MD, the final solvation free energy equation is:<sup>15</sup>

$$\begin{aligned}
\Delta\Delta\mu^{(HNC/HS)} &= \Delta\mu^{(HNC)} - \Delta\mu^{HS} \\
&= \frac{\rho}{\beta} \sum_{\alpha\gamma} \int 4\pi r^2 \frac{u_{\alpha\gamma}^C(r) + u_{\alpha\gamma}^{(2)}(r)}{2} (1 + h_{\alpha\gamma}(r)) + \frac{u_{\alpha\gamma}^{(2)}(r)}{2} (1 + h_{\alpha\gamma}^{HS}(r)) dr + \\
&(2\pi)^{-3} \frac{\rho}{\beta} \sum_{\alpha\gamma} \int_{k=k_0}^{\infty} 4\pi k^2 \left[ \hat{h}_{\alpha\gamma}(k) + \hat{h}_{\alpha\gamma}^{HS}(k) \right] \left[ \frac{c_{\alpha\gamma}^{RISMsr}(k) - c_{\alpha\gamma}^{RISM(HS)}(k)}{2} + \frac{h_{\alpha\gamma}(k) - h_{\alpha\gamma}^{HS}(k)}{2} \right] \\
&+ 4\pi k^2 \left[ \frac{u_{\alpha\gamma}^{(2)}(k)}{2} (\hat{h}_{\alpha\gamma}(k) - \hat{h}_{\alpha\gamma}^{HS}(k)) + \hat{h}_{\alpha\gamma}^{HS}(k) \hat{c}'_{\alpha\gamma}(k) - \hat{h}_{\alpha\gamma}(k) \hat{c}'_{\alpha\gamma}(k) \right] dk
\end{aligned} \tag{3.1}$$

Besides HNC/HS, two other solvation free energy expressions were tested here including the Gaussian Fluctuation (GF) expression,<sup>44</sup> the partial wave (PW) expression<sup>45, 46</sup> (by replacing the term  $\frac{h_{\alpha\gamma}(k) - h_{\alpha\gamma}^{HS}(k)}{2}$  in equation 3.1), and finally the LRA formulation.

### 3.1.3 Methodology assessment

The approach using the hard-sphere reference combined with the four above-mentioned expressions (HNC, GF, PW and LRA) was employed for solvation free energy calculation of several organic molecules. Freedman et al.<sup>35</sup> suggested that reasonable predictions were obtained when they used a low-k cutoff of 1.5 for  $c^{RISM}$ . In this study, results derived from three values of k (1.3, 1.4 and 1.5) were reported in order to obtain information about the numerical error associated with different cutoffs. It was seen that k=1.4 performs best in this study. After that, the results were compared with those from TI or FEP calculations using the same force fields (either OPLS or GAFF) as well as experimental data.

Results presented as follows were obtained by adding each of the variants HNC/HS, GF/HS, and PW/HS as well as LRA/HS to free energy of cavity formation terms associated with using a hard sphere potential.

LRA/HS predicts well the solvation free energy of some small cyclic molecules, even better than FEP calculations in the case of phenol and cyclohexane, as shown in Table 3.1.

**Table 3.1.** Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/MD compared to experimental values, FEP values, and stand-alone RISM-HNC results of Sato et al.

Solute	$k_0$	$\Delta\mu_S^{\text{HNC/H}}$	$\Delta\mu^{\text{GF/HS}}$	$\Delta\mu^{\text{PW/HS}}$	$\Delta\mu^{\text{LRA/HS}}$	$\Delta\mu^{\text{PW(RISM/HNC)*}}$	$\Delta\mu^{\text{FEP}^\dagger}$	Expt.
Cyclohexane	1.3	3.60	1.25	1.10	1.24	4.29		1.2
	1.4	3.64	1.30	2.30				
	1.5	3.50	1.36	3.08				
Benzene	1.3	0.89	-1.73	-1.00	-3.47	3.08	-1.3	-0.87
	1.4	0.41	-2.11	-0.60				
	1.5	-0.17	-2.54	-0.59				
Toluene	1.3	2.28	-1.45	-1.08	-2.28	3.10	-2.2	-0.89
	1.4	1.71	-1.71	-0.43				
	1.5	1.16	-1.97	0.20				
Pyridine	1.3	-2.74	-5.94	-6.20	-6.75	0.30	-3.4	-4.7
	1.4	-3.02	-6.15	-5.46				
	1.5	-3.27	-6.30	-3.89				
Phenol	1.3	-3.00	-7.03	-5.19	-7.97	-3.80	-3.7	-6.62
	1.4	-3.30	-7.26	-4.51				
	1.5	-3.41	-7.26	-3.51				

\*The results of Sato *et al.*<sup>47</sup>

†The results of Mobley *et al.*<sup>48</sup>

Reasonable results were also obtained using PW/HS and GF/HS while those from the HNC/HS method are least accurate (markedly too positive values for all molecules).

This approach was also validated on methylamine and N-methylacetamide (NMA) and the results are shown in Table 3.2. GF/HS yields values close to those of FEP, TI as well as experimental data, while HNC/HS and PW/HS produce least accurate results. Particularly, the energy difference between HNC/HS and experiment is more than 4 kcal/mol for the case of NMA for all values of cutoffs. LRA/HS also produces reasonable results, however, not as good as GF/HS in terms of the solvation free energy difference between the two compounds.

Finally, the approach was tested on some small organic molecules using two different force fields GAFF and OPLS, and the results are shown in Tables 3.3 and 3.4, respectively. For the OPLS force-field and HF/6-31G(d,p) charges, LRA/HS performs much better than HNC/HS, PW/HS, GF/HS and even slightly better than TI in comparison with experimental data. For the GAFF force field and the HF/3-21G charges, LRA/HS produces reasonable results except for Acetic acid. In general, LRA/HS performs better with OPLS than with the GAFF force field in this study.

**Table 3.2** Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/ MD compared to experimental values, and those of Wan et al.

Solute	$k_0$	$\Delta\mu^{\text{HNC/HS}}$	$\Delta\mu^{\text{GF/HS}}$	$\Delta\mu^{\text{PW/HS}}$	$\Delta\mu^{\text{LRA/HS}}$	$\Delta\mu^{\text{FEP}^*}$	$\Delta\mu^{\text{TI}^*}$	Expt.
Methylamine	1.3	-3.51	-5.10	-3.36	-5.71	-4.7	-5.4	-4.6
	1.4	-3.45	-4.93	-3.26				
	1.5	-3.01	-4.55	-2.90				
NMA	1.3	-5.64	-10.24	-7.33	-9.28	-10.4	-11.3	-10.1
	1.4	-5.69	-10.08	-7.23				
	1.5	-5.37	-9.63	-6.75				

\*The results of Wan *et al*<sup>49</sup>.

**Table 3.3** Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/ MD with the GAFF force field and HF/3-21G charges, compared to experimental values.

Solute	$k_0$	$\Delta\mu^{\text{HNC/HS}}$	$\Delta\mu^{\text{GF/HS}}$	$\Delta\mu^{\text{PW/HS}}$	$\Delta\mu^{\text{LRA/HS}}$	Expt.
Acetic acid	1.3	-10.29	-13.07	-9.92	-11.36	-6.7
	1.4	-10.18	-13.05	-9.80		
	1.5	-9.44	-12.48	-9.18		
Ammonia	1.3	-2.39	-3.77	-2.27	-6.59	-4.3
	1.4	-2.51	-3.58	-2.28		
	1.5	-2.34	-3.10	-1.95		
Ethane	1.3	0.61	-0.37	0.67	0.69	1.8
	1.4	0.69	-0.41	0.73		
	1.5	-0.39	-0.37	0.86		
Methane	1.3	1.03	-0.45	0.37	0.47	2.0
	1.4	0.91	-0.42	0.31		
	1.5	1.04	-0.36	0.39		
Methanethiol	1.3	0.53	-0.32	0.57	-1.48	-1.2
	1.4	0.61	-0.27	0.66		
	1.5	0.92	-0.13	0.89		
Methanol	1.3	-5.05	-7.34	-5.09	-6.59	-5.1
	1.4	-5.24	-7.42	-5.25		
	1.5	-5.02	-7.24	-5.10		

### 3.1.4 Conclusion and outlook

The present study shows that applying LRA/HS for solvation free energy produces a reasonably accurate prediction in comparison to experiment or other MD methods as well as other free energy expressions. Considering the accuracy and computational cost, the method is a good tool for solvation free energy calculations. However, all molecules used in this study are smaller than a typical drug-like compound. Therefore, more testing is needed to see if convergence problems appear as the molecular sizes increase before the method can be applied widely for quantitative analysis in rational drug design.

**Table 3.4** Free energies of solvation (kcal/mol) calculated by coupled RISM-HNC/MC with the OPLS force field and HF/6-31G(d,p) charges, compared to experimental values and results of TP.

Solute	$k_0$	$\Delta\mu^{\text{HNC/HS}}$	$\Delta\mu^{\text{GF/HS}}$	$\Delta\mu^{\text{PW/HS}}$	$\Delta\mu^{\text{LRA/HS}}$	$\Delta\mu^{\text{TP*}}$	Expt.
Acetic acid	1.3	-3.09	-7.29	-3.02	-6.03		-6.7
	1.4	-3.34	-7.58	-3.22			
	1.5	-3.40	-7.78	-3.40			
Ammonia	1.3	-1.09	-2.63	-0.70	-4.11		-4.3
	1.4	-1.78	-2.93	-1.24			
	1.5	-1.74	-2.65	-1.10			
Ethane	1.3	4.02	1.09	3.56	0.86		1.8
	1.4	3.54	0.59	3.12			
	1.5	3.64	0.50	3.11			
Methane	1.3	4.69	1.54	3.68	1.08	2.24	2.0
	1.4	3.12	0.24	2.26			
	1.5	2.67	-0.10	1.89			
Methanethiol	1.3	1.71	-0.49	1.52	-1.84	-0.55	-1.2
	1.4	1.70	-0.54	1.58			
	1.5	1.74	-0.58	1.63			
Methanol	1.3	-1.32	-4.14	-1.29	-4.84	-4.51	-5.1
	1.4	-2.71	-5.36	-2.90			
	1.5	-2.59	-5.24	-2.75			

\*The results of Shirts and Pande<sup>50</sup>

### 3.2 Explicitly-solvated ligand contribution to continuum

#### solvation models for binding free energies: Selectivity of

#### theophylline binding to an RNA aptamer

##### 3.2.1 Introduction of MM-PB/LRA-SA

The MM-PBSA method evaluates solvation free energy, by using a dielectric continuum model, which represents the solvent implicitly as a continuum environment<sup>51-</sup>  
<sup>53</sup>. The method has been widely used because it provides adequate accuracy at low computational cost. However, in certain biological systems in which solvent plays an important role in ligand binding, neglecting explicit solvent interactions might leads to unreliable results. To provide more accurate computational estimates of binding free

energies in solution from MD simulations, Freedman et al. proposed to use a separate solvation contribution for the binding ligand, determined from a linear response treatment. The new approach is called Molecular Mechanics, Poisson Boltzmann, and Surface Area contributions (MM-PB/LRA-SA).

In the approach taken here, the solvation part is decomposed into ligand and receptor contributions as follow

$$\begin{aligned} \Delta G_{bind}^{sol} &= \left( G_{rec(complex)}^{sol} - G_{rec(free)}^{sol} \right) + \left( G_{lig(complex)}^{sol} - G_{lig(free)}^{sol} \right) \\ &= \left[ G_{rec(complex)}^{sol} - G_{rec(free)}^{sol} \right]^{PB(orGB)-SA} + \left[ G_{lig(complex)}^{sol} - G_{lig(free)}^{sol} \right]^{LRA-SA} \end{aligned} \quad (3.2)$$

The contribution to the solvation free energy difference upon ligand binding from the receptor is calculated by the PBSA (or the GBSA) approximations; ligand contributions to solvation free energy are calculated from the LRA expression and summed to the receptor contribution to obtain total contributions of solvation to binding free energies. LRA solvation free energy components were calculated as averages over values determined from radial distribution functions<sup>15</sup> as described in equation 3.3.

$$\Delta\Delta\mu^{(LRA/HS)} = \frac{\rho}{\beta} \sum_{\alpha\gamma} \int 4\pi r^2 \left( \frac{u_{\alpha\gamma}^c(r)}{2} (g_{\alpha\gamma}(r)) + \frac{u_{\alpha\gamma}^{(2)}(r)}{2} (g_{\alpha\gamma}(r) + g_{\alpha\gamma}^{HS}(r)) \right) dr \quad (3.3)$$

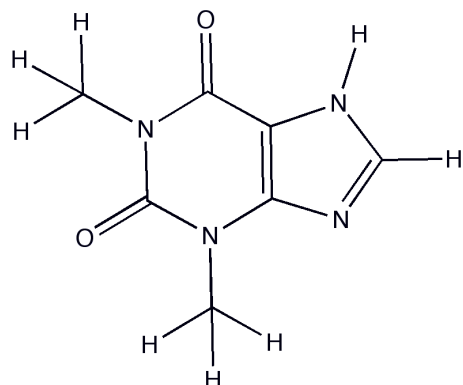


### 3.2.2 Using MM-PB/LRA-SA to study selectivity of Theophylline

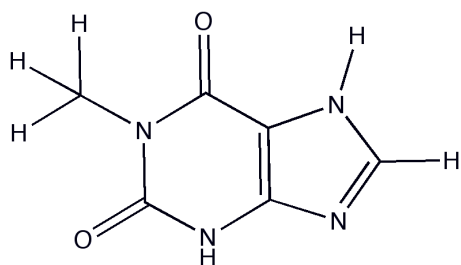
binding to an RNA aptamer

Characteristics of Theophylline to an RNA aptamer. Theophylline and its derivatives (shown in Figure 3.1) bound to an RNA aptamer are interesting cases for assessment of the MM-PB/LRA-SA method. First, the systems have important applications in biomedicine. Second, solvation usually plays an important role in the thermodynamics of ligand binding to RNA aptamers due to its highly anionic nature and relatively solvent-accessible binding site. Third, a theoretical study of these systems by Gouda *et al.*<sup>54</sup> using the MM-PBSA method has been done. However, this continuum model ranked several of the analogues in the wrong order (theophylline > 1-methylxanthine > 3-methylxanthine > xanthine > caffeine) compared to experimental binding free energies (theophylline > 3-methylxanthine > xanthine > 1-methylxanthine > caffeine)<sup>55</sup>. The differences in binding free energies of these analogs between MM-PBSA and experimental studies are also significant thus make the systems an interesting case for assessment of the MM-PB/LRA-SA method. In addition, the systems are also interesting cases for studying how interlocking structural motifs can be assembled into a highly specific ligand-binding site that has strong affinity and molecular discrimination in the RNA aptamer. The initial structure of the theophylline-bound RNA aptamer used in this study (Figure 3.2) was provided by Gouda *et al.*, which was originally derived from the crystal structure in the Protein Data Bank (PDB entry 1EHT).

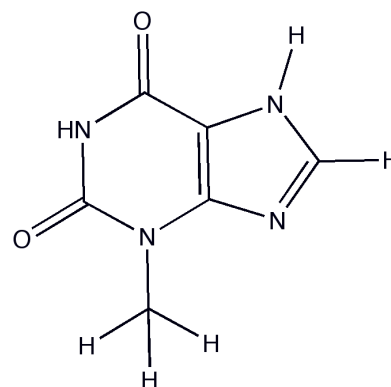
(a)



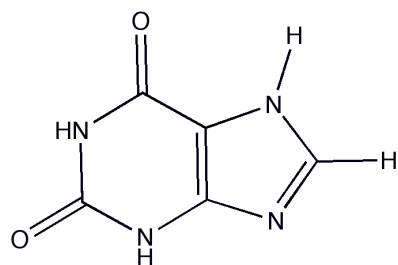
(b)



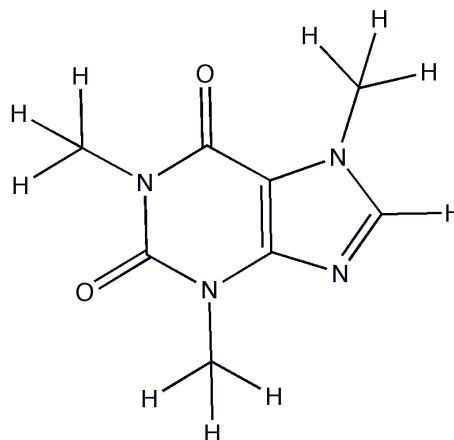
(c)



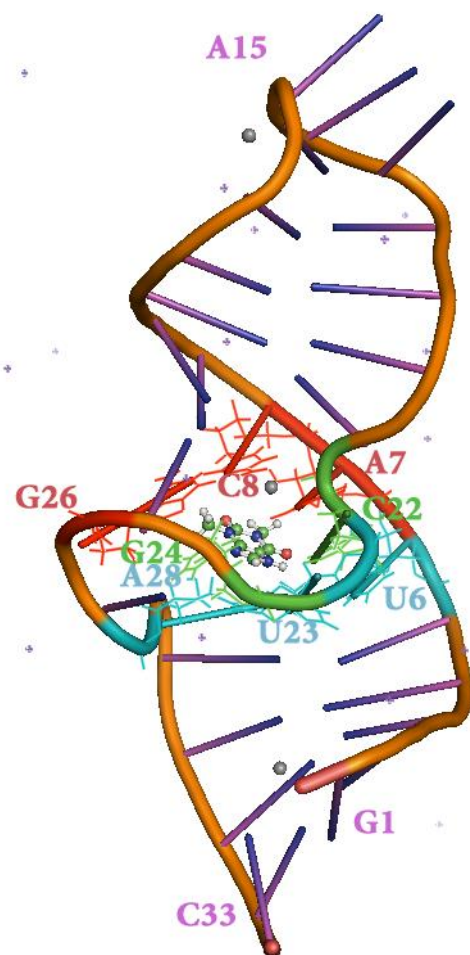
(d)



(e)

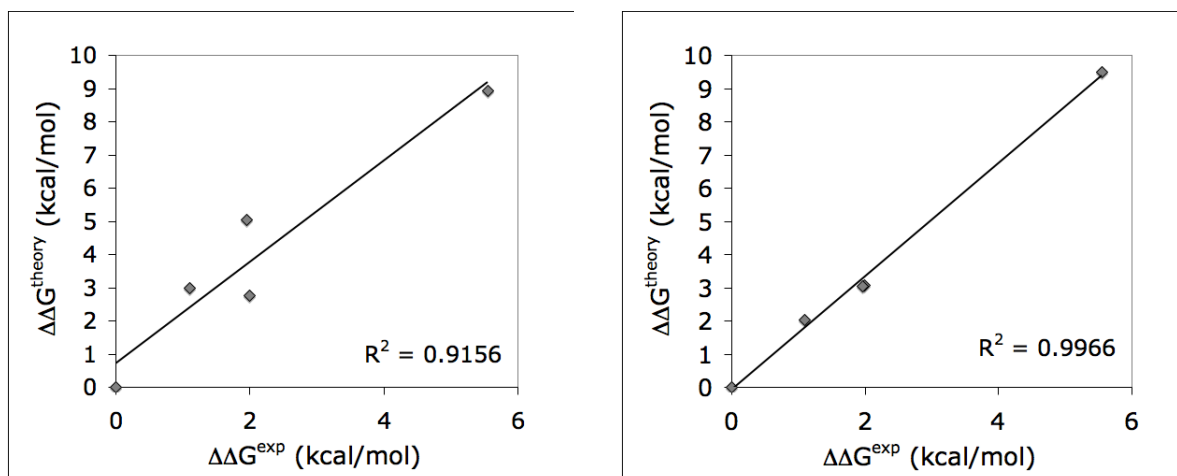


**Figure 3.1** Chemical structures of theophylline (A) and derivatives, 1-methylxanthine (B), 3-methylxanthine (C), xanthine (D) and caffeine (E).



**Figure 3.2** Conformation of theophylline-bound RNA aptamer after MD simulation, with triples labeled by residue and distinguished by separate colors, visualized using the Pymol viewing program.<sup>56</sup> Grey balls represent Mg<sup>2+</sup> ions.

Relative binding energies of the five analogs from MM-PBSA and MM-PB/LRA-SA vs. those from experiments are shown in Figure 3.3 A and B, respectively. The squared correlation coefficients of the linear fits in the two graphs show that MM-PB/LRA-SA (0.9966) outperforms MM-PBSA (0.9156). Total Binding Free Energies are shown in Table 3.5. MM-PB/LRA-SA calculates the preferential binding order of the five ligands in the order, theophylline > 3-methylxanthine > 1-methylxanthine > xanthine > caffeine. These results agree well with those of experiment and TI (except for the



**Figure 3.3** Relative binding free energies calculated by (A) MM-PBSA and (B) MM-PB/LRA-SA for the 5 theophylline analogues compared with experimental data, showing linear fits of data with  $R^2$  values.

relative order of xanthine and 1-methylxanthine) while MM-PBSA and MM-GBSA completely misranked the binding affinities of these compounds.

Table 3.5 shows relative binding energy of the four analogs compared to theophylline. Compared to experiment and TI, results derived from the LRA-SA method, which require an additional simulation, are more reasonable than MMPB-SA or MMGB-SA alone, both in the values of binding energy and the ranks of all compounds of interest. In all methods, there is a significant difference in binding affinity between caffeine and theophylline. The reason might be that this compound occupies a different binding site from the other three analogs.<sup>56</sup>

**Table 3.5** Calculated binding free energies of theophylline and analogues, compared between different methods and with experimental values. Values are given relative to the value for theophylline binding, which is indicated in parentheses.

	One simulation		Two simulations each of complex and ligand		Experiment and TI simulations	
	$\Delta\Delta G^{\text{MM-CB-SA}}$	$\Delta\Delta G^{\text{MM-PB-SA}}$	$\Delta\Delta G^{\text{MM-CB/LRA-SA}}$	$\Delta\Delta G^{\text{MM-PB/LRA-SA}}$	$\Delta\Delta G^{\text{Exp.‡}}$	$\Delta\Delta G^{\text{TI*}}$
Theophylline	0 (-30.6)	0 (-27.6)	0 (-42.1)	0 (-39.2)	0 (-8.92)	0
3-Methylxanthine	3.4	3.0	0.0	0.4	1.10±0.05	1.36±1.04
Xanthine	4.8	5.1	1.3	3.5	1.96±0.28	1.64±0.83
1-Methylxanthine	2.9	2.8	0.4	1.7	1.99±0.29	1.82±0.33
Caffeine	7.0	8.9	5.2	7.7	5.55±0.02	6.63±0.82

‡ ref [55], \* ref [54]

### 3.2.3 Conclusion and outlook

The MM-PB/LRA-SA method produces quite accurate relative binding free energies for 5 theophylline and its four derivatives to the RNA strand with reasonable computational cost. In comparison to MM-PBSA alone the MM-PB/LRA-SA method requires about twice the computational cost. The method requires running the same simulations with all nonbonded interactions of the ligand with the water molecules turned off. Whereas, conventional MD simulation methods (e.g. FEP) would require a computing approximately 10 times more than that of MM-PB/LRA-SA considering the cost of testing to find suitable restraints to keep ligand binding modes unchanged during simulations and both forward and backward simulations for validation.

Theoretically, the approach is straightforward and feasible to be implemented in the available MD packages (e.g AMBER or NAMD) and, in our opinion, thus is applicable for quantitative analysis in lead optimization. However, the level of accuracy achieved in this study may be due to the high similarity in the compounds of interest, leading to cancellation of error in calculations. To apply the method to sets of ligands with diverse structures, nonlinear correction needs to be taken into account and is still under development.

## CHAPTER 4

### MOLECULAR MODELING OF MUTATION-INDUCED OSELTAMIVIR DRUG RESISTANCE FOR INFLUENZA N1 NEURAMINIDASES

In Chapter 4, we discuss our observations, based on drug-protein endpoint interactions on how the mutations H274Y and N294S induce oseltamivir resistance in neuraminidase N1 subtypes. Part of Chapter 4 originally appeared in *PLoS Curr Influenza*, RRN1015 and in a manuscript submitted to *J. Mol. Med.* The work was done in collaboration with Eric Lee. David Hardy also contributed to the calculations of electrostatic surface potential.

#### **4.1 Literature review of oseltamivir-resistance and our proposed study**

Genetics study of influenza H5N1 virus isolated from patients who died despite being given oseltamivir showed that mutations H274Y or N294S confer high-level resistance to oseltamivir.<sup>11, 57, 58</sup> There is even emerging evidence that these drug-resistant mutants pose the same risk with H1N1pdm, as shown by reported cases of the H274Y mutation of H1N1pdm.<sup>59-62</sup> The rapid emergence of oseltamivir resistance in avian flu has already motivated numerous studies, both experimental and theoretical, to

uncover the mechanisms of how point mutations in neuraminidase alter drug binding. Despite initial inroads, the current understanding of drug resistance remains incomplete and some conclusions are conflicting. For example, in one study, it was reported that the H274Y mutation disrupts the E276-R224 salt bridges, but in a separate study the same salt bridges were observed to be stable.

In this study we characterize the drug-protein interactions of oseltamivir bound forms of wild type avian H5N1 and swine H1N1pdm neuraminidases and how their mutations, H274Y and N294S, rupture these interactions to confer drug resistance through molecular modeling and all-atom molecular dynamics simulations (MD).

## 4.2 Computational details

The coordinates for the H5N1 neuraminidase bound with oseltamivir was taken from a monomer of the Protein Data Bank (PDB) structure 2HU4 (tetramer), while those of mutants H274Y and N294S were taken from structures 3CL0 (monomer) and 3CL2 (monomer) respectively.<sup>63</sup> Even though the biological form of neuraminidase is tetrameric, its monomer contains a functionally complete active site and yields reasonable results in a prior study using MD simulations.<sup>64</sup> The position for oseltamivir bound to H1N1pdm was adopted from its corresponding location in H5N1, as the two proteins' binding pockets differ only by residue 347 (which is Y in H5N1 and N in H1N1pdm), located on a loop at the periphery of the active site. Oseltamivir-mutant complexes of H1N1pdm were built by mutating H274Y and N294S of the H1N1pdm wild type model.

In total, 6 systems were modeled and simulated for oseltamivir bound H5N1, and H1N1pdm wild type and H274Y and N294S mutants. Crystallographically resolved



water molecules and a structurally relevant calcium ion near the native binding site for SA were retained and modeled in all simulated systems. The protein complexes were then solvated in a TIP3P<sup>23</sup> water box and ionized by NaCl (0.152M) to mimic physiological conditions.

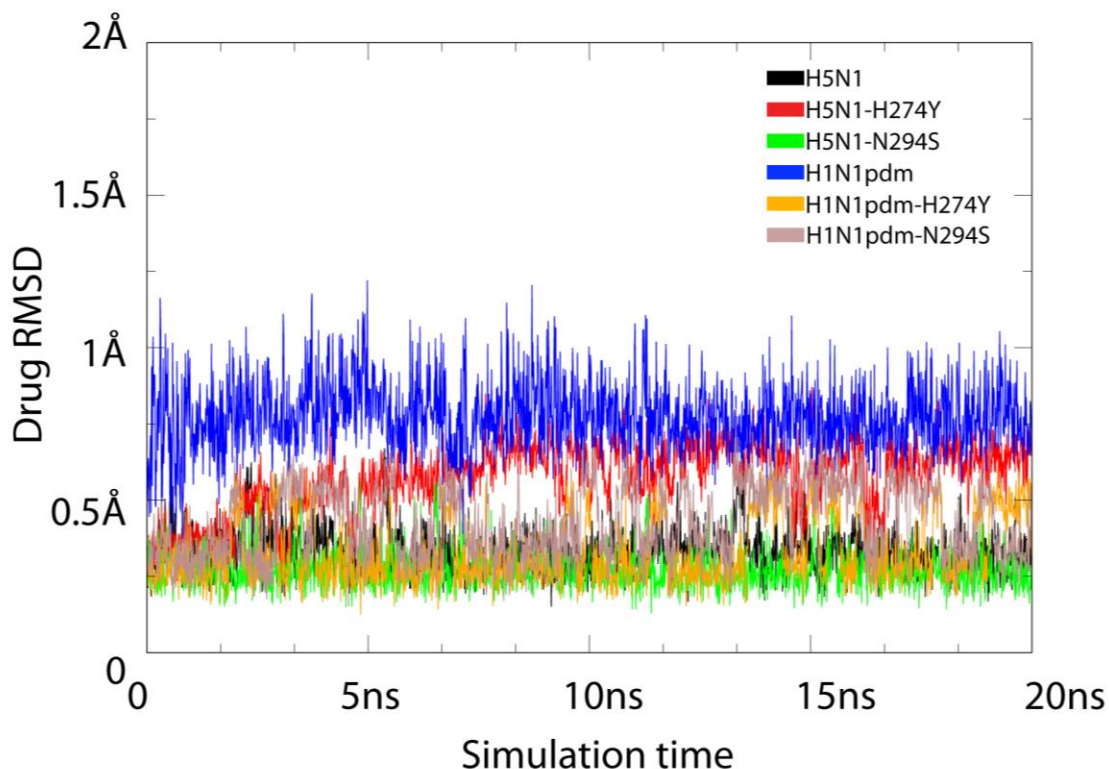
All simulations were performed using NAMD 2.7<sup>65</sup> and the CHARMM31 force field with CMAP correction.<sup>66, 67</sup> The ionized systems were minimized for 10,000 integration steps and equilibrated for 20 ns with 1 fs time steps. Following this, a 20 ns unconstrained equilibration production run was performed for subsequent trajectory analysis, with frames stored after each picosecond (every 1000 time steps). Constant temperature ( $T = 300$  K) was enforced using Langevin dynamics with a damping coefficient of  $1 \text{ ps}^{-1}$ . Constant pressure ( $p = 1 \text{ atm}$ ) was enforced using the Nosé-Hoover Langevin piston method. Van der Waals interaction cutoff distances were set at  $12 \text{ \AA}$  (smooth switching function beginning at  $10 \text{ \AA}$ ) and long-range electrostatic forces were computed using the particle-mesh Ewald (PME) summation method.

Analysis included the calculation of an averaged electrostatic potential field over all frames of the trajectory using RMSD-aligned structures. Maps of the electrostatic potential field were calculated on a three-dimensional lattice. The long-range contributions to the electrostatics were approximated using the multilevel summation method (MSM), which uses nested interpolation of the smoothed pairwise interaction potential, with computational work that scales linearly with the size of the system. The calculation was performed using the molecular visualization program VMD that provides a GPU-accelerated version of MSM to produce the electrostatic potential map.<sup>68</sup> The GPU acceleration of MSM provided a significant speedup over conventional electrostatic

summation methods such as the Adaptive Poisson Boltzman Solver (APBS), achieving a benchmark processing time of 0.2s per frame versus 180 seconds per frame (on a conventional CPU) using APBS<sup>69</sup> for a 35,000 atom system, offering a speedup factor of about 900. The use of GPU acceleration enabled averaging the electrostatic potential field over all frames of the simulation trajectories. Hydrogen bond analysis utilized a distance and angle cutoff of 3.5 Å and 60 degrees, respectively.

### 4.3 Results and discussion

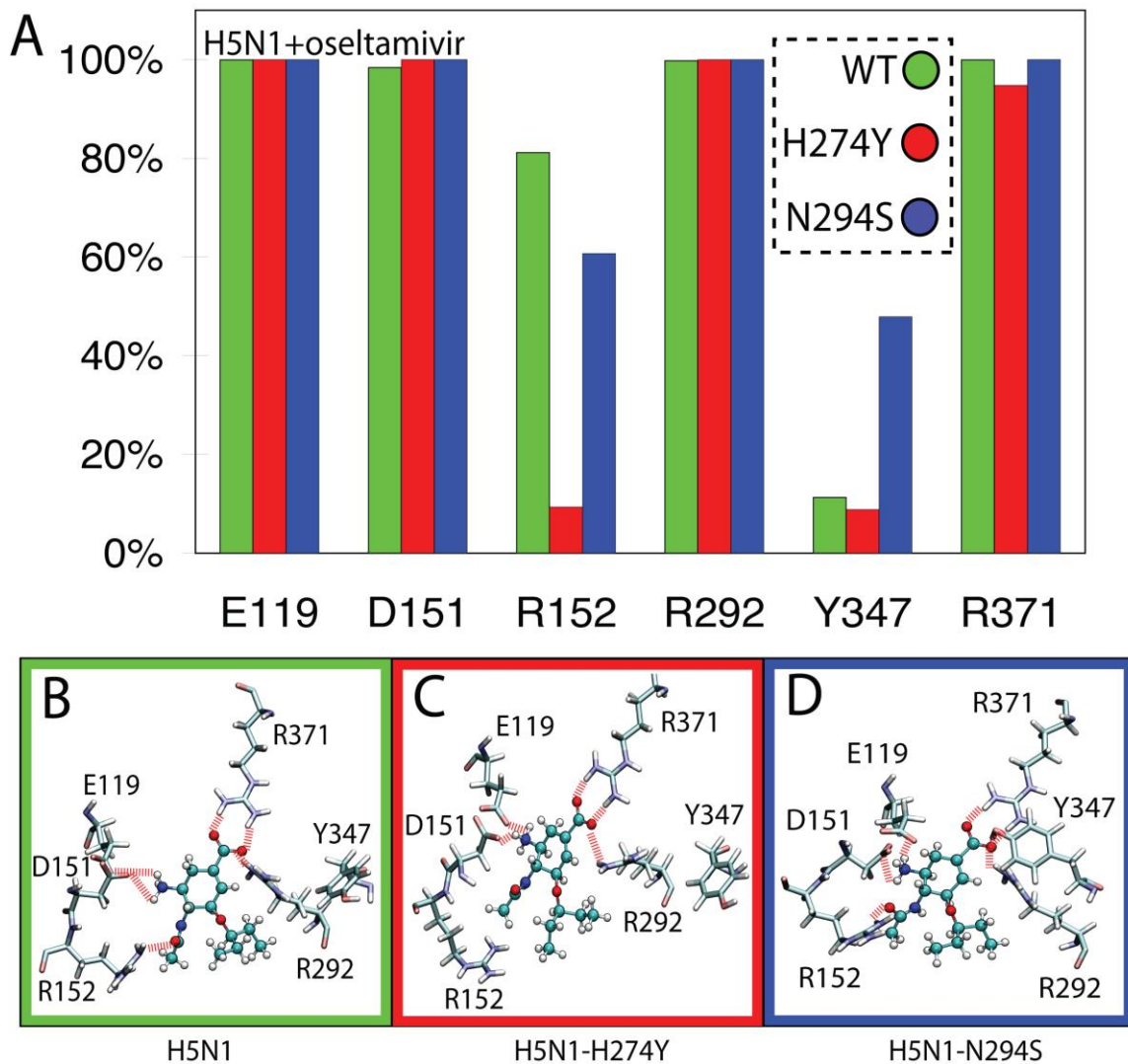
Hydrogen bonds form the primary interactions through which many drugs bind to neuraminidases. While the SA binding sites of H5N1 and H1N1pdm appear to differ mainly in the sequence of loop residue 347, it is not well understood whether antiviral drugs bind to each protein in the same manner, or if the drug resistant H274Y and N294S mutations disrupt critical hydrogen bonds. To address this question, the hydrogen bonds which form between oseltamivir and the residues lining the SA binding pockets of H5N1 and H1N1pdm were calculated for all simulation trajectories. The drug RMSD (Figure 4.1) of six simulated systems show that the positions of the drugs remain fairly stable within the binding pocket, thereby permitting the characterization of the specific drug-protein interactions responsible for binding oseltamivir to the neuraminidase active site. The results of these calculations are illustrated in Figures 4.2A and 4.3A, which show histograms listing hydrogen bond frequency, with schematic views of the specific residues involved in drug-protein hydrogen bond pairings in Figures 4.2B to 4.2D for simEQ4, and simEQ6, respectively. In both wild type and mutant simulations, hydrogen bonds between oseltamivir and binding site residues were well conserved, including



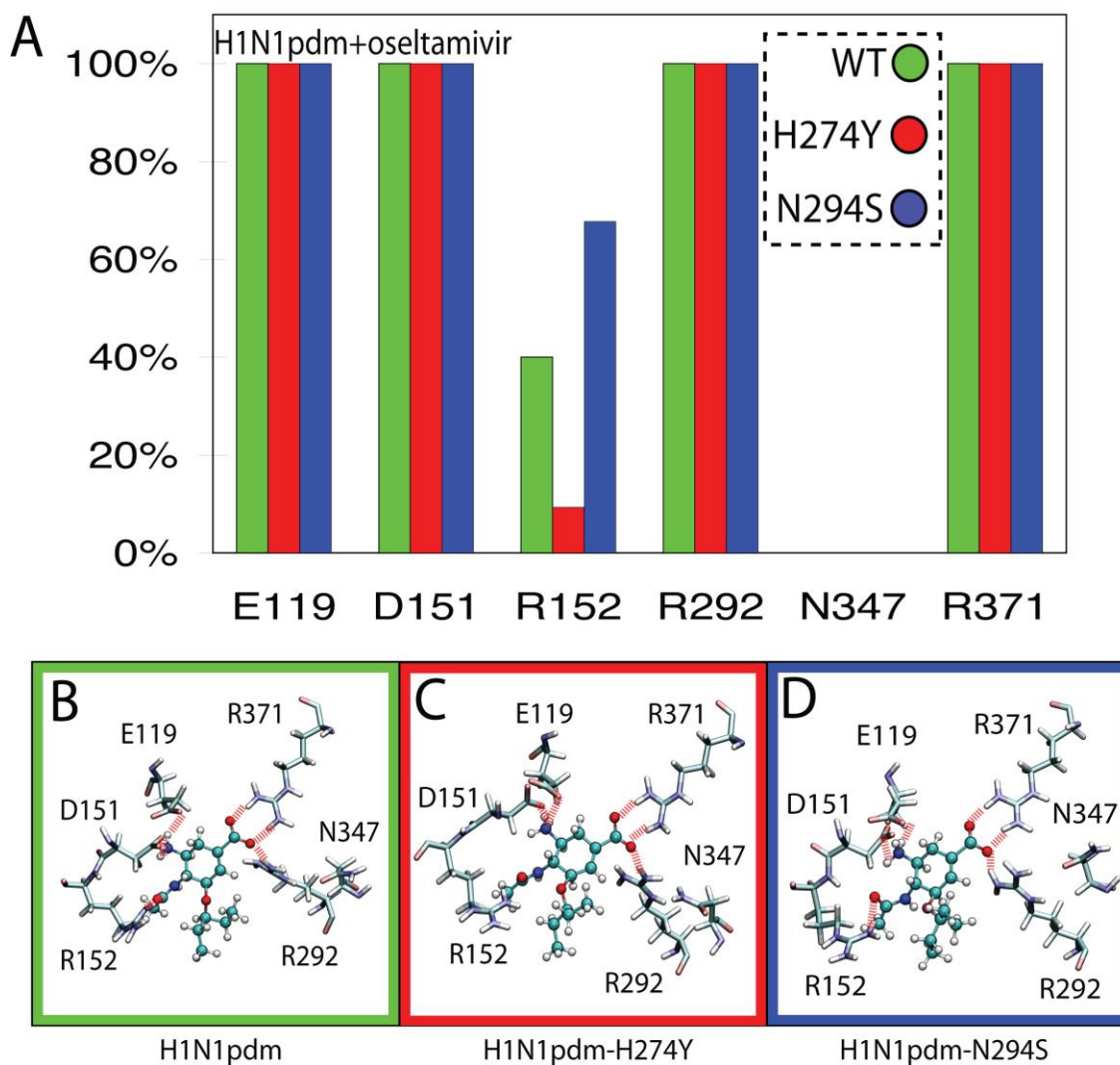
**Figure 4.1** Root mean squared deviation (RMSD) of WT and mutant avian H5N1 and swine H1N1pdm neuraminidases across six 20 ns production simulations (simEQ1 to simEQ6). The values reflect the equilibration of each of the neuraminidase systems.

D151, R292, and R371. Specifically, R292 and R371 were observed to hydrogen bond with oseltamivir's carboxyl moiety, and E119 and D151 with oseltamivir's amino group ( $\text{NH}_3^+$ ). The H274Y mutation, however, appeared to disrupt the hydrogen bonding of oseltamivir's acetyl group with R152, an interaction which was seen in the wild type and N294S systems for both simEQ1 and simEQ2.

Prior analyses of crystallographic data alone suggested that Y347 forms a stable hydrogen bond with oseltamivir's carboxyl group and is the source of oseltamivir-resistance in the N294S mutant.<sup>63</sup>



**Figure 4.2** Network and occupancy of hydrogen bonds stabilizing oseltamivir in the SA binding pocket of wild type and drug-resistant mutant avian H5N1 neuraminidases, in simEQ1, simEQ3, and simEQ5. (A) shows histograms of the percent of hydrogen-bond occupancies for interactions between oseltamivir and residues E119, D151, R152, R292, Y347, and R371 across each simulation run. (B) through (D) are schematic views depicting the orientation of protein sidechains which form protein-drug hydrogen bonds. Hydrogen bonds in all three simulations were conserved for residues E119, D151, R292, and R371. The H274Y mutation was observed to disrupt hydrogen bonding to R152. Despite the increased interaction with Y347 in the case of the N294S mutant, the hydrogen bonds between oseltamivir and Y347 were not stable in any of the simulated systems.



**Figure 4.3** Network and occupancy of hydrogen bonds stabilizing oseltamivir in the SA binding pocket of wild type and drug-resistant mutant avian H1N1pdm neuraminidases, in simEQ2, simEQ4, and simEQ6. (A) are histograms of the percent of hydrogen-bond occupancies for interactions between oseltamivir and residues E119, D151, R152, R292, N347, and R371 across each simulation run. (B) through (D) are schematic views depicting the orientation of protein sidechains which form protein-drug hydrogen bonds. Hydrogen bonds in all three simulations were conserved for residues E119, D151, R292, and R371. The H274Y mutation was observed to disrupt hydrogen bonding to R152. Of note, residue 347, which distinguishes the binding pocket of H1N1pdm from H5N1, makes no contribution to the drug-protein hydrogen-bond network in the case of H1N1pdm.

Our MD simulations however, reflect statistics collected from long timescale simulations which produce a dynamic picture of molecular interactions in greater detail and resolution than can be seen from a static crystal structure. In our simulations oseltamivir's carboxyl group primarily forms hydrogen bonds with R292 and R371, having little involvement with Y347. In fact, residue Y347 undergoes rotation to interact strongly with residue W295. Therefore, the speculation from previous studies, that the N294S mutation in the case of H5N1 actually destabilizes the hydrogen bonding between oseltamivir and Y347 to induce drug resistance, is not supported in our simulations.

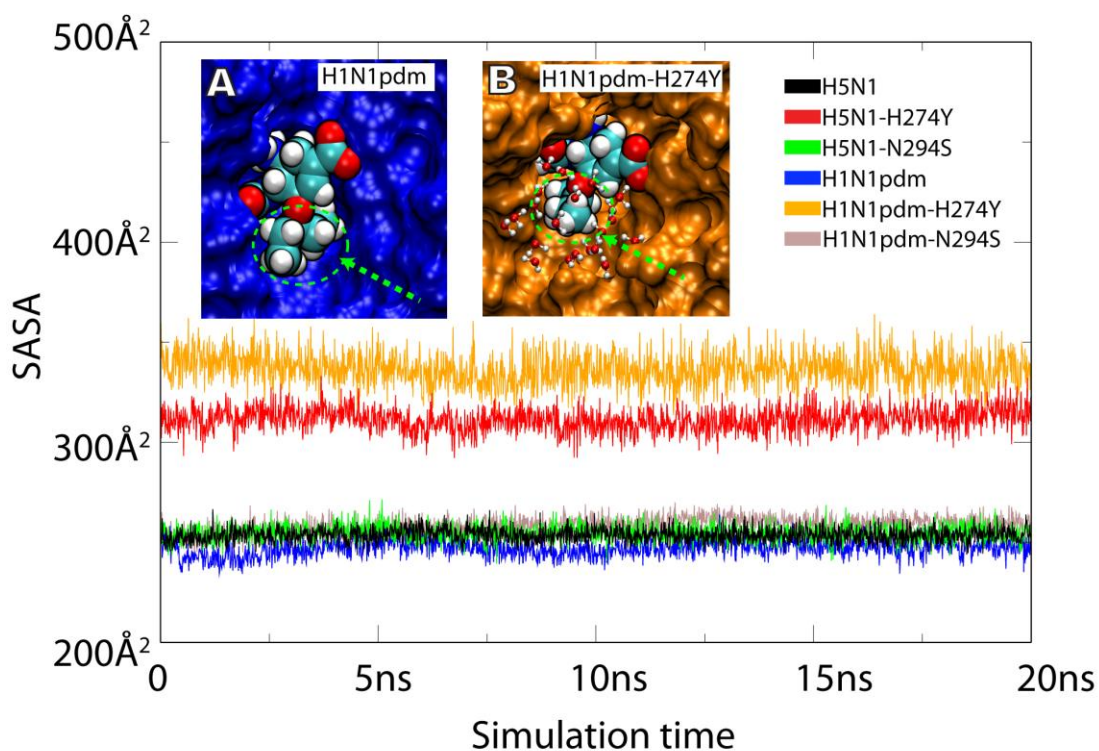
The notable difference between H5N1 and H1N1pdm neuraminidases is the replacement of Y347 by N347 at the drug binding pocket.<sup>14</sup> No conserved drug-protein hydrogen bond was observed for N347 in any of the three H1N1pdm simulations. Given the transient nature of even the N294S mutant induced hydrogen bond involving residue 347 in the case of H5N1, and the lack of interaction with residue 347 in any of the other simulations, it is highly unlikely that the single residue change (Y347 to N347), between the H5N1 and H1N1pdm strains significantly alters the drug-protein stability in regard to the hydrogen bond network involved. H274Y mutation induced disruption of the stable hydrophobic packing of oseltamivir's pentyl group in both H5N1 and H1N1pdm neuraminidases

Beyond disrupting the drug-protein hydrogen-bonding network, another mechanism through which protein mutations may induce drug resistance is by disruption of the hydrophobic packing of the drug into the protein binding pocket. Through inspection of the static crystal structures of the H274Y and N294S mutants of H5N1, it has been speculated that the mutations disrupt favorable hydrophobic packing

interactions necessary for strong binding of oseltamivir. In our wild type simulations for both the H5H1 and H1N1pdm systems, the packing of oseltamivir's pentyl moiety tended to favor close association with residues I222, R224, A246, and E276. To test the effect of mutations H274Y and N294S on hydrophobic interactions of oseltamivir's pentyl group with the proteins, we calculated the solvent accessible surface area (SASA) of oseltamivir's pentyl group for all simulation trajectories. While there was no significant change to the pentyl group SASA (henceforth referred to as PG-SASA) in the wild type and N294S mutant for either H5N1 or H1N1pdm neuraminidases, an outward rotation of the pentyl group was observed in the H274Y mutant simulations, visible in oseltamivir's binding pose and evident in a notably higher calculated PG-SASA. The PG-SASAs for all simulated systems are plotted in Figure 4.4, with inset images of oseltamivir's binding pose in simEQ2 (Figure 4.4A) and simEQ4 (Figure 4.4B) illustrating the rotation of the pentyl group towards the open mouth of the binding pocket.

Previously published MD simulations performed over relatively short time scales (3 to 6 ns) have suggested two possible mechanisms: 1) that the H247Y mutation reduces the size of the hydrophobic pocket within the SA binding pocket near Oseltamivir's pentyl moiety,<sup>70</sup> and 2) that the H274Y mutation breaks a critical salt bridge between E276 and R224 to disrupt drug binding.<sup>71</sup> Our longer (40ns) simulations were able to corroborate the former suggested mechanism (as shown by the increase in PG-SASA in the case of the H274Y mutant simulations) but not the latter. In fact, in all six of our simulations, E276 maintains stable charge-charge interactions (salt bridging) with R224 despite displacement of the drug from the protein I222-R224-A246-E276 pocket in the case of

H274Y mutants. This drug displacement increases water penetration into the I222-R224-A246-E276 pocket (Figure 4.4B). Evidence from our simulations therefore supports predictions from earlier studies of a possible mechanism for the H274Y mutation-induced drug resistance through water infiltration and destabilization of favorable drug packing.<sup>72</sup> However, because no change in PG-SASA was observed during our simulations in the case of N294S for either protein, N294S induced drug resistance is probably due to a different mechanism.



**Figure 4.4.** The solvent accessible surface area of oseltamivir's pentyl group (PG-SASA) in H5N1 and H1N1pdm wild type and mutant simulations. The PG-SASA in the N294S mutant simulation (simEQ5 and simEQ6) did not vary significantly from wild type simulations (simEQ1 and simEQ2). However, there was a huge loss of hydrophobic interaction between drug and protein in the case of the H274Y mutant simulations (simEQ3 and simEQ4), resulting in an increase in measured PG-SASA.



While alterations in hydrophobic sidegroup packing may explain in part the mechanism behind H274Y mutation-induced drug resistance, they fail to shed light on the role that the N294S mutation plays for oseltamivir binding inhibition. Initial investigations of the electrostatic surface potentials of drug-bound N1 neuraminidases have proposed that drug binding affinity is closely related to favorable charge-charge interactions with the electrostatic potential of the binding pocket wall, which exhibits a weak negative charge.<sup>12</sup> In order to understand whether charge-charge interactions may play a role in mutation induced drug resistance, the electrostatic potentials were calculated and mapped onto the surfaces of the proteins.

Even though prior studies have investigated the possible role of the electrostatic surface potential for drug binding in neuraminidases, the extensive electrostatic calculations required for fully understanding role of this phenomenon over a long simulation trajectory has not yet been done. The electrostatic surface potentials of the equilibrated systems were calculated and averaged across every trajectory frame in six equilibrium simulations described in Table 4.1 using the GPU accelerated multilevel summation method (described in the Methods section).<sup>68</sup> Extensive electrostatic calculations for the six systems show that the mouth of the binding pocket is positively charged, except for a narrow pathway of negative surface charge that seems to direct a possible binding pathway through which the drug may access the binding pocket. In Figure 4.5, the electrostatic surface potentials calculated from all six simulations is shown for the SA binding pocket with oseltamivir bound, with Figures 4.5A-C corresponding to

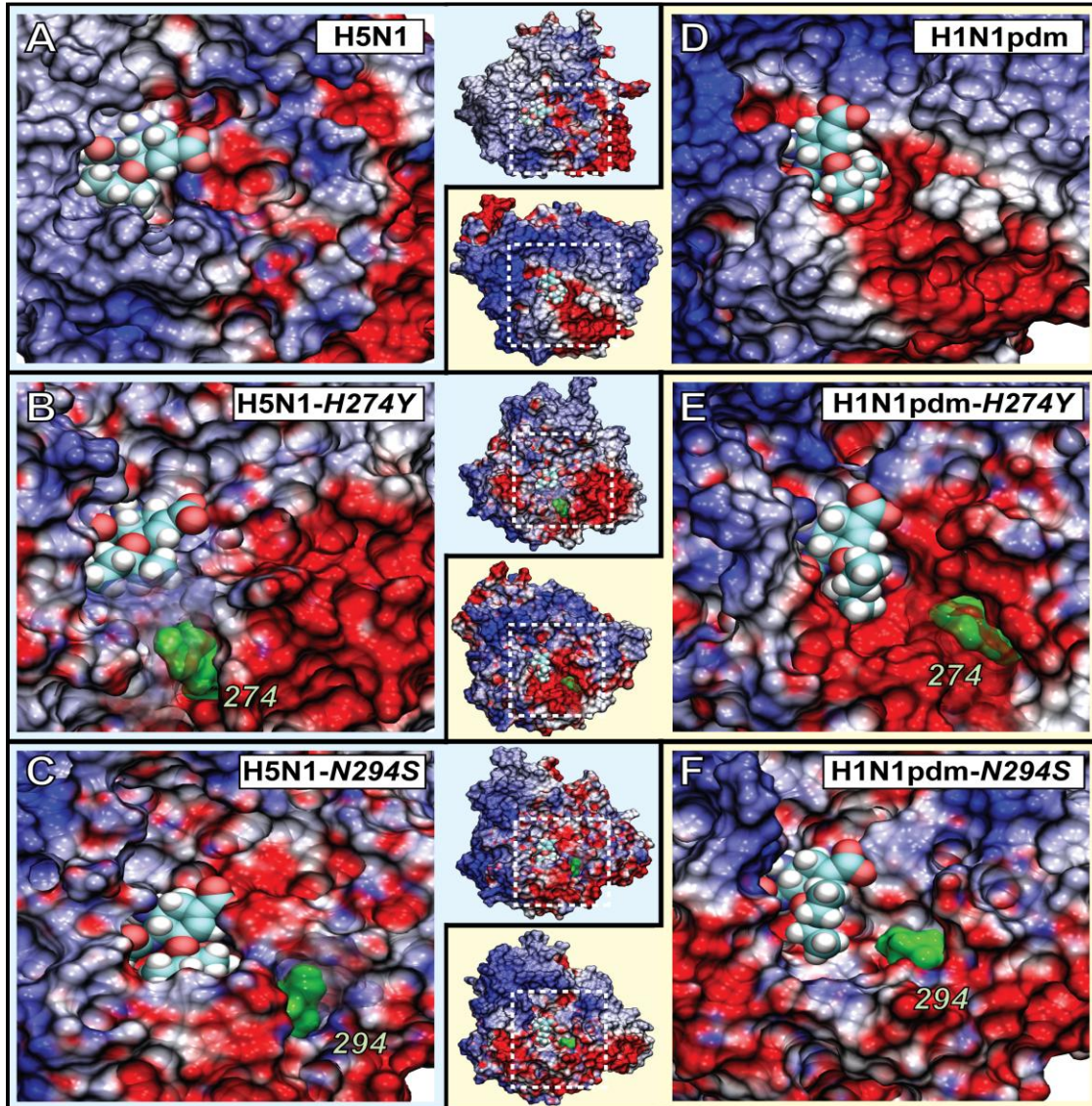
**Table 4.1** Summary of simulations. The “ensemble” column lists the variables held constant during the simulations; N, p and T correspond to the number of atoms, pressure, and temperature, respectively. The simulation times reflect a 20 ns setup/equilibration run followed by a 20 ns production run.

Name	Structure	Atoms	Ensemble	Time (ns)
simEQ1	H5N1 + oseltamivir	34860	NpT	40
simEQ2	H1N1pdm+ oseltamivir	34707	NpT	40
simEQ3	H274Y-H5N1 + oseltamivir	34880	NpT	40
simEQ4	H274Y-H1N1pdm+ oseltamivir	34729	NpT	40
simEQ5	N294S-H5N1 + oseltamivir	34873	NpT	40
simEQ6	N294S-H1N1pdm+ oseltamivir	34727	NpT	40

snapshots from simEQ1, simEQ3, and simEQ5, respectively, and Figures 4.5D-F corresponding to simEQ2, simEQ4, and simEQ6, respectively.

#### 4.4 Conclusion and outlook

Mutation H274Y in the H1N1pdm virus shares the same sources of oseltamivir resistance as H5N1, including loss of hydrogen bonds with R152 and reduction of hydrophobic interaction of oseltamivir’s pentyl group. This leads to our suggestion to replace the pentyl group by a more hydrophilic one by adding a few hydroxyl groups on this bulky hydrophobic group. Our results do not support previous suggestions that the N294S mutation of H5N1 destabilizes the hydrogen bonding between oseltamivir and Y347, or that it disrupts the hydrophobic pocket, leading to drug resistance. Sources of drug resistance in mutant N294S remain unclear. Quantitative analysis of drug binding to



**Figure 4.5** Electrostatic surface potential of avian H5N1 and swine H1N1pdm neuraminidases in oseltamivir-bound simulations, revealing a positively charged pathway into the binding pocket. (A) through (C) show oseltamivir bound to H5N1 wild type, H274Y, and N294S drug-resistant mutant structures, respectively. (D) through (F) show oseltamivir bound to wild type, H274Y, and N294S drug-resistant mutant structures, respectively. The outer columns show a close-up view of the binding pocket, highlighted as a subset of the enter protein shown in the central column. The positions of the mutant residues are shown in green for residue 274 and 294 in the mutant systems. Simulations revealed the presence of a negatively charged pathway at the mouth of the binding pocket which may play a role for drug binding and mutation induced resistance, as the position of residue 294 maps directly onto the pathway while position of residue 274 is positioned on or adjacent.

WT and the mutants will be importantly complementary to obtained results in this chapter. Since the two mutations are nonactive-site, endpoint interactions alone cannot account for all drug resistance. Kinetics of drug binding is therefore also important for drug resistant mechanism. Especially, based on the decrease in association rate constants ( $k_{on}$ ) of oseltamivir with H5N1 neuraminidase mutants ( $2.52 \mu\text{M}^{-1}\text{s}^{-1}$  in the WT,  $0.24$   $2.52 \mu\text{M}^{-1}\text{s}^{-1}$  in H274Y and  $1.1$   $2.52 \mu\text{M}^{-1}\text{s}^{-1}$  in N294S mutants), we speculate that mutations also prevent drugs from entering the SA binding site. It is hoped that our observation of a possible binding pathway for oseltamivir will encourage further investigations that test the hypothesis posed here, by identifying whether the actual drug binding pathway follows this route.

## CHAPTER 5

### AN ELECTROSTATIC FUNNEL DIRECTS BINDING OF OSELTAMIVIR TO INFLUENZA N1 SUBTYPE NEURAMINDASES

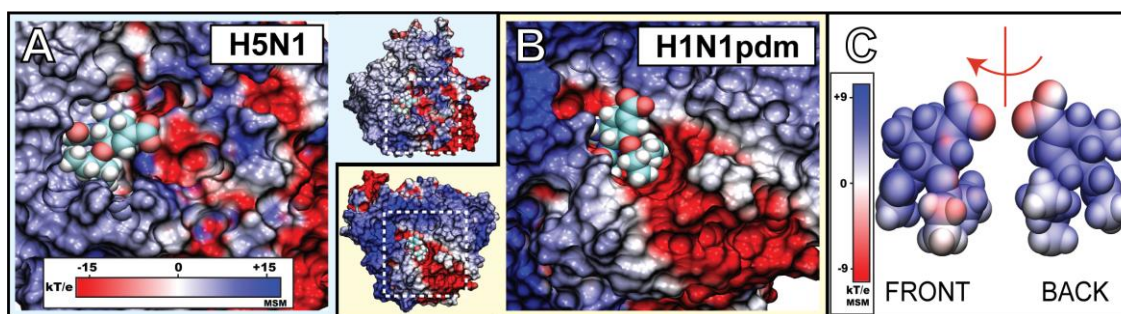
Chapter 5 is a follow-up study based on the observation in Chapter 4. This chapter reflects what was reported in a manuscript submitted to *PLoS Computational Biology* with the same title.<sup>17</sup> The work was done in collaboration with Eric Lee. Klaus Schulten also made a significant contribution in writing the manuscript.

#### **5.1 Characteristics of the electrostatic surface potential of N1 neuraminidases shed light for further study on drug binding and drug resistance**

Up to this point, all proposed mechanisms for oseltamivir resistance have focused on the effects of the mutations on the SA binding site but little has been known about the effects of those mutations on the kinetics of drug binding. Since both H274Y and N294S are nonactive site mutations, prior studies which focused on endpoint interactions between drug and proteins have been unable to provide a full understanding of how these mutations directly impact drug binding. In fact, given that the drug binding kinetics of

H5N1 mutants are significantly diminished, it is possible that these mutations alter the binding process, and not necessarily just the specific endpoint interactions.

It is well known that electrostatic surface potential of a protein can be an important driving force directing the diffusion of ligands into a protein's active site.<sup>73, 74</sup> The resulting electrostatic maps shown in Figure 5.1A for H5N1 and Figure 5.1B for H1N1pdm reveal a highly negatively charged column of residues that forms a pathway 10 Å in length between the SA binding site and the edge of the binding cavity mouth. Electrostatic calculations also reveal that oseltamivir has a highly positive electrostatic surface potential, illustrated in Figure 5.1C. The question arises whether the negatively charged surface column plays a role in the binding and unbinding of oseltamivir, given a possible mutual attraction between oseltamivir and this column. To answer this question, we employed SMD simulations (described in the Methods section) to pull oseltamivir out of the SA binding site and probe possible unbinding pathways. Such simulations have



**Figure 5.1** Electrostatic surface potential of the SA binding pocket of H1N1pdm and oseltamivir. Shown in A) and B) are closeup views of the SA binding pocket with drug bound H1N1pdm and avian H5N1 neuraminidase, respectively. The region of the binding pocket where the drug bound possesses a highly negative potential (colored red), whereas the opening of the pocket is surrounded by a highly positive potential ring (colored blue). In C), a detailed surface electrostatic potential for oseltamivir. Shown are the “front” side facing the annulus of the binding pocket, and “back” side facing the interior of the binding pocket is shown.

## 5.2 Computational details

A summary of simulations is shown in Table 5.1. In total, 6 systems were modeled and simulated for oseltamivir bound H5N1, and H1N1pdm wild type and H274Y and N294S mutants.

In simSMD1, steered molecular dynamics (SMD) simulations were used to remove oseltamivir from its stable binding site in H5N1 neuraminidase. In simFEQ1-10, equilibration simulations used a starting point generated from simSMD1 in which

**Table 5.1** Summary of simulations. The “Structure” column lists the type of neuraminidase, associated mutations, and drug modeled in each system.

Name	Structure	Atoms	Type	Ensemble	Time (ns)
simSMD1	H5N1 + oseltamivir	34860	SCV	NV	15
simFEQ1	H5N1 + oseltamivir	34860	EQ	NpT	15
simFEQ2	H5N1 + oseltamivir	34860	EQ	NpT	10
simFEQ3	H5N1 + oseltamivir	34860	EQ	NpT	15
simFEQ4	H5N1 + oseltamivir	34860	EQ	NpT	50
simFEQ5	H5N1 + oseltamivir	34860	EQ	NpT	100
simFEQ6	H5N1 + oseltamivir	34860	EQ	NpT	10
simFEQ7	H5N1 + oseltamivir	34860	EQ	NpT	15
simFEQ8	H5N1 + oseltamivir	34860	EQ	NpT	50
simFEQ9	H5N1 + oseltamivir	34860	EQ	NpT	50
simFEQ10	H5N1 + oseltamivir	34860	EQ	NpT	50

oseltamivir has undergone an axial rotation such that it is partially displaced from its binding site. In total, 680 ns of simulations were carried out on system sizes of approximately 35,000 atoms.

The “Ensemble” column lists the variables held constant during the simulations; N, p, T, and V correspond to the number of atoms, pressure, temperature, and volume respectively. Under “Type”, EQ denotes equilibration, and SCV denotes constant velocity SMD simulation with the speed of 0.5 Å/ns. SimSMD1 is a steered MD simulation with the starting structure from equilibrated simEQ1 (“preflip” drug position). In simFEQ1 to simFEQ10, the starting structure reflected a “flipped” position of oseltamivir from simSMD1 after 7.5ns of simulation, whereas several of its main stabilizing hydrogen bonds to the protein had already been ruptured. Crystallographically resolved water molecules and a structurally relevant calcium ion near the native binding site for SA were retained and modeled in all simulated systems. The protein complexes were then solvated in a TIP3P water box<sup>23</sup> and ionized by NaCl (0.152M) to mimic physiological conditions.

All simulations were performed using NAMD 2.7 and the CHARMM31 force field with CMAP correction. The ionized systems were minimized for 10,000 integration steps and equilibrated for 20 ns with 1 fs time steps. Following this, a 20 ns unconstrained equilibration production run was performed for subsequent trajectory analysis, with frames stored after each picosecond (every 1000 time steps). Constant temperature ( $T = 300$  K) was enforced using Langevin dynamics with a damping coefficient of 1 ps<sup>-1</sup>. Constant pressure ( $p = 1$  atm) was enforced using the Nose-Hoover Langevin piston method with a decay period of 100 fs and a damping time

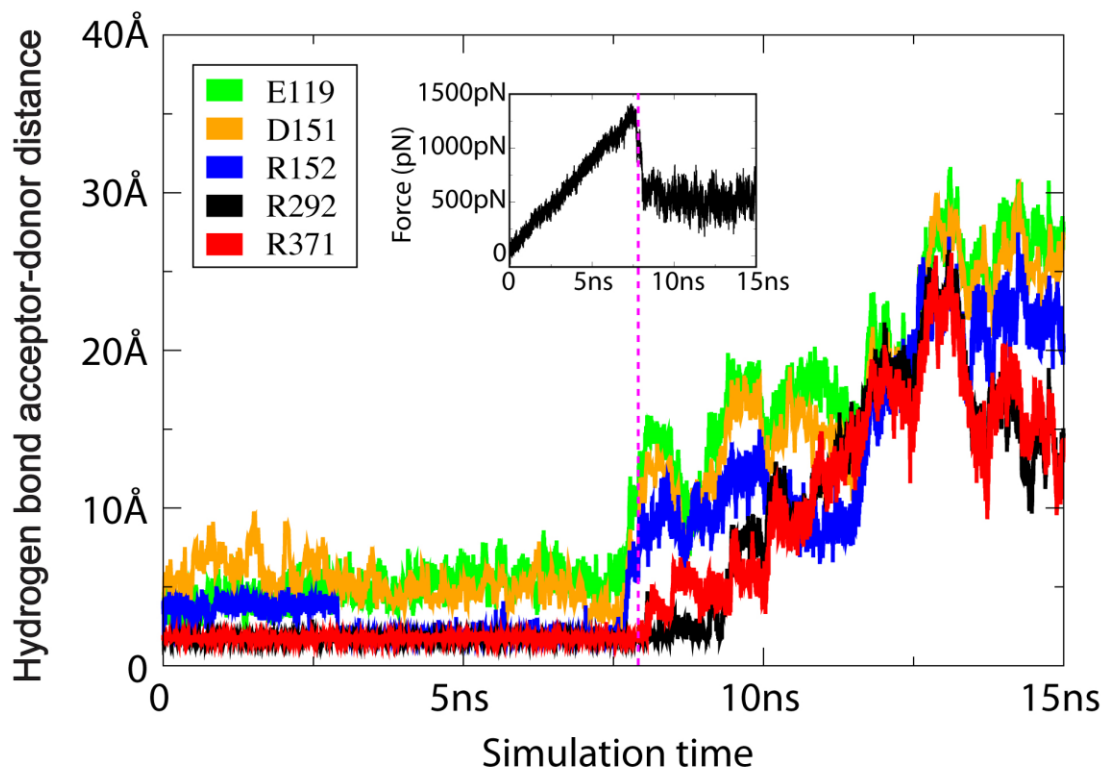


constant of 50 fs. Van der Waals interaction cutoff distances were set at 12Å, (smooth switching function beginning at 10 Å) and long-range electrostatic forces were computed using the particle-mesh Ewald (PME) summation method with a grid size of less than 1Å, along with the pencil decomposition protocol where applicable. SMD simulations fixed the center of mass of neuraminidase -carbons and applied a force to the center of mass of oseltamivir, along a vector connecting the two center of masses.

In simSMD1, a constant velocity protocol was employed, with stretching velocity of 0.5Å/ns. For the SMD spring constant,<sup>78, 79</sup> we chose  $k_{\text{SMD}} = 3k_{\text{B}}T/\text{Å}^2$  which corresponds to an RMSD value of  $(pk_{\text{B}}T/k_{\text{SMD}})^{1/2} < 0.6 \text{ Å}$ .

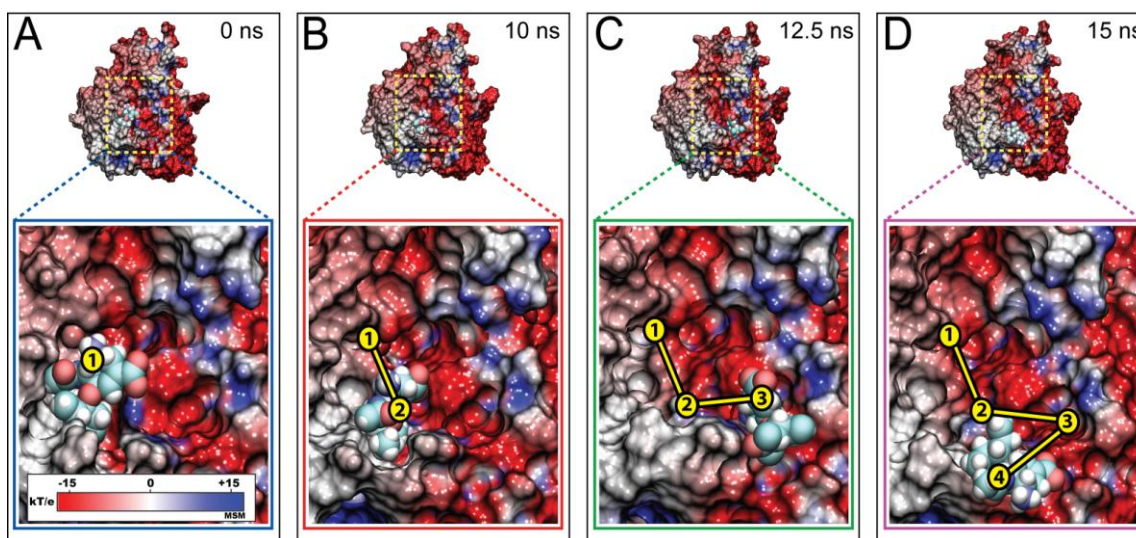
### 5.3 Results

To simulate the binding of the drug would be the most natural approach, but the computations would be rather expensive. Unbinding simulations however are feasible and may also reveal features that are characteristic for binding. In simSMD1, a pulling force was applied to rupture all the stabilizing hydrogen bonds between H5N1 and oseltamivir, in order to draw the drug away from the SA binding site. The results of simSMD1 show that the behavior of oseltamivir under effect of a pulling force can be divided into three distinct stages: 1) from 0 to 8 ns, a buildup of forces during which hydrogen bonds between oseltamivir with E119, D151 and R152 are destabilized; 2) at 8 ns where the remaining stable hydrogen bonds with R292 and R371 rupture; 3) after 8 ns during which the drug is pulled out of the binding pocket. Figure 5.2 shows the force dependent rupture of hydrogen bonds in simSMD1 by plotting both hydrogen bond stability and (in inset) the force vs. time curve. To our surprise, it was observed that



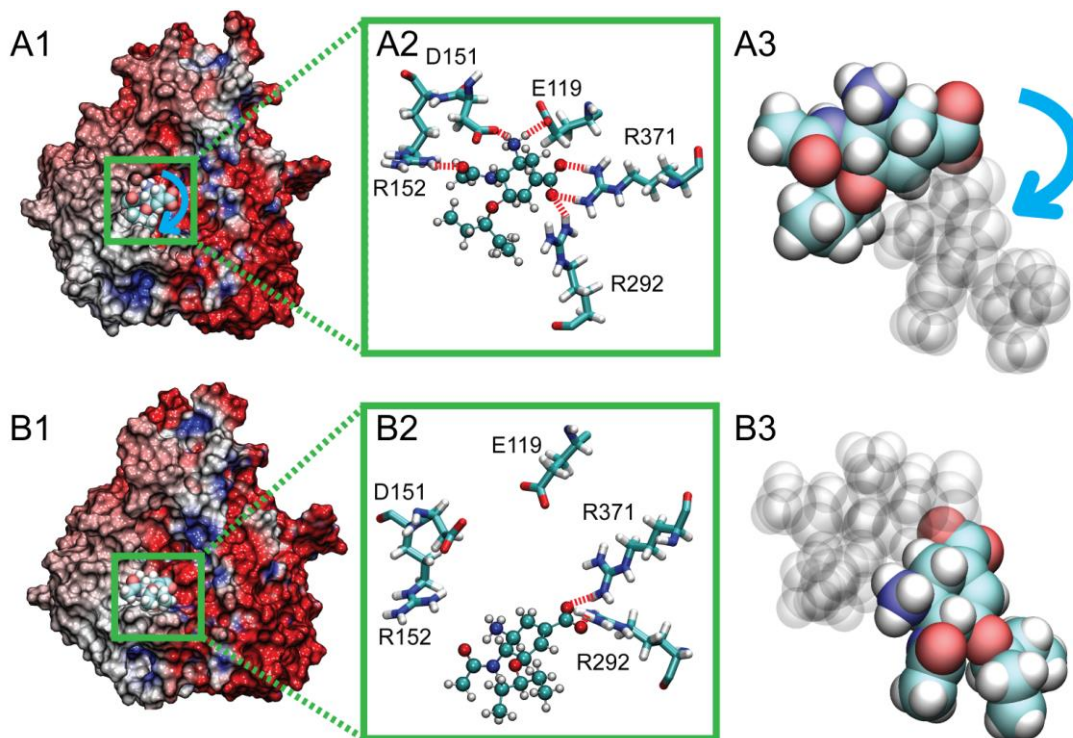
**Figure 5.2** Distances between hydrogen bond acceptor-donor pairs between oseltamivir and active site amino acids vs. simulation time in simSMD1. Most hydrogen bonds were quickly broken by the pulling force except for those with R292 and R371, which fully ruptured only after 8 ns, corresponding to the peak of the curve of the applied force vs simulation time (shown in inset). position of oseltamivir and all its possible hydrogen bonding pairs with residues located along this pathway, it was seen that nonspecific electrostatic attractions formed the predominant interactions between drug and protein.

oseltamivir followed a lateral escape path through strong interaction with the negatively charged column of residues identified via electrostatic mapping above, despite application of force to pull the drug straight out of the binding pocket. This divergent path taken by the drug is shown in Figure 5.3A-D. A second notable observation in simSMD1 was that just before rupture of hydrogen bonds between oseltamivir's



**Figure 5.3** Forced unbinding of oseltamivir from H5N1 neuraminidase. Shown here are the relative positions of oseltamivir on the electrostatic surface of avian H5N1 neuraminidase during simSMD1. At 0 ns (A), oseltamivir was stably bound within the active site, as seen in simEQ1. Application of force ruptured the stabilizing hydrogen bonds between H5N1 and oseltamivir (see Figure 5.2), drawing the drug away from its stable binding site within 10 ns, as shown in B. Over the next 2.5 ns of pulling, oseltamivir followed the charged binding funnel (shown in C) until it was completely free of the protein binding pocket after 15 ns, as shown in D).

carboxyl functional group and residues R292 and R371 of the binding pocket, the drug underwent a rather significant rotation due to the destabilization of hydrogen bonds with E119, D151 and R152. It is likely that this rotation or “flip” is crucial for placing oseltamivir in a position which permits it to leave the SA binding site. Despite application of force to pull oseltamivir straight out of its binding pocket, the drug followed a somewhat lateral path through the electrostatically charged funnel. The “preflip” or stably bound state for oseltamivir in H5N1 is shown in Figure 5.4A1, 5.4A2, 5.4A3, and the flipped transition state shown in Figure 5.4B1 and 5.4B2, 5.4B3.



**Figure 5.4** Oseltamivir in “flipped” position (position of oseltamivir at 7.5ns in simSMD1) in comparison with its stable equilibrium position, shown with the electrostatic surface potential of the protein. In A1, the stable binding pose of oseltamivir is shown prior to application of force in simSMD1, with stable hydrogen bonds to E119, D151, R152, R292, and R371 (shown in A2). After 7.5 ns, however, oseltamivir had adopted a new binding pose (shown in B1), where hydrogen bonds with E119, D151, and R152 were ruptured, leaving only stable interactions with R292 and R371 (shown in B2). The blue arrow in A1 denotes orientation change of the drug from “preflip” to “flipped” positions, shown in A3 and B3 respectively with contrast position in transparent mode. The weakly bound “flipped” position of oseltamivir in B1, B2 and B3 was used as the starting point for simulations simFEQ1-10.

It was observed in simSMD1 that following the transition to the flipped state, very little force was then required to subsequently draw oseltamivir out of its binding pocket in neuramidase. This result suggests that one may be able to probe the unbinding pathway via diffusion, if the starting state consists of oseltamivir already in its flipped orientation. We therefore performed additional equilibrium simulations (simFEQ1-10) with oseltamivir already in this flipped state. From these simulations, we were able to observe

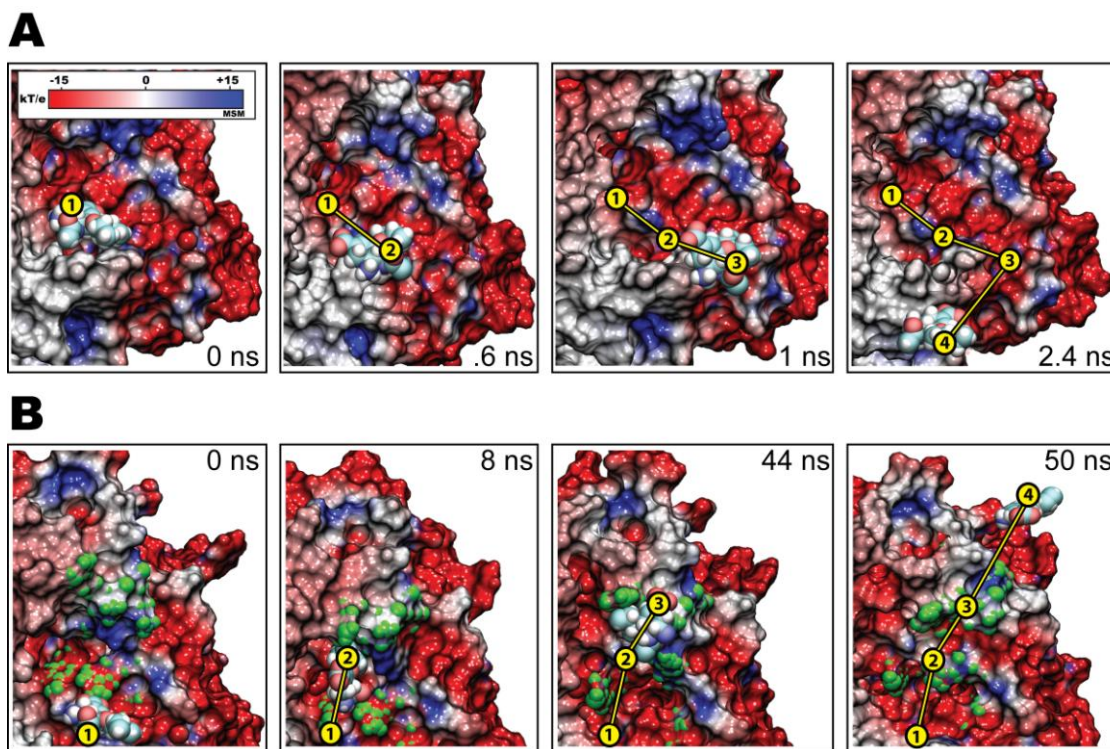
two distinct outcomes: 1) oseltamivir is able to escape the binding pocket by interacting favorably with the charged binding funnel and 2) the drug returns, not unexpectedly, to its stably bound “preflip” state. Each simulation for FEQ1-10 was run long enough to observe either outcome, with the exception of FEQ5. SimFEQ5 was a special case in which the drug, after following the binding funnel to escape the protein, actually rebound to the SA active site through the same binding funnel after we extended the simulation to follow its movement. A summary of observed outcomes from these simulations is shown in Table 5.2.

In simFEQ1-5, oseltamivir successfully escaped the SA binding site, whereas in simFEQ6-10, oseltamivir returned to its stably bound “preflip” state. Oseltamivir was observed to diffuse out of the SA active site after strong interaction with the electrostatically charged binding funnel (described above) in five out of ten equilibrium

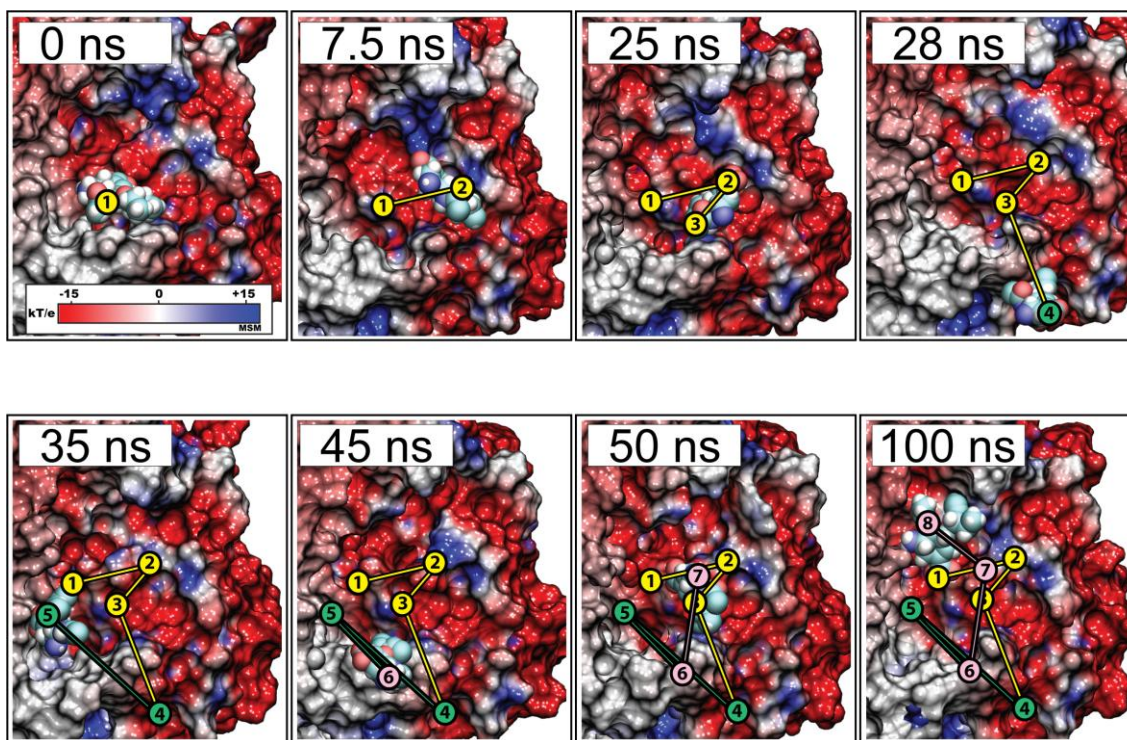
**Table 5.2** Summary of FEQ1-10 simulations starting from “flipped” position of oseltamivir taken from simSMD1 at 7.5 ns.

Name	Result	Time (ns)
simFEQ1	Drug escape via binding funnel	15
simFEQ2	Drug escape via binding funnel	10
simFEQ3	Drug escape via binding funnel	15
simFEQ4	Drug interaction with binding funnel but escape via 430-cavity	50
simFEQ5	Drug escape and rebind into SA pocket via binding funnel	100
simFEQ6	Drug returned to “preflip” position	10
simFEQ7	Drug returned to “preflip” position	15
simFEQ8	Drug returned to “preflip” position	50
simFEQ9	Drug returned to “preflip” position	50
simFEQ10	Drug returned to “preflip” position	50

simulations we performed (simFEQ1-5). In four of the cases (simFEQ1-3, 5) oseltamivir was observed to diffuse along the full length of the binding funnel before disassociating with neuraminidase. Snapshots from a representative simulation (in this case simFEQ1) illustrating the trajectory oseltamivir follows along our proposed binding funnel are shown in Figure 5.5A. This trajectory and those from simFEQ2,3, and 5 can be visualized even more clearly in Movies S1, S2, S3, and S5 in Appendix. In simFEQ4, oseltamivir was observed to briefly interact with the binding funnel, but diffused out of neuraminidase through an alternate path, namely, the “430-cavity” identified in earlier computational study.<sup>12</sup> Specifically, the simulation revealed that hydrogen bonds between oseltamivir carboxyl group and the guanidino group of R430 appear to stabilize the transition of the drug along this alternate pathway. The 430-cavity is believed to function as a secondary binding site for SA, and our simulations here show that it may (not surprisingly) also serve as a viable conduit for the binding/unbinding of oseltamivir as well. Snapshots from simFEQ4 illustrating the trajectory of oseltamivir through the 430-cavity are shown in Figure 5.5B, with 430-cavity-specific residues colored in green. In simFEQ1-4, once oseltamivir disassociated from neuraminidase, the drug diffused freely in the solvent environment away from the protein. However in simFEQ5 we observed not only a diffusion of oseltamivir through the charged binding funnel, but also the re-entry. Snapshots from each of these events in FEQ5 are shown in Figure 5.6. Analysis of interactions of the newly rebound oseltamivir with active site residues from 50 to 100 ns revealed that the drug was stabilized by hydrogen bonds with Y406, R292, D151, E119, and R118, even though the pentyl group had not yet moved to its requisite hydrophobic pocket (I222-R224- A246-E276).<sup>81</sup> A full trajectory taken by the drug can be seen in



**Figure 5.5** Escape of oseltamivir from H5N1 neuraminidase in equilibrium simulations from simFEQ1 and simFEQ4. The initial system for simulation reflected the “flipped” orientation of oseltamivir, as illustrated in Figure 4.4. Shown in A) are snapshots from simFEQ1, showing the interaction of oseltamivir with H5N1’s electrostatically charged binding funnel as the drug diffuses out of the protein’s SA binding pocket. Results were similar for simFEQ2-3 and simFEQ5 (see text and Table 5.2 for specific timescale for those events), demonstrating a key role for this charged funnel for directing drug transition into and out of influenza neuraminidases. In snapshots from simFEQ4, (shown in B), oseltamivir briefly interacted with the binding funnel but took a different path out of the binding pocket, interacting instead with residues from the so-called “430-cavity”, predicted in a prior study<sup>12</sup> to be a secondary SA binding site. of the drug through the same pathway after it had freely diffused away from neuraminidase. Specifically, the sequence we observed in simFEQ5 was as follows: 1) between 0 to 25 ns oseltamivir diffuses out of neuraminidase’s SA active site by following the charged binding funnel, 2) between 25 and 35 ns the drug unsuccessfully attempted to rebind from an unsuitable direction through hydrogen bond interactions with R152 in the so-called flexible 150-loop<sup>80</sup>; 3) between 35 and 45 ns, the drug again diffuses away from neuraminidase; 4) between 45 and 50 ns, the drug rediscovers the binding funnel, drawing it back into and binding at the SA active site for 5) at least the next 50 ns (100ns total simulation time).



**Figure 5.6** Escape and rebinding of oseltamivir through the electrostatic binding funnel in H5N1 neuraminidase during simFEQ5. Shown here are snapshots of simFEQ5, in which oseltamivir first diffused out of the SA active site through interaction with the electrostatic binding funnel (see Figure 5.1) similar to that seen in simFEQ1-3 (see Figure 5.5A) within the first 25 ns of simulation. Between 28 and 35 ns, oseltamivir diffuses and approaches the periphery of the binding pocket away from the binding funnel, but is prohibited from entering due to electrostatic repulsion (45 ns). However, between 45 and 50 ns, oseltamivir was observed to approach and enter the neuraminidase binding pocket through the binding funnel, adopting a stable position within the SA binding pocket through hydrogen bonds with Y406, R292, D151, E119, R118.

Movie S5, illustrating the strongest evidence observed thus far that the electrostatic funnel serves a crucial role in both the binding and the unbinding of oseltamivir.

#### 5.4 Discussion and future studies

Our study has shed light on the important role of the electrostatic surface potentials in directing the diffusion of oseltamivir into the SA binding site. From our simulations, it is clear that the negatively charged funnel serves as a prominent binding



and unbinding pathway for oseltamivir in the wild type systems investigated with SMD and followup equilibrium simulations in simSMD1 and simFEQ1-10. It turns out also that the binding funnel may possibly play a crucial role in drug resistance caused by mutations. The conspicuous location of residue 294, which maps directly onto this negatively charged pathway may play a key role in the N294S mutation for disrupting the proper guidance of the drug into its binding pocket. Thus, it is possible that the 274 and 294 mutations may confer drug resistance by not only disrupting the end-point interactions of oseltamivir but also its entry into the binding pocket by interfering with the binding funnel. While sources of end-point interactions, including hydrogen bonds, hydrophobic packing and solvent infiltration, of oseltamivir resistance<sup>70-72</sup> have been thoroughly studied, little is known about the kinetics of the drug binding in mutants at atomic level. The idea that drug resistant mutants actually disrupt entry of oseltamivir into the SA active site of neuraminidase through disruption of an electrostatic binding funnel is in part supported by experiments which have noted reduced drug binding kinetics in H5N1 H274Y and N294S mutants. Even though the oseltamivir-resistant mutations were seen located in or adjacent to the funnel, clearly additional study is still needed for a full understanding of how the H274Y and N294S mutations weaken the binding of the drug. Furthermore, the mutations might not only affect the electrostatic gradient but also the geometry of the drug binding and unbinding path. Future or followup studies should therefore focus on the specific drug entry/exit pathways for oseltamivir-resistant mutant systems, sampling timescales great enough (such as in the case of simFEQ5 described above) to capture both the binding.

The role of flexible loop-150 and loop-430 in mediating binding for oseltamivir. Previous studies have found that the presence of two flexible loops (termed “150” and “430” due to their residue positions) may play a role in guarding drug access to the SA binding pocket of N1 neuraminidases.<sup>80</sup> Our simulations suggest that a loss of hydrogen bonds to loop 150 (namely between oseltamivir’s acetyl group and R152) in the H274Y mutant systems (simEQ3 and simEQ4). However, based on our simulations, the disruption of endpoint interactions between oseltamivir and loop 150 do not appear to appreciably destabilize overall drug binding within the primary SA active site. This observation, however, does not rule out what role loop 150 (and its associated D151 and R152 hydrogen bonds to oseltamivir) or the other flexible loop 430 may play if the drug is occupying an observed secondary SA binding site seen in previous studies<sup>12, 82</sup> and in one of our simulation trajectories (simFEQ4). While beyond the scope of this study, it remains to be seen therefore, whether these flexible loops are key players involved in mutation-induced drug resistance. Our electrostatic maps favor a drug binding/unbinding pathway for oseltamivir along the charged binding funnel to the primary SA binding site discussed in the results section above, involving possible association with the positions occupied by residues 274 and 294. However, simFEQ4 did capture drug unbinding through the secondary “430-cavity” binding pocket (via interaction of oseltamivir’s carboxyl group and R430’s guanidino group) and subsequently through an alternate pathway different from the primary charged binding funnel that we have presented. It is possible therefore that because our simulations from which the electrostatic maps were derived modeled oseltamivir in its primary binding site, a secondary charged binding channel/exit path may emerge when the drug occupies its secondary site; verifying

whether such a charged pathway exists would therefore be a suitable endeavor for a subsequent studies.

Our simulations here were restricted to studying six subtypes of N1 neuraminidase; expanding the scope of investigation to encompass the N2-9 neuraminidases may potentially reveal similar electrostatic funnels that direct drug binding to those proteins. Finally, it is well known that the H274Y and N294S mutants do not alter the binding affinity for another antiviral drug, zanamivir,<sup>18</sup> or certainly not its natural substrate, SA. Zanamivir, which shares a higher similarity in structure with SA than oseltamivir, might enter SA binding pocket in the same way as the SA but different from that of oseltamivir. It is hoped that our observations of a possible charged pathway for oseltamivir will encourage further investigations into whether other drug candidates follow this route to enter the SA binding site and how they interact with H274Y and N294S mutations during the binding/unbinding processes.

## 5.5 Supplementary information

Movies located at <http://www.ks.uiuc.edu/~lyle/Tamiflu-binding-pathway-movies.html>

Movie S1. (S1.avi) Trajectory from simSMD1, where a force is applied to oseltamivir perpendicular to the plane of the SA binding pocket. Despite the direction of force, oseltamivir interacts with and follows the charged electrostatic pathway identified and discussed in Figure 5.1, with snapshots shown in Figure 5.3.

Movies S2-S4. (S2, S3, and S4.avi) Trajectories from simFEQ1-3, respectively, for movies S2-S4, showing the diffusion of oseltamivir out of the neuraminidase binding

pocket via interaction with the electrostatic binding funnel (see Figure 5.3 for snapshots from simFEQ1).

Movie S5. (S5.avi) Trajectory from simFEQ4, where oseltamivir briefly interacts with the electrostatic binding funnel before diffusing out of the neuraminidase binding pocket via an alternate pathway in the region of the “430-loop”<sup>12</sup> shown in green as shown in Figure 5.3.

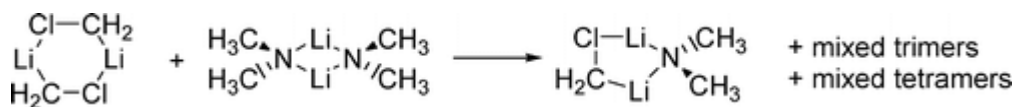
Movie S6. (S6.avi) Trajectory from simFEQ5, where oseltamivir diffuses out of the neuraminidase active site via interaction with the electrostatic binding funnel, fails to enter the active site at a different location at periphery of the binding pocket due to electrostatic repulsion, then rebinds stably to neuraminidase through entry via the binding funnel. See Figure 5.6 for snapshots.

## APPENDIX

This section presents an abstract of published/presented results that are not included in this dissertation.

### A.1. A Computational study of mixed aggregates of chloromethyl lithium with lithium dialkylamides<sup>83</sup>

DFT calculations were performed to examine the possible formation of mixed aggregates between chloromethyl lithium carbenoids and lithium dimethylamide (LiDMA). In the gas phase mixed aggregates were readily formed and consisted of mixed dimers, mixed trimers, and mixed tetramers. THF solvation disfavored the formation of mixed tetramers and resulted in less exergonic free energies of mixed dimer and mixed trimer formation.



## A.2. Structure and mechanism of formation of an important ion in doping control<sup>84</sup>

An ion with  $m/z$  143 serves as a biomarker that is often continuously monitored in urine samples undergoing screening by electron ionization gas chromatography/mass spectrometry (EI GC/MS) for banned anabolic agents. The ion is known to arise from trimethylsilyl (TMS)-derivatized synthetic 17-hydroxy, 17-methyl steroids. The purpose of this work was to characterize, in detail, the origin(s), structure(s), and mechanism(s) of formation of such ions with  $m/z$  143. High resolution mass spectrometry (HRMS) data revealed the elemental composition of the D-ring derived  $m/z$  143 ion to be  $C_7H_{15}OSi$ . Analysis of dihydrotestosterone (DHT) and its 2-methyl substituted analog dromostanolone by HRMS revealed that an elementally equivalent ion of  $m/z$  143 could be derived from the A-ring of TMS-derivatized 3-keto-enol steroids demonstrating that an abnormally intense peak in the  $m/z$  143 extracted ion chromatogram of urine samples undergoing screening for banned anabolic agents does not necessarily indicate the presence of a 17-hydroxy, 17-methyl steroid. To gain information on ion structure, breakdown curves for the most abundant product ions of the  $m/z$  143 ion were generated using both native and perdeutero-TMS derivatives, providing structures for second, third, and fourth generation product ions. An EI-mass spectrum of  $[16,16,17-^2H_3]$ -DHT (DHT-d3) demonstrated that one of the C-16 hydrogen atoms is removed prior to the formation of an ion that is highly analogous to the ion with  $m/z$  143 strongly suggesting, in accord with all other evidence, one particular fragmentation pathway and resulting product: a resonance stabilized 3-(*O*-trimethylsilyl)but-1-ene ion.

### **A.3. Liquid-liquid immiscibility in supercooled aqueous solutions driven by water's polyamorphism<sup>85</sup>**

It has long been known that the addition of salts to water favors vitrification. The identification of two distinct glass transitions in calorimetric experiments of LiCl-water glasses prompted the hypothesis of a liquid-liquid phase separation in supercooled water. These experiments, however, could not confirm the existence of two phases, nor their composition or dimensions of phase segregation.

We present direct evidence by computer simulations of the liquid-liquid immiscibility that develops in supercooled aqueous solutions. Two water nanophases form on hyperquenching of some water solutions: a pure low-density liquid phase that vitrifies to LDA, and a higher density water-solute mixture. The immiscibility of supercooled aqueous solutions is driven by the transformation of water from a high to a low density liquid (LDL): the hydrophilic solutes cannot be incorporated into the tetrahedral LDL network and are expelled to form an ion-rich water nanophase. The simulation results agree with the interpretation of raman spectra and quantitatively reproduce the relative amounts of low and high density water as a function of solute content of their experiment. Our simulations yield further insight into the mechanism of the process and the dimensions of phase segregation.

The supercooled solutions are studied through coarse-grained molecular dynamics simulations, employing the mW monatomic water model of Molinero et al. The formation of ice, that can be controlled by the cooling rate of the solution, is also discussed in relation of a minimum LDL cluster size needed to produce ice nucleation in these systems.

## REFERENCES

1. Berman HM, B. T., Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C., The Protein Data Bank, Acta Crystallogr D Biol Crystallogr. 2002 Jun;58(Pt 6 No 1):899-907. Epub 2002 May 29.
2. Holbeck, S. L., Update on NCI in vitro drug screen utilities. *European journal of cancer (Oxford, England : 1990)* **2004**, 40 (6), 785-793.
3. Irwin, J.; Shoichet, B., ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling* **2005**, 45 (1), 177-182.
4. Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M., DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucl. Acids Res.* **2008**, 36 (suppl\_1), D901-906.
5. Le, Q. M.; Kiso, M.; Someya, K.; Sakai, Y. T.; Nguyen, T. H.; Nguyen, K. H. L.; Pham, N. D.; Ngyen, H. H.; Yamada, S.; Muramoto, Y.; Horimoto, T.; Takada, A.; Goto, H.; Suzuki, T.; Suzuki, Y.; Kawaoka, Y., Avian flu: Isolation of drug-resistant H5N1 virus. *Nature (London, U. K.)* **2005**, 437 (7062), 1108.
6. Laver, G., Antiviral drugs for influenza: Tamiflu past, present and future. *Future Virol.* **2006**, 1 (5), 577-586.
7. Ford Stephen, M.; Grabenstein John, D., Pandemics, avian influenza A (H5N1), and a strategy for pharmacists. *Pharmacotherapy* **2006**, 26 (3), 312-22.
8. Butler, D., Swine flu goes global. *Nature.* **2009**, (458), 1082-1083.
9. Abed, Y.; Baz, M.; Boivin, G., Impact of neuraminidase mutations conferring influenza resistance to neuraminidase inhibitors in the N1 and N2 genetic backgrounds. *Antiviral Ther.* **2006**, 11 (8), 971-976.
10. Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol* 11: 224-230.



11. Le, Q. M.; Kiso, M.; Someya, K.; Sakai, Y. T.; Nguyen, T. H.; Nguyen, K. H. L.; Pham, N. D.; Ngyen, H. H.; Yamada, S.; Muramoto, Y.; Horimoto, T.; Takada, A.; Goto, H.; Suzuki, T.; Suzuki, Y.; Kawaoka, Y., Avian flu: Isolation of drug-resistant H5N1 virus. *Nature (London, U. K.)* **2005**, *437* (7062), 1108.
12. Cheng, L. S.; Amaro, R. E.; Xu, D.; Li, W. W.; Arzberger, P. W.; McCammon, J. A., Ensemble-Based Virtual Screening Reveals Potential Novel Antiviral Compounds for Avian Influenza Neuraminidase. *J. Med. Chem.* **2008**, *51* (13), 3878-3894.
13. Nguyen HT, L. L., Truong TN., Top-hits for H1N1pdm identified by virtual screening using ensemble-based docking: top-hits for A/H1N1 identified by virtual screening using ensemble-based docking PLoS Curr Influenza. 2009. RRN1030.
14. Le L, L. E., Schulten K, Truong TN., Molecular modeling of swine influenza A/H1N1, Spanish H1N1, and avian H5N1 flu N1 neuraminidases bound to Tamiflu and Relenza. PLoS Curr Influenza. RRN1015.
15. Freedman, H.; Le, L.; Tuszynski, J. A.; Truong, T. N., Improving the Performance of the Coupled Reference Interaction Site Model-Hyper-netted Chain (RISM-HNC)/Simulation Method for Free Energy of Solvation. *Journal of Physical Chemistry B* **2008**, *112* (8), 2340-2348.
16. Freedman H, H. L., Le L, Cheatham TE 3rd, Tuszynski JA, Truong TN., Explicitly solvated ligand contribution to continuum solvation models for binding free energies: selectivity of theophylline binding to an RNA aptamer. *J Phys Chem B.* 2010 Feb 18;114(6):2227-37.
17. Ly Le, E. H. L., David J. Hardy, Thanh N. Truong, and Klaus Schulten, Electrostatic Funnel Directs Binding of Tamiflu to H5N1/H1N1pdm Neuraminidase. *submitted to PLOS Computational Biology.*
18. Colman, P. M., Zanamivir: an influenza virus neuraminidase inhibitor. *Expert Rev. Anti-Infect. Ther.* **2005**, *3* (2), 191-199.
19. 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 (London, U. K.)* **2006**, *443* (7107), 45-49.
20. Humphrey, W.; Dalke, A.; Schulten, K., VDM: visual molecular dynamics. *Journal of Molecular Graphics* 1996, *14* (1), 33-8, plates, 27-28.
21. M.A. Marti-Renom, A. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Sali. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291-325, 2000.

22. John, B. a. A. S., Comparative Protein structure modelling by iterarivealignment, model building and model assessment. *Nucleic acid Res.*, 31: 3982-3992. **2003**.
23. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, 79 (2), 926-35.
24. MacKerell AD, J., Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *Journal of Physical Chemistry B* 102: 3586-3616. **1998**.
25. MacKerell AD, J., . Feig, M., Brooks, C.L., III, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* 25: 1400-1415. **2004**.
26. Saam J, Ivanov I, Walther M, Holzhütter H, and Kuhn H. (2007) Molecular dioxygen enters the active site of 12/15-lipoxygenase via dynamic oxygen access channels. *Proc. Natl. Acad. Sci.*, 104(33), 13319-13324.
27. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
28. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, R. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 2008, 4 (3), 435-447.
29. Lin, J.-H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. Computational drug design accommodating receptor flexibility: The relaxed complex scheme. *J. Am. Chem. Soc.* 2002, 124 (20), 5632-5633.

30. Lin, J.-H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers* 2003, 68 (1), 47-62.
31. Amaro, R. E.; Baron, R.; McCammon, J. A. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput.-Aided Mol. Des.* 2008, 22 (9), 693-705.
32. Raffa, R. B. Harmonic mean relationship between affinity for wild-type receptors and alanine-scan mutants. *J. Theor. Biol.* 2002, 218 (2), 207-214.
33. Zhang, Z. Y.; Thieme-Seffler, A. M.; Maclean, D.; McNamara, D. J.; Dobrusin, E. M.; Sawyer, T. K.; Dixon, J. E. Substrate specificity of the protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. U.S.A.* 1993, 90 (10), 4446-4450.
34. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* 1998, 19 (14), 1639-1662.
35. Freedman, H.; Truong, T. N., A coupled RISM/MD or MC simulation methodology for solvation free energies. *Chemical Physics Letters* **2003**, 381 (3,4), 362-367.
36. Freedman, H.; Truong, T. N., Coupled reference interaction site model/simulation approach for thermochemistry of solvation: Theory and prospects. *Journal of Chemical Physics* **2004**, 121 (5), 2187-2198.
37. Kinoshita, M.; Okamoto, Y.; Hirata, F., First-Principle Determination of Peptide Conformations in Solvents: Combination of Monte Carlo Simulated Annealing and RISM Theory. *Journal of the American Chemical Society* **1998**, 120 (8), 1855-1863.
38. Kinoshita, M.; Okamoto, Y.; Hirata, F., Analysis on conformational stability of C-peptide of ribonuclease A in water using the reference interaction site model theory and Monte Carlo simulated annealing. *Journal of Chemical Physics* **1999**, 110 (8), 4090-4100.
39. Smith, P. E.; Pettitt, B. M.; Karplus, M., Stochastic dynamics simulations of the alanine dipeptide using a solvent-modified potential energy surface. *Journal of Physical Chemistry* **1993**, 97 (26), 6907-13.
40. Pettitt, B. M.; Karplus, M.; Rossky, P. J., Integral equation model for aqueous solvation of polyatomic solutes: application to the determination of the free energy surface for the internal motion of biomolecules. *Journal of Physical Chemistry* **1986**, 90 (23), 6335-45.

41. Freedman, H.; Truong, T. N., A Coupled Reference Interaction Site Model/Molecular Dynamics Study of the Potential of Mean Force Curve of the SN2 Cl<sup>-</sup>+ CH<sub>3</sub>Cl Reaction in Water. *Journal of Physical Chemistry B* **2005**, *109* (10), 4726-4730.
42. Freedman, H.; Truong, T. N., An application of coupled reference interaction site model/molecular dynamics to the conformational analysis of the alanine dipeptide. *Journal of Chemical Physics* **2004**, *121* (24), 12447-12456.
43. Freedman, H.; Nguyen, H. N.; Truong, T. N., A Study of the Tautomeric Equilibria of 2-Hydroxypyridine/2-Oxopyridine and of Cytosine in Water Using the Coupled Reference Interaction Site Model(RISM)/Molecular Dynamics (MD) Approach. *Journal of Physical Chemistry B* **2004**, *108* (49), 19043-19048.
44. Chandler, D.; Singh, Y.; Richardson, D., *J. Chem. Phys.* **1984**, *81*, 1975.
45. Ten-no, S., Free energy of solvation for the reference interaction site model: Critical comparison of expressions. *J. Chem. Phys.* **2001**, *115* (8), 3724-31.
46. Ten-no, S.; Iwata, S., On the connection between the reference interaction site model integral equation theory and the partial wave expansion of the molecular Ornstein-Zernike equation. *J. Chem. Phys.* **1999**, *111* (11), 4865-8.
47. Sato, K.; Chuman, H.; Ten-no, S., Comparative Study on Solvation Free Energy Expressions in Reference Interaction Site Model Integral Equation Theory. *Journal of Physical Chemistry B* **2005**, *109* (36), 17290-17295.
48. Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A., Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. *Journal of Physical Chemistry B* **2007**, *111* (9), 2242-2254.
49. Wan, S.; Stote, R. H.; Karplus, M., Calculation of the aqueous solvation energy and entropy, as well as free energy, of simple polar solutes. *Journal of Chemical Physics* **2004**, *121* (19), 9539-9548.
50. Shirts, M. R., Rande, V. S., *Journal of Chemical Physics* **2005**, *122*, 134508.
51. Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A., Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *Journal of the American Chemical Society* **1998**, *120* (37), 9401-9409.
52. 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., III, Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Accounts of Chemical Research* **2000**, *33* (12), 889-897.

53. Orozco, M.; Luque, F. J., Theoretical Methods for the Description of the Solvent Effect in Biomolecular Systems. *Chemical Reviews (Washington, D. C.)* **2000**, *100* (11), 4187-4225.
54. Gouda, H.; Kuntz, I. D.; Case, D. A.; Kollman, P. A., Free energy calculations for theophylline binding to an RNA aptamer: comparison of MM-PBSA and thermodynamic integration methods. *Biopolymers* **2003**, *68* (1), 16-34.
55. Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B., High-resolution molecular discrimination by RNA. *Science (Washington, DC, United States)* **1994**, *263* (5152), 1425-9.
56. Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J. P.; Pardi, A., Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nature structural biology* **1997**, *4* (8), 644-9.
57. de Jong, M. D.; Thanh, T. T.; Khanh, T. H.; Hien, V. M.; Smith, G. D.; Nguyen, V. C.; Cam, B. V.; Qui, P. T.; Ha, D. Q.; Guan, Y.; Peiris, J. S. M.; Phil, D.; Hien, T. T.; Farrar, J., Oseltamivir resistance during treatment of influenza A (H5N1) infection. *N. Engl. J. Med.* **2005**, *353* (25), 2667-2672.
58. Chutinimitkul, S.; Suwannakarn, K.; Chieochansin, T.; Mai, L. Q.; Damrongwatanapokin, S.; Chaisingh, A.; Amonsin, A.; Landt, O.; Songserm, T.; Theamboonlers, A.; Poovorawan, Y., H5N1 Oseltamivir-resistance detection by real-time PCR using two high sensitivity labeled TaqMan probes. *J. Virol. Methods* **2007**, *139* (1), 44-49.
59. *Oseltamivir-resistant 2009 pandemic influenza A (H1N1) virus infection in two summer campers receiving prophylaxis--North Carolina, 2009*; Centers for Disease Control and Prevention (CDC): United States, 2009; pp 969-72.
60. *Oseltamivir-resistant novel influenza A (H1N1) virus infection in two immunosuppressed patients - Seattle, Washington, 2009*; Centers for Disease Control and Prevention (CDC): United States, 2009; pp 893-6.
61. Guo, L.; Garten, R. J.; Foust, A. S.; Sessions, W. M.; Okomo-Adhiambo, M.; Gubareva, L. V.; Klimov, A. I.; Xu, X., Rapid identification of oseltamivir-resistant influenza A(H1N1) viruses with H274Y mutation by RT-PCR/restriction fragment length polymorphism assay. *Antiviral Res.* **2009**, *82* (1), 29-33.
62. Pandemic (H1N1) 2009 briefing note 1 Viruses resistant to oseltamivir (Tamiflu) identified. *MMWR Morb Mortal Wkly Rep* **2009**, *84*, 299-399.
63. Collins, P. J.; Haire, L. F.; Lin, Y. P.; Liu, J.; Russell, R. J.; Walker, P. A.; Skehel, J. J.; Martin, S. R.; Hay, A. J.; Gamblin, S. J., Crystal structures of oseltamivir-

resistant influenza virus neuraminidase mutants. *Nature (London, U. K.)* **2008**, *453* (7199), 1258-1261.

64. Chachra R, Rizzo RC (2008) Origins of resistance conferred by the R292K neuraminidase mutation via molecular dynamics and free energy calculations. *J Chem Theor Comp* 4: 1526-1540.

65. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable molecular dynamics with NAMD. *J Comp Chem* 26: 1781-1802.

66. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *Journal of Physical Chemistry B* **1998**, *102* (18), 3586-3616.

67. Boresch, S.; Karplus, M., The Role of Bonded Terms in Free Energy Simulations. 2. Calculation of Their Influence on Free Energy Differences of Solvation. *Journal of Physical Chemistry A* **1999**, *103* (1), 119-136.

68. Hardy DJ, Stone JE, Schulten K (2009) Multilevel summation of electrostatic potentials using graphics processing units. *J Paral Comp* 35: 164-177.

69. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98: 10037-10041.

70. Malaisree M, Rungrotmongkol T, Nunthaboot N, Aruksakunwong O, Intharathep P, et al. (2008). Source of oseltamivir resistance in avian influenza H5N1 virus with the h274y mutation. *Amino Acid* 37: 725-732.

71. Wang Nick, X.; Zheng Jie, J., Computational studies of H5N1 influenza virus resistance to oseltamivir. *Protein Sci* **2009**, *18* (4), 707-15.

72. Park J, Jo W (2009) Infiltration of water molecules into the oseltamivir-binding site of H274Y neuraminidase mutant causes resistance to oseltamivir. *J Chem Inf Model* 29: 2735-2741.

73. Ripoll DR, Faerman CH, Axelsen PH, Silman I, Sussman JL (1993) An electrostatic mechanism for substrate guidance down the aromatic gorge of acetylcholinesterase. *Proc Natl Acad Sci USA* 90: 5138-5132.

74. Wang Y, Tajkhorshid E (2008) Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc Natl Acad Sci USA* 105: 9598-9603.

75. Yang L, Zou J, Xie H, Li L, Wei Y, et al. (2009) Steered molecular dynamics simulations reveal the likelier dissociation pathway of imatinib from its targeting kinases c-kit and abl. *PLoS One* 4:e8470.
76. Liu X, Xu Y, Wang X, Barrantes FJ, Jiang H (2008) Unbinding of nicotine from the acetylcholine binding protein: steered molecular dynamics simulations. *J Phys Chem B* 112: 408793.
77. Shen J, Li W, Liu G, Tang Y, Jiang H (2009) Computational insights into the mechanism of ligand unbinding and selectivity of estrogen receptors. *J Phys Chem B* 113: 10436-44.
78. Evans E, Ritchie K (1997) Dynamic strength of molecular adhesion bonds. *Biophys J* 72: 1541-1555.
79. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K (1997) Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys J* 72: 1568-1581.
80. Amaro, R. E. M., David D. L.; Cheng, Lily S.; Lindstrom, William M., Jr.; Olson, Arthur J.; Lin, Jung-Hsin; Li, Wilfred W.; McCammon, J. Andrew. , Remarkable Loop Flexibility in Avian Influenza N1 and Its Implications for Antiviral Drug Design. *Journal of the American Chemical Society* **2007**, 129 (25), 7764-7765.
81. Yen H, Hoffmann E, Taylor G, Scholtissek C, Monto AS, et al. (2006) Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. *J Virol* 80: 87878795.
82. Sung JC, Wynsberghe AWV, Amaro RE, Li WW, McCammon JA (2010) Role of secondary sialic acid binding sites in influenza N1 neuraminidase. *J Am Chem Soc* 132: 2883-2885.
83. Pratt, L. M.; Le, L. T.; Truong, T. N., A Computational Study of Mixed Aggregates of Chloromethylithium with Lithium Dialkylamides. *The Journal of Organic Chemistry* **2005**, 70 (21), 8298-8302.
84. Borges, C. R.; Taccogno, J.; Crouch, D. J.; Le, L.; Truong, T. N., Structure and mechanism of formation of an important ion in doping control. *International Journal of Mass Spectrometry* **2005**, 247 (1-3), 48-54.
85. Ly Le, Valeria. Molinero., Liquid-liquid immiscibility in supercooled aqueous solutions driven by water's polyamorphism. *Short talk section, the ESF-FWF Conference "Water Interfaces in Physics, Chemistry and Biology: A Multi-Disciplinary Approach", Austria* **2007**.