



**THESIS APPROVAL**  
**GRADUATE SCHOOL, KASETSART UNIVERSITY**

Master of Science (Chemistry)

DEGREE

Chemistry

FIELD

Chemistry

DEPARTMENT

**TITLE:** Effect of Antigenic Site Mutation on Cell Receptor Binding of Influenza A virus (H5N1) Hemagglutinin

**NAME:** Miss Krongkaew Navakul

**THIS THESIS HAS BEEN ACCEPTED BY**

THESIS ADVISOR

( Assistant Professor Chak Sangma, Ph.D. )

THESIS CO-ADVISOR

( Assistant Professor Kiattawee Choowongkomon, Ph.D. )

THESIS CO-ADVISOR

( Mr. Songwut Suramitr, Ph.D. )

DEPARTMENT HEAD

( Assistant Professor Noojaree Prasitpan, Ph.D. )

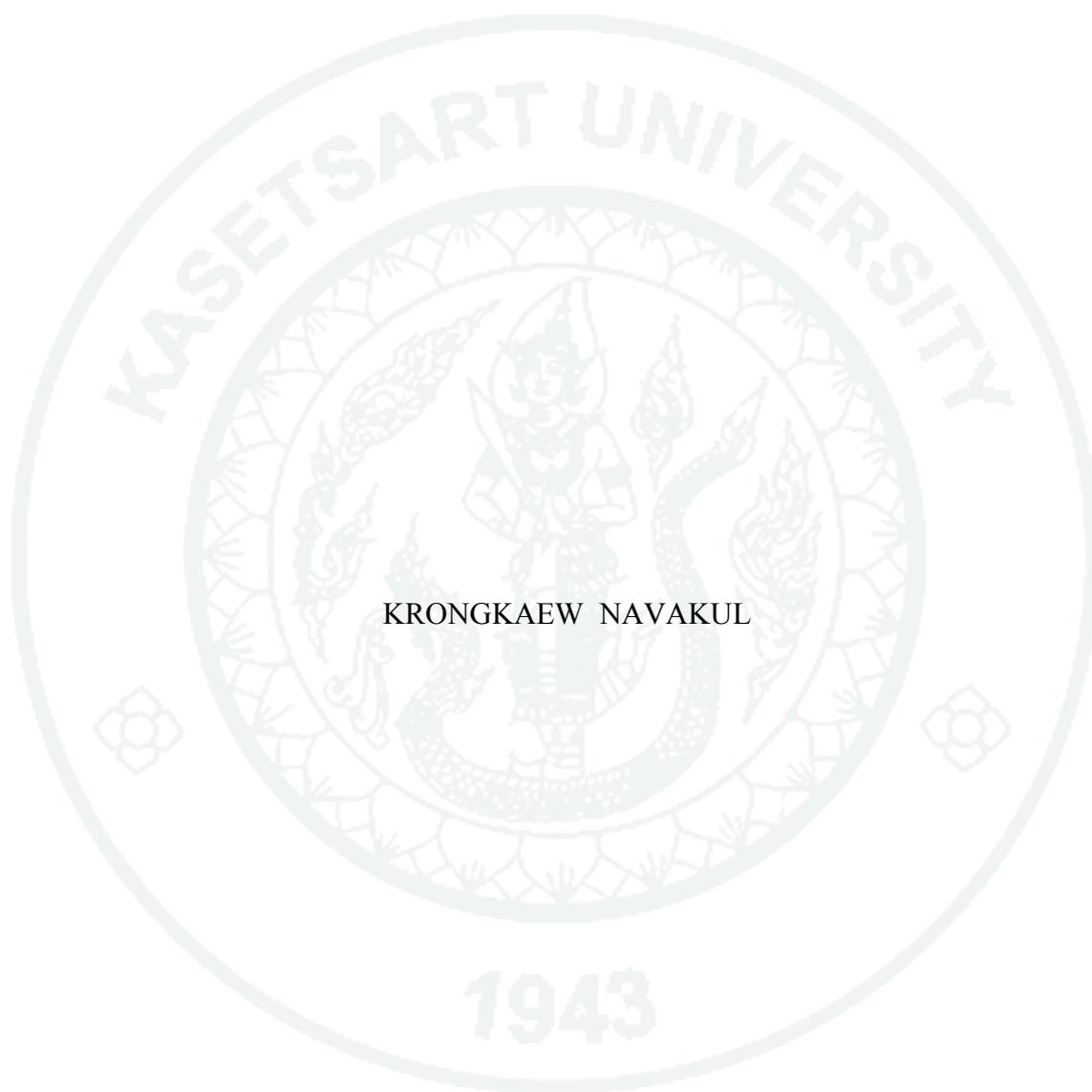
**APPROVED BY THE GRADUATE SCHOOL ON** \_\_\_\_\_

DEAN

( Associate Professor Gunjana Theeragool, D.Agr. )

THESIS

EFFECT OF ANTIGENIC SITE MUTATION ON CELL RECEPTOR  
BINDING OF INFLUENZA A VIRUS (H5N1) HEMAGGLUTININ



KRONGKAEW NAVAKUL

A Thesis Submitted in Partial Fulfillment of  
the Requirements for the Degree of  
Master of Science (Chemistry)  
Graduate School, Kasetsart University  
2010

Krongkaew Navakul 2010: Effect of Antigenic Site Mutation on Cell Receptor Binding of Influenza A virus (H5N1) Hemagglutinin. Master of Science (Chemistry), Major Field: Chemistry, Department of Chemistry. Thesis Advisor: Assistant Professor Chak Sangma, Ph.D. 120 pages.

Hemagglutinin (HA) is an antigenic glycoprotein found on the surface of avian influenza, responsible for binding the virus to sialosaccharides on host cell surfaces. It facilitates the release of progeny viruses from infected cells. Previous study revealed that mutations along the protein loop 130 (H5 numbering) are important for host cell selectivity of HA (Auewarakul et al.,2007). In other HA subtype, beyond residue 137, this loop is also an antigenic site. Interestingly, in swine influenza A virus (H9/Swine), this antigenic site is absent and, coincidentally human and avian receptor can bind with the H9 HA. This work studied whether or not this loop deletion or mutations could cause the change in host type selectivity in avian influenza. Molecular dynamics simulations were used to build H5 HA model from the sequences where amino acids were removed or changed between residues 137 and 141. The result showed more selectivity toward human cell receptor if certain mutation of this loop occurs.

---

Student's signature

---

Thesis Advisor's signature

/ /

## ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Assistant Professor Chak Sangma for his kindness, valuable suggestions, encouragement, continuous support and constant inspiration throughout the course of my degree. I would sincerely like to thank Assistant Professor Kiattawee Choowongkomon and Assistant Professor Arunee Thitithanyanont for kindly providing valuable suggestions and helpful guidance. I also wish to express my appreciation to Drs. Songwut Suramitr and Sissades Tongsimma for valuable suggestions and comments to make my thesis complete.

I am deeply appreciated to Miss Nipa Jongkon and Miss Daungmanee Chuakheaw who took the time to help and suggest in my thesis. Furthermore, I am also thankful to Mr. Kun Silprasit and Prasit naake who I spent time discussing in my experimental work.

This work was supported in part by the Thailand Research Fund, the National Center of Excellence in Petroleum, Petrochemicals Technology, Thailand Graduate Institute of Science and Technology (TGIST) and National Center for Genetic Engineering and Biotechnology (BIOTEC).

Special thanks to the Laboratory for Computational and Applied Chemistry (LCAC) and Cheminformatics Research Unit (CRU) at the Kasetsart University. I acknowledge for the use of computational resource and software facilities. I would like to thank Department of Biochemistry for providing experimental research facilities.

I appreciate my colleagues at LCAC and Biochemistry Department for their helpful assistance and supporting during my study. Finally, I would like to thank my parents, my sister, my dream and my bosom friends, who encouraged me to achieve success in my education.

Krongkaew Navakul

February, 2010.

**TABLE OF CONTENTS**

	<b>Page</b>
TABLE OF CONTENTS	i
LIST OF TABLES	ii
LIST OF FIGURES	iv
LIST OF ABBREVIATIONS	vii
INTRODUCTION	1
OBJECTIVES	21
LITERATURE REVIEW	22
MATERIALS AND METHODS	29
Materials	29
Methods	34
RESULTS AND DISCUSSION	52
CONCLUSIONS	74
LITERATURE CITED	75
APPENDICES	
Appendix A	98
Appendix B Poster Presentation to Conferences	114
CURRICULUM VITAE	119

**LIST OF TABLES**

<b>Table</b>		<b>Page</b>
1	Gene assignments for influenza A virus segments	3
2	Pandemics of influenza A and other important influenza events in the last one hundred years	4
3	Hemagglutinin amino acid substitutions involved in receptor specificity	7
4	Cycling Parameters for the QuikChange <sup>®</sup> Site-Directed Mutagenesis Method	30
5	Sequences of oligonucleotide primer designed from nucleotide sequence of HA mutant gene	51
6	The relative torsion angle monitoring and binding free energy of avian (SA $\alpha$ 2,3-Gal) and human (SA $\alpha$ 2,6-Gal) cell receptor analogs complexes with A/Duck/Singapore/3/97:H5, mutation on loop 130	56

## LIST OF FIGURES

Figure		Page
1	The structure of the influenza A virus	2
2	The principle of polymerase chain reaction (PCR)	16
3	Electrophoresis of the PCR fragments on 1 % agarose gel	18
4	Mechanism of the cytokine storm evoked by influenza virus	27
5	Schematic diagram of molecular dynamics simulations method	46
6	The receptor bound conformations were investigated by torsion angle ( $\Phi$ ). The torsion angles were defined that the cis and trans conformation.	53
7	The amount of time that each receptor analog spent having a particular $\Phi$ angle for SA $\alpha$ 2,6-Gal (red line) and SA $\alpha$ 2,3-Gal (black line) in the binding sites of the hemagglutinin	54
8	The swine influenza A virus (H9) of A/Swine/Hong Kong/98:H9 from Protein Data Bank [PDB] accession number 1jsi (a). The avian influenza A virus (H5) of A/Duck/Singapore/3/97:H5 from Protein Data Bank [PDB] accession number 1jso (b).	55
9	The average structural comparison of A/Duck/Singapore/3/97: H5 (green ribbon) and mutated of deletion loop 130 (137-140 Res., H5 no.) (blue ribbon) binding with SA $\alpha$ 2,6-Gal-linked saccharides, LSTc (green stick), and SA $\alpha$ 2,6-Gal (blue stick). The superimposition of LSTc and SA $\alpha$ 2,6-Gal around the receptor binding domain generated through computation	57

## LIST OF FIGURES (Continued)

Figure		Page
10	The amount of time that each receptor analog spent having a particular $\Phi$ angle for SA- $\alpha$ -2, 3-Gal in the binding sites of the template H5 (A/Duck/Singapore/3/97), the Q222L mutant, the Q222L/G224S mutant and the G224S mutant	58
11	The amount of time that each receptor analog spent having a particular $\Phi$ angle for SA- $\alpha$ -2, 6-Gal in the binding sites of the template H5 (A/Duck/Singapore/3/97), the Q222L mutant, the Q222L/G224S mutant and the G224S mutant	61
12	The superimposition of SA- $\alpha$ -2, 3-Gal and SA- $\alpha$ -2, 6-Gal around the receptor binding domain. The average structural comparison of the double mutated Q222L/G224S (H5 no.) binding with SA- $\alpha$ -2, 3-Gal and the double mutated Q222L/G224S binding with SA- $\alpha$ -2, 6-Gal as the template H5 (A/Duck/Singapore/3/97)	61
13	The amount of time that each receptor analog spent having a particular $\Phi$ angle for SA- $\alpha$ -2, 3-Gal in the binding sites of the template H1 (A/Puerto/Rico/8/34), the Q222L (H5 no.) mutant, the Q222L/G224S mutant and the G224S mutant	62
14	The superimposition of LSTa and SA- $\alpha$ -2, 3-Gal around the receptor binding domain. The average structural comparison of A/Puerto/Rico/8/34:H1 and mutated Q222L/G224S (H5 no.)	63
15	The amount of time that each receptor analog spent having a particular $\Phi$ angle for SA- $\alpha$ -2, 6-Gal in the binding sites of the template H1 (A/Puerto/Rico/8/34), the Q222L (H5 no.) mutant, the Q222L/G224S mutant and the G224S mutant	63

**LIST OF FIGURES (Continued)**

<b>Figure</b>		<b>Page</b>
16	The superimposition of SA- $\alpha$ -2, 3-Gal and SA- $\alpha$ -2, 6-Gal around the receptor binding domain. The average structural comparison of A/Puerto/Rico/8/34:H1 and double mutated Q222L/G224S (H5 no.) binding with SA- $\alpha$ -2, 3-Gal and SA- $\alpha$ -2, 6-Gal	64
17	Full length amplification of mutant HA gene deleted loop 130 by PCR RNA from A/Thailand/1(KAN-1)2004: H5, The PCR reactions were subjected to electrophoresis on a 1% agarose gel and stained with ethidium bromide	65

## LIST OF ABBREVIATIONS

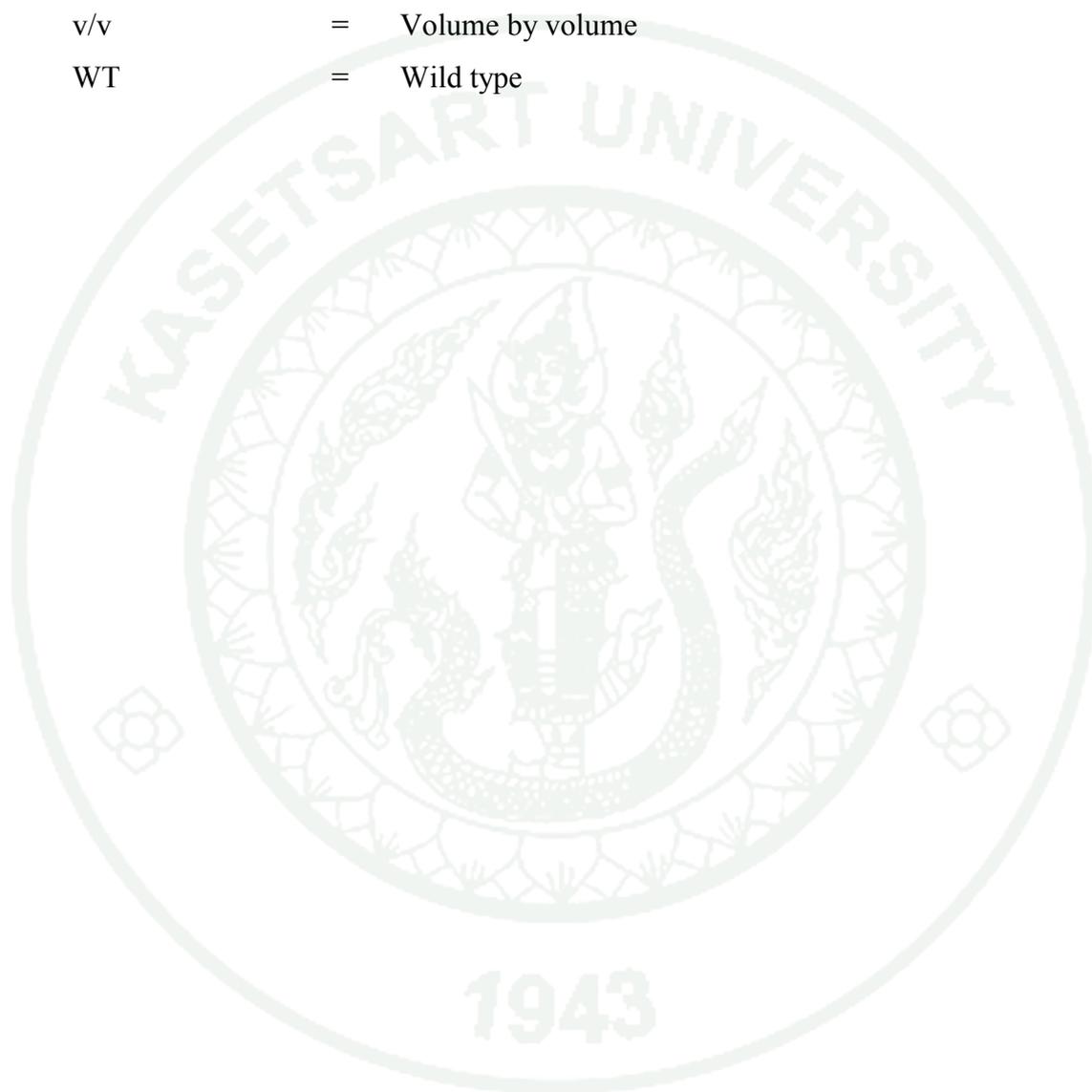
Å	=	Angstrom
Ala (A)	=	Alanine
Arg (R)	=	Arginine
Asn (N)	=	Asparagine
Asp (D)	=	Aspartic acid
a.u.	=	Arbitrary units
AZT	=	Azidothymidine
BSA	=	Bovine serum albumin
bp	=	base pair(s)
°C	=	Degree Celsius
Cys (C)	=	Cysteine
Da	=	Dalton
DNA	=	Deoxyribonucleic acid
dNTP	=	Deoxynucleoside triphosphates
dsDNA	=	Double stranded DNA
<i>E. Coli</i>	=	<i>Escherichia coli</i>
EDTA	=	Ethylenediaminetetraacetic acid disodium salt
FDA	=	The U.S. Food and Drug Administration
Gln (Q)	=	Glutamine
Gly (G)	=	Glycine
HA	=	Hemagglutinin
His (H)	=	Histidine
IPTG	=	Isopropyl β-D-1-thiogalactopyranoside
kb	=	Kilo base pairs
kDA	=	Kilo Dalton
LB	=	Luria-Bertani broth
Leu (L)	=	Leucine
Lys (K)	=	Lysine

**LIST OF ABBREVIATIONS (Continued)**

MD	=	Molecular dynamics simulations
mM	=	Millimolar
mRNA	=	Messenger RNA
mg	=	Milligram
NA	=	Neuraminidase
ng	=	Nanogram
NP	=	Nucleoprotein
NS	=	Non-structural protein
PCR	=	Polymerase Chain Reaction
PDB	=	Protein Data Bank
Pfu	=	Plaque Forming Units
Phe (F)	=	Phenylalanine
Pro (P)	=	Proline
RMSD	=	Root mean square deviation
RNA	=	Ribonucleic acid
Res.	=	Residue
SA $\alpha$ 2, 3-Gal	=	Sialic acid alpha2, 3-galactose
SA $\alpha$ 2, 6-Gal	=	Sialic acid alpha2, 6-galactose
Ser (S)	=	Serine
SOC	=	Super optimal culture
ssRNA	=	Single stranded RNA
Trp (W)	=	Tryptophan
Tyr (Y)	=	Tyrosine
U	=	Unit
$\mu$ g	=	Microgram
$\mu$ l	=	Microliter

**LIST OF ABBREVIATIONS (Continued)**

$\mu\text{M}$	=	Micromolar
Val (V)	=	Valine
v/v	=	Volume by volume
WT	=	Wild type

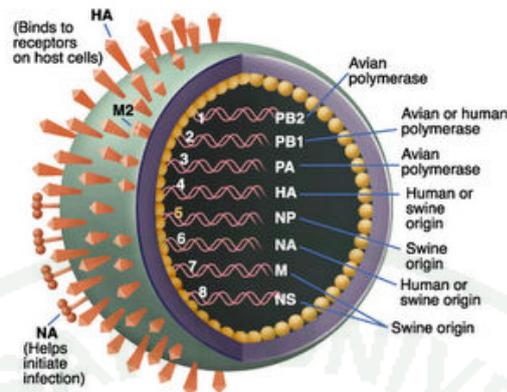


# **EFFECT OF ANTIGENIC SITE MUTATION ON CELL RECEPTOR BINDING OF INFLUENZA A VIRUS (H5N1) HEMAGGLUTININ**

## **INTRODUCTION**

Influenza A virus subtype H5N1, also known as A (H5N1), which can cause illness in humans and many other animal species. A bird-adapted strain of H5N1, called HPAI A (H5N1) for “highly pathogenic avian influenza virus of type A of subtype H5N1”, is the causative agent of H5N1 flu, commonly known as “avian influenza” or “bird flu”. Low pathogenic avian influenza virus (LPAI) or HPAI is determined according to laboratory tests of pathogenicity. The Office International des Epizootics (OIE) establishes that an avian influenza virus with the subtype of H5 or H7 with an intravenous pathogenicity index of greater than 1.2 in 6 week-old chicks is an HPAI virus. Alternatively, an avian influenza virus that causes 75% mortality in 4-8 week-old is also considered to be an HPAI virus. LPAI viruses are those of the H5 or H7 subtype that are not HPAI. Nevertheless, LPAI can mutate into HPAI viruses. Note that the OIE calls HPAI, HPNAI and LPAI, LPNAI. The added N stands for “notifiable” meaning that if these avian influenza viruses are detected, they should be reported to the OIE as well as corresponding local animal and human health authorities.

HA (H5) coded for hemagglutinin, an antigenic glycoprotein found on the surface of the influenza viruses, is responsible for binding the virus to the infected cell. NA (N1) coded for neuraminidase is an antigenic glycosylated enzyme found on the surface of the influenza viruses. It facilitates the release of progeny viruses from infected cells. The hemagglutinin (HA) and neuraminidase (NA) RNA strands specify the structure of proteins that are most medically relevant as targets for antiviral drugs and vaccines.



**Figure 1** The structure of the influenza A virus.

**Source:** Raymond *et al.* (2005)

The viruses are classified in the family Orthomyxoviridae, genus Influenza virus A (Cox *et al.*, 2000). Orthomyxoviridae viruses are enveloped, segmented, single-stranded negative sense RNA virus. The eight segments of RNA have the size between 890 bases to 2,341 bases. The virion is approximately 100 nm in diameter. These segments have been sequenced and the viral protein that each encodes have been determined by genetic mean (Table 1). Influenza virus has eight RNA segments that encode 10 different proteins including PB2, PB1, PA, HA, NP, NA, M1, M2, NS1, and NS2 proteins. Three of these proteins are surface glycoproteins such as HA, NA, and M2, that are embedded in a viral envelope and elicit an antibodies response to prevent or decrease infection for the host. The virion had another six internal including PB2, PB1, PA, NP, M1, NS1, and NS2 proteins, especially PB2, PB1, PA, and NP that form the polymerase complex necessary for the viral genome transcription. The M1 protein is associated with the viral RNA and the NS2 protein is also present in small quantities. Only the NS1 protein is thought not to be packaged in the virion (Figure1).

**Table 1** Gene assignments for influenza A virus segments

Segment	Polypeptide	Size (bp)	Name : Function
1	PB2	2341	Transcriptase : cap binding
2	PB1	2341	Transcriptase : elongation
3	PA	2233	Transcriptase : protease activity
4	HA	1778	Hemagglutinin : viral attachment
5	NP	1565	Nucleoprotein : ribonucleoprotein
6	NA	1413	Neuraminidase : viral releasing
7	M1,M2	1027	Matrix proteins : major component of viral envelope and ion channel
8	NS1,NS2	890	Nonstructural proteins : effects on cellular RNA transport, splicing, translation

### Origin of H5N1 virus

Influenza A viruses are distributed throughout the world, causing symptomatic and asymptomatic infections in many vertebrate species, including a large variety of birds such as ducks, geese, terns, shearwaters, gulls, chickens, quail, turkeys, pheasants, starlings and budgerigars, as well as in swine, horses, seals whales, gibbons, baboons, chimpanzees and humans (Hinshaw *et al.*, 1980). Phylogenetic studies of influenza A viruses reveal that aquatic birds are the source of all influenza viruses in other species (Webster *et al.*, 1992).

Over the past 150 years at least four pandemics of influenza occurred at irregular intervals, including three in the 20<sup>th</sup> century (Table 2). These have caused high attack rates in all susceptible age groups, with high morbidity and mortality. The most lethal influenza pandemic in modern history was the H1N1 Spanish flu, which killed approximately 100 million people around the world between 1918 and 1919. The origin of the 1918 pandemic remains an enigma, but it is now clear that the virus

had features of an avian virus (Reid *et al.*, 2003), and it appears that an intermediate host, such as swine was involved (Taubenberger *et al.*, 2005). Swine are known to be susceptible to both avian and human viruses, and could have served as hosts for additional drift resulting in the accumulation of changes from the original avian strain (Taubenberger *et al.*, 2005). The pandemics of "Asian flu" (H2N2) in 1957 and the "Hong Kong flu" (H3N2) in 1968 caused an estimated 1 to 3 million deaths (Arruda *et al.*, 2006).

**Table 2** Pandemics of influenza A and other important influenza events in the last one hundred years.

Year of appearance	Time of circulation (year)	Virus subtype	Common designation	Estimated deaths
1918	39	H1N1	Spanish	100 million
1957	11	H2H2	Asian	86,000 (USA)
1968	Present	H3H2	Hong Kong	34,000 (USA)
1997	Present	H5H1	Chicken or bird	148 deaths from 2003 until October 2006

**Source:** Arruda *et al.* (2006)

A highly pathogenic H5N1 avian influenza virus was first isolated in 1996, from a goose in China and in 1997 it caused death in poultry in Asia. In general, avian influenza A viruses do not cause disease in humans and prior to 1997 only two cases of natural human infections by animal influenza virus had been documented (Kurtz *et al.*, 1996). However, since 1997, avian influenza virus H5N1 has caused many cases of influenza in humans. In January 1998, Subbarao and collaborators (Subbarao *et al.*, 1998) characterized an H5N1 virus isolated from a child who died from a fatal respiratory illness on May 21, 1997 in Hong Kong. It was shown that genes encoding all internal and nonstructural proteins were closely related to known genes of avian origin and that this virus retained a preference for binding to SA $\alpha$ 2,3-Gal-terminated sialylglycoconjugates receptors, a feature typical of avian influenza viruses (Matrosovich *et al.*, 1999). It has also been demonstrated that the HA of H5N1 has

basic amino acids adjacent to the cleavage site, allowing it to be cleaved by proteases other than trypsin-like ones (Maines *et al.*, 2005). This augmented susceptibility of HA to cleavage by proteases widely distributed in tissues and organs helps to explain the broad tissue tropism of this virus and, consequently, it enhanced pathogenicity. In fact, H5N1 virus in mammals is not restricted to the lungs, but extends to other organs, including the brain (Maines *et al.*, 2005).

These H5N1 influenza viruses have continued to reassort and undergo antigenic drift, as well as extend their host range, what is reflected by the detections of infections in swine and feline (Butler *et al.*, 2006), raising the possibility of a possible participation of these animals in the spread of avian influenza virus. In 2004, 8 (0.25%) of 3.175 vietnamese pigs tested were positive for H5N1 influenza viruses by serum neutralization assay and Western blot (Chio *et al.*, 2005). Pigs possess receptors (NeuAca2,3Gal and NeuAca2,6Gal) with affinity for both influenza A viruses (Ito *et al.*, 1998), which led to the hypothesis that swine may serve as a "mixing vessel" for the generation of reassortants (Ito *et al.*, 1998).

Felines have also become new targets for H5N1 research. In Germany, in February 2006, 8 of 111 apparently healthy cats coming from Thailand tested positive for antibodies to H5N1 viruses (Butler *et al.*, 2006). This finding has raised the question whether cats can spread avian influenza virus. Rimmelzwaan and collaborators (2006) showed that experimentally infected cats shed virus in sputum and feces (Rimmelzwaan *et al.*, 2006).

While other animals have been found to be susceptible to either natural or experimental infection by influenza avian influenza, aquatic birds, mainly ducks, are the most important reservoirs of the virus throughout the world (Webster *et al.*, 2006). In ducks, avian viruses replicate in the lung and in cells lining the intestinal tract, and are shed in high concentrations in feces (Webster *et al.*, 1978). Water contamination with duck feces is a logical mechanism for the transmission of avian influenza viruses among aquatic birds and from them to domestic avian and mammalian species. The

majority of H5N1 genotypes are avirulent in ducks, indicating that these animals can ensure perpetuation of these viruses (Lipatov *et al.*, 2004).

From 1997 through 2001, the HA of the various H5N1 genotypes remained antigenically homogeneous, but in 2002 marked antigenic drift was reported (Guan *et al.*, 2004) and a strain of the so called Z genotype emerged, which is highly pathogenic for ducks and other aquatic birds (Li K. *et al.*, 2004). In February 2003, two members of a family in Hong Kong died of infection caused by an H5N1 influenza virus strain (Peiris *et al.*, 2004), antigenically and molecularly similar to the one that was highly pathogenic for ducks and chickens (Guan *et al.*, 2004).

Human influenza virus is generally transmitted from person to person by large droplets, small-particle aerosols, and possibly by fomites, with hand contamination and subsequent self-inoculation (Arruda *et al.*, 2006). Recent evidence of limited human-to-human transmission of H5N1 was reported during a large family cluster in Indonesia (Normile *et al.*, 2006). Epidemiologic and genetic sequencing data suggest that a 10-year-old boy contracted the virus from his aunt and then passed it on to his father. In total, six members of the family died of H5N1 infection (Normile *et al.*, 2006).

Virtually all people who became infected with H5N1 influenza virus had been exposed to an infected bird. Plucking and preparing diseased birds; handling fighting cocks; playing with poultry, particularly asymptomatic infected ducks; and consumption of duck's blood or, possibly, undercooked poultry, have all been implicated (Beigel *et al.*, 2005). This relative lack of adaptation of avian H5N1 to be transmitted from person to person has so far prevented the establishment of an overt avian influenza pandemic.

Thus, this can restrict the virus to transmit efficiently from human to human. Nonetheless, for H5N1, there were two studies determined the amino acid mutations which were responsible for the binding of the virus to human type receptor (Table 3). These mutations, Gln182Arg, Asn192Lys, Leu129Val and Arg134Val demonstrate

that avian influenza H5N1 virus could be adapted to the human-type receptor (Ghedini *et al.*, 2005). Besides, the virus isolated from the 1918, 1957 and 1968 pandemics preferentially recognized  $\alpha$ 2,6-linked sialic acids, although their HA were from an avian virus. This indicates that conversion of receptor specificity to  $\alpha$ 2,6 linkage is an alteration that might be necessary for the generation of a virus with pandemic potential.

**Table 3** Hemagglutinin amino acid substitutions involved in receptor specificity.

Strain	Position (H5 No.)	HA specificity	
		$\alpha$ 2,3	$\alpha$ 2,6
H5N1	129 and 134	L and A	V and V
	182	Q	R
	192	N	K
H3N2	222	Q	L
	224	G	S

Addition, Influenza viruses attach to host cells by binding of the hemagglutinin to sialosaccharides on the host cell surface. Human influenza viruses prefer sialic acid (SA) $\alpha$ 2,6-Gal-terminated saccharides receptor, whereas avian influenza viruses prefer those terminating in SA $\alpha$ 2,3-Gal (Baigent *et al.*, 2003). It is believed that this receptor binding property is the major factor preventing the H5N1 virus from efficiently transmitting from person to person and causing a pandemic (Suzuki *et al.*, 2005).

The receptor binding preference of H5N1 viruses can be altered by only a few amino acid substitutions in the hemagglutinin protein. Mutations that change the receptor binding preference from the avian to the human type could potentially enable the virus to transmit efficiently in the human population and cause a catastrophic pandemic (Auewarakul *et al.*, 2007).

A previous study showed that mutations at positions 222 and 224 (H5 numbering) (Q222L and G224S), which are the adaptive mutations for H2 and H3, could reduce the binding affinity to sialic acid  $\alpha$ 2,3-galactose (SA $\alpha$ 2,3-Gal) of an

H5N1 virus isolated in 1997 (Harvey *et al.*, 2004). Human H5N1 isolates from Hong Kong that were isolated in 2003, which contain a mutation at position 223 (S223N) (H5 numbering), were shown to have a reduced binding affinity toward SA $\alpha$ 2,3-Gal and an increased affinity toward SA $\alpha$ 2,6-Gal (Gambaryan *et al.*, 2006). Another previous study it was found that mutations along the loop 130 at the 137 to 141 residues (H5 numbering) were important for host cell selectivity of HA (Auewarakul *et al.*, 2007).

In other HA subtypes, beyond residue 137 this loop is also an antigenic site. Interestingly, in swine influenza A virus (H9), the antigenic site is absent and, coincidentally, H9 HA can bind with the human receptor. In this study we try to study whether or not this deletion can also cause the change in host type selectivity in avian influenza. We used molecular dynamics simulations (MD) to build H5 HA model from the sequences where amino acid were removed or changed between residues 137 to 141. Additionally, we have cloned hemagglutinin mutants by the polymerase chain reaction (PCR) technique, expression and purification proteins in the laboratory for testing the activity and the ability of the protein to bind to a human-type receptor.

## Molecular Dynamics Simulations

Molecular dynamics (MD) is a form of computer simulations in which atoms and molecules are allowed to interact for a period of time by approximations of known physics, giving a view of the motion of the particles. This kind of simulation is frequently used in the study of proteins and biomolecules, as well as in materials science. It is tempting, though not entirely accurate, to describe the technique as a "virtual microscope" with high temporal and spatial resolution. Whereas it is possible to take "still snapshots" of crystal structures and probe features of the motion of molecules through NMR, no conventional experiment allows access to all the time scales of motion with atomic resolution, recent developments in atto-second lasers might give an opportunity in the near future. Richard Feynman once said that "If we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jiggings and wiggings of atoms." MD lets scientists peer into the motion of individual atoms in a way which is not possible in laboratory experiments.

MD is a specialized restraint of molecular modeling and computer simulation based on statistical mechanics; the main justification of the MD method is that statistical ensemble averages are equal to time averages of the system, known as the ergonomics hypothesis. MD has also been termed "statistical mechanics by numbers" and "Laplace's vision of Newtonian mechanics" of predicting the future by animating nature's forces (Schlick *et al.*, 1996 and de Laplace *et al.*, 1820) and allowing insight into molecular motion on an atomic scale. However, long MD simulations are mathematically ill-conditioned, generating cumulative errors in numerical integration that can be minimized with proper selection of algorithms and parameters, but not eliminated entirely. Furthermore, current potential functions are, in many cases, not sufficiently accurate to reproduce the dynamics of molecular systems, so the much more computationally demanding Ab Initio MD method must be used. Nevertheless, MD techniques allow detailed time and space resolution into representative behavior in phase space for carefully selected systems.

Before it became possible to simulate MD with computers, some undertook the hard work of trying it with physical models such as macroscopic spheres. The idea was to arrange them to replicate the properties of a liquid. J.D. Bernal in 1962 said that "... I took a number of rubber balls and stuck them together with rods of a selection of different lengths ranging from 2.75 to 4 inches. I tried to do this in the first place as casually as possible, working in my own office, being interrupted every five minutes or so and not remembering what I had done before the interruption." (Bernal *et al.*, 1964). Fortunately, now computers keep track of bonds during a simulation. Because molecular systems generally consist of a vast number of particles, it is in general impossible to find the properties of such complex systems analytically. When the number of particles interacting is higher than two, the result is chaotic motion (see n-body problem).

MD simulations circumvent the analytical intractability by using numerical methods. It represents an interface between laboratory experiments and theory, and can be understood as a "virtual experiment". MD probes the relationship between molecular structure, movement and function. Molecular dynamics is a multidisciplinary method. Its laws and theories stem from mathematics, physics, and chemistry, and it employs algorithms from computer science and information theory. It was originally conceived within theoretical physics in the late 1950s (Alder *et al.*, 1959) and early 1960s (Rahman *et al.*, 1964), but is applied today mostly in materials science and the modeling of biomolecules.

### Software for MD simulations

AMBER (an acronym for Assisted Model Building with Energy Refinement) is a family of force fields for molecular dynamics of biomolecules originally developed by the late Peter Kollman's group at the University of California, San Francisco. AMBER is also the name for the molecular dynamics software package that simulates these force fields. It is maintained by an active collaboration between David Case at Rutgers University, Tom Cheatham at the University of Utah, Tom

Darden at NIEHS, Ken Merz at Florida, Carlos Simmerling at Stony Brook University, Ray Luo at UC Irvine, and Junmei Wang at Encysive Pharmaceuticals.

The term "AMBER force field" generally refers to the functional form used by the family of AMBER force fields. This form includes a number of parameters; each member of the family of AMBER force fields provides values for these parameters and has its own name.

### Functional form

The functional form of the AMBER force field is (Cornell *et al.*, 1995)

$$\begin{aligned}
 V(r^N) = & \sum_{\text{bonds}} \frac{1}{2} k_b (l - l_0)^2 + \sum_{\text{angles}} k_a (\theta - \theta_0)^2 \\
 & + \sum_{\text{torsions}} \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)] + \sum_{j=1}^{N-1} \sum_{i=j+1}^N \left\{ \epsilon_{i,j} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}
 \end{aligned}
 \tag{1}$$

Note that despite the term force field, this equation defines the potential energy of the system; the force is the derivative of this potential with respect to position.

The meanings of right hand side terms are:

- First term (summing over bonds): represents the energy between covalently bonded atoms. This harmonic (ideal spring) force is a good approximation near the equilibrium bond length, but becomes increasingly poor as atoms separate.
- Second term (summing over angles): represents the energy due to the geometry of electron orbitals involved in covalent bonding.
- Third term (summing over torsions): represents the energy for twisting a bond

due to bond order (e.g. double bonds) and neighboring bonds or lone pairs of electrons. Note that a single bond may have more than one of these terms, such that the total torsional energy is expressed as a Fourier series.

- Fourth term (double summation over  $i$  and  $j$ ): represents the non-bonded energy between all atom pairs, which can be decomposed into van der Waals (first term of summation) and electrostatic (second term of summation) energies.

The form of the van der Waals energy is evinced by the equilibrium distance ( $\sigma$ ) and well depth ( $\epsilon$ ). The factor of 2 ensures that the equilibrium distance is  $\sigma$ . The form of the electrostatic energy used here assumes that the charges due to the protons and electrons in an atom can be represented by a single point charge. (or in the case of parameter sets that employ lone pairs, a small number of point charges.)

### **Parameter sets**

To use the AMBER force field, it is necessary to have values for the parameters of the force field (e.g. force constants, equilibrium bond lengths and angles, charges). A fairly large number of these parameter sets exist, and are described in detail in the AMBER software user manual. Each parameter set has a name, and provides parameters for certain types of molecules.

- Peptide, protein and nucleic acid parameters are provided by parameter sets with names beginning with "ff" and containing a two digit year number, for instance "ff99".
- GAFF (Generalized AMBER force field) provides parameters for small organic molecules to facilitate simulations of drugs and small molecule ligands in conjunction with biomolecules.
- The GLYCAM force fields have been developed by Rob Woods for simulating carbohydrates.

The AMBER software suite provides a set of programs for applying the AMBER force fields to simulations of biomolecules. It is written in Fortran 90 and C

with support for most major Unix-like systems and compilers. Development is conducted by a loose association of mostly academic labs. New versions are generally released in the spring of even numbered years; AMBER 10 was released in April 2008.

### Programs

- LEaP is used for preparing input files for the simulation programs.
  - Antechamber automates the process of parameterizing small organic molecules using GAFF.
  - SANDER (Simulated Annealing with NMR-Derived Energy Restraints) is the central simulation program and provides facilities for energy minimization and MD with a wide variety of options.
  - pmemd is a somewhat more feature-limited reimplementation of sander by Bob Duke. It was designed with parallel processing in mind and has significantly better performance than sander when running on more than 8-16 processors.
  - nmode calculates normal modes.
  - ptraj provides facilities for numerical analysis of simulation results.
- AMBER does not include visualization capabilities; visualization is commonly performed with VMD. A new visualization alternative is Sirius.
- MM-PBSA allows for implicit solvent calculations on snap shots from MD simulations.

### Polymerase Chain Reaction Technique

In molecular biology, the polymerase chain reaction (PCR) is a technique to amplify a single or few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. The method relies on thermal cycling, consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA. Primers (short DNA fragments) containing sequences complementary to the target region along with a DNA polymerase (after which the method is named) are key components to enable

selective and repeated amplification. As PCR progresses, the DNA generated is itself used as a template for replication, setting in motion a chain reaction in which the DNA template is exponentially amplified. PCR can be extensively modified to perform a wide array of genetic manipulations.

Almost all PCR applications employ a heat-stable DNA polymerase, such as Taq polymerase, an enzyme originally isolated from the bacterium *Thermus aquaticus*. This DNA polymerase enzymatically assembles a new DNA strand from DNA building blocks, the nucleotides, by using single-stranded DNA as a template and DNA oligonucleotides (also called DNA primers), which are required for initiation of DNA synthesis. The vast majority of PCR methods use thermal cycling, i.e., alternately heating and cooling the PCR sample to a defined series of temperature steps. These thermal cycling steps are necessary first to physically separate the two strands in a DNA double helix at a high temperature in a process called DNA melting. At a lower temperature, each strand is then used as the template in DNA synthesis by the DNA polymerase to selectively amplify the target DNA. The selectivity of PCR results from the use of primers that are complementary to the DNA region targeted for amplification under specific thermal cycling conditions.

Developed in 1983 by Kary Mullis, PCR is now a common and often indispensable technique used in medical and biological research labs for a variety of applications (Saiki *et al.*, 1985 and 1988). These include DNA cloning for sequencing, DNA-based phylogeny, or functional analysis of genes; the diagnosis of hereditary diseases; the identification of genetic fingerprints (used in forensic sciences and paternity testing); and the detection and diagnosis of infectious diseases. In 1993, Mullis was awarded the Nobel Prize in Chemistry for his work on PCR.

PCR is used to amplify a specific region of a DNA strand (the DNA target). Most PCR methods typically amplify DNA fragments of up to ~10 kilo base pairs (kb), although some techniques allow for amplification of fragments up to 40 kb in size (Cheng *et al.*, 1994). A basic PCR set up requires several components and reagents (Joseph *et al.*, 2001). These components include:

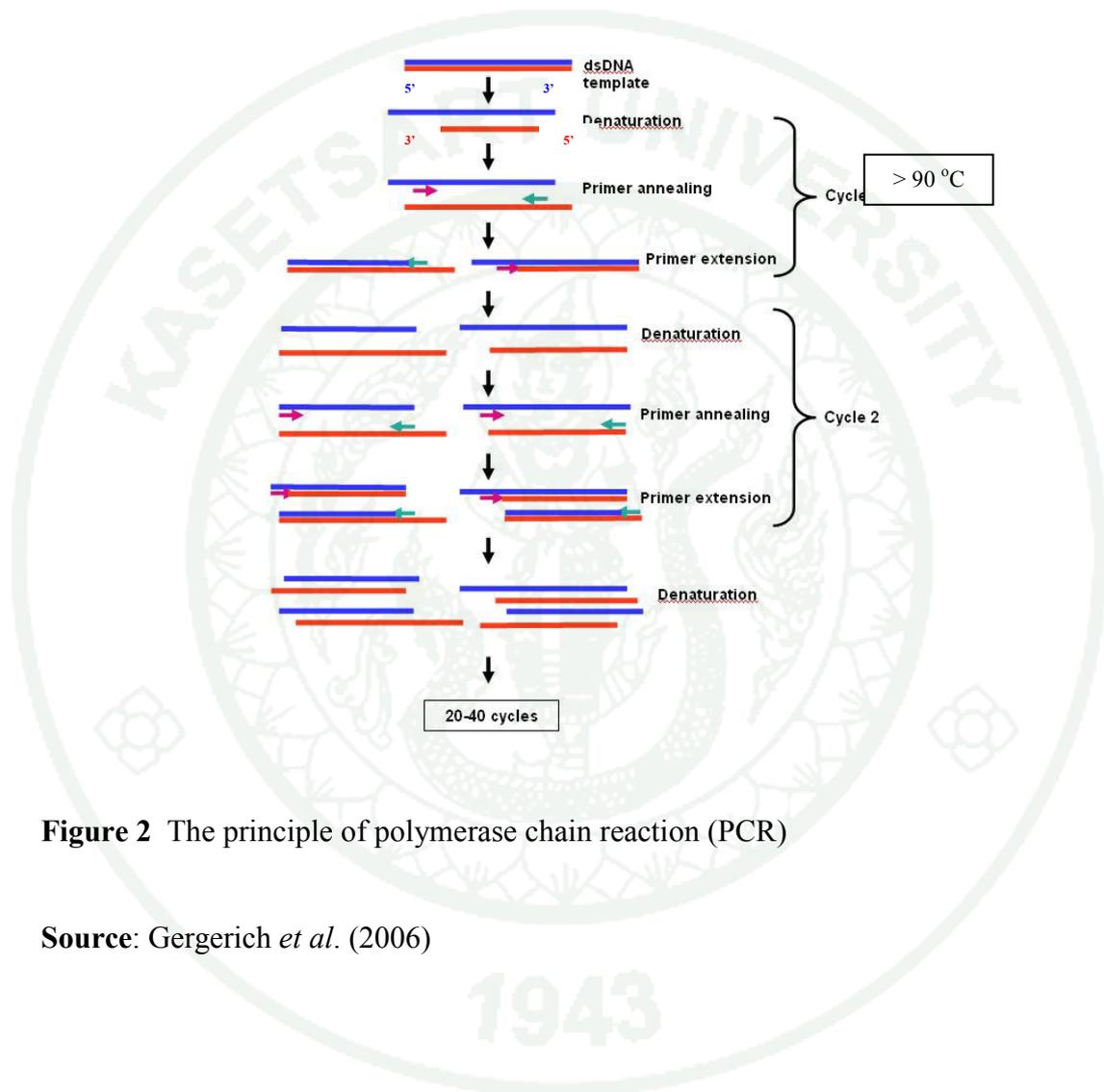
- DNA template that contains the DNA region (target) to be amplified.
- Two primers that are complementary to the 3' (three prime) ends of each of the sense and anti-sense strand of the DNA target.
  - Taq polymerase or another DNA polymerase with a temperature optimum at around 70 °C.
  - Deoxynucleoside triphosphates (dNTPs) the building blocks from which the DNA polymerase synthesizes a new DNA strand.
  - Buffer solution, providing a suitable chemical environment for optimum activity and stability of the DNA polymerase.
  - Divalent cations, magnesium or manganese ions; generally  $Mg^{2+}$  is used, but  $Mn^{2+}$  can be utilized for PCR-mediated DNA mutagenesis, as higher  $Mn^{2+}$  concentration increases the error rate during DNA synthesis (Pavlov *et al.*, 2004).
  - Monovalent cation potassium ions.

The PCR is commonly carried out in a reaction volume of 10-200  $\mu$ l in small reaction tubes (0.2-0.5 ml volumes) in a thermal cycler. The thermal cycler heats and cools the reaction tubes to achieve the temperatures required at each step of the reaction. Many modern thermal cyclers make use of the Peltier effect which permits both heating and cooling of the block holding the PCR tubes simply by reversing the electric current. Thin-walled reaction tubes permit favorable thermal conductivity to allow for rapid thermal equilibration. Most thermal cyclers have heated lids to prevent condensation at the top of the reaction tube. Older thermocyclers lacking a heated lid require a layer of oil on top of the reaction mixture or a ball of wax inside the tube.

### **PCR Procedure**

The PCR usually consists of a series of 20-40 repeated temperature changes called cycles; each cycle typically consists of 2-3 discrete temperature steps. Most commonly PCR is carried out with cycles that have three temperature steps (Figure 2). The cycling is often preceded by a single temperature step (called hold) at a high temperature ( $>90^{\circ}C$ ), and followed by one hold at the end for final product extension

or brief storage. The temperatures used and the length of time they are applied in each cycle depend on a variety of parameters. These include the enzyme used for DNA synthesis, the concentration of divalent ions and dNTPs in the reaction, and the melting temperature ( $T_m$ ) of the primers (Rychlik *et al.*, 1990).



**Figure 2** The principle of polymerase chain reaction (PCR)

Source: Gergerich *et al.* (2006)

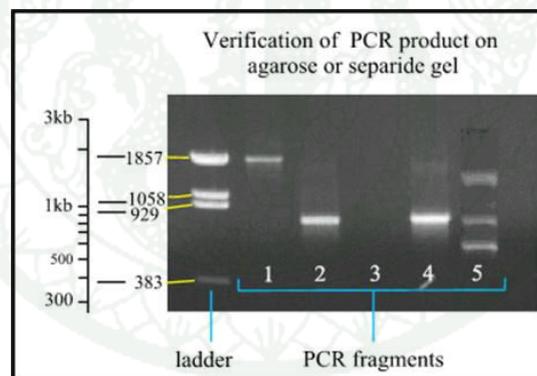
### The step of PCR

1. Initialization step: This step consists of heating the reaction to a temperature of 94-96 °C (or 98 °C if extremely thermostable polymerases are used), which is held for 1-9 minutes. It is only required for DNA polymerases that require heat activation by hot-start PCR (Sharkey *et al.*, 1994).
2. Denaturation step: This step is the first regular cycling event and consists of heating the reaction to 94-98 °C for 20-30 seconds. It causes DNA melting of the DNA template by disrupting the hydrogen bonds between complementary bases, yielding single strands of DNA.
3. Annealing step: The reaction temperature is lowered to 50-65 °C for 20-40 seconds allowing annealing of the primers to the single-stranded DNA template. Typically the annealing temperature is about 3-5 degrees Celsius below the  $T_m$  of the primers used. Stable DNA-DNA hydrogen bonds are only formed when the primer sequence very closely matches the template sequence. The polymerase binds to the primer-template hybrid and begins DNA synthesis.
4. Extension/elongation step: The temperature at this step depends on the DNA polymerase used; Taq polymerase has its optimum activity temperature at 75-80 °C (Chien *et al.*, 1976 and Lawyer *et al.*, 1993) and commonly a temperature of 72 °C is used with this enzyme. At this step the DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand by adding dNTPs that are complementary to the template in 5' to 3' direction, condensing the 5'-phosphate group of the dNTPs with the 3'-hydroxyl group at the end of the nascent (extending) DNA strand. The extension time depends both on the DNA polymerase used and on the length of the DNA fragment to be amplified. As a rule-of-thumb, at its optimum temperature, the DNA polymerase will polymerize a thousand bases per minute. Under optimum conditions, i.e., if there are no limitations due to limiting substrates or reagents, at each extension step, the amount of DNA target is doubled, leading to exponential (geometric) amplification of the specific DNA fragment.

5. Final elongation: This single step is occasionally performed at a temperature of 70-74 °C for 5-15 minutes after the last PCR cycle to ensure that any remaining single-stranded DNA is fully extended.
6. Final hold: This step at 4-15 °C for an indefinite time may be employed for short-term storage of the reaction.

### PCR products analysis

To check whether the PCR generated the anticipated DNA fragment (also sometimes referred to as the amplicon or amplicon), agarose gel electrophoresis is employed for size separation of the PCR products. The sizes of PCR products is determined by comparison with a DNA ladder (a molecular weight marker), which contains DNA fragments of known size, run on the gel alongside the PCR products (Figure 3).



**Figure 3** Electrophoresis of the PCR fragments on 1 % agarose gel.

**Source:** Linzhao *et al.* (2001)

**The PCR process can be divided into three stages:**

1. Exponential amplification: At every cycle, the amount of product is doubled (assuming 100% reaction efficiency). The reaction is very sensitive: only minute quantities of DNA need to be present (Campbell Biology, 7<sup>th</sup> edition).
2. Levelling off stage: The reaction slows as the DNA polymerase loses activity and as consumption of reagents such as dNTPs and primers causes them to become limiting.
3. Plateau: No more product accumulates due to exhaustion of reagents and enzyme.

**PCR optimization**

In practice, PCR can fail for various reasons, in part due to its sensitivity to contamination causing amplification of spurious DNA products. Because of this, a number of techniques and procedures have been developed for optimizing PCR conditions. Contamination with the extraneous DNA is addressed with lab protocols and procedures that separate pre-PCR mixtures from potential DNA contaminants (Joseph *et al.*, 2001). This usually involves spatial separation of PCR-setup areas from areas for analysis or purification of PCR products, use of disposable plastic ware, and thoroughly cleaning the work surface between reaction setups. Primer-design techniques are important in improving PCR product yield and in avoiding the formation of spurious products, and the usage of alternate buffer components or polymerase enzymes can help with amplification of long or otherwise problematic regions of DNA.

### **Application of PCR**

PCR allows isolation of DNA fragments from genomic DNA by selective amplification of a specific region of DNA. This use of PCR augments many methods, such as generating hybridization probes for Southern or northern hybridization and DNA cloning, which require larger amounts of DNA, representing a specific DNA region. PCR supplies these techniques with high amounts of pure DNA, enabling analysis of DNA samples even from very small amounts of starting material.

Other applications of PCR include DNA sequencing to determine unknown PCR-amplified sequences in which one of the amplification primers may be used in Sanger sequencing, isolation of a DNA sequence to expedite recombinant DNA technologies involving the insertion of a DNA sequence into a plasmid or the genetic material of another organism. Bacterial colonies (*E.coli*) can be rapidly screened by PCR for correct DNA vector constructs (Pavlov *et al.*, 2006). PCR may also be used for genetic fingerprinting; a forensic technique used to identify a person or organism by comparing experimental DNAs through different PCR-based methods.

Some PCR “fingerprints” methods have high discriminative power and can be used to identify genetic relationships between individuals, such as parent-child or between siblings, and are used in paternity testing. This technique may also be used to determine evolutionary relationships among organisms.

## OBJECTIVES

1. To predict the binding mechanism to human and avian receptor analogs of influenza A virus (H5N1) hemagglutinin by using molecular dynamics simulations (MD).
2. To study the effect of antigenic site mutation on cell receptor binding of hemagglutinin protein.
3. To study the binding affinity of the receptor analogs in their complexes were also calculated and compared with HA from wild-type and mutant-type of avian influenza A H5N1.
4. To clone H5 HA gene for antigenic site (loop 130) mutation in the experiment.
5. To understand the structural properties of viral resistance in order to improve the vaccine strategies trends in the future.

## LITERATURE REVIEW

Of the three influenza pandemics of the last century, the 1957 (H2N2) and 1968 (H3N2) pandemic viruses were avian-human reassortments in which three and two of the eight avian gene segments, respectively, were reassorted into an already circulating, human-adapted virus (Kawaoka *et al.*, 1989). The origin of the genes of the 1918 influenza virus (H1N1), which killed about 50 million people worldwide (Johnson *et al.*, 2002), is unknown. The extinct pandemic virus from 1918 has recently been reconstructed in the laboratory and was found to be highly virulent in mice and chicken embryos (Taubenberger *et al.* and Tumpey *et al.*, 2005). With continued outbreaks of the H5N1 virus in poultry and wild birds, further human cases are likely, and the potential for the emergence of a human-adapted H5 virus, either by reassortment or mutation, is a threat to public health worldwide.

Clayton *et al.* (1984) proposed two mutations in the receptor-binding site of a human hemagglutinin (HA) at residue 226 and 228 (H3 numbering), receptor binding site, had effect to allowed replication in ducks.

Nicholas *et al.* (1989) reported that mutations at amino acid 226 change the specificity of hemagglutinin for  $\alpha$ 2,6 and  $\alpha$ 2,3-glycosidic-linkages. The NMR line broadening of sialyloligosaccharides suggested that sialic acid is the only component that contacts the protein. Saccharides containing two sialic acid residues appear to have two separate binding modes. Hemagglutinin that has undergone a low pH induced conformational change retains the ability to bind sialic acid.

Studying mutation in hemagglutinin proteins, are one of the most important mechanisms for producing variation in influenza viruses (Robert G. W., 1992), consider position of mutation in protein sequence characterize by associated with glycosylation sites, cleavage site, residues of receptor binding site and antigenic site of protein sequence (Shiuh-Ming *et al.*, 1992). Thus, HA is the main determinant of the host range of virus (Kanta *et al.* 1998).

Kanta *et al.* (1998) isolated an avian H5N1 influenza A virus (A/Hong Kong/156/97) from a tracheal aspirate obtained from a 3-year-old child in Hong Kong with a fatal illness consistent with influenza. They found that the hemagglutinin protein contained multiple basic amino acids adjacent to the cleavage site. A feature characteristic of highly pathogenic avian influenza A viruses.

Couceiro *et al.* (1998) reported that avian, human, and swine upper respiratory tract epithelia preferentially express 2-3 linkages, 2-6 linkages, and both 2-3 and 2-6 linkages, respectively. Not shown, a subset of human respiratory epithelial cells of the lower tract that are ciliated mainly express 2-3 linkages. These expression patterns account for the ability of avian influenza A strains, which preferentially bind via HA to 2-3 sialic acid residues, to infect both birds and pigs, and for pigs to serve as a site of antigenic shift when they are infected with both human and avian influenza A strains. Since the HA of human isolates preferentially binds to 2-6 sialic acid linkages, this also accounts for the usual restriction of infection by avian influenza isolates to birds and pigs but not humans, at least in cases of upper respiratory tract infection. A large body of data shows the multigenic character of influenza virus pathogenicity. All human and avian isolates from live bird markets in Hong Kong 1997 contained multiple basic amino acids at the cleavage site of the HA, a feature known to be associated with high virulence among avian influenza viruses (Claas *et al.*, 1998; Suarez *et al.*, 1998).

Suzuki *et al.* (2000) reported that the resolution of the crystal structure of HA derived from A/Vietnam/1203/2004 (H5N1) virus. It is believed that a switch from  $\alpha$ 2,3 to  $\alpha$ 2,6 receptor specificity is a critical step in the adaptation of avian viruses to a human host, while  $\alpha$ 2,3 specificity alone appears to be one of the reasons that most avian influenza viruses, including current avian H5 strains, are not easily transmitted from human to human after avian-to-human infection.

Ya *et al.* (2001) investigated four new three-dimensional structures of avian H5 and swine H9 influenza hemagglutinins (HAs). They found that closely related to those that caused outbreaks of human disease in Hong Kong in 1997 and 1999 were determined bound to avian and human cell receptor analogs. Form structures show that HA binding sites specific for human receptors appear to be wider than those preferring avian receptors and how avian and human receptors are distinguished by atomic contacts at the glycosidic linkage. They compare new structures with previously reported crystal structures of HA/sialoside complexes of the H3 subtype that caused the 1968 Hong Kong influenza virus pandemic and analyzed in relation to HA sequence of all 15 subtypes and to receptor affinity.

Ya *et al.* (2002) determined the three-dimensional structures of the HAs from H5 avian and H9 swine viruses closely related to the viruses isolated from humans in Hong Kong. They compared it with known structures of the H3 HA from the virus that caused the 1968 H3 pandemic and of the HAesrerase-fusion (HEF) glycoprotein from an influenza C virus. The result suggest that HA subtypes may have originated by diversification of properties that affected the metastability of HAs requires for their membrane fusion activities in viral infection.

Harvey *et al.* (2004) showed that mutations at positions 226 and 228 (H3 numbering) (Q226L and G228S), which are the adaptive mutations for H2 and H3, could reduce the binding affinity to sialic acid $\alpha$ 2,3-galactose (SA $\alpha$ 2,3-Gal) of an H5N1 virus isolated in 1997 (Hong Kong flu).

Gambaryan *et al.* (2004) suggested that Ser-to-Ile substitution at position 227 (H3 numbering) of HA can reduce the virulence potential of the virus. This amino acid is in the receptor-binding pocket on the distal tip of the HA and therefore may affect the ability of the virus to recognize and bind to target cells.

Enrique T. M. and Michael W. D. (2005) proposed method, from statistical mechanics and probabilistic statistics, to quantify the non-monotonic immune response that results from antigenic drift in the epitope of the hemagglutinin and

neuraminidase protein. They found that the results, compare epitope sequences of the hemagglutinin protein A/Fujian/411/2002 and A/Panama/2007/99, explain the ineffectiveness of the 2003-3004 influenza vaccine in the United States and provide an accurate measure by which to optimize the effectiveness of future annual influenza vaccines.

James *et al.* (2006) investigated relation between HA structure from highly pathogenic Vietnamese H5N1 influenza virus (Viet04) and 1918 and other human H1 HAs influenza A virus. They found that Viet04 more related to 1918 and other human H1 HAs than 1997 duck H5 HA by studying variation in antigenic site and receptor binding site effected to  $\alpha 2,3$  and  $\alpha 2,6$  receptor specificity only enhanced or reduced affinity for avian type receptors. Mutations that can convert avian H2 and H3 HAs to human receptor specificity, when inserted on to the Viet04 H5 HA framework, permitted binding to a natural human receptor.

Minyong *et al.* (2006) have determined how the SA $\alpha 2,3$ -Gal and SA $\alpha 2,6$ -Gal bind with H5N1 HA using ab initio quantum calculation, molecular docking, molecular mechanics, and molecular dynamics simulation. They found that the results presented, it indicates that the SA $\alpha 2,3$ -Gal-HA complex has strong multiple hydrophobic and hydrogen bond interactions whereas the SA $\alpha 2,6$ -Gal-HA complex only shows weak interactions.

Yang *et al.* (2006) reported that adaptation of avian viruses to humans is associated with HA specificity for  $\alpha 2,6$ - rather than  $\alpha 2,3$ -linked SA receptors. They defined mutations in influenza A subtype H5N1 (avian) HA that alter its specificity for SA either by decreasing  $\alpha 2,3$ - or increasing  $\alpha 2,6$ -SA recognition. The receptor binding domain (RBD) mutants were used to develop vaccines and monoclonal antibodies that neutralized new variants. Structure-based modification of HA specificity can guide the development of preemptive vaccines and therapeutic monoclonal antibodies (MAbs) that can be evaluated before the emergence of human-adapted H5N1 strains.

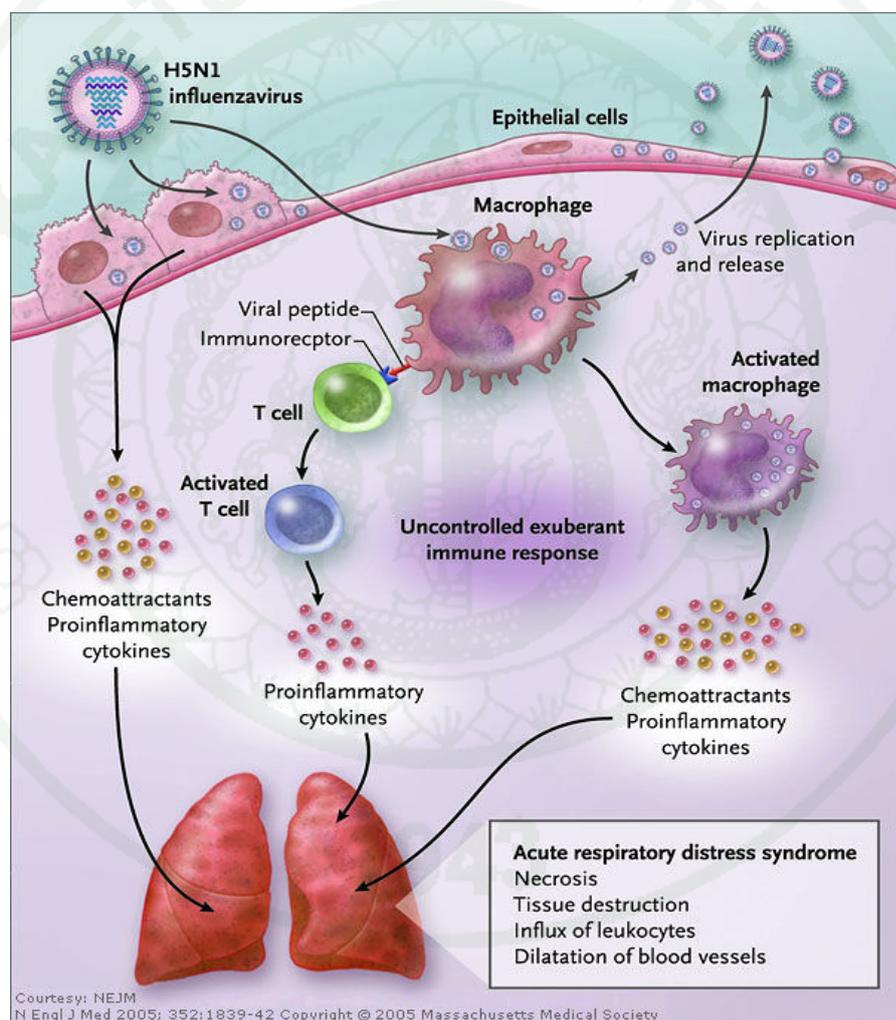
Stevens *et al.* (2006) reported that neither these mutations nor the mutations that could adapt H1 viruses to the human receptor (E190D and G225D) (H3 numbering) could completely convert a Vietnam H5N1 virus isolated in 2004 to the 2,6-type receptor specificity. Except for S227N, these mutations have not been found in H5N1 viruses isolated from humans or animals. N182K and Q192R mutations were shown to enhance the binding of a Vietnam H5N1 HA isolated in 2004 to the SA<sub>2,6</sub>-Gal receptor (Yamada *et al.*, 2006).

Auewarakul *et al.* (2007) found substitutions at position 129 and 134 identified in a virus isolated from a fatal human case that could change the receptor-binding preference of HA of H5N1 virus from SA<sub>2,3</sub>-Gal to both SA<sub>2,3</sub>-Gal and SA<sub>2,6</sub>-Gal. Molecular modeling demonstrated that the mutation (L129V and A134V) may stabilize SA<sub>2,6</sub>-Gal in its optimal cis conformation in the binding pocket. The mutation was found in approximately half of the viral sequences directly amplified from a respiratory specimen of the patient.

Nikolai *et al.* (2007) suggested that the HA antigenic structure differs substantially between A/Vietnam/1203/04 (H5N1) virus and the low-pathogenic A/Mallard/Pennsylvania/10218/84 (H5N2) virus they previously characterized (Kaverin *et al.*, 2002). The hemagglutination inhibition reactions of the MAbs with recent highly pathogenic H5N1 viruses were consistent with the antigenic-site amino acid changes but not with clades and subclades based on H5 phylogenetic analysis. These results provide information on the recognition sites of the MAbs widely used to study H5N1 viruses and demonstrate the involvement of the HA antigenic sites in the evolution of highly pathogenic H5N1 viruses, findings that can be critical for characterizing pathogenesis and vaccine design.

In 2008, Angeline and coworker studied the inhibitory effects of these five antibodies were similar to those observed with a previously described neutralizing antibody specific for the 140s antigenic loop present within HA1 and highlight the exciting possibility that these antibodies may be efficacious against multiple H5N1 strains.

In 2009, Nipa and coworker studied the interactions between receptor models and hemagglutinin proteins from H5N1 strains A/Duck/Singapore/3/97, mutated A/Duck/Singapore/3/97 (Q222L, G224S, Q222L/G224S), A/Thailand/1(KAN-1)/2004, and mutated A/Thailand/1(KAN-1)/2004 (L129V/A134V) by MD simulations. The avian receptor was represented by SA $\alpha$ 2,3-Gal substructure and human receptor by SA $\alpha$ 2,6-Gal. The glycoside binding conformation was monitored throughout the simulations, some mutations preference toward human-type receptors.



**Figure 4** Mechanism of the cytokine storm evoked by influenza virus

**Source:** Michael *et al.* (2005)

Infection of virus host can be described in following steps. First, influenza virus binds through HA onto sialic acid sugars on the surfaces of epithelial cells, typically in the nose, throat and lungs of mammals and intestines of birds. Second, the cell imports the virus by endocytosis. In the acidic endosome part of the HA protein fuses the viral envelop with the vacuole's membrane, releasing the viral RNA molecule into the cytoplasm (Melike *et. al.*, 2003). Next, protein and viral RNA form a complex that is transported into cell nucleus (Cros, J and Palese P, 2003) and synthesized viral protein and form new viral genome particle by using inhibiting translation of host cell mRNAs (Kash *et. al.*, 2006). Finally, the mature virus depart from the cell in a sphere of host phospholipids membrane It detach once their neuraminidase has cleaved sialic acid residues from the host cell after cell died.

Sepsis is a severe systemic inflammatory response and is one example of a pathologic condition associated with "cytokine storm" (Figure 4). Sepsis is an often lethal hemodynamic collapse which is usually the result of a super infection by gram-negative bacterial endotoxins. Sepsis is also classified as septic shock syndrome (SSS).

Michael *et al.* (2005) reported that cytokine storm can also result from viral infections such as influenza, and an exaggerated systemic immune response to that particular viral infection (designated a type A, subtype "H5N1" virus) may have been the cause of high lethality seen in the influenza pandemic of 1918 to 1919. The great influenza pandemic was the most destructive pandemic in recorded world history, and killed more people (estimated between 20 to 50 million) than all casualties resulting from the first World War. Although the Spanish Flu pandemic affected an enormous percentage of the world wide population (up to 20% of the world population according to some sources), and killed between 20 and 50 million persons, no more than 5% of the people who contracted the Spanish Flu died (Brown *et. al* reported the highest death rate in India at 50 deaths per 1000 persons contracting the disease, or a five percent fatality rate). After 218 human cases of bird flu have been confirmed world-wide (as of May, 2006) the lethality rate stands at 57%. Should this strain develop into a pandemic, and should it keep its current mortality rate, it has the potential to be 10 times more lethal than the 1918 pandemic.

## MATERIALS AND METHODS

### Materials

#### 1. Molecular dynamics simulations part

##### 1.1 Hardware

1.1.1 Computer Cluster of the Center of Cheminformatics  
(Chemistry Department, Faculty of Science, Kasetsart University, Bangkok).

1.1.2 Eclipse.biotech.or.th of the National Center for Genetic  
Engineering and Biotechnology (BIOTEC).

1.1.3 Thai National Grid Center (TNGC).

##### 1.2 Software

1.2.1 AMBER9 program

1.2.2 Swiss-pdb Viewer 3.7

1.2.3 SWISS-MODEL

1.2.4 PyMOL Viewer

1.2.5 Weblab Viewerpro 4.0

1.2.6 SSH Secure File Transfer

1.2.7 FileZilla

1.2.8 EditPlus

1.2.9 Putty

1.2.10 Glycam06 parameter

1.2.11 VMD: visual molecular dynamics

1.2.12 Xmgrace

1.2.13 EndNote version 9.0

1.2.14 X-Win32

## 2. Laboratory part

QuikChange<sup>®</sup> Site-Directed Mutagenesis Kit.

### 2.1 Mutant Strand Synthesis Reaction. (Thermal Cycling).

This material was contained with

2.1.1 Synthesize two complimentary oligonucleotides containing the desired mutation, flanked by unmodified nucleotide sequence.

2.1.2 Prepare the sample reactions as indicated below:

5  $\mu$ l of 10X reaction buffer

1  $\mu$ l (5–50 ng) of dsDNA template

1  $\mu$ l (125 ng) of oligonucleotide forward primer

1  $\mu$ l (125 ng) of oligonucleotide reverse primer

1  $\mu$ l of dNTP mix

Add H<sub>2</sub>O to a final volume of 50  $\mu$ l

Then add 1  $\mu$ l of PfuTurbo DNA polymerase (2.5 U/ $\mu$ l)

**Table 4** Cycling Parameters for the QuikChange<sup>®</sup> Site-Directed Mutagenesis Method

Segment	Cycles	Temperature	Time
1	1	95 °C	30 seconds
2	12 - 18	95 °C	30 seconds
		55 °C	1 minute
		68 °C	1 minute/kb of plasmid length

2.1.3 Cycle each reaction using the cycling parameters outlined in Table 4. (For the control reaction, use a 5-minute extension time and run the reaction for 18 cycles.)

2.1.4 Add 1  $\mu$ l of the Dpn I restriction enzyme (10 U/ $\mu$ l) at 37 °C for 1 hour.

## 2.2 Transformation of XL1-Blue Supercompetent Cells

This material was contained with

- 2.2.1 XL1-Blue supercompetent cells 50 U/ $\mu$ l
- 2.2.2 Transfer 1  $\mu$ l of the Dpn I-treated DNA
- 2.2.3 Add 0.5 ml of SOC medium
- 2.2.4 For the mutagenesis and transformation controls, spread cells on LB–Ampicillin agar plates containing 80  $\mu$ g/ml X-gal and 20 mM IPTG

## 2.3 Chemicals

- 2.3.1 Acetic acid (J. T. Baker, Thailand)
- 2.3.2 Acrylamide (Bio Basic, Canada)
- 2.3.3 Agar bacteriologico americano (Pronadisa, Spain)
- 2.3.4 Ammonium sulfate (Bio Basic, Canada)
- 2.3.5 Ammonium persulfate (Ajax Finechem, Australia)
- 2.3.6 Ampicillin
- 2.3.7 Bovine serum albumin (BSA) (Fluka Biochemika, USA)
- 2.3.8 Ethylenediaminetetraacetic acid disodium salt (EDTA) (Univar, Australia)
- 2.3.9 Glucose (Ajax Finechem, Australia)
- 2.3.10 Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (Fermentus, USA)
- 2.3.11 2–mercaptoethanol (Merck, Germany)
- 2.3.12 Magnesium Chloride (J. T. Baker, Malaysia)
- 2.3.13 Magnesium Sulfate (J. T. Baker, Malaysia)
- 2.3.14 Methanol (Ajax Finechem, Australia)
- 2.3.15 Potassium Chloride (J. T. Baker, Malaysia)

- 2.3.16 Sodium dodecyl sulfate (SDS) (Bio Basic, USA)
- 2.3.17 Sodium chloride (J. T. Baker, Malaysia)
- 2.3.18 Tetramethylethylenediamine (TEMED) (Bio Basic, USA)
- 2.3.19 Triton<sup>®</sup> X-100 (USB Corporation, USA)
- 2.3.20 Tris (hydrochloride) (Research Organic, USA)
- 2.3.21 Tryptone type-I (Himedia, India)
- 2.3.22 Yeast extract powder (Himedia, India)

## 2.4 Buffers and Solutions

### 2.4.1 Luria-Bertani (LB) Agar / 1,000 ml

NaCl	10	g
Tryptone	10	g
Yeast extract	5	g
Agar	20	g
Add deionized H <sub>2</sub> O 1 liter		
Adjust pH to 7.0 with 5 N NaOH		

The mixture was autoclaved for 3 hours

### 2.4.2 LB-Ampicillin Agar / 1,000 ml

1 liter of LB agar, autoclaved

Cool to 55°C

Add 10 ml of 10-mg/ml filter-sterilized ampicillin

Pour into petri dishes (~25 ml/100-mm plate)

### 2.4.3 Buffer of Reaction

10X reaction buffer

100 mM KCl

100 mM(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>

200 mM Tris-HCl (pH 8.8)

20 mM MgSO<sub>4</sub>

1% Triton® X-100

1 mg/ml nuclease-free bovine serum albumin

(BSA) reaction buffer

TE Buffer

1 mM Tris-HCl (pH 7.5)

1 mM EDTA

### 2.4.4 SOC Media Recipe / 1000 ml

2.4.4.1 Add the following to 900ml of distilled H<sub>2</sub>O

20 g Bacto Tryptone

5 g Bacto Yeast Extract

2 ml of 5M NaCl

2.5 ml of 1M KCl

10 ml of 1M MgCl<sub>2</sub>

10 ml of 1M MgSO<sub>4</sub>

20 ml of 1M glucose

2.4.4.2 Adjust to 1 L with distilled H<sub>2</sub>O

2.4.4.3 Sterilize by autoclaving

## Methods

### 1. Molecular dynamics simulations

1.1 The crystal structures of HA from A/Duck/Singapore/3/97 (H5N1) (Stevens *et al.*, 2005) were used as templates for the simulations of HA binding to SA $\alpha$ 2,3-Gal (Protein Data Bank [PDB] accession number 1JSN) and SA $\alpha$ 2,6-Gal (PDB accession number 1JSO). In the structural template under PDB accession number 1JSO, the sialic residue has no galactose unit connected; therefore, we added it with the torsion angle of 55°. Both glycosides in the two structures were terminated with a methoxy group and used as the input for MD.

1.2 In homology modeling, the wild-type and mutant HA (deleted loop130/S141D) and other mutants were three-dimensionally aligned via the SWISS-MODEL server (Schwede *et al.*, 2003) using PDB accession numbers 1JSN and 1JSO and used as the initial input structures for MD.

1.3 All structures were solvated using the TIP5P water model (Mahoney *et al.*, 2000), and we performed energy minimization to relieve bad contacts caused by unreasonable distances in the initial structures and then equilibrated for 100 ps before a 3-ns productive run at 300 K using the SANDER module in the AMBER9 program (University of California, San Francisco) with a Glycam06 parameter (<http://glycam.ccruc.uga.edu>). Xmgrace (<http://plasma-gate.weizmann.ac.il/Grace/>), VMD (Humphrey *et al.*, 1996), and AMBER tools running on UNIX were exploited to visualize and manipulate all figures. Sietraj was used to calculate the free energy of binding (Naim *et al.*, 2007, 52. Åqvist *et al.*, 1994).

## **Molecular Modeling Work Flow**

### *Homology modeling with SWISS-MODEL server.*

Homology modeling is used to predict the three-dimensional (3D) structures of unknown proteins or mutant HA based on the known structure of a similar protein (X-ray crystal structures of HA from PDB). During evolution, sequence changes much faster than structure. It is possible to identify the 3D-structure by looking at a molecule with some sequence identity.

SWISS-MODEL (<http://swissmodel.expasy.org>) is a server for automated comparative modeling of 3D protein structures. It pioneered the field of automated modeling starting in 1993 and is the most widely-used free web-based automated modeling facility today. In 2002 the server computed 120,000 user requests for 3D protein models. SWISS-MODEL provides several levels of user interaction through its World Wide Web interface: in the “first approach mode” only an amino acid sequence of a protein is submitted to build a 3D model. Template selection, alignment and model building are done completely automated by the server. In the “alignment mode”, the modeling process is based on a user-defined target-template alignment. Complex modeling tasks can be handled with the “project mode” using DeepView (Swiss-PdbViewer), an integrated sequence-to-structure workbench. All models are sent back via email with a detailed modeling report. What check analyses and ANOLEA evaluations are provided optionally. The reliability of SWISS-MODEL is continuously evaluated in the EVA-CM project. The SWISS-MODEL server is under constant development to improve the successful implementation of expert knowledge into an easy-to-use server.

### *Calculation of partial atomic charges*

All quantum mechanical calculations were performed with the Gaussian94 program. Electrostatic potentials were computed at the Hartree-Fock (HF) level with the 6-31G\* basis set, at points around the solvent accessible surface of the molecules as determined using the CHELPG protocol. MD simulations were performed with the SANDER module of the AMBER 9.0 program using the all atom GLYCAM06 force field parameters set for oligosaccharides and glycoproteins.

xLeap and tLeap: These programs perform the same function with the difference being that xLeap opens in an X-window interface and tLeap operates from a terminal prompt. The principal function of these programs is to prepare the AMBER coordinate and topology files.

Leap to solvate the unit in a truncated octahedral box using a spacing distance of 9.0 angstroms around the molecule. Ideally, you should set the spacing at no less than 8.5 Å (~ 3 water layers) to avoid periodicity artifacts (Weber *et al.*, 2000). For particle-mesh ewald electrostatics, (Darden *et al.*, 1993 and Essmann *et al.*, 1995). Our box side length must be > 2 X nonbonded cutoff. We will use a 10.0 Å cutoff in our solvated system; therefore, our box side must be > 20 Å. Our box side length will be (2 X 9) peptide dimension, which should easily be greater than 20 Å. The system must be neutral in terms of overall formal charge. Fortunately, our system is neutral as is. If this had not been the case then we would have used the add ions command to neutralize the charge (use Na<sup>+</sup> to counter a negative charge or Cl<sup>-</sup> to counter a positive charge). The save AmberParm command saves the parameter file (prmtop or top file) and the initial coordinate file (inpcrd or crd file). We did this before solvating the system so we could perform an in vacuo dynamics simulation for comparisons to the solvated system.

## Molecular Dynamics for the AMBER9 program

### Molecular Dynamics in a water box

This job will be accomplished in 4 steps:

**Step 1. Restrained Minimization** – relieve bad vander Waals contacts in the surrounding solvent while keeping the proteins restrained.

Using the Sander program. The SANDER program is the number crunching juggernaut of the AMBER software package. SANDER will perform minimization, dynamics and NMR refinement calculations. We must specify an input file to tell SANDER what computations we want to perform and how we would like to perform those computations. Study the input file (\*.in) for minimization below.

#### Contents of min1.in

```

1jsn: unmutated initial minimization solvent + ions
&cntrl
  imin      = 1,
  maxcyc    = 2500,
  ncyc      = 1000,
  ntb       = 1,
  ntr       = 1,
  cut       = 10
/
Hold the protein fixed
500.0
RES 1 325
END
END

```

*Note that:*

**&cntrl and /** - Most if not all of your instructions must appear in the “control” block (hence&cntrl).

**cut** = nonbonded cutoff in angstroms.

**ntb** = 1 constant volume dynamics

**ntr** = Flag used to perform position restraints (1 = on, 0 = off)

**imin** = Flag to run energy minimization (if = 1 then perform minimization; if = 0 then perform molecular dynamics).

**maxcyc** = maximum 2500 of cycles

**ncyc** = After ncyc cycles the minimization method will switch from steepest descents to conjugate gradient.

**Hold the peptide fixed** = 500.0 (This is the force in kcal/mol used to restrain the atom positions.)

**RES 1 325** (Tells AMBER to apply this force to residue's 1 to 325).

**Step 2. Unrestrained Minimization** – Relieve bad contacts in the entire system.

*Contents of min2.in*

```

1jsn: unmutated initial minimization solvent + ions
&cntrl
  imin      = 1,
  maxcyc    = 3000,
  ncyc      = 1000,
  ntb       = 1,
  ntr       = 0,
  cut       = 10
/

```

**Step 3. Restrained Dynamics** – Relax the solvent layers around the solute while gradually bringing the system temperature from 0 K to 300 K.

This initial dynamics run is performed to relax the positions of the solvent molecules. In this dynamics run, we will keep the macromolecule atom positions restrained (not fixed, however). In a position-restrained run, we apply a force to the specified atoms to minimize their movement during the dynamics. The solvent we are using in our system, water, has a relaxation time of 10 ps; therefore we need to perform at least > 10 ps of position restrained dynamics to relax the water in our periodic box.

*Contents of mdl.in*

```

1jsn: 100ps MD with small res on protein
&cntrl
  imin      = 0,
  irest     = 0,
  ntx       = 1,
  ntb       = 1,
  cut       = 10,
  ntr       = 1,
  ntc       = 2,
  ntf       = 2,
  tempi      = 0.0,
  temp0     = 300.0,
  ntt       = 3,
  gamma_ln  = 1.0,
  tautp     = 0.1,
  nstlim    = 50000, dt = 0.002,
  ntp      = 100, ntwx = 500, ntwr = 1000
/
Keep protein fixed with weak restraints
10.0
RES 1 325
END
END

```

*Note that:*

**imin** = 0 Switch to indicate that we are running a dynamics.

**nstlim** = 50000 of steps limit.

**dt** = 0.002 time step in ps (2 fs)

**temp0** = 300 reference temp (in degrees K) at which system is to be kept.

**tempi** = 100 initial temperature (in degrees K)

**gamma\_ln** = 1 collision frequency in ps<sup>-1</sup> when ntt = 3

**ntt** = 3 temperature scaling switch (3 = use langevin dynamics)

**tautp** = 0.1 Time constant for the heat bath (default = 1.0) smaller constant gives tighter coupling.

**ntc** = 2 Flag for the Shake algorithm (1 – No Shake is performed; 2 – bonds to hydrogen are constrained; 3 – all bonds are constrained).

**Step 4. Production Run** – Run the production dynamics at 300 K and 1 bar pressure.

*Contents of md2.in*

```

1jsn: 1000ps MD
&cntrl
  imin      = 0,
  irest     = 1,
  ntx       = 7,
  ntb       = 2,
  cut       = 10,
  ntr       = 0,
  ntp       = 1,
  taup      = 2.0,
  pres0     = 1.0,
  ntc       = 2,
  ntf       = 2,
  tempi     = 300.0,
  temp0    = 300.0,
  ntt       = 3,
  gamma_ln = 1.0,
  nstlim   = 500000, dt = 0.002,
  ntpr     = 100, ntwx = 500, ntwr = 1000
/

```

*Note that:*

**ntb** = 2 Constant pressure dynamics.

**ntp** = 1 md with isotropic position scaling.

**taup** = 2.0 pressure relaxation time in ps

**pres0** = 1 reference pressure in bar.

## Analysis

A. The RMSD plot. We will use the ptraj program.

### Contents of rms.in

```
trajin ljsn_3u_md2.mdcrd
trajin ljsn_3u_md3.mdcrd
trajin ljsn_3u_md4.mdcrd
rms first out ljsn_3u.rms @N,C,CA time 1.0
```

### Note that:

**trajin** – specifies trajectory file to process.

**rms** – computed RMS fit to the first structure of the first trajectory read in.

**out** – specifies name of output file.

**@N, C, CA** – Atom mask specifier (backbone atoms) (Note: The @ symbol is the atom specifier; alternately or in combination, you may use the colon : to specify residue ID # as well) For example if you only desired to examine the backbone atoms of residue #23, use : 23@N,C,CA

**time** – tells ptraj that each frame represents 1 ps.

B. Computing an average structure. From the RMSD plot indicates a region of stability. We will use these time points to compute an average. Average structures are good when you need a structure that is representative of a specific time range in the trajectory (e.g. a region of stability or equilibrium).

### Contents of avg.in

```
trajin jsn_md2.mdcrd 2950 3000
center :1-9
image center familiar
rms first mass out av_rms.dat :1-9
average jsn_avg.pdb pdb
```

The first steps involve centering the box followed by mass weighted RMS fit analysis, which is required for computation of the final average structure. The water

molecule positions are highly variable in the average structure and will be essentially worthless. Remove the water molecules.

C. Analysis of hydrogen bonds over the course of the trajectory. Use the following input file (hbond.in) to ptraj.

*Contents of hbond.in*

```
trajin oxy_md2.mdcrd
donor TYR O
acceptor ASN N H
acceptor CYX N H
hbond distance 3.5 angle 120.0 includeself
donor acceptor neighbor 2 series \
hbond
```

*Note that:*

**donor/acceptor** – use to specify the donor acceptor atoms.

**distance**– use to specify the cutoff distance in angstroms between the heavy atom participating in the interaction. (3.0 is the default)

**includeself** – include intramolecular H-bond interactions if any.

**angle** – The H-bond angle cutoff (H-donor-acceptor) in degrees. (0 is the default)

**series** – Directs output of H-bond data summary to STDOUT.

The statistical analysis in the hbond.dat file will be most important. Look for those H-bonds with a high % occupancy (> 60%). These are the more stable H-bonds. Throw out results with angles less than 120 degrees. The higher the % occupancy; the better. When analyzing H-bond data, it is best to establish reasonable guidelines for the distance and angle cutoffs. A recent paper by Chapman *et al.* provides a nice discussion of hydrogen bond criteria.

## D. Torsion angle measurements of the peptide backbone

### *Contents of torsions.in*

```

trajin ljsn_3u_md2.mdcrd
trajin ljsn_3u_md3.mdcrd
trajin ljsn_3u_md4.mdcrd
dihedral ljsn_3u :233@O6 :233@C2 :232@O6 :232@C6 out ljsn_3u.dat
dihedral ljsn_3u_O1B :233@O1B :233@C1 :233@C2 :232@O6 out ljsn_3u_O1B.dat
dihedral ljsn_3u_C5N :233@C5N :233@N5 :233@C5 :233@C6 out ljsn_3u_C5N.dat
dihedral ljsn_3u_C2O6_phi :233@C1 :233@C2 :232@O6 :232@C6 out ljsn_3u_C2O6_phi.dat
dihedral ljsn_3u_C6O6_psi :233@C2 :232@O6 :232@C6 :232@C5 out ljsn_3u_O6C6_psi.dat
dihedral ljsn_3u_C6C5_ome :232@O6 :232@C6 :232@C5 :232@O5 out ljsn_3u_C6C5_ome.dat
dihedral ljsn_3u_C1O4_u23 :232@O5 :232@C1 :232@O4 :232@C4 out ljsn_3u_C1O4_u23.dat
dihedral ljsn_3u_O4C4_u23 :232@C1 :232@O4 :232@C4 :231@C5 out ljsn_3u_O4C4_u23.dat

```

The individual files can be plotted with xmgrace or Microsoft Excel to view the dihedral angle fluctuation with time in the simulation.

## E. Visual Molecular Dynamics (VMD)

VMD is a molecular visualization and analysis program designed for biological systems such as proteins, nucleic acids, lipid bilayer assemblies, etc. It is developed by the Theoretical and Computational Biophysics Group at the University of Illinois at Urbana-Champaign. Among molecular graphics programs, VMD is unique in its ability to efficiently operate on multi-gigabyte molecular dynamics trajectories, its interoperability with a large number of molecular dynamics simulation packages, and its integration of structure and sequence information.

### *General molecular visualization*

VMD is a general application for displaying molecules containing any number of atoms and is similar to other molecular visualization programs in its basic capabilities. VMD reads data files using an extensible plugin system, and supports Babel for conversion of other formats. User-defined atom selections can be displayed in any of the standard molecular representations. Displayed graphics can be exported

to an image file, to a scene file usable by ray tracing programs, or to a geometry description file suitable for use with 3-D printers.

#### *Visualization of dynamic molecular data*

VMD can load atomic coordinate trajectories from AMBER, Charmm, DLPOLY, Gromacs and many other simulation packages. The data can be used to animate the molecule or to plot the change in molecular properties such as angles, dihedrals, interatomic distances, or energies over time.

#### *Visualization of volumetric data*

VMD can load, generate, and display, volumetric maps. Supported map formats include CryoEM maps, electrostatic potential maps, electron density maps, and many other map file formats.

#### *Interactive molecular dynamics simulations*

VMD can be used as a graphical front-end to a live MD program running on a remote supercomputer or high-performance workstation. VMD can interactively apply and visualize forces in an MD simulation as it runs.

#### *Molecular analysis commands*

Many commands are provided for molecular analysis. These include commands to extract information on sets of atoms and molecules, vector and matrix routines for coordinate manipulation, and functions for computing values such as center of mass and radius of gyration.

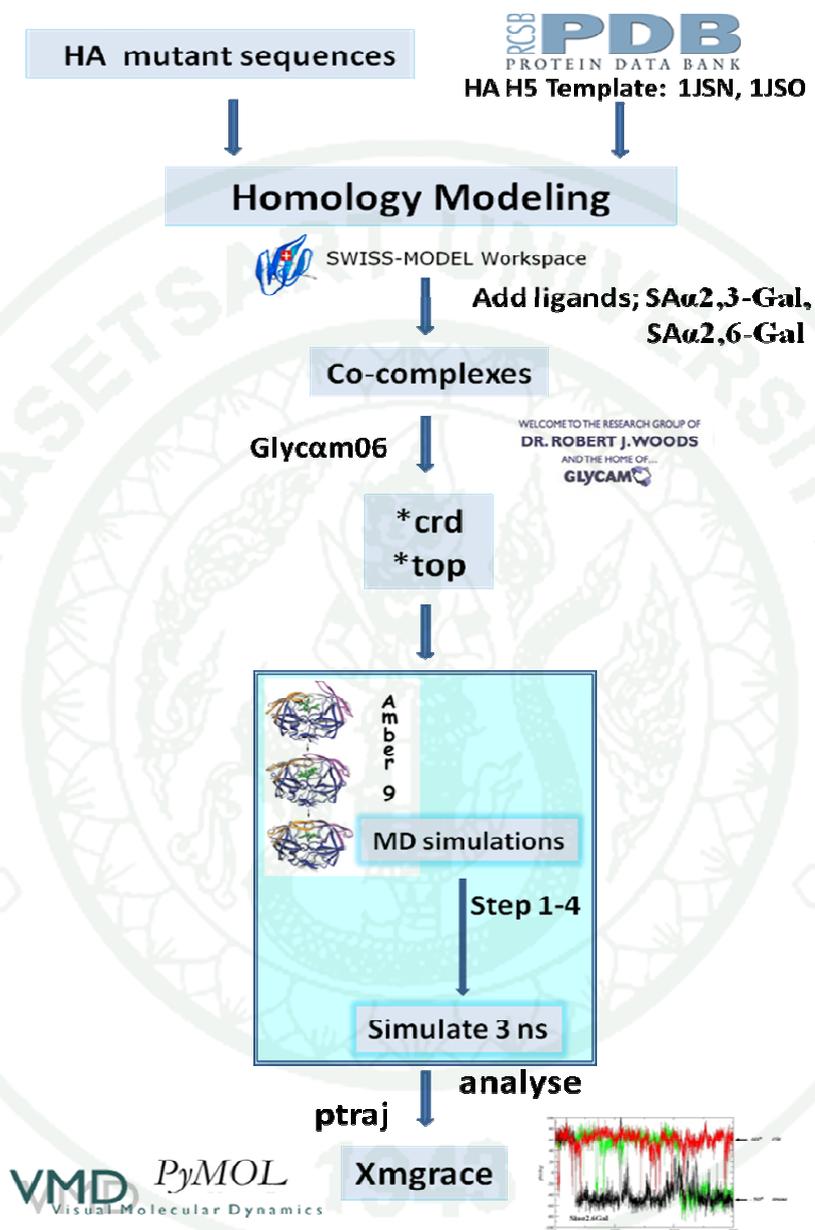
*PyMOL: Molecular visualization program*

PyMOL ([www.pymol.org](http://www.pymol.org)) is a molecular graphics system with an embedded Python interpreter designed for real-time visualization and rapid generation of high-quality molecular graphics images and animations. It can also perform many other valuable (such as editing PDB files).

Two unique and valuable features of this program over some other visualization program are the use of the powerful programming language (Python) and an emphasis on high-quality graphics. At the present time the program is still undergoing active development, but it is mature enough to be used for learning and research. During the early development, emphasis was placed on providing functionality, and not on intuitive user interface. As a result, it takes a little time to learn the program. We can interact with the program via four complimentary ways

- Pull-down menus provide access to the file functions, the molecular editor, the program options, and a few more advanced tasks. For example, you can change the display quality settings, the background color, the size of your spheres, or the transparency of the surface from the pull-down menus.
- A task bar on the right side of the screen allows one to select which objects are displayed in which way. For example, you can show your protein as a gray surface and the inhibitor as green sticks if you have separated them into two different objects.
- The molecular display area allow to one view and interactively manipulate molecules. Ours three-button mouse allows to rotate, move, and zoom in or out by holding down the left, middle, or right button, respectively. Furthermore, you can select and perform specific tasks via mouse and keyboard shortcuts. For example, you can select two atoms (first by CTRL-SHIFT/left-clicking, second by CTRL-SHIFT/right-clicking) and create a covalent by pressing CTRL-T.
- We can create dashed lines (a good way to show hydrogen bonds) between oxygens and nearby amide hydrogens in the alpha-helix by using a command in the PyMOL manual.

### Molecular Modeling Work Flow



**Figure 5** Schematic diagram of molecular dynamics simulations method

**Source:** Schwede *et al.* (2003)

## **2. Cloning and generation of mutants in the experimental.**

2.1 Construction and expression of recombinant HA gene. The cDNA of HA genome was kindly provided from the reposition of St. Jude Children's Research Hospital and grown in embryonated eggs. The HA gene will be isolated from cDNA of the HA genome by using the Polymerase Chain Reaction (PCR) technique. The primers will be designed from sequence of H5 HA gene for the wild-type and mutant HA (deleted loop130) in the database and added enzyme restriction sites into the sequence for cloning.

2.2 The PCR product is purified by ethanol precipitation and cut for ligation into the expression vector, pHW2000 by cutting out the DNA, a restriction enzyme which cut with same enzymes. The recombinant plasmid harboring HA gene is transformed into *Escherichia coli* (*E. coli*) BL21 for high expression.

2.3 For the sequencing, The Hartwell Center for Bioinformatics and Biotechnology at St. Jude Children's Research Hospital determined the sequence of template DNA by using synthetic oligonucleotides and rhodamine or dRhodamine Dye-Terminator Cycle Sequencing Ready Reaction Kits with AmpliTaq DNA polymerase FS (Perkin-Elmer, Applied Biosystems, Inc. [PE/ABI], Foster City, CA). Samples were subjected to electrophoresis, detection, and analysis on PE/ABI model 373, model 373 Stretch, or model 377 DNA sequencers.

*Mutagenic primer design.*

The mutagenic oligonucleotide primers for use in this protocol must be designed individually according to the desired mutation. The following considerations should be made for designing mutagenic primers:

◆ Both of the mutagenic primers must contain the desired mutation and anneal to the same sequence on opposite strands of the plasmid.

◆ Primers should be between 25 and 45 bases in length, with a melting temperature ( $T_m$ ) of  $\geq 78$  °C. Primers longer than 45 bases may be used, but using longer primers increases the likelihood of secondary structure formation, which may affect the efficiency of the mutagenesis reaction. The following formula is commonly used for estimating the  $T_m$  of primers for primers intended to introduce insertions or deletions:

$$T_m = 81.5 + 0.41(\%GC) - 675/N$$

For calculating  $T_m$ :

- N is the primer length in bases
- values for % GC is whole numbers

◆ The desired mutation (deletion or insertion) should be in the middle of the primer with ~10–15 bases of correct sequence on both sides.

◆ The primers optimally should have a minimum GC content of 40% and should terminate in one or more C or G bases.

*Mutant strand synthesis reaction (Thermal Cycling): QuikChange<sup>®</sup> Site-Directed Mutagenesis Kit.*

Amplification of H5 gene by PCR technique, The PCR mixture comprising of 10X reaction buffer 5 µl, dNTP mix 1 µl, each of forward (F) and reverse (R) primers 1 µl, PfuTurbo DNA polymerase 2.5 U/µl and DNA template 1 µl, was amplified by using Primus96 plus (Hybaid) thermocycler. The PCR condition was pre-denaturation at 95°C for 30 seconds, 18 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for 1 min, an extension at 68°C for 7 min, and additional final extension at 68°C for 10 min. The PCR products were subjected to 1.0 % agarose gel, 100 volt for 45 min by electrophoresis, the gel was stained with ethidium bromide solution (2.5µg/ml) for 2-5 min, destained with water for 10 min and visualized under UV transillumination.

If desired, amplification may be checked by electrophoresis of 10 µl of the product on a 1.0 % agarose gel. A band may or may not be visualized at this stage. In either case proceed with Dpn I digestion and transformation. Adding the Dpn I restriction enzyme to the reaction tubes during the digestion step or when transferring the 1 µl of Dpn I treated DNA to incubate at 37 °C for 1 hour in the transformation reaction.

*Transformation of XL1-Blue supercompetent cells*

The amplified H5 gene was purified using QIA quick gel extraction kit (QIAGEN<sup>®</sup>), there after, it was digested and ligated to pHW2000 plasmids. The ligated plasmids were used to transform XL1-Blue supercompetent cells (50 µl). Transfer 1 µl of the Dpn I-treated DNA from each control and sample reaction to separate aliquots of the supercompetent cells. The transformation reactions gently to mix, incubate the reactions on ice for 30 minutes, heat pulse the transformation reactions for 45 seconds at 42°C, then place the reactions on ice for 2 minutes and add 0.5 ml of SOC medium preheated to 42°C and incubate the transformation reactions

at 37°C for 1 hour with shaking at 225-250 rpm. The positive clone was selected by white-blue colony screening in LB agar plates containing ampicillin (50 µg /ml) when incubated the transformation plates at 37°C for >16 hours. The presence of H5 gene encoding for HA was confirmed by PCR.

**Table 5** Sequences of oligonucleotide primer designed from nucleotide sequence of HA mutant gene.

Primers name	Sequences	T <sub>m</sub> (°C)	Primer length (bp)	Application
HA mutant deleted loop 130	F 5' ggg gtg agc tca gca tgt cca tec tec ttt ttc aga aat gtg 3' R 5' cac att tct gaa aaa gga gga tgg aca tgc tgc gct cac ccc 3' (Restriction enzyme: NcoI/ECorI)	74.3	42	Primer for cloning into pHW2000 vector

## Sequence HA full range

1 atctgtcaaa atggagaaaa tagtgcttct tttgcaata gtcagtcttg ttaaaagtga  
 61 tcagatttgc attggttacc atgcaaaaa ctcgacagag cagggtgaca caataatgga  
 121 aaagaacggt actgttacac atgcccga/ga c/ata/ctg/gaa /aag/aca/cac/a ac/ggg/aag/ct  
 181 ctgcatccta gatggagtga agcctctaata tttgagagat ttagtgtag ctggatggct  
 241 actcggaaac ccaatgtgtg acgaattcat caatgtgccg gaatggtctt acatagtgga  
 301 gaaggccaat ccagtcaatg acctctgta cccaggggat ttcaatgact atgaagaatt  
 361 gaaacaccta ttgagcagaa taaaccattt tgagaaaatt cagatcatcc caaaagtgc  
 421 ttggtccagt catgaagcct cattaggggt g <sup>F 5' → 3'</sup> **age tca gcat gtc ca** **tac cag gga aag tc**  
 481 **ctc ctt ttt cag** aaatgtgg tatggcttat caaaaagaac agtacatacc caacaataaa  
 541 gaggagctac aataatacca accaagaaga tcttttgta ctgtggggga ttaccatcc  
 601 taatgatgcg gcagagcaga caaagcteta tcaaaacca accacctata ttccgttg  
 661 gacatcaaca ctaaaccaga gattgttacc aagaatagct actagatcca aagtgaacgg  
 721 gcaaagtgga aggatggagt tcttctggac aatttataaa ccgaatgatg caatcaact  
 781 cgaagtaat ggaaattca ttgctcaga atatgcatac aaaattgtca agaaagggga  
 841 ctcaacaatt atgaaaagtg aattggaata tgtaactgc aataccaagt gcaaaactcc  
 901 aatgggggcg ataaacteta gtatgccatt ccacaatata caccctctca ccacgggga  
 961 atgcccataa tatgtgaaat caaacagatt agttcttgcg actgggctca gaaatagccc  
 1021 tcaaatagag agaagaagaa aaaagagagg attatttga gctatagcag gttttataga  
 1081 gggaggatgg caggaatgg tagatggtg gtatgggtac caccatagca atgagcaggg  
 1141 gagtgggtac gctgcagaca aagatccac tcaaaaggca atagatggag tcaccaataa  
 1201 ggtcaactcg atcattaaca aatgaacac tcagtttgag gccgttgga gggaatttaa  
 1261 caacttagaa aggagaatag agaatttaa caagaagatg gaagatgggt tcttagatgt  
 1321 ctggacttat aatgctgaac ttctggttct catggaaaat gagagaactc tagactttca  
 1381 tgactcaaat gtcaagaacc ttacgacaa ggtccgacta cagcttaggg ataatgaaa  
 1441 ggagctgggt aacggttgtt tcgatttcta tcataaatgt gataatgaat gtatggaaag  
 1501 tgtaagaaac ggaacgtatg actaccgca gtattcagaa gaagcaagac taaaagaga  
 1561 ggaataagt ggagtaaat tggaatcaat aggaattac caaactgt caattatc  
 1621 tacagtggcg agttccctag cactggcaat catgtagct ggtctatcct tatggatgtg  
 1681 ctccaatggg tcgttaca tgc aga att tgc att taa att tgtgagtca gatg

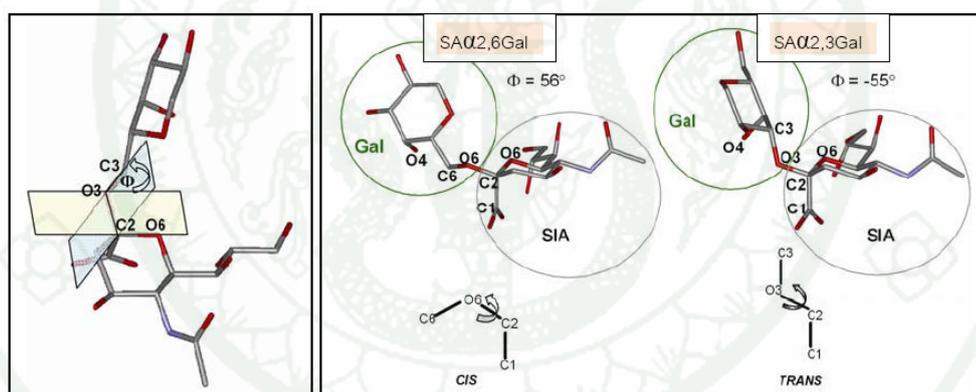
Loop  
130

## RESULTS AND DISCUSSION

### Part 1 Molecular dynamics simulations.

#### 1. The torsion angle

In the hemagglutinin binding pocket, the SA $\alpha$ 2,3-Gal and SA $\alpha$ 2,6-Gal receptors were shown to have a specific conformation, either *cis* or *trans*. Torsion angle ( $\Phi$ ) is the angle between two planes containing O6, C1 of the SA unit, and O3 (or O6), C3 of Gal unit (Figure 6). The angle indicates whether the glycoside is in the *cis* ( $\Phi = 56^\circ$ ) or *trans* ( $\Phi = -55^\circ$ ) conformation (Figure 6). Within its bound state to H5 in an X-ray cocrystal structure, SA $\alpha$ 2,3-Gal was found in the *trans* conformation.



**Figure 6** The receptor bound conformations were investigated by torsion angle ( $\Phi$ ). The torsion angles were defined that the *cis* and *trans* conformation.

#### 2. Binding Free Energy Expression and Models

MD simulations can provide not only plentiful structural–dynamical information on protein complex structures in solution but also a wealth of energetic information, including the free energy of binding between protein partners. The most rigorous MD-based approaches to estimate binding free energy are the free energy perturbation and thermodynamic integration methods.

Because of long convergence time, they are computationally intensive and prohibitive on large systems such as protein–protein complexes. A more commonly used and tractable approach is the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method. In this first-principle based method, the gas-phase energy, calculated using conventional molecular mechanics force fields such as AMBER is combined with a continuum model of solvation that includes a surface-area based nonpolar contribution and a polar solvation free energy calculated with the Poisson–Boltzmann (PB) model. Solute entropy can be incorporated from statistical thermodynamics with normal mode analysis.

The target-ligand binding free energy is approximated using the solvated interaction energy (SIE) formalism (Naim *et al.*, 2007 and Åqvist *et al.*, 1994).

$$\Delta G_{\text{bind}}^{\text{calc}} = \underbrace{E_{\text{inter}}^{\text{C}} + \Delta G_{\text{bind}}^{\text{R}}}_{\Delta G_{\text{bind}}^{\text{elec}}} + \underbrace{E_{\text{inter}}^{\text{vdw}} + \Delta G_{\text{bind}}^{\text{np}}}_{\Delta G_{\text{bind}}^{\text{np}}} \quad \text{----- (2)}$$

$E_{\text{inter}}^{\text{C}}$  and  $E_{\text{inter}}^{\text{vdw}}$  are the intermolecular Coulomb and van der Waals interaction energies in the bound state, respectively. The electrostatic contribution of the solvation free energy to binding,  $\Delta G_{\text{bind}}^{\text{R}}$ , is the change in the reaction field energy between the bound and free states. The nonpolar contribution of the solvation free energy to binding,  $\Delta G_{\text{bind}}^{\text{np}}$ , is the change in the nonelectrostatic solvation free energy (i.e., the solute-water van der Waals energy plus the cavitation cost in water) between the bound and free states. The target ligand complex represents the bound state of the system, whereas the free, unbound state is obtained by infinite rigid separation of the ligand and target from the complex. The terms in eq 2 can be grouped into the electrostatic and nonpolar contributions to binding,  $\Delta G_{\text{bind}}^{\text{elec}}$  and  $\Delta G_{\text{bind}}^{\text{np}}$ , respectively. Not included in eq 2 is the change in solute entropy upon binding, that is, changes in conformational, translational, and rotational entropy. Also, because we are using a rigid infinite separation model of binding, internal energies of neither the protein nor the ligand are taken into account.

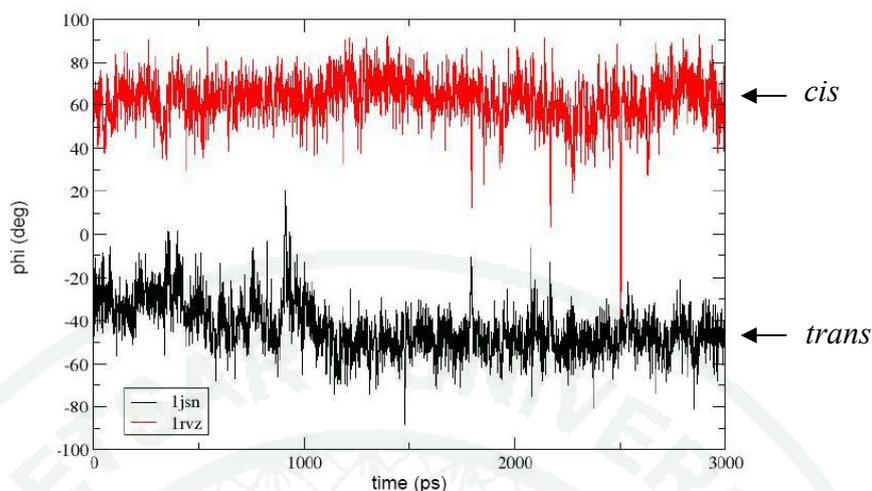
### 3. LIGPLOT program (Wallace *et al.*, 1995)

The LIGPLOT program automatically generates schematic 2-D diagrams of protein-ligand interactions for a given PDB file input. The interactions shown are those mediated by hydrogen bonds and by hydrophobic contacts. Hydrogen bonds are indicated by dashed lines between the atoms involved, while hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact. The contacted atoms are shown with spokes radiating back. The output is a color, or black and white, PostScript file giving a simple and informative representation of the intermolecular interactions and their strengths, including hydrogen bonds, hydrophobic interactions and atom accessibilities. The program is completely general for any ligand and can also be used to show other types of interaction in proteins and nucleic acids. It was designed to facilitate the rapid inspection of many enzyme complexes, but has found many other applications.

### 4. The mutation H5 HA

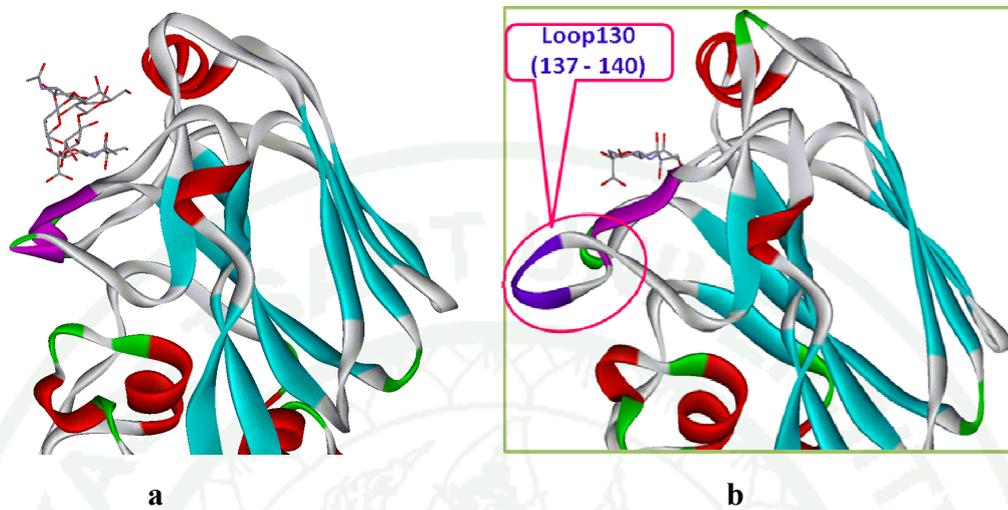
#### 4.1 The deletion loop 130 (137-140 res.) of mutation H5 HA

For a human H1 virus (1rvz.pdb), SA $\alpha$ 2,6-Gal showed an average  $\Phi$  angle of 63°, indicating a *cis* conformation. While the avian H5 virus (1jsn.pdb), SA $\alpha$ 2,3-Gal showed an average  $\Phi$  angle of -55°, indicating a *trans* conformation in the MD simulations (Figure 7).

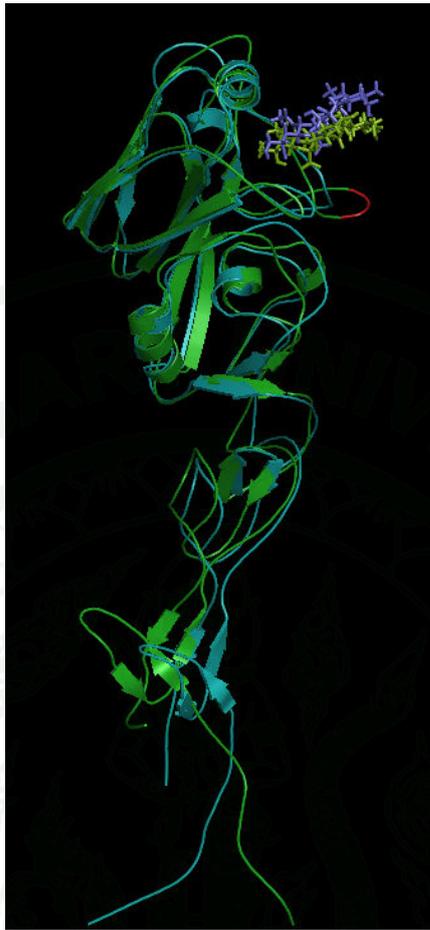


**Figure 7** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,6-Gal (red line) and SA $\alpha$ 2,3-Gal (black line) in the binding sites of the HA.

In this study the MD simulations, because the swine influenza A virus (A/Swine/Hong Kong/98: H9, 1jsi.pdb), the antigenic site is absent and, coincidentally, the H9 HA can bind both the human and avian receptor (figure 8). So we try to study whether or not this deletion can also cause the change in host type selectivity in avian influenza. We used MD simulations to build H5 HA model from the sequences where amino acid were removed between residues 137 to 140. The results of the MD simulations show that torsion angle ( $\Phi$ ) and binding free energy or  $\Delta G_{\text{bind}}$  in kcal/mol (Table 5). Calculation the H5 between two ligands (SA $\alpha$ 2,3-Gal and SA $\alpha$ 2,6-Gal) and HA protein complexes, when bound to the solvated receptor binding site and in water.



**Figure 8** The swine influenza A virus of A/Swine/Hong Kong/98: H9; 1jsi.pdb (a).  
The avian influenza A virus of A/Duck/Singapore/3/97: H5; 1jso.pdb (b).



**Figure 9** The average structural comparison of wild-type H5 (green ribbon) has loop 130 (137-140 res; red line) and deletion loop 130 of mutant H5 (blue ribbon) binding with SA $\alpha$ 2,6-Gal-linked saccharides (green stick and blue stick respectively).

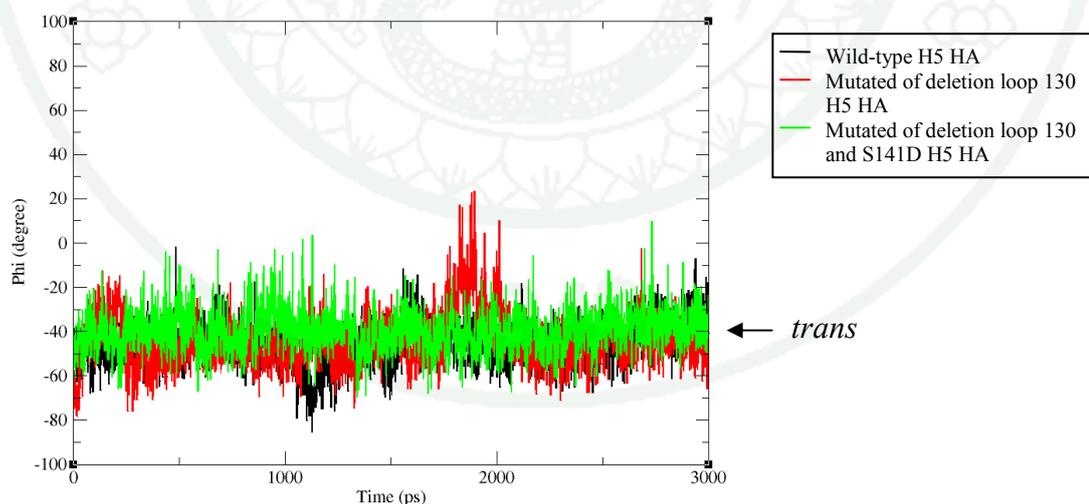
#### 4.2 The deletion loop 130 and S141D of mutation H5 HA

The sequences comparison of the swine influenza A virus H9 HA and the avian influenza A virus H5 HA (Figure 10). Besides this antigenic site is absent and we try to change amino acid from serine (S) to Aspartic acid (D) at the residue 141. The deletion loop 130 and S141D of mutation H5 HA can also cause the change in host type selectivity in avian influenza. We used MD simulations to build H5 HA model from the sequences where amino acid were removed and changed between

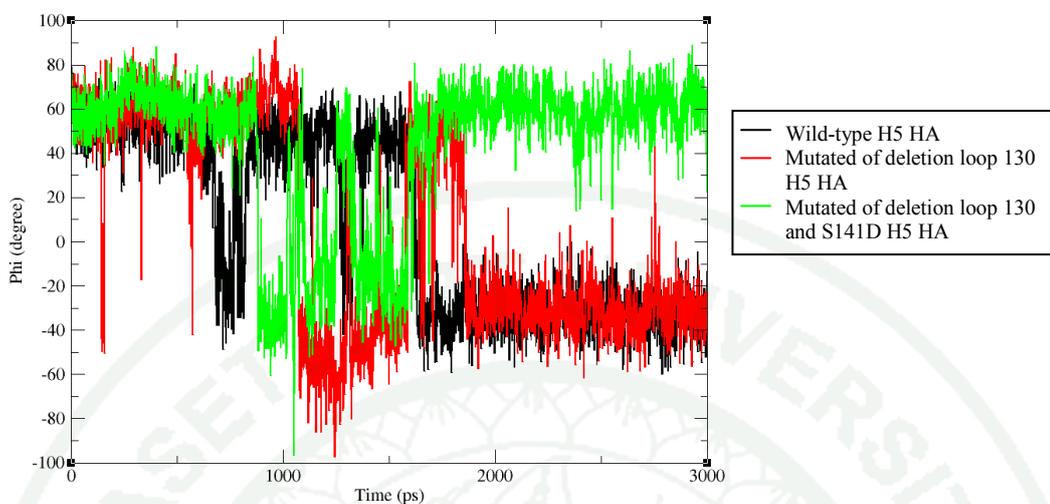
residues 137 to 141. The results of the MD simulations show that  $\Phi$  angle (Figure 11) and  $\Delta G_{\text{bind}}$  in kcal/mol (Table 6).



**Figure 10** Amino acids alignments of H9 HA gene from Protein Data Bank [PDB] (accession 1jsi) and H5 HA gen (accession 1js0) by using BioEdit Sequence Alignment Editor program.



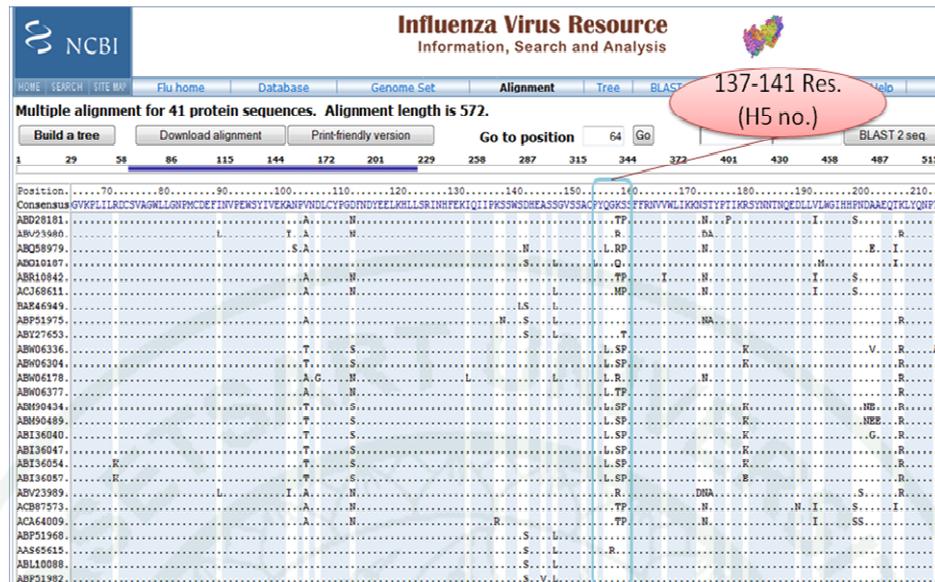
**Figure 11** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,3-Gal in the binding sites of the template H5 HA (A/Duck/Singapore/3/97) (black), the mutated of deletion loop 130 (red) and the mutated of deletion loop 130 and S141D (green).



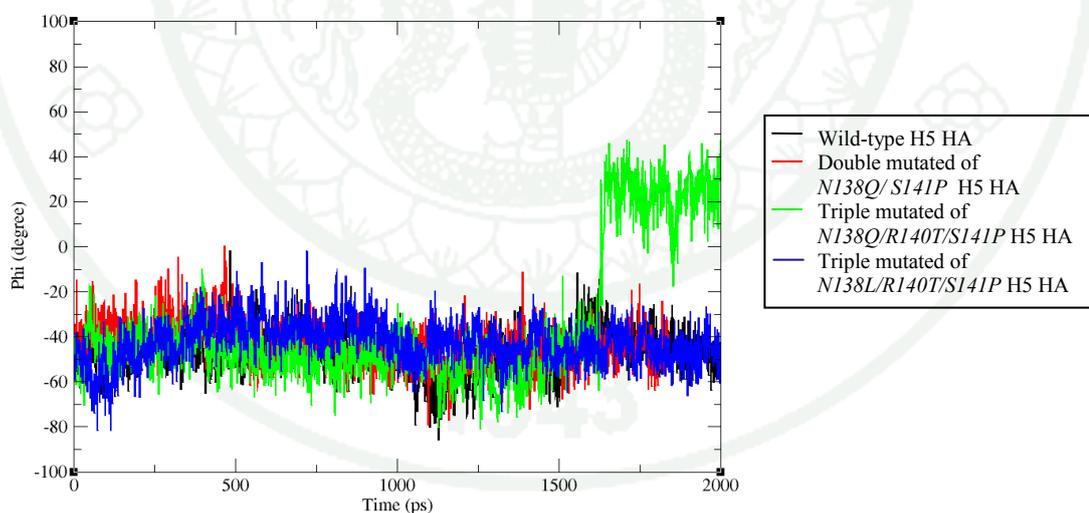
**Figure 12** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,6-Gal in the binding sites of the template H5 HA (A/Duck/Singapore/3/97) (black), the mutated of deletion loop 130 (red) and the mutated of deletion loop 130 and S141D (green).

#### 4.3 The mutated loop 130 of mutation H5 HA

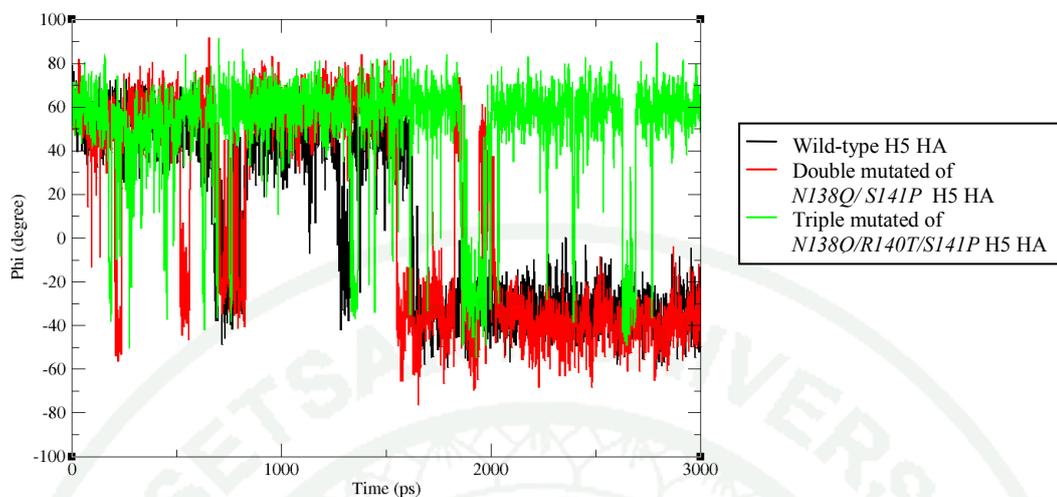
The H5 HA sequences from the GenBank at The National Center for Biotechnology Information (NCBI). The GenBank consists of several divisions, most of which can be accessed through the nucleotide database. We observe at the loop 130 region of the influenza virus A H5N1 HA protein on the human host cell in Asia from 2003 to 2009 (figure 13). So the amino acids altered from the abundant mutations. We built the mutant H5 HA models, the first model is the double mutant N138Q/S141P and the models are the triple mutant N138Q/R140T/S141P and N138L/R140T/S141P, then we studied MD simulations.



**Figure 13** The H5N1 HA proteins from the GenBank database at NCBI, focus on loop 130 region of the influenza virus A for the human host cell in Asia from 2003 to 2009.



**Figure 14** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,3-Gal in the binding sites of the template H5 HA (A/Duck/Singapore/3/97) (black), the double mutated of N138Q/S141P (red) and the triple mutants are N138Q/R140T/S141P (green) and N138L/R140T/S141P (blue).



**Figure 15** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,6-Gal in the binding sites of the template H5 HA (A/Duck/Singapore/3/97) (black), the double mutated of N138Q/S141P (red) and the triple mutated of N138Q/R140T/S141P (green).

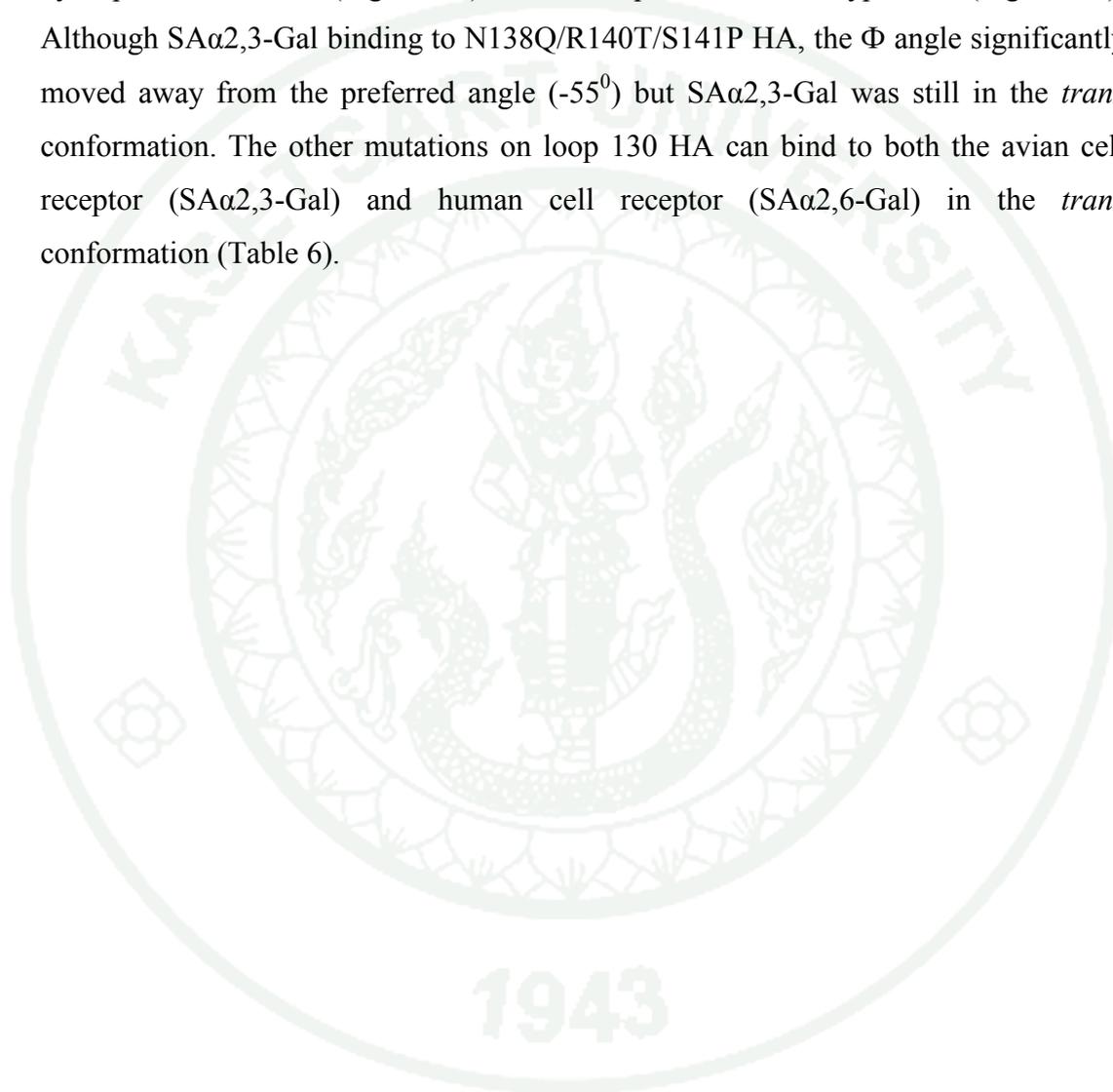
**Table 6** The relative torsion angle monitoring and binding free energy of avian (SA $\alpha$ 2,3-Gal) and human (SA $\alpha$ 2,6-Gal) cell receptor analogs complexed with A/Duck/Singapore/3/97:H5, mutation on loop 130.

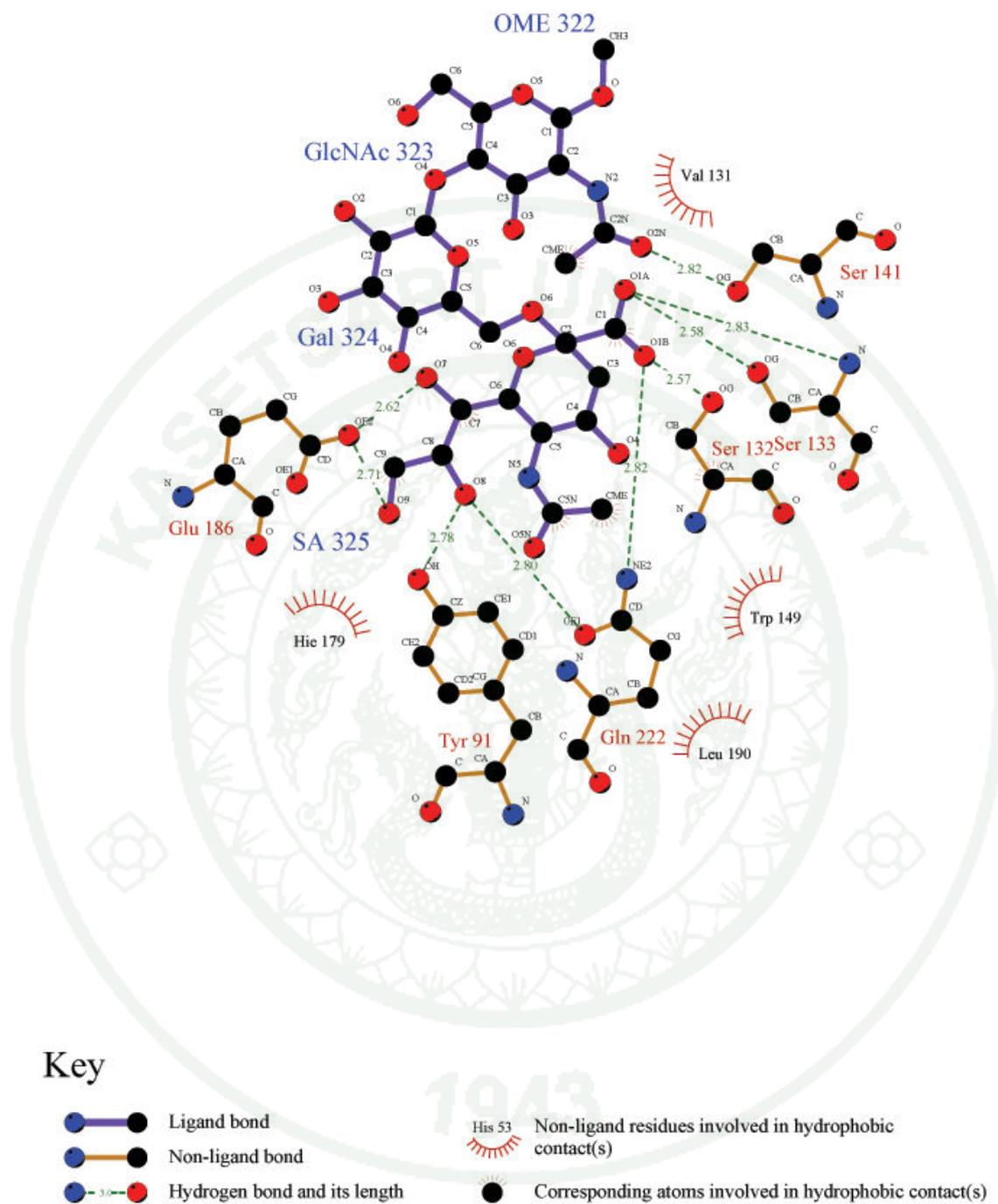
Protein	Avian (SA $\alpha$ 2,3-Gal)		Human (SA $\alpha$ 2,6-Gal)	
	Torsion angle ( $\Phi$ )	$\Delta G_{\text{bind}}$ (kcal/mol)	Torsion angle ( $\Phi$ )	$\Delta G_{\text{bind}}$ (kcal/mol)
Wild-type H5 HA	<i>trans</i>	-7.92	<i>trans</i>	-6.18
Mutated of deletion loop 130 (137-140 Res.) (YNGR)	<i>trans</i>	-5.90	<i>trans</i>	-6.09
Mutated of deletion loop 130 and S141D	<i>trans</i>	-7.83	<i>cis</i>	-6.70
Double Mutated on loop130 (H5) N138Q/S141P	<i>trans</i>	-7.11	<i>trans</i>	-6.93
Triple Mutated on loop130 (H5), N138Q/R140T/S141P	<i>trans</i>	-6.99	<i>cis</i>	-7.78

This provides a positive control for SA $\alpha$ 2,6-Gal binding in the simulation and indicates that the  $\Phi$  angle can be used as an indicator of the receptor preference. In order to have structural insight into the altered receptor specificity of the mutant HA, we performed MD simulations using two sialic acid-H5 cocrystals as reference structures for the two types of glycosidic linkages, SA $\alpha$ 2,3-Gal (PDB accession number 1JSN) and SA $\alpha$ 2,6-Gal (PDB accession number 1JSO).

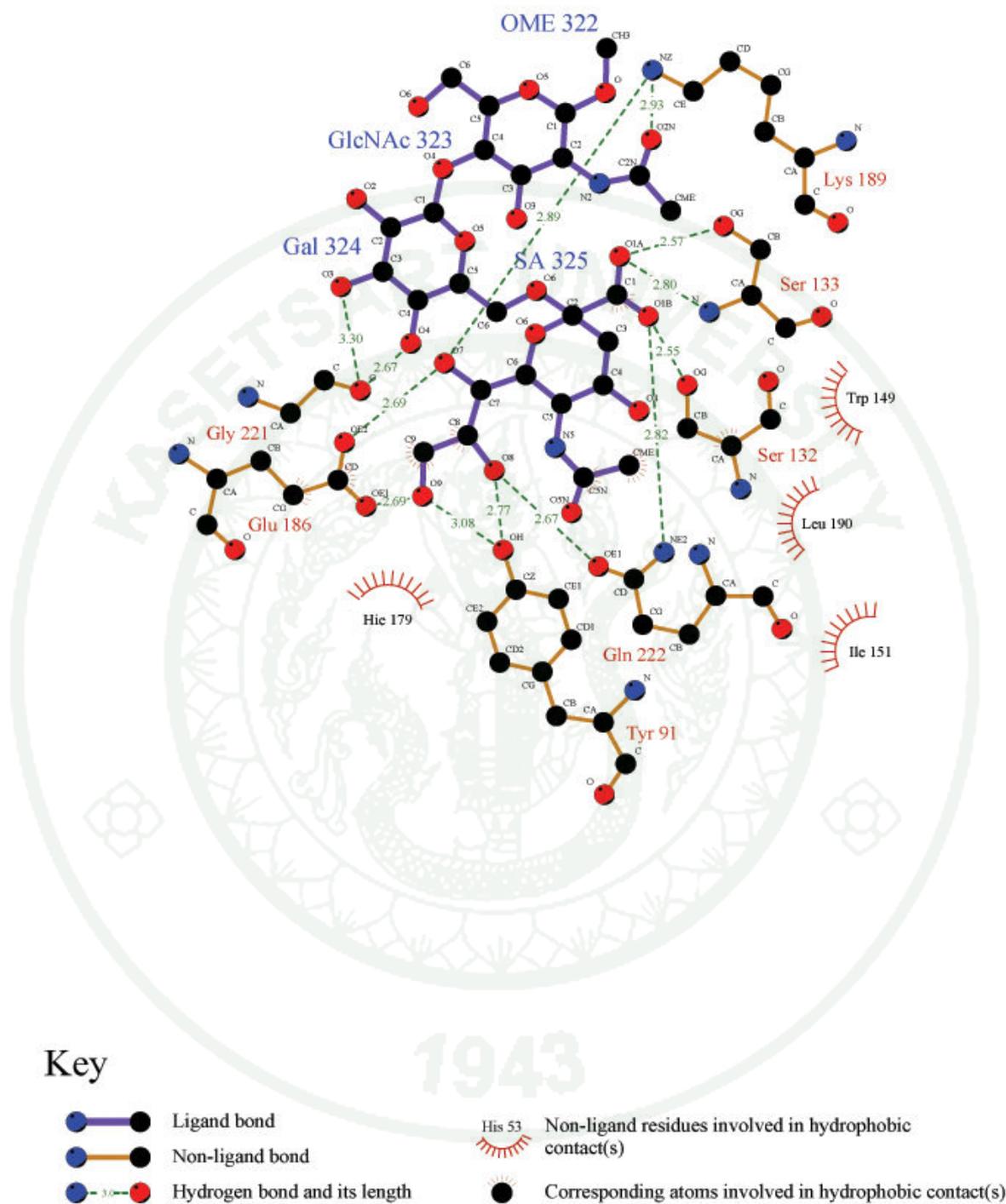
From the table 5 show that in the modeling, both the wild-type and the mutant on loop 130 (137-141 res.), HAs shared similar binding patterns within the sialic binding pocket as previously reported for both SA $\alpha$ 2,3-Gal and SA $\alpha$ 2,6-Gal binding. The  $\Phi$  angle indicated that SA $\alpha$ 2,6-Gal in the triple mutant on loop130, N138Q/R140T/S141P H5 HA spent most of the time in the lower binding free energy was -7.78 kcal/mol and *cis* conformation. Although SA $\alpha$ 2,6-Gal, the wild-type HA

was *trans* conformation at the end of the simulation period. The increased stability in the *cis* conformation of the binding to N138Q/R140T/S141P HA was due to an alternative interaction between many amino acids near the binding pocket and ligand (SA $\alpha$ 2,6-Gal) with the strong hydrogen bonding (distances  $< 3.0 \text{ \AA}$ ) and the strong hydrophobic contacts (Figure 17) when compared the wild-type HA (Figure 16). Although SA $\alpha$ 2,3-Gal binding to N138Q/R140T/S141P HA, the  $\Phi$  angle significantly moved away from the preferred angle ( $-55^\circ$ ) but SA $\alpha$ 2,3-Gal was still in the *trans* conformation. The other mutations on loop 130 HA can bind to both the avian cell receptor (SA $\alpha$ 2,3-Gal) and human cell receptor (SA $\alpha$ 2,6-Gal) in the *trans* conformation (Table 6).

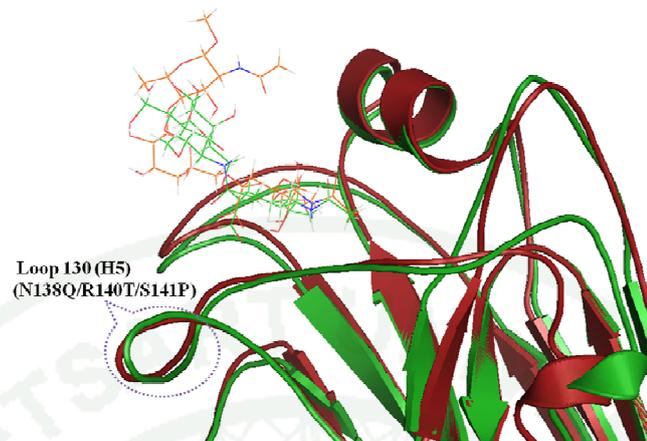




**Figure 16** A schematic illustration of the interactions of SA $\alpha$ 2,6-Gal with the residues around the receptor site of wild-type H5 HA.



**Figure 17** A schematic illustration of the interactions of SA $\alpha$ 2,6-Gal with the residues around the receptor site of the triple mutated of N138Q/R140T/S141P H5 HA.



**Figure 18** The superimposition of SA $\alpha$ 2,6-Gal around the receptor binding domain. The average structural comparison of A/Duck/Singapore/3/97:H5 (green ribbon) and mutated N138Q/R140T/S141P (red ribbon) binding with SA $\alpha$ 2,6-Gal-linked.

## 4.4 The other single mutations of mutation H5 HA

**Table 7** The relative torsion angle monitoring and binding free energy of avian and human cell receptor analogs complexes with KAN-1 HA H5, mutation on loop 130 and other single mutations.

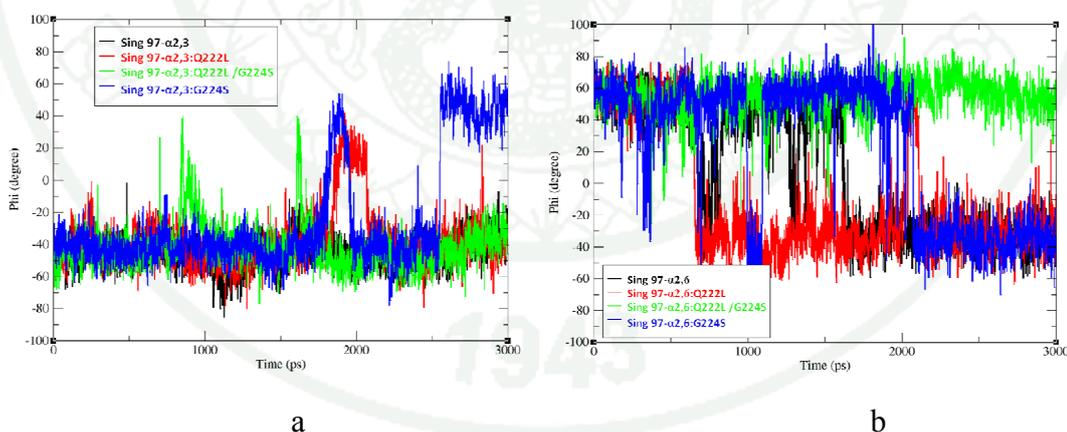
Protein	Avian (SA $\alpha$ 2,3-Gal)		Human (SA $\alpha$ 2,6-Gal)	
	Torsion angle ( $\Phi$ )	$\Delta G_{\text{bind}}$ (kcal/mol)	Torsion angle ( $\Phi$ )	$\Delta G_{\text{bind}}$ (kcal/mol)
Wild-type KAN-1 HA	<i>trans</i>	-7.79	<i>cis</i>	-7.62
E75G	-	-	<i>trans</i>	-7.32
Y91C	-	-	<i>trans</i>	-7.25
P92L	-	-	<i>trans</i>	-5.23
G130E	-	-	<i>trans</i>	-8.27
A134V	-	-	<i>cis</i>	-6.26
Q138A	-	-	<i>trans</i>	-7.64

The results from the table 7, the wild-type KAN-1 HA and the mutant A134V HA similar binding patterns within the sialic binding pocket as previously reported for both avian and human cell receptor binding. However, it was observed that over the period of simulations, hydrophobicity and spatial constraint changes residue 134 due to the different alkyl side chains, especially the A134V mutation, which happened near the glycosidic linkage, caused some changes in glycoside binding patterns. The  $\Phi$  angle indicated that SA $\alpha$ 2,6-Gal in the A134V HA spent most of the time in the low-energy *cis* conformation, whereas those in the wild-type Kan-1 HA and the template H5 were forced to change from their *cis* conformation in the initial input structures to the *trans* conformation. Although SA $\alpha$ 2,6-Gal in both wild-type Kan-1 HA and the template HA was in the *trans* conformation at the end of the simulation period, SA $\alpha$ 2,6-Gal in Kan-1 HA stayed longer in the *cis* conformation than that in the template HA. This suggested that wild-type Kan-1 HA might have a slightly

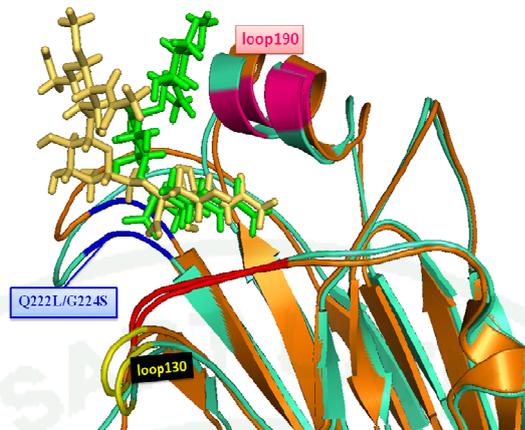
increased affinity for SA $\alpha$ 2,6-Gal compared to the template HA of A/Duck/Singapore/3/97. The increased stability in the *cis* conformation of the binding to A134V HA was due to an alternative interaction between Gly221 and 4-OH or 3-OH (O3) of Gal (as figure 17) caused by a displacement of Gly221 by the A134V mutation.

#### 4.5 The mutation of HAs subtype (H1, H5) effect on loop 130

Besides that we study in other HAs subtype, the Q222L single mutants (H5 numbering), the Q222L/G224S double mutants and the G224S single mutants. Of all the mutants HA and the template H5 (A/Duck/Singapore/3/97) bind to the avian cell receptor were represented the SA $\alpha$ 2,3-Gal in the *trans* conformation (Figure 19a). The results of the modeling indicated an altered binding conformation in the Q222L/G224S double mutant. It showed that SA $\alpha$ 2,6-Gal in a *cis* conformation in the Q222L/G224S mutant HA (Figure 19b). We suggested that the result showed more selectivity toward human cell receptor for the Q222L/G224S mutant HA.

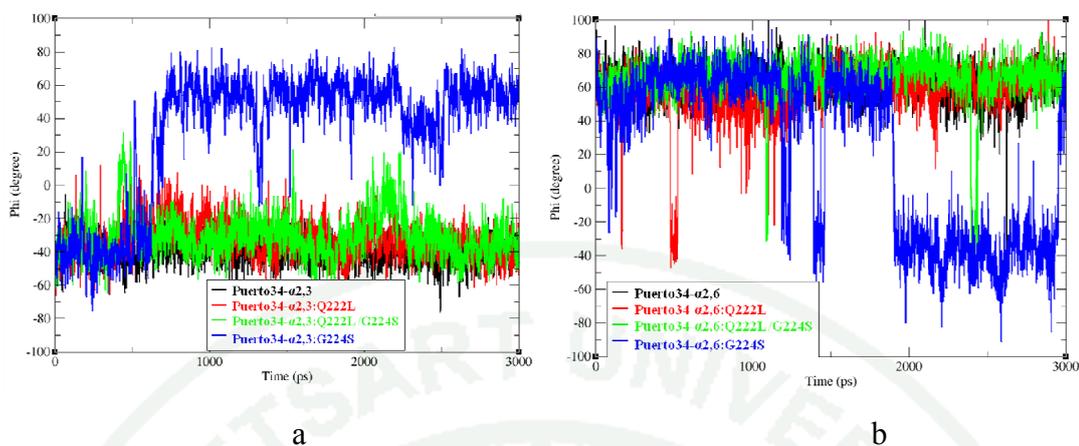


**Figure 19** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,3-Gal (a) and SA $\alpha$ 2,6-Gal (b) in the binding sites of the template H5 (A/Duck/Singapore/3/97) (black), the Q222L mutant (red), the Q222L/G224S mutant (green) and the G224S (blue).

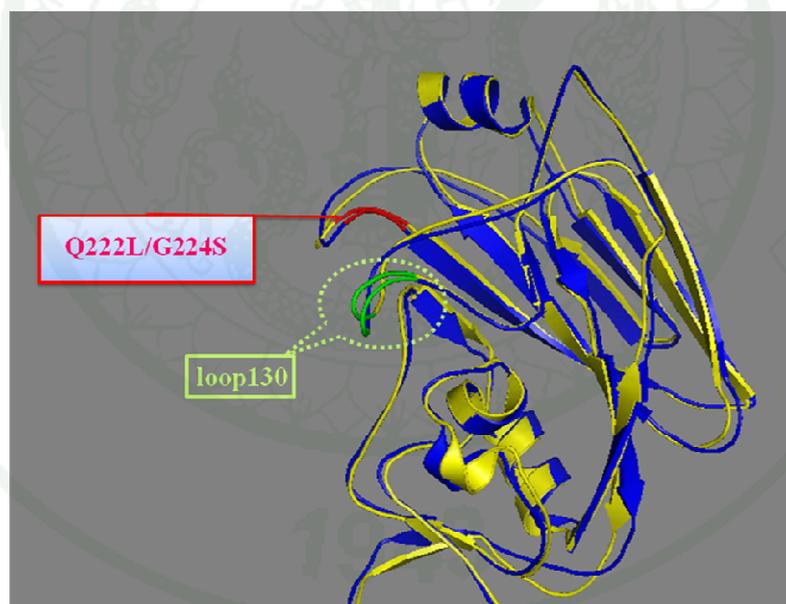


**Figure 20** The superimposition of SA $\alpha$ 2,3-Gal (green stick) and SA $\alpha$ 2,6-Gal (orange stick) around the receptor binding domain. The average structural comparison of the double mutated Q222L/G224S (green ribbon) binding with SA $\alpha$ 2,3-Gal and the double mutated Q222L/G224S (orange ribbon) binding with SA $\alpha$ 2,6-Gal as the template H5 (A/Duck/Singapore/3/97).

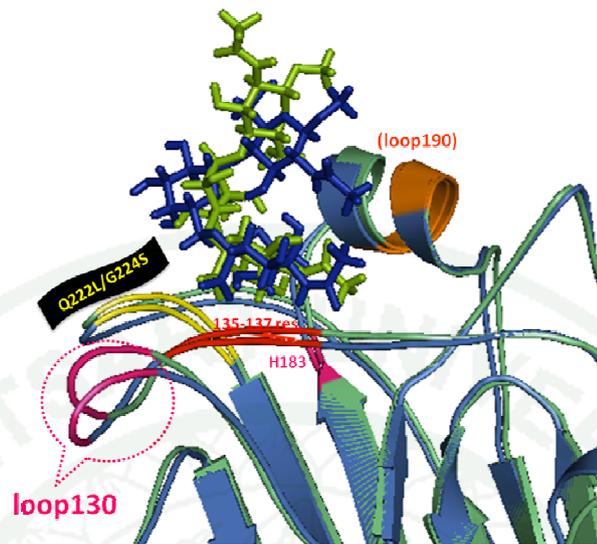
Although the double mutated Q222L/G224S HA prefer to bind to a human-type receptor were represented the SA $\alpha$ 2,6-Gal in the *cis* conformation at the end of the simulation period but the structure HA were distorted on loop 130 by the Q222L/G224S HA double mutation (figure 19).



**Figure 20** The amount of time that each receptor analog spent having a particular  $\Phi$  angle for SA $\alpha$ 2,3-Gal (a) and SA $\alpha$ 2,6-Gal (b) in the binding sites of the template H1 HA (A/Puerto/Rico/8/34) (black), the Q222L mutant (red), the Q222L/G224S mutant (green) and the G224S (blue).



**Figure 21** The superimposition of SA $\alpha$ 2,3-Gal around the receptor binding domain. The average structural comparison of A/Puerto/Rico/8/34:H1 HA (yellow ribbon) and mutated Q222L/G224S (blue ribbon).



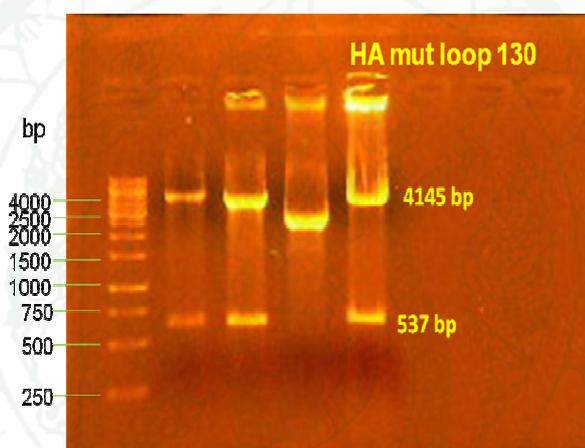
**Figure 22** The superimposition of SA $\alpha$ 2,3-Gal (green stick) and SA $\alpha$ 2,6-Gal (blue stick) around the receptor binding domain. The average structural comparison of A/Puerto/Rico/8/34:H1 HA (green ribbon) and double mutated Q222L/G224S (blue ribbon) binding with SA $\alpha$ 2,3-Gal and SA $\alpha$ 2,6-Gal respectively.

Finally, the results of the modeling of the other HA subtype, the Q222L single mutants, the Q222L/G224S double mutants and the G224S single mutants. Of all the mutants HA and the template A/Puerto/Rico/8/34:H1 HA bind to the avian cell receptor were represented the SA $\alpha$ 2,3-Gal in the *trans* conformation, but the G224S single mutants had not stable the *cis* conformation (Figure 20). From the figure 20 showed that SA $\alpha$ 2,6-Gal in a *cis* conformation in the Q222L single mutants the Q222L/G224S double mutants, the G224S single mutants and as the template H1 HA.

Thus, the major of mutants HA H1 were ability to bind to both the avian and human cell receptor. Although the mutations of HA H1 can bind to two cell receptor analogs but the structure HA were distorted on loop 130 by the mutations (figure 20, 21, 22).

## Part 2 Cloning and generation of mutants in the experimental.

PCR analysis of the HA gene deleted loop 130 (137-140 Res., H5 no.) mutation was introduced the HA gene of the A/Thailand/1(KAN-1)2004 cloned into the pHW2000 expression vector, using a Quick Change site directed mutagenesis kit according to the manufacturer's instructions. The presence of the desired mutation and the absence of the unwanted mutation were confirmed by DNA sequencing the full length of the cloned.



**Figure 23** Full length amplification of mutant HA gene deleted loop 130 by PCR RNA from A/Thailand/1(KAN-1)2004: H5, The PCR reactions were subjected to electrophoresis on a 1% agarose gel and stained with ethidium bromide.

The PCR products after gel electrophoresis, the gel also shows a positive control, and a DNA ladder containing DNA fragments of defined length mutant deleted loop 130 HA was 1660 bp and molecular weight was 142,665.14 Daltons for sizing the bands in the experimental PCRs used to NcoI and EcoRI restriction enzyme for checking the mutant HA into the pHW2000 expression vector (figure 23).



## CONCLUSIONS

The molecular dynamics simulations results showed that the mutants HA no change in binding pattern of deleted loop 130 HA compared with the wild type H5 HA. The mutants HA of the N138Q/R140T/S141P (H5 no.) triple mutation as the template H5 (A/Duck/Singapore/3/97) binding with the human cell receptor (SA $\alpha$ 2,6-Gal) in the *cis* conformation with the strong hydrogen interaction. The modeling suggested that more selectivity toward human cell receptor for the N138Q/R140T/S141P mutants HA.

The double mutated Q222L/G224S (H5 no.) HA prefer to bind to a human-type receptor were represented the SA $\alpha$ 2,6-Gal in the *cis* conformation but the structure HA were distorted on loop 130 by the Q222L/G224S HA double mutation as the template H1 (A/Puerto/Rico/8/34).

The major of mutants H1 HA were ability to bind to both the avian and human cell receptor. Although the mutations of HA H1 can bind to two cell receptor analogs but the structure HA were distorted on loop 130 by the mutations.

The results from the modeling suggested that mutation along the loop 130 (H5 numbering) were important for host cell selectivity of the other HA subtype.

In this study, we try to clone DNA and generate of the HA gene deleted loop 130 mutation in the laboratory. The result confirmed by DNA sequencing the full length of the cloned.

## LITERATURE CITED

- Ahn, S.J., J. Costa and J.R. Emanuel. 1996. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. **Nucleic Acids Res.** 24: 2623-5.
- Alexander, D. J. 1982. Avian influenza: Recent developments. **Vet. Bull.** 52: 341-359.
- Alexander, D. J. and G. Parsons. 1980. Protection of chickens against challenge with virulent influenza A viruses of H5 subtype conferred by prior infection with influenza A viruses of Hsw1 subtype. **Arch. Virol.** 66: 265-269.
- Alexander, D.J. 1987. Criteria for the definition of pathogenicity of avian influenza viruses, pp. 228-245. *In* B. C. Easterday (ed.). **Proceedings of the second international symposium on avian influenza**. U.S. Animal Health Association. Virginia.
- Alexander, D.J. 1993. Orthomyxocirus Infections, pp. 149-151. *In* J.B. McFerran and M.S. Mc Nulty (eds.). **Virus Infection of Birds**. Elsevier Science. London.
- Alexander, D.J. and R.E. Gough. 1986. Isolations of avian influenza virus from birds in Great Britain. **Vet. Rec.** 118: 537-538.
- Allwright, D.M., W.P., Burger, A. Geyer and A.W. Terblanche. 1993. Isolation of an influenza A virus from ostriches. **Avian Pathol.** 22: 59-65.
- Amonsin, A., S. Payungporn, A. Theamboonlers, R. Thanawongnuwech, S. Suradhat and Y. Poovorawan. 2006. Genetic characterization of H5N1 influenza A viruses isolated from zoo tigers in Thailand. **Virology.** 20; 344(2): 480-91.

- Arnold, S. and M.D. Monto. 2005. The threat of an avian influenza pandemic. **J. Med.** 352(4): 323-324.
- Auewarakul, P., O. Suptawiwat, A. Kongchanagul, C. Sangma, Y. Suzuki, K. Ungchusak, S. Louisirootchanaikul, H. Lerdsamran, P. Pooruk, A. Thitithanyanont, C. Pittayawonganon, C.-T. Guo, H. Hiramatsu, W. Jampangern, S. Chunsutthiwat and P. Puthavathana. 2007. An Avian Influenza H5N1 Virus That Binds to a Human-Type Receptor. **J. Virol.** 81(18):9950-9955.
- Beard, C.W. 1970. Demonstration of type-specific influenza antibody in mammalian and avian sera by immunodiffusion. **Bull World Health Organ.** 42: 779-785.
- Beard, C.W. 1987. To vaccinate or not to vaccinate, pp. 258-263. *In* C.W. Beard (ed.). **Proceedings of the Second International Symposium on Avian Influenza.** U.S. Animal Health Association. Richmond.
- Beard, C.W., Schnitzlein, W.M. and D.N. Tripathy. 1991. Protection of chickens against highly pathogenic avian influenza virus (HSN2) by recombinant fowlpox viruses. **Avian Dis.** 35: 356-359.
- Beby-Defaux, A., Giraudeau, G., Bouguermouh, S. and G. Agius. 2003. Influenza: virological aspects, epidemiology and virological diagnosis. **Medicine et Maladies Infectieuses.** 33: 134-142.
- Boten, E and A. Azzo. 2000. Lysosomal neuraminidase catalytic activation in insect cells is controlled by the protective protein/cathepsin A. **The Journal of Biological Chemistry.** 275(48): 37657-37663.
- Boyce, W.M., C. Sandrock, C. Kreuder-Johnson, T. Kelly and C. Cardona. 2009. Avian influenza viruses in wild birds: A moving target. **Comparative Immunology, Microbiology and Infectious Diseases.** 32(4):275-286.

- Brugh, M. and D.C. Johnson. 1987. Epidemiology of avian influenza in domestic poultry, pp. 177-186. *In Proceedings of the Second International Symposium on Avian Influenza*. U.S. Animal Health Association. Richmond.
- Buxton B.C., Katz, J.M., Seto, W.H., Chan, P.K., Tsang, D. and W. Ho. 2000. Risk of influenza A (H5N1) infection among health care workers exposed to patients with influenza A (H5N1), Hong Kong. *J. Infect Dis.* 181: 344-348.
- Capua, L. F., Mutinelli, M.A. Bozza, Terregino, C. and G. Cattoli. 2000. Highly pathogenic avian influenza (H7N1) in ostriches (*Struthio camelus*). *Avian Pathol.* 29: 643-646.
- Capua, L., Mutinelli, F., Marangon, S. and D.J. Alexander. 2000. H7N1 avian influenza in Italy (1999 to 2000) in intensively reared chickens and turkeys. *Avian Pathol.* 29: 537-543.
- Case, D.A., T.A. Darden, T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, W.S. Ross and P.A. Kollman. 2006. **WYSIWYG 2D plotting tool**. AMBER 9.
- Cauthen, A.N., Swayne, D.E., Schultz-Cherry, S., Perdue, M.L. and D.L. Suarez. 2000. Continued circulation in China of highly pathogenic avian influenza virus encoding the hemagglutinin gene associated with the 1997 H5N1 outbreak in poultry and humans. *J. Virol.* 74(14): 6592-6599.
- Chen, H., Deng, G., Li Z., Tian, G, Li, Y., and P. Jiao. 2004. The evolution of H5N1 influenza viruses in ducks in southern China. *Proc Natl Acad Sci USA.* 101:10452–10457.

- Chen, H., Smith, G.J., Li, K.S., Wang, J., Fan X.H., and J.M. Rayner. 2006. Establishment of multiple sublineages of H5N1 influenza virus in Asia: implications for pandemic control. **Proc Natl Acad Sci.** 103: 2845-2850.
- Claas, E.C., Osterhaus, A.D., Van Beek, R., De Jong, J.C., Rimmelzwaan, G.F., Senne, D.A., Krauss, S., Shortridge, K.F. and R.G. Webster. 1998. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. **THE LANCET.** 351(9101): 472-477.
- Collins, R.A., Ko, L.S., So, K.L., Ellis, T., Lau, L.T. and A.C.H. Yu. 2002. Detection of highly pathogenic and low pathogenic avian influenza subtype H5 (Eurasian lineage) using NASBA. **J. Virol. Methods.** 103: 213-225.
- Cox. N.J., Fuller, F., Kaverin, N., Klenk, H.D., Lamb, R.A., Mahy, B.W., McCauley, J.W., Nakamura, K., Palese, P. and R.G. Webster. 2000. Orthomyxovirid, pp. 585-597. *In* Van Regenmor, M. H., Fauquet, C. M., Bishop, D.H.L., Caretens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R. and R.B. Wickner (eds.). **Virus taxonomy Seventh report of the international committee on taxonomy of viruses.** Academic Press. San Deigo.
- De Marco, M.A., Campitelli, L., Foni, E., Raffini, E., Barigazzi, G., Delogu, M., Guberti, V., Di Trani L., Tollis, M. and I. Donatelli. 2004. Influenza surveillance in birds in Italian wetlands (1992-1998): is there a host restricted circulation of influenza viruses in sympatric ducks and coots?. **Vet. Microbiol.** 98: 197-208.
- Dowdle, W.R. and G.C. Schild. 1975. Laboratory propagation of human influenza viruses, experimental host range and isolation from clinical materials, pp. 257-261. *In* E.D. Kilbourne (ed.). **The influenza viruses and influenza.** Academic Press. New York.

- Easterday, B.C. and B. Tumova. 1978. Avian influenza, pp. 549-573. *In* Hofstad, M.S., Calnek, B.W., Helmbolt, C.F., Reid, W.M. and H.W. Yoder (eds.). **Diseases of poultry.**, 7<sup>th</sup> ed. Iowa State University Press. Los Angelis.
- Easterday, B.C., Hinshaw, V.S. and D.A. Halvorson. 1997. Influenza, pp. 583-605. *In* B.W. Calnek, H.J., Barnes, C.W., Beard, McDougald, L.R. and Y.M. Saif (eds.). **Diseases of Poultry.**, 10th ed. Iowa State University Press. Los Angelis.
- Englund, L., Klingeborn, B. and T. Mejerland. 1986. Avian influenza A virus causing an outbreak of contagious interstitial pneumonia in mink. **Acta. Vet. Scand.** 27: 497-504.
- European Commission. 2000. The definition of avian influenza and the use of vaccination against avian influenza. **Sanco/B3/AH/R17/ 2000:** 1-35.
- Fauci A.S. 2006. Emerging and re-emerging infectious diseases: influenza as a prototype of the host-pathogen balancing act. **Cell.** 124: 665-670.
- Fouchier, R.A., Bestebroer, T.M., Herfst, S., Van Der Kemp, L., Rimmelzwaan, G.F. and A.D. Osterhaus. 2000. Detection of influenza A viruses from different species by PCR amplification of conserved sequences in the matrix gene [In Process Citation]. **J. Clin. Microbiol.** 38: 4096-4101.
- Fouchier R.A., Munster, V., Wallensten, A., Bestebroer, T.M., Herfst, S., Smith, D., Rimmelzwaan, G.F., Olsen, B. and A.D. Osterhaus. 2005. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. **J. Virol.** 79(5): 2814-2822.

- Frank, M. and C.W. Lieth. 1997. Comparison of the Conformational Behavior of Sialyllactose Complexed with the two Viral Attachment Proteins Influenza A Hemagglutinin and the Murine Polyomavirus. **J Mol Model**. 3408 - 414.
- Gambaryan, A.S., A.B. Tuzikov, G.V. Pazynina, R.G. Webster, M.N. Matrosovich and N.V. Bovin. 2004. H5N1 chicken influenza viruses display a high binding affinity for Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-4(6-HSO<sub>3</sub>)GlcNAc-containing receptors. **Virology**. 326(2):310-316.
- Gambaryan, A.S., Tuzikov, A.B., Bovin, N.V., Yamnikova, S.S., Lvov, D.K., Webster, R.G. and M.N. Matrosovich. 2003. Differences between influenza virus receptors on target cells of duck and chicken and receptor specificity of the 1997 H5N1 chicken and human influenza viruses from Hong Kong. **Avian Dis**. 47: 1154-1160.
- Gamblin, S.J., L.F. Haire, R.J. Russell, D.J. Stevens, B. Xiao, Y. Ha, N. Vasisht, D.A. Steinhauer, R.S. Daniels, A. Elliot, D.C. Wiley and J.J. Skehel. 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. **Science**. 303:1838 - 1842.
- Govorkova, E.A., Rehg, J.E., Krauss, S., Yen, H.L., Guan, Y. and M. Peiris. 2005. Lethality to ferrets of H5N1 influenza viruses isolated from humans and poultry in 2004. **J. Virol**. 79: 2191–2198.
- Gregory, V., Bennett, M., Orkhan, M.H., Hajjar, A.I., Varsano, S., Mendelson, N., Zambon, E., Ellis, M., Hay, J. and Y.P. Lin. 2002. Emergence of influenza A H1N2 reassortant viruses in the human population during 2001. **Virology**. 300(1): 1-7.

- Guan, Y., Pieris, J.S., Lipatov, A.S., Ellis, T.M., Dyrting, K.C., Krauss, S., Zhang, L.J., Webster, R.G. and K.F. Shortridge. 2002. Emergence of multiple genotypes of H5N1 avian influenza viruses in Hong Kong SAR. **Proc. Natl. Acad. Sci.** 99(13): 8950-8955.
- Guan, Y., Poon, L.L., Cheung, C.Y., Ellis, T.M., Lim, W. and A.S. Lipatov. 2004. H5N1 influenza: a protean pandemic threat. **Proc. Natl. Acad. Sci. USA.** 101: 8156-8161.
- Guan, Y., Shortridge, K.F., Krauss, S. and R.G. Webster. 1999. Molecular characterization of H9N2 influenza viruses; were they the donors of the “internal” genes of H5N1 viruses in Hong Kong? **Proc. Natl. Acad. Sci. USA.** 96: 9363–9367.
- Ha, Y., D. Stevens, J.J. Skehel and D.C. Wiley. 2002. H5 avian and H9 swine influenza virus haemagglutinin structures: possible origin of influenza subtypes. **EMBO J.** 21:865 - 875.
- Ha, Y., D.J. Stevens, J.J. Skehel and D.C. Wiley. 2001. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. **Proc Natl Acad Sci USA.** 98:11181 - 11186.
- Ha, Y., D.J. Stevens, J.J. Skehel and D.C. Wiley. 2002. H5 avian and H9 swine influenza virus haemagglutinin structures: possible origin of influenza subtypes. **Embo Journal.** 21(5):865-875.
- Halvorson, D.A. 2000. The control of avian influenza, pp. 131-138. **Proceedings of the third international symposium on turkeys diseases.** Germany Veterinary Medical Society. Berlin.

- Halvorson, D.A., Frame, D.D., Friendshuh, K.A.J. and D.P. Shaw. 1998. Outbreaks of low pathogenicity avian influenza in U.S.A., pp 36-46. *In* D.E. Swayne and R.D. Slemons (eds.). **Proceedings of the Fourth International Symposium on Avian Influenza**. U.S. Animal Health Association. Richmond.
- Hampton, T. 2006. Avian flu researchers make strides. **JAMA**. 295: 1107-1108.
- Harvey, R., A.C.R. Martin, M. Zambon and W.S. Barclay. 2004. Restrictions to the Adaptation of Influenza A Virus H5 Hemagglutinin to the Human Host. **J. Virol**. 78(1):502-507.
- Hay, A.J., Zambon, M.C., Wolstenholme, A.J., Skehel, J.J. and M.H. Smith. 1986. Molecular basis of resistance of influenza A viruses to amantadine. **J. Antimicrob. Chemother**. 18 Suppl B: 19-29.
- Hinshaw, V.S., Webster, R.G., Easterday, B.C. and W.J.J. Bean. 1981. Replication of avian influenza A viruses in mammals. **Infect. Immun**. 34: 354-361.
- Hoffmann, E., Stech, J., Leneva, I., Krauss, S., Scholtissek, C., Chin, P.S., Peiris, M., Shortridge, K.F. and R.G. Webster. 2000. Characterization of the influenza A virus gene pool in avian species in southern China: was H5N1 a derivative or a precursor of H5N1?. **J. Virol**. 74(14): 6309-6315.
- Horimoto T, Takada, A., Fujii, K., Goto, H., Hatta, M., and S. Watanabe. 2006. The development and characterization of H5 influenza virus vaccines derived from a 2003 human isolate. **Vaccine**. 24: 3669- 3676.
- Huang, Q., C.-L. Chen and A. Herrmann. 2004. Bilayer Conformation of Fusion Peptide of Influenza Virus Hemagglutinin: A Molecular Dynamics Simulation Study. **Biophysical Journal**. 87(1):14-22.

- Hulse, D.J., Webster, R.G, Russell, R.J. and D.R. Perez. 2004. Molecular determinants within the surface proteins involved in the pathogenicity of H5 N1 influenza viruses in chickens . **J . Virol.** 78(18): 9954-9964.
- Hulse-Post D.J., Sturm-Ramirez, K.M., Humberd, J., Seiler, P., Govorkova, E.A., and S. Krauss. 2005. Role of domestic ducks in the propagation and biological evolution of highly pathogenic H5N1 influenza viruses in Asia. **Proc. Natl. Acad. Sci. USA.** 102: 10682-10687.
- Humphrey, W., A. Dalke and K. Schulten. 1996. VMD: Visual molecular dynamics. **Journal of Molecular Graphics** 14(1):33-38.
- Ilyushina, N.A., Bovin, N.V., Webster, R.G. and E.A. Govorkova. 2006. Combination chemotherapy, a potential strategy for reducing the emergence of drug-resistant influenza A variants. **Antiviral Res.** 70: 121-131.
- Ilyushina, N.A., I.A. Rudneva, A.S. Gambaryan, A.B. Tuzikov and N.V. Bovin. 2004. Receptor specificity of H5 influenza virus escape mutants. **Virus Res.** 100(2):237-241.
- Isoda, N., Sakoda, Y., Kishida, N., Bai, G.R., Matsuda, K. and T. Umemura. 2006. Pathogenicity of a highly pathogenic avian influenza virus, A/chicken/Yamaguchi/7/04 (H5N1) in different species of birds and mammals. **Arch. Virol.** 151: 1267-1279.
- Ito, T., Couceiro, J.N., Kelm, S., Baum, L.G., Krauss, S. and M.R. Castrucci. 1998. Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. **J. Virol.** 72: 7367-7373.

- Ito, T., Y. Suzuki, T. Suzuki, A. Takada, T. Horimoto, K. Wells, H. Kida, K. Otsuki, M. Kiso, H. Ishida and Y. Kawaoka. 2000. Recognition of N-glycolylneuraminic acid linked to galactose by the alpha2,3 linkage is associated with intestinal replication of influenza A virus in ducks. **J Virol.** 74:9300 - 9305.
- Iwata, T., K. Fukuzawa, K. Nakajima, S. Aida-Hyugaji, Y. Mochizuki, H. Watanabe and S. Tanaka. 2008. Theoretical analysis of binding specificity of influenza viral hemagglutinin to avian and human receptors based on the fragment molecular orbital method. **Computational Biology and Chemistry.** 32(3):198-211.
- Jongkon, N., W. Mokmak, D. Chuakheaw, P. Shaw, S. Tongsimma and C. Sangma. 2009. Prediction of avian influenza A binding preference to human receptor using conformational analysis of receptor bound to hemagglutinin. **BMC Genomics.** 10(Suppl 3):S24.
- Kasson, P.M. and V.S. Pande. 2008. Structural Basis for Influence of Viral Glycans on Ligand Binding by Influenza Hemagglutinin. **Biophys J.** 95:L48 - L50.
- Kaverin, N.V., I.A. Rudneva, N.A. Ilyushina, A.S. Lipatov, S. Krauss and R.G. Webster. 2004. Structural differences among hemagglutinins of influenza A virus subtypes are reflected in their antigenic architecture: Analysis of H9 escape mutants. **Journal of Virology.** 78(1):240-249.
- Kendal, A.P. 1982. New techniques in antigenic analysis with influenza viruses, pp. 51-78. In A.S. Beare (ed.). **Basic and applied influenza research.** CRC Press, Inc. Boca Raton.

- Kilbourne, E.D. 1987. **Influenza**. Plenum. New York. 359 p.
- Kitts, P.A., Ayres, M.D. and R.D. Possee. 1990. **Nucleic Acids Res.** 18, 5667.
- Klenk, H.D., Keil, W., Niemann, H., Geyer, R. and R.T. Schwarz. 1983. The characterization of influenza A viruses by carbohydrate analysis. **Curr. Top. Microbiol. Immunol.** 104: 247-257.
- Kodihalli, S., Kobasa, D.L. and R.G. Webster. 2000. Strategies for inducing protection against avian influenza A virus subtypes with DNA vaccines. **Vaccine.** 18: 2592-2599.
- Krug, R.M. 2006. Clues to the virulence of H5N1 viruses in humans. **Science.** 311:1562-1563.
- Kuiken, T., Rimmelzwaan, G., van Riel, D., van Amerongen, G., Baars, M. and R. Fouchier. 2004. Avian H5N1 influenza in cats. **Science.** 306:241.
- Kuroda, K., Geyer, H., Geyer, R., Doerfler, W. and H.D. Klenk. 1990. The oligosaccharides of influenza virus hemagglutinin expressed in insect cells by a baculovirus vector. **Virology.** 174: 418-429.
- Kuroda, K., Hauser, C., Rott, R., Klenk, H.D. and W. Doerfler. 1986. Expression of the influenza virus hemagglutinin in insect cells by a baculovirus vector. **The EMBO Journal.** 5(6): 1359-1365.
- Kuroda, K., Veit, M. and H.D. Klenk. 1991. Retarded processing of influenza virus hemagglutinin in insect cells. **Virology.** 180: 159-165.
- Lamb, R.A. and R.M. Krug. 1996. Orthomyxoviridae: The viruses and their replication, pp. 1353-1395. In B.N. Field, D.M. Knipe and P.M. Howley (eds.). **Fields Virology**. Lippincott Raven. New York.

- Lau, L.T., Bank, J., Aherne, R., Brown, I.H., Dillon, N., Collins, R.A., Chan, K.Y. and Y.W.W. Fung. 2004. Nucleic acid sequence based amplification methods to detect avian influenza virus. **Biochemical and Biophysical Research Communications**. 313: 336-342.
- Le, Q.M., Kiso, M., Someya, K., Sakai, Y.T., Nguyen, T.H. and K.H.N. Guyen. 2005. Isolation of drug-resistant H5N1 virus. **Nature**. 437:1108.
- Lee, C.W., D.L. Suarez, T.M. Tumpey, H.W. Sung, Y.K. Kwon, Y.J. Lee, J.G. Choi, S.J. Joh, M.C. Kim, E.K. Lee, J.M. Park, X. Lu, J.M. Katz, E. Spackman, D.E. Swayne and J.H. Kim. 2005. Characterization of highly pathogenic H5N1 avian influenza A viruses isolated from South Korea. **J Virol**. 79:3692 - 3702.
- Lee, M.S., Chang, P.C., Shien, J.H., Cheng, M.C. and H.K. Shieh. 2001. Identification and subtyping of avian influenza viruses by reverse transcription-PCR. **J. Virol. Method**. 97: 13-22.
- Lewis, D.B. 2006. Avian flu to human influenza. **Annu. Rev. Med.** 57: 139-154.
- Li, M. and B. Wang. 2006. Computational studies of H5N1 hemagglutinin binding with SA- $\alpha$ -2, 3-Gal and SA- $\alpha$ -2, 6-Gal. **Biochemical and Biophysical Research Communications**. 347(3):662-668.
- Lieth, C.W. and T. Kozar. 1996. Towards a better semiquantitative estimation of binding constants: molecular dynamics exploration of the conformational behavior of isolated sialyllactose and sialyllactose complexed with influenza A hemagglutinin. **J Mol Struct. (Theochem)** 368:213 - 222.
- Mahoney, M.W. and W. Jorgenson. 2000. GLYCAM force field parameters A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. **J Chem Phys**. 112:8910 - 8922.

- Manvell, R.J., McKinney, P., Wernery, U. and K.M. Frost. 2000. Isolation of a highly pathogenic influenza A virus of subtype H7N3 from a peregrine falcon (*Falco peregrinus*). **Avian Pathol.** 29: 635-637.
- Martin, J., S.A. Wharton, Y.P. Lin, D.K. Takemoto, J.J. Skehel, D.C. Wiley and D.A. Steinhauer. 1998. Studies of the Binding Properties of Influenza Hemagglutinin Receptor-Site Mutants. **Virology.** 241(1):101-111.
- Matrosovich, M., Zhou, N., Kawaoka, Y. and R. Webster. 1999. The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties. **J. Virol.** 73:1146-1155.
- Minyong, L. and W. Binghe. 2006. Computational studies of H5N1 hemagglutinin binding with SA-alpha-2, 3-Gal and SA-alpha-2, 6-Gal. **Biochem Biophys Res Commun.** 347662 - 668.
- Nicholson, K.G., Wood, J.H. and M. Zambon. 2003. Influenza. **THE LANCET.** 362: 1733-1745.
- Nobusawa, E., T. Aoyama, H. Kato, Y. Suzuki, Y. Tateno and K. Nakajima. 1991. Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. **Virology.** 182475 - 485.
- Normile, D. 2004. Ducks may magnify threat of avian flu virus. **Science.** 306: 953.
- Normile, D. 2006. Evidence points to migratory birds in H5N1 spread. **Science.** 311: 1225.

- Office International des Epizootics (OIE). 1992. **Code Commission. International animal health code**. 6th ed. Office International des Epizooties. Paris. 550 p.
- Office International des Epizootics (OIE). 2004. **OIE Manual of Diagnostic Test and Vaccines for Terrestrial Animals (Highly Pathogenic Avian Influenza)**. OIE Terrestrials Manual. Paris. 12 p.
- Palese, P. and A. Garcia-Sastre. 1999. Influenza Viruses (Orthomyxoviruses): Molecular Biology, pp. 830-836. *In* Granoff, A. and R.G. Webster (eds.). Encyclopedia of Virology: Vol. 2, 2nd ed. **Academic Press**. San Diego.
- Panigrahy, B., D.A. Senne and J.C. Pedersen. 2002. Avian influenza virus subtypes inside and outside the live bird markets, 1993-2000: A spatial and temporal relationship. **Avian Diseases**. 46(2):298-307.
- Panigrahy, B., Senne, D.A. and J.E. Pearson. 1995. Presence of avian influenza virus (AIV) subtypes H5N2 and H7N1 in emus (*Dromaius novaehollandiae*) and rheas (*Rhea americana*): Virus isolation and serologic findings. **Avian Dis**. 39: 64-67.
- Perkins, L.E.L. and D.E. Swayne. 2001. Pathobiology of A/chicken/Hong Kong/220/97 (H5N1) avian influenza virus in seven gallinaceous species. **Vet. Pathol**. 38: 149-164.
- Philpott, M., C. Hioe, M. Sheerar and V.S. Hinshaw. 1990. Hemagglutinin mutations related to attenuation and altered cell tropism of a virulent avian influenza A virus. **J. Virol**. 64(6):2941-2947.
- Puthavathana, P., P. Auewarakul, P.C. Charoenying, K. Sangsiriwut, P. Pooruk, K. Boonnak, R. Khanyok, P. Thawachsupa, R. Kijphati and P. Sawanpanyalert. 2005. Molecular characterization of the complete genome of human influenza H5N1 virus isolates from Thailand. **J Gen Virol**. 86:423 - 433.

- Rabadan, R., A.J. Levine and H. Robins. 2006. Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. **J Virol.** 80:11887 - 11891.
- Riedel, S. 2006. Crossing the species barrier: the threat of an avian influenza pandemic. **Proc (Bayl Univ Med Cent).** 19:16-20.
- Rowe, T., Abernathy, R.A., Primmer, J.H., Thompson, W.W., Lu, X., Lim, W., Fukuda, K., Cox, N.J. and J.M. Katz. 1999. Detection of antibody to avian influenza A (H5N1) virus in human serum by using a combination of serologic Assays. **J. Clin. Microb.** 37: 937-943.
- Russell, C.J. and R.G. Webster. 2005. The genesis of a pandemic influenza virus. **Cell.** 10: 368-371.
- Rychaert, J.P., G. Ciccotti and H.J.C. Berendsen. 1997. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. **J Comput Phys.** 23327 - 341.
- Sambrook, J. and D.W. Russell. 2001. **Molecular Cloning: A Laboratory Manual.** 3<sup>rd</sup> ed. Spring Harbor Laboratory Press. New York. 631 p.
- Sammalkorpi, M. and T. Lazaridis. 2007. Configuration of influenza hemagglutinin fusion peptide monomers and oligomers in membranes. **Biochimica Et Biophysica Acta-Biomembranes.** 1768(1):30-38.
- Sangma, C., P. Nunrium and S. Hannongbua. 2006. The Analysis of Binding Patterns on Different Receptors Bound to Hemagglutinin of Avian and Avian-like Influenza Virus Using Quantum Chemical Calculations. **J Theor and Comp Chem.** 5753 - 768.

- Scholtissek, C., Quack, G., Klenk, H.D. and R.G. Webster. 1998. How to overcome resistance of influenza A viruses against adamantane derivatives. **Antiviral Res.** 37: 83-95.
- Schwede, T., J. Kopp, N. Guex and M.C. Peitsch. 2003. SWISS-MODEL: an automated protein homology-modeling server. **Nucl. Acids Res.** 31(13):3381-3385.
- Schweiger, B., Zadow, I., Heckler, R., Timm, H. and G. Paul. 2000. Application of a fluorogenic PCR assay for typing and subtyping of influenza viruses in respiratory samples. **J. Clin. Microb.** 38: 1552-1558.
- Shinya, K., M. Ebina, S. Yamada, M. Ono, N. Kasai and Y. Kawaoka. 2006. Avian flu: Influenza virus receptors in the human airway. **Nature.** 440(7083):435-436.
- Skehel, J.J. and D.C. Wiley. 2000. RECEPTOR BINDING AND MEMBRANE FUSION IN VIRUS ENTRY: The Influenza Hemagglutinin. **Annual Review of Biochemistry.** 69(1):531.
- Slemons, R.D. and D.E. Swayne. 1990. Replication of a waterfowl -origin influenza virus in the kidney and intestine of chickens. **Avian Dis.** 34: 277-284.
- Smirnov, Y.A., A.K. Gitelman, E.A. Govorkova, A.S. Lipatov and N.V. Kaverin. 2004. Influenza H5 virus escape mutants: immune protection and antibody production in mice. **Virus Res.** 99(2):205-208.
- Songserm, T., Amonsin, A., Jam-on, R., and Y. Poovarawan. 2006. Fatal avian influenza A H5N1 in a dog. **Emerg Infect Dis.** 12: 1744-1747.

- Songserm, T., Amonsin, A., Jam-on, R., Sae-Heng, N., Meemak, N., Pariyothorn, N., et al. 2006. Avian influenza H5N1 in naturally infected domestic cat. **Emerg Infect Dis.** 12:681-3.
- Staliknecht, D.E. and S.M. Shane. 1988. Host range of avian influenza virus in freeliving birds. **Veterinary Research Communications.** 12: 125-141.
- Starick, E., Romer-Oberdorfer, A. and O. Werner. 2000. Type-and subtype specific RT-PCR assays for avian influenza A virus. **J. Vet. Med.** 47: 295-301.
- Stevens, J., O. Blixt, L. Glaser, J.K. Taubenberger, P. Palese, J.C. Paulson and I.A. Wilson. 2006. Glycan Microarray Analysis of the Hemagglutinins from Modern and Pandemic Influenza Viruses Reveals Different Receptor Specificities. **Journal of Molecular Biology.** 355(5):1143-1155.
- Stevens, J., O. Blixt, T.M. Tumpey, J.K. Taubenberger, J.C. Paulson and I.A. Wilson. 2006. Structure and Receptor Specificity of the Hemagglutinin from an H5N1 Influenza Virus. **Science.** 312(5772):404-410.
- Stone, H., Mitchell, B. and M. Brugh. 1997. *In ovo* vaccination of chicken embryos with experimental Newcastle disease and avian influenza oil-emulsion vaccines. **Avian Dis.** 41: 856-863.
- Stubbs, E.L. 1948. Fowl pest, pp. 603-614. *In* H. E. Biester and L. H. Schwarte (eds.). **Diseases of Poultry**, 2<sup>nd</sup> ed. Iowa State University Press. Los Angelis.
- Sturm-Ramirez, K.M., Ellis, T., Bousfield, B., Bissett, L., Dyrting, K. and J.E. Rehg. 2004. Reemerging H5N1 influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. **J. Virol.** 78:4892-4901.

- Su, Y., H.Y. Yang, B.J. Zhang, H.L. Jia and P. Tien. 2008. Analysis of a point mutation in H5N1 avian influenza virus hemagglutinin in relation to virus entry into live mammalian cells. **Archives of Virology**. 153(12):2253-2261.
- Suarez, D.L. and C. S. Schultz. 2000. Immunology of avian influenza virus: A review. **Dev. Comp. Immunol.** 24: 269-283.
- Subbarao, K. and J. Katz. 2000. Avian influenza viruses infecting humans. **Cell. Mol. Life. Sci.** 57(12): 1770-1784.
- Subbarao, K., Klimov, A., Katz, J., Regnery, H., Lim, W., Hall, H., Perdue, M., Swayne, D., Bender, C., Huang, J., Hemphill, M., Rowe, T., Shaw, M., Xu, X., Fukuda, K. and N. Cox. 1998. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. **Science** 279(5349): 393-396.
- Suptawiwat, O., A. Kongchanagul, W. Chan-It, A. Thitithanyanont, W. Wiriyarat, K. Chaichuen, T. Songserm, Y. Suzuki, P. Puthavathana and P. Auewarakul. 2008. A simple screening assay for receptor switching of avian influenza viruses. **J Clin Virol.** 42:186 - 189.
- Swayne, D.E., Beck, J.R. and T.R. Mickle. 1997. Efficiency of recombinant fowlpox virus vaccine in protecting chickens against a highly pathogenic Mexican-origin H5N2 avian influenza virus. **Avian Dis.** 41: 910-922.
- Swayne, D.E., Perdue, M.L., Beck, J.R., Garcia, M. and D.L. Suarez. 2000b. Vaccines protect chickens against H5 highly pathogenic avian influenza in the face of genetic changes in field viruses over multiple years. **Vet. Microbiol.** 74: 165-172.

- Swayne, D.L., Garcia, M., Beck, J.R., Kinney, N. and D.L. Suarez. 2000a. Protection against diverse highly pathogenic H5 avian influenza viruses in chickens immunized with a recombinant fowlpox vaccine containing an H5 avian influenza hemagglutinin gene insert. **Vaccine**. 18: 1088-1095.
- Swayne, J.R., Glasson, M.W., Jackwood, Pearson, J.E. and W.M. Reed (eds.). **Isolation and Identification of Avian Pathogens**, 4<sup>th</sup> ed. American Association of Avian Pathologist. Kennett Square.
- Taubenberger, J.K., Reid, A.H., Lourens, R.M., Wang, R., Jin, G., Fanning, T.G. 2005. Characterization of the 1918 influenza virus polymerase genes. **Nature**. 437:889-893.
- Tumpey, T.M., Suarez, D.L., Perkins, L.E., Senne, D.A., Lee, J.G., Lee, Y.J., Sung, I.P., Mo, H.W. and D.E. Swayne. 2002. Characterization of a highly pathogenic H5N1 avian influenza A virus isolated from duck meat. **J. Virol**. 76(12): 6344-6355.
- Ungchusak, K., Auewarakul, P., Dowell, S.F., Kitphati, R., Auwanit, W. and P. Puthavathana. 2005. Probable person-to-person transmission of avian influenza A (H5N1). **N. Engl. J. Med.** 352: 333-340.
- van Riel, D., V.J. Munster, E. de Wit, G.F. Rimmelzwaan, R.A.M. Fouchier, A.D.M.E. Osterhaus and T. Kuiken. 2006. H5N1 Virus Attachment to Lower Respiratory Tract. **Science**. 312(5772):399.
- Vines, A., Wells, K., Matrosovich, M., Castrucci, M.R., Ito, T. and Y. Kawaoka 1998. The role of influenza A virus hemagglutinin residues 226 and 228 in receptor specificity and host range restriction. **J. Virol**. 72: 7626-7631.

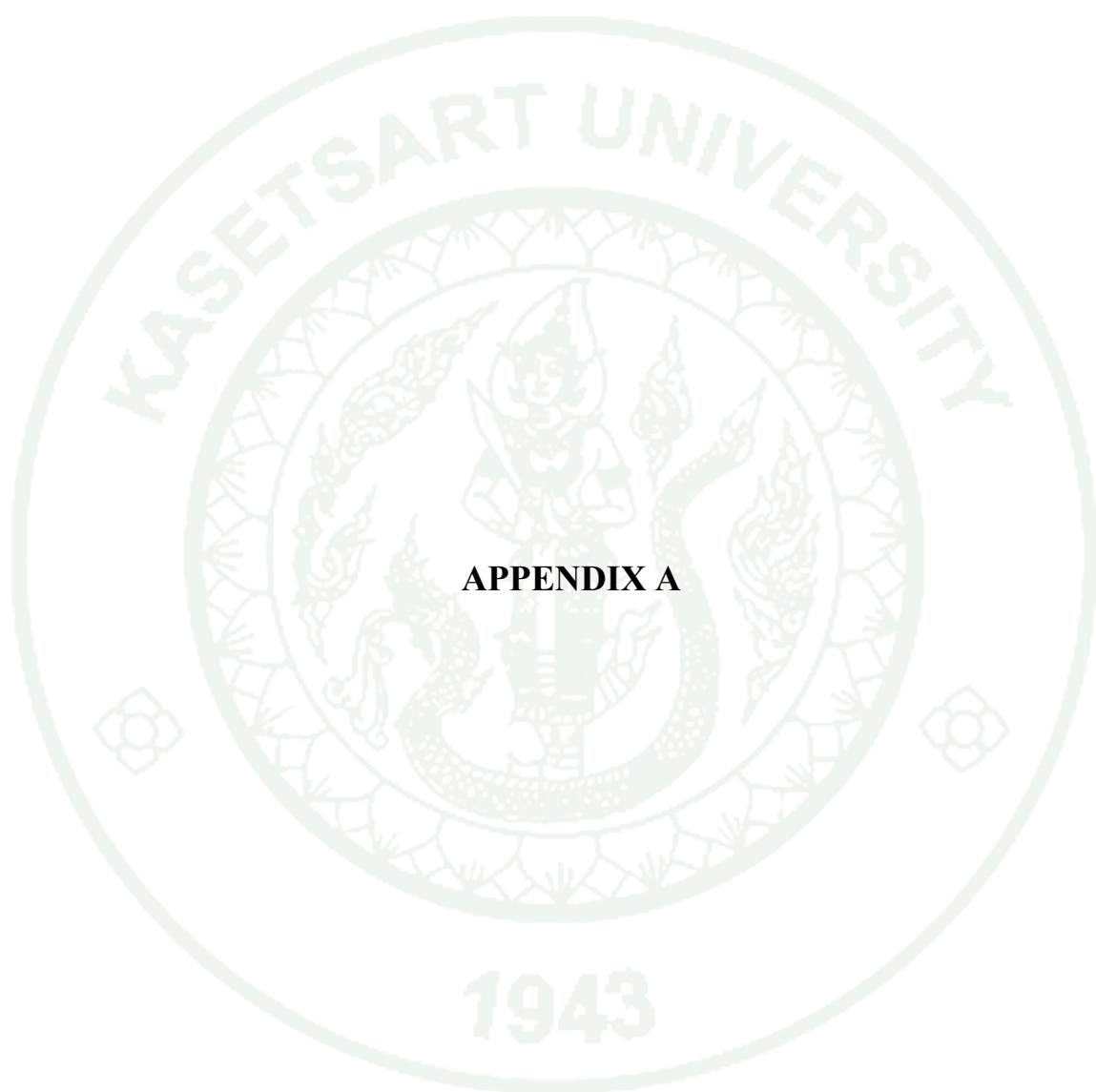
- Viseshakul, N., R. Thanawongnuwech, A. Amonsin, S. Suradhat, S. Payungporn, J. Keawchareon, K. Oraveerakul, P. Wongyanin, S. Plitkul, A. Theamboonlers and Y. Poovorawan. 2004. The genome sequence analysis of H5N1 avian influenza A virus isolated from the outbreak among poultry populations in Thailand. **Virology**. 328:169 - 176.
- Walls, H.H., Harmon, M.W., Slangue, J.J., Stocksdales, C. and A.R. Kendal. 1986. Characterization and evaluation of monoclonal antibodies developed for typing influenza A and influenza B viruses. **J. Clin. Microbiol.** 23: 240-245.
- Wan, H. and D.R. Perez. 2006. Quail carry sialic acid receptors compatible with binding of avian and human influenza viruses. **Virology**. 346(2):278-286.
- Webster, R.G., Bean, W. J., Gorman, O.T., Chambers, T.M. and Y. Kawaoka. 1992. Evolution and ecology of influenza A viruses. **Microbiol. Rev.** 56: 152-179.
- Webster, R.G., Guan, Y., Peiris, M., Walker, D., Krauss, S. and N.N. Zhou. 2002. Characterization of H5N1 influenza viruses that continue to circulate in geese in southeastern China. **J. Virol.** 76: 118-126.
- Webster, R.G., Kawaoka, Y., Bean, W.J., Beard, C.W. and M. Brugh. 1995. Chemotherapy and vaccination: A possible strategy for the control of highly virulent influenza virus. **J. Virol.** 55: 173-176.
- Weiss, R.A. 2003. Cross-species infections. **Curr Top Microbiol Immunol.** 278:47 - 71.
- Westbury, H.A. 1998. History of highly pathogenic avian influenza in Australia, pp. 23-30. In D.E. Swayne and R.D. Slemons (eds.). **Proceedings of the fourth international symposium on avian influenza**. U.S. Animal Health Association. Richmond.

- Wood, J.M. and J.S. Robertson. 2004. From lethal virus to life-saving vaccine: developing inactivated vaccines for pandemic influenza. **Nat. Rev. Microbiol.** 2: 842-847.
- Wu, G. and S. Yan. 2006. Prediction of possible mutations in H5N1 hemagglutinins of influenza A virus by means of logistic regression. **Comparative Clinical Pathology.** 15(4):255-261.
- Wu, G. and S.M. Yan. 2006. Determination of mutation trend in hemagglutinins by means of translation probability between RNA codons and mutated amino acids. **Protein and Peptide Letters.** 13(6):601-609.
- Wu, G. and S.M. Yan. 2007. W improvement of prediction of mutation positions in H5N1 Hemagglutinins of influenza A virus using neural network with distinguishing of arginine, leucine and serine. **Protein and Peptide Letters.** 14(5):465-470.
- Wu, W.L., Y. Chen, P. Wang, W. Song, S.-Y. Lau, J.M. Rayner, G.J.D. Smith, R.G. Webster, J.S.M. Peiris, T. Lin, N. Xia, Y. Guan and H. Chen. 2008. Antigenic Profile of Avian H5N1 Viruses in Asia from 2002 to 2007. **J. Virol.** 82(4):1798-1807.
- Xu, D., E.I. Newhouse, R.E. Amaro, H.C. Pao, L.S. Cheng, P.R.L. Markwick, J.A. McCammon, W.W. Li and P.W. Arzberger. 2009. Distinct Glycan Topology for Avian and Human Sialopentasaccharide Receptor Analogues upon Binding Different Hemagglutinins: A Molecular Dynamics Perspective. **J Mol Biol.** 387465 - 491.
- Xu, X., Cox, N.J. and Y. Guo. 1999. Genetic characterization of the pathogenic influenza A/Goose/Guangdong/ 1/96 (H5N1) virus: similarity of its hemagglutinin gene to those of H5N1 viruses from the 1997 outbreaks in Hong Kong. **Virology.** 261: 15-19.

- Yamada, S., Y. Suzuki, T. Suzuki, M.Q. Le, C.A. Nidom, Y. Sakai-Tagawa, Y. Muramoto, M. Ito, M. Kiso, T. Horimoto, K. Shinya, T. Sawada, M. Kiso, T. Usui, T. Murata, Y. Lin, A. Hay, L.F. Haire, D.J. Stevens, R.J. Russell, S.J. Gamblin, J.J. Skehel and Y. Kawaoka. 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. **Nature**. 444(7117):378-382.
- Yang, Z.-Y., C.-J. Wei, W.-P. Kong, L. Wu, L. Xu, D.F. Smith and G.J. Nabel. 2007. Immunization by Avian H5 Influenza Hemagglutinin Mutants with Altered Receptor Binding Specificity. **Science**. 317(5839):825-828.
- Yao, Y., Mingay, L.J., Mccauley, J.W. and W.S. Barclay. 2001. Sequences in influenza A virus PB2 protein that determine productive infection for an avian influenza virus in mouse and human cell lines. **J. Virol**. 75: 5410-5415.
- Zhou, E.M., Chan, M., Heckert, R.A., Riva, J. and M.F. Cantin. 1998. Evaluation of a competitive ELISA for detection of antibodies against avian influenza virus nucleoprotein. **Avian Dis**. 43: 517-522
- Zhou, H., M. Jin, H. Chen, Q. Huag and Z. Yu. 2006. Genome-sequence Analysis of the Pathogenic H5N1 Avian Influenza A Virus Isolated in China in 2004. **Virus Genes**. 3285 - 95.



**APPENDICES**



**APPENDIX A**

## Theory of Molecular Dynamics Simulations

One of the principal tools in the theoretical study of biological molecules is the method of molecular dynamics simulations (MD). This computational method calculates the time dependent behavior of a molecular system. MD simulations have provided detailed information on the fluctuations and conformational changes of proteins and nucleic acids. These methods are now routinely used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes. They are also used in the determination of structures from x-ray crystallography and from NMR experiments.

Biological molecules exhibit a wide range of time scales over which specific processes occur; for example

- Local motions (0.01 to 5 Å,  $10^{-15}$  to  $10^{-1}$  s)
  - Atomic fluctuations
  - Side chain motions
  - Loop motions
- Rigid body motions (1 to 10Å,  $10^{-9}$  to 1s)
  - Helix motions
  - Domain motions (hinge bending)
  - Subunit motions
- Large-Scale motions ( $> 5$ Å,  $10^{-7}$  to  $10^4$  s)
  - Helix coil transitions
  - Dissociation or association
  - Folding and unfolding

## 1. Main principle

The main principle of M.D. simulation is as follows: given the system state  $S(t_0)$ , that is, the position  $\mathbf{r}$  and velocity  $\mathbf{v}$  of every particle (atom) in the system at time  $t_0$ , subsequent states  $S(t_0 + \Delta t)$ ,  $S(t_0 + 2\Delta t)$ , ..., are calculated by using Newton's law  $F = ma$ . For accurate results small timesteps  $\Delta t$  have to be used. To calculate  $S(t_0+(n+1)\Delta t)$  from  $S(t_0+n\Delta t)$ , first for every particle  $i$ ,  $F_i(t_0+n\Delta t)$  is calculated.  $F_i(t_0 + n\Delta t)$  is the sum of the forces on  $i$  as exerted by the other particles of the system at time  $t_0 + n\Delta t$ . For every particle  $i$  the force  $F_i(t_0+n\Delta t)$  is then integrated to get the new velocity  $\mathbf{v}_i(t_0 + n\Delta t)$ . Using this velocity, for every particle  $i$  the new position  $\mathbf{r}_i(t_0 + (n+1)\Delta t)$  can be calculated (Figure 4A).

## 2. Integration

A widely used, simple and numerically stable, integration algorithm is the *leapfrog* algorithm. In this algorithm particle positions are calculated at times  $t_0 + n\Delta t$  and velocities at midpoints, i.e. at  $t_0 + (n + \frac{1}{2})\Delta t$  (this differs from the scheme in the previous paragraph where particle positions and velocities were both evaluated at  $t_0 + n\Delta t$ ). With  $t_n \equiv t_0 + n\Delta t$ , the formulas for leap-frog integration are

$$\mathbf{v}_i(t_n + \Delta t/2) = \mathbf{v}_i(t_n - \Delta t/2) + m^{-1}\mathbf{F}_i(t_n)\Delta t \quad (2.1)$$

$$\mathbf{r}_i(t_n + \Delta t) = \mathbf{r}_i(t_n) + \mathbf{v}_i(t_n + \Delta t/2)\Delta t \quad (2.2)$$

## 3. Interaction forces

In an M.D. simulation the forces between particles only depend on particle positions, not on velocities. Usually, interactions are specified by giving an expression for the potential energy of the interaction, hence the force can be written as a gradient of the potential.

#### 4. Non-bonded interactions

Two classes of interactions may be distinguished: non-bonded interactions and bonded interactions. Non-bonded interactions model flexible interactions between particle pairs. Two well-known non-bonded interaction potentials are the Coulomb potential and the Lennard-Jones potential. Using the convention that the distance between the particles  $i$  and  $j$  is defined as  $r_{ij} \equiv |r_{ij}| \equiv |r_i - r_j|$ , the Coulomb potential can be written as

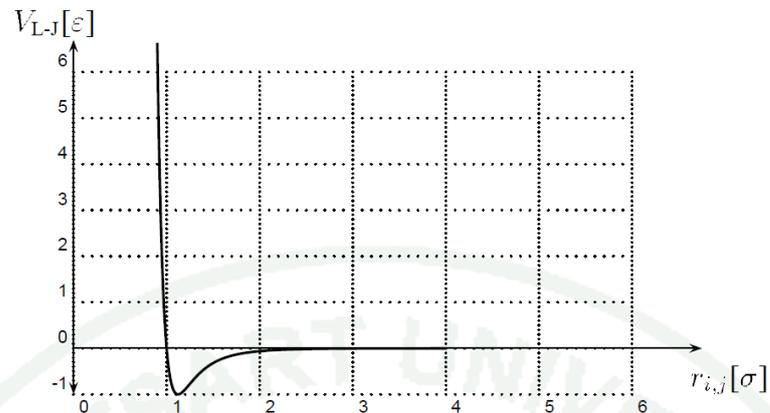
$$V_{Coul} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \quad (4.1)$$

Here  $q_i$  and  $q_j$  are the charges of particles  $i$  resp.  $j$ . The Coulomb force on particle  $i$ , due to particle  $j$  is given by

$$\mathbf{F}_{ij \text{ Coul.}} = \frac{q_i q_j}{4\pi\epsilon_0} \frac{\mathbf{r}_{ij}}{r_{ij}^3} \quad (4.2)$$

The Lennard-Jones potential is given by

$$V_{L-J} = 4\epsilon \left( \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right) \quad (4.3)$$



**Figure 1A** Lennard-Jones potential (equation 4.3). The depth of the potential well is determined by  $\epsilon$ , and the diameter of the particle by  $\sigma$ .

Where  $\epsilon$  is a constant determining the depth of the potential well, and where  $\sigma$  determines the diameter of the particle (see Figure 1A). The term  $\left(\frac{\sigma}{r_{ij}}\right)^{12}$  models a strong repelling force at very short distances, and the term  $-\left(\frac{\sigma}{r_{ij}}\right)^6$  models an attracting force with a longer interaction range. Adding these two terms gives a potential well. The force of the Lennard-Jones interaction exerted on particle  $i$  by particle  $j$  is given by

$$\mathbf{F}_{ij} = -\nabla_i V_{L-J}(r_{ij}) = 4\epsilon \left( 12 \left(\frac{\sigma}{r_{ij}}\right)^{12} - 6 \left(\frac{\sigma}{r_{ij}}\right)^6 \right) \left(\frac{\mathbf{r}_{ij}}{r_{ij}^2}\right) \quad (4.4)$$

Due to the ‘action = - reaction’ principle (Newton’s third principle), the forces between two interacting particles are related by

$$\mathbf{F}_{ji} = -\mathbf{F}_{ij} \quad (4.5)$$

So, the force between every particle pair  $i, j$  has to be evaluated only once instead of twice. The usual way to evaluate every interaction only once is to use the

$j > i$  criterion, i.e., to calculate explicitly the force on the particle with the highest particle number. The force on the other particle is calculated with (4.5).

## 5. Bonded interactions

Bonded interactions model rather strong chemical bonds, and are not created or broken during a simulation. For this reason, these interactions may be evaluated by running through a *fixed list* of groups of particle numbers, where each group represents a bonded interaction between two or more particles. The three most widely used bonded interactions are the covalent interaction, the bond-angle interaction, and the dihedral interaction.

The covalent interaction is a bonded interaction between two particles  $i$  and  $j$  with interaction potential

$$V = \frac{1}{2} K_b (r_{ij} - b_0)^2 \quad (5.1)$$

This interaction may be thought of as a very stiff linear spring between  $i$  and  $j$ . The spring has a neutral length  $b_0$  and a spring constant  $K_b$ . The force of this interaction is given by

$$\mathbf{F}_i = -K_b (r_{ij} - b_0) (\mathbf{r}_{ij} / r_{ij}) \quad (5.2)$$

The bond-angle interaction is a three particle interaction between  $i, j, k$  (Figure 2A; a), with interaction potential

$$V = \frac{1}{2} K_\Theta (\Theta - \Theta_0)^2 \quad (5.3)$$

with

$$\Theta \equiv \arccos \left( \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{r_{ij} r_{kj}} \right) \quad (5.4)$$

(Valid for small  $(\Theta - \Theta_0)$ .) This interaction may be thought of as a torsion spring between the lines  $i, j$  and  $k, j$ . The spring has a neutral angle  $\Theta_0$  and a spring constant  $K_\Theta$ .

The dihedral-angle interaction  $V(\phi)$  is a four particle interaction between  $i, j, k, l$  (Figure 2A; b) with an interaction potential  $V = V(\phi)$ . Two often used expressions for  $V$  are

$$V = K_\phi (1 + \cos(n\phi - \delta)) \text{ and } V = \frac{1}{2} K_\phi (\phi - \phi_0)^n \in \mathbb{Z} \quad (5.5)$$

where  $\delta$  and  $\phi_0$  are constants. The definition of the dihedral angle  $\phi$  is given by

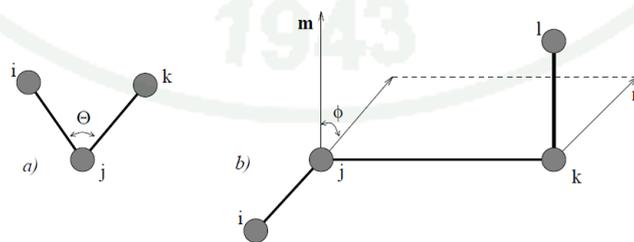
$$\phi \equiv \text{sign}(\phi) \arccos(\hat{\mathbf{m}} \cdot \hat{\mathbf{n}}) \quad (5.6)$$

$$\mathbf{m} \equiv \mathbf{r}_{ij} \times \mathbf{r}_{kj} \quad (5.7)$$

$$\mathbf{n} \equiv \mathbf{r}_{kj} \times \mathbf{r}_{lj} \quad (5.8)$$

$$\text{sign}(\phi) \equiv \frac{\mathbf{r}_{ij} \cdot \mathbf{n}}{|\mathbf{r}_{ij} \cdot \mathbf{n}|} \quad (5.9)$$

where  $\hat{\mathbf{r}} \equiv \frac{\mathbf{r}}{r}$



**Figure 2A** a: Bond-angle interaction.  $\Theta$  is the angle between the lines  $i, j$  and  $k, j$ .

b: Dihedral interaction with negative  $\phi$ .  $\mathbf{m}$  is normal to the  $i, j, k$  plane,  $\mathbf{n}$  is normal to the  $j, k, l$  plane.

## 6. Cut-off radius and neighbour searching

In principle, a non-bonded interaction exists between every particle pair. Because most of the CPU (Central Processing Unit) time of M.D. simulation is spent in non-bonded force calculations, the greatest gain in performance can be achieved by efficiency improvements in this part. Two optimisations are widely applied: the use of a cut-off radius, and the use of neighbour lists.

Earlier we proposed to evaluate all pair interactions, no matter how far the particles are separated. However, the main contribution to the total force on a particle is from neighbouring particles. Therefore, only a small error is introduced when only interactions are evaluated between particles with a distance less than a cut-off radius  $R_{co}$ . Choosing  $R_{co}$  so that for each particle 100 to 300 other particles are within cut-off radius gives a good balance between correct physics and efficiency. For a system of  $10^4$  particles this makes the non-bonded force computation a factor  $\frac{10^4}{100}$  to  $\frac{10^4}{300}$  faster.

Still, when using a cut-off radius, all pairs have to be inspected to see if their separation is less than  $R_{co}$ . This is called neighbour searching. However, the timesteps made, are so small that particles travel only a very small distance during one timestep. In other words, the set of particles within  $R_{co}$  of a given particle hardly changes during one timestep. Therefore, only a small error is made when only every 10 or 20 timesteps all pairs are inspected to see if their distance is less than  $R_{co}$ . Pairs with a distance less than  $R_{co}$  are stored in so called neighbour lists which are used the next 10 or 20 timesteps. In this way the all-pairs inspection every timestep is replaced by an all-pairs inspection every 10 or 20 timesteps at the cost of some memory space to store the neighbour list of every particle. Using  $R_{co}$  and a neighbour list, non-bonded force calculations still take about 70% of the total CPU time.

## 7. Periodic boundary conditions

Because of the limited CPU power, in current M.D. simulations typically  $10^3 \dots 10^5$  particles are involved. Single systems of this size suffer strongly from finite system effects. For example, due to the surface tension of water (or Laplace pressure), the pressure in a spherical droplet of water consisting of  $2 \times 10^4$  molecules, will be approximately 275 bar. Many other anomalies are introduced by using small, single systems. Therefore, most simulations are done with periodic boundary conditions. This means that the simulation takes place in a computational box, which is virtually surrounded by an infinite number of identical replica boxes, stacked in a space filling way, all with exactly the same contents (Figure 3A). Only the behaviour of one box, the 'central box', has to be simulated; other boxes behave in the same way. When periodic boundary conditions are used particles may freely cross box boundaries. For each particle leaving the box, at the same instant an identical particle from an adjacent replica box enters the box at the opposite side. In M.D. system with periodic boundary conditions particles are influenced by particles in their own box and particles in surrounding boxes.

The shape of the computational box should be such that it can be stacked in a space filling way. For reasons of efficiency only convex boxes are used. In 3-D space there exist five box types with these properties: the triclinic box, the hexagonal prism, two types of dodecahedrons, and the truncated octahedron. A system with periodic boundary conditions is an infinite system, but has a crystal-like long range order. Ideally, one would like to have a system without this long range order. By choosing  $R_{co}$  not too large, the long range order effects are limited.

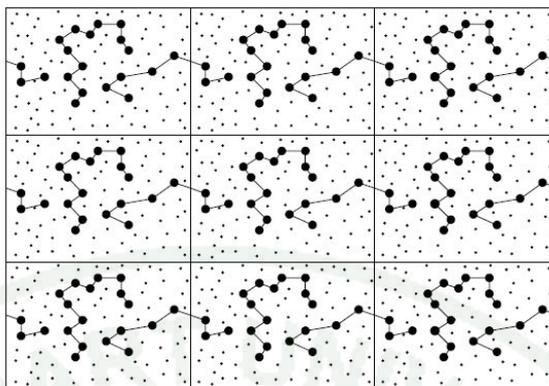


Figure 3A 2-D Periodic Boundary Conditions. One box is surrounded by eight identical boxes.

## 8. Constraint dynamics

Every timestep, during the force calculations, many types of interaction forces are evaluated: Coulomb forces, Lennard-Jones forces, covalent forces, etc. Some of these interactions are very rigid. The most rigid interaction in M.D. simulation is the covalent interaction. This means that two particles having a covalent interaction, have an almost constant distance, or put in another way, two particles with a covalent bond vibrate with a high frequency. The maximally allowed timestep used in M.D. simulation is dictated by the allowed numerical drift of the integration algorithm, so it is dictated by the highest frequency in the system, and should be approximately  $1/(40 \times \text{highest frequency})$ . However, the behaviour of covalent interactions is not part of the physics of interest of an M.D. simulation because covalent vibrations are only weakly coupled to the other vibrations of the system. Leaving out frequencies above  $1/4$  to  $1/2$  the highest covalent eigenfrequency does not influence the outcome of M.D. simulation. So, it is a waste of computer time to use a timestep based on covalent eigenfrequencies. For that reason, nowadays in most M.D. programs, the covalent interactions are handled using *constraint dynamics*, which means that the distance between particles with a covalent bond is kept constant. Then the timestep may be as high as  $1/20$  to  $1/10$  ( $\times$ highest frequency). In this way, the same time span of physics can be simulated two to four times faster.

Because an atom may have covalent interactions with a number of atoms, substituting covalent interactions with length constraints will in general result in a set of connected length constraints with a, possibly cyclic, graph-like structure. In a typical M.D. system the number of constraints is of the same order as the number of particles.

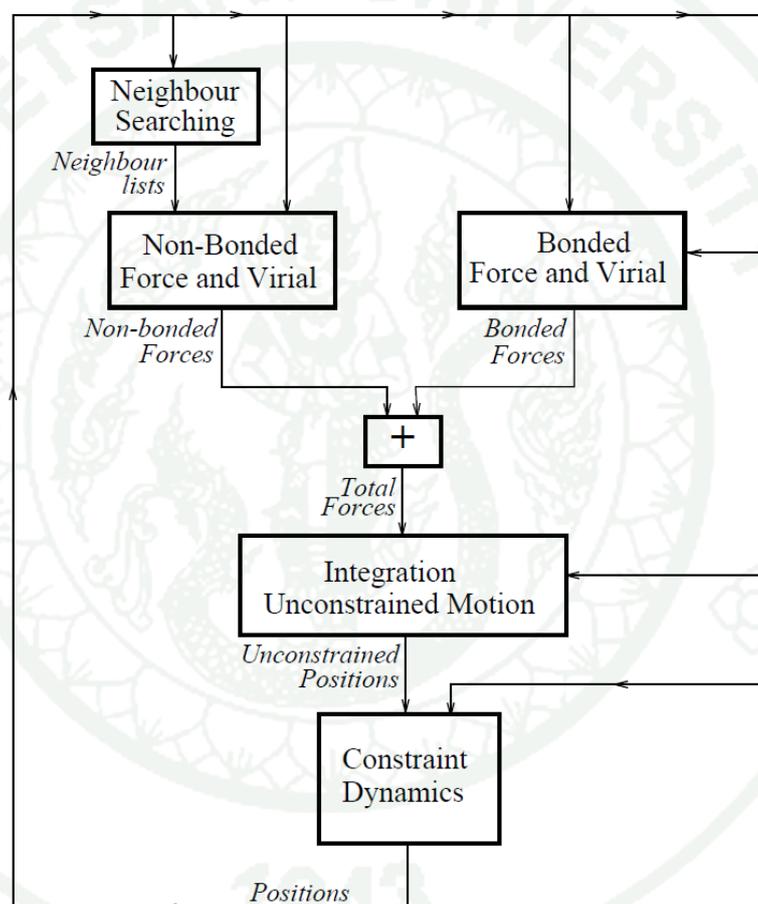
The introduction of length-constraints has no consequences for the force calculations, except of course that the forces of covalent interactions are not calculated. However, the introduction of length-constraints has severe consequences for the algorithm in which Newton's law is integrated, resulting in a matrix equation. As the rank of the matrix is the number of constraints in the system, for systems with many constraints, solving this equation directly is complex. There exists however a fast, iterative method, called SHAKE to solve the matrix equation. The special thing about SHAKE is that its iterative way of solving the matrix equation is directly reflected in iterative adjustments of pairs of particle positions.

SHAKE is used as follows. Every timestep, the interaction forces, the new velocities, and new positions are calculated as if no constraints exist, except that no covalent interaction forces are evaluated. Particle positions obtained in this way do not fulfill the distance constraints between particles. Then SHAKE is invoked. In SHAKE, particle positions are iteratively corrected until all length-constraints are fulfilled within a predefined tolerance. So, at the end of every timestep many SHAKE iterations have to be done.

## 9. Applications

M.D. simulations are used to complement and to replace experiments in physics and chemistry. As such, M.D. has been used to study simple gases, liquids, polymers, crystals, liquid crystals, proteins, proteins in liquids, membranes, DNA-protein interactions, etc. For example, the equation of state (the  $p$ ,  $T$ ,  $V$  diagram) or transport phenomena, such as thermal conductivity of a gas, may be calculated by M.D. simulation. For polymers M.D. has been used to calculate mechanical properties

like compressibility and tensile strength. In the area of drug design, M.D. is used to calculate the free energy of a reaction. Nuclear Magnetic Resonance experiments give incomplete information about inter-atom distances; M.D. is used to refine these data. Many of the physical properties mentioned are not derived from one system state but as a time average over a long sequence of consecutive states (Gunsteren *et al.*, 1990).



**Figure 4A** Main components of the M.D. simulation algorithm.

## GLYCAM PARAMETERS

(FOR AMBER 8.0, RESP 0.010), COPYRIGHT CCRC 2004

File: Parm\_Docs/Glycam\_06.txt

The following valence bond/angle parameters were replaced in the first peer-review submitted version of GLYCAM06 (Glycam\_06.dat\_FirstSubmission\_JCC) with those from AMBER's Parm94/Parm99 parameter files

From:			
C -OS	409.0	1.33	
To:			
C -OS	450.0	1.323	Parm99
From:			
CG-HC	360.0	1.095	Ethane
CG-H1	410.0	1.092	Methanol
CG-H2	440.0	1.105	1,1-Dimethoxyethane
To:			
CG-HC	340.0	1.090	Parm94
CG-H1	340.0	1.090	Parm94
CG-H2	340.0	1.090	Parm94
From:			
C -O	999.0	1.220	N-Methylethanamide
C -N	490.0	1.360	N-Methylethanamide
C -O2	730.0	1.260	2-Methylpropanoate
To:			
C -O	570.0	1.229	Parm94
C -N	490.0	1.335	Parm94
C -O2	656.0	1.250	Parm94
From:			
OH-HO	700.0	0.960	Methanol
To:			
OH-HO	553.0	0.960	Parm94
From:			
N -H	600.0	1.010	N-Methylethanamide
To:			
N -H	434.0	1.010	Parm94

From:  
 OS-C -O 132.0 122.20 K calculated from methyl acetate (eqm value  
 from crystal average)

To  
 OS-C -O 80.0 125.00 Parm99

From:  
 O2-C -O2 100.00 130.00 Acetate, 2-Methylpropanoate

To:  
 O2-C -O2 80.00 126.00 Parm94

From:  
 N -C -O 80.00 122.70 Force constant taken from Parm94, Angle set to  
 HF/6-31G(d) value N-Methylethanamide

To:  
 N -C -O 80.00 122.90 Parm94

From:  
 C -N -H 60.00 122.00 Ethanamide  
 H -N -H 30.00 119.00 Force constant taken from Parm94, Angle set to  
 agree with C -N -H in ethanamide

To:  
 C -N -H 30.00 122.00 Parm94  
 H -N -H 35.00 120.00 Parm94

From:  
 H1-CG-N 35.00 109.50 Parm91 H-CG-N

H2-CG-N 35.00 109.50 Parm91 H-CG-N

To:  
 H1-CG-N 50.00 109.50 Parm94 H1-CT-N  
 H2-CG-N 50.00 109.50 Parm94 H1-CT-N

**Notes:** This force field may be employed independently for simulating carbohydrates or may be used in conjunction with Parm94 without introducing any conflict. Correct rotational behavior for O-C-C-O fragments requires SCEE=SCNB=1.0. This is in contrast to "standard" AMBER, in which it is normal to set SCEE=1.2 and SCNB=2.0. Unless you are attempting to generate rotamer populations, it is OK to use the "standard" values. Using non-standard values (SCEE=SCNB=1.0) may be unacceptable when a protein is also present. To obtain the current version of this force field and/or carbohydrate prep files go to the website <http://www.glycam.ccruc.uga.edu>.

### The Genetic Code

	U	C	A	G	
First position (5' end)	U UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA } Stop UAG } Stop	UGU } Cys UGC } UGA } Stop UGG } Trp	U C A G
	C CUU } Leu CUC } CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }	U C A G
	A AUU } Ile AUC } AUA } AUG }	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }	U C A G

#### Amino acid names:

Ala = alanine

Arg = arginine

Asn = asparagine

Asp = aspartate

Cys = cysteine

Gln = glutamine

Glu = glutamate

Gly = glycine

His = histidine

Ile = Isoleucine

Leu = leucine

Lys = lysine

Met = methionine

Phe = phenylalanine

Pro = proline

Ser = serine

Thr = threonine

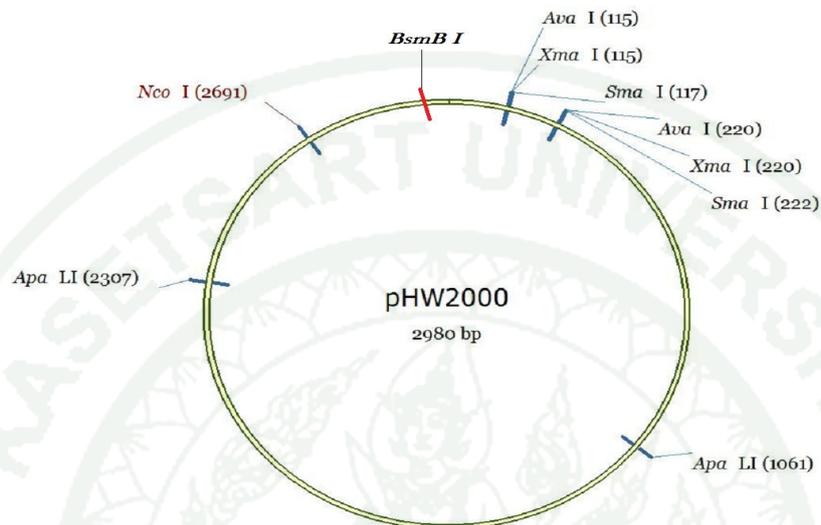
Trp = tryptophan

Tyr = Tyrosine

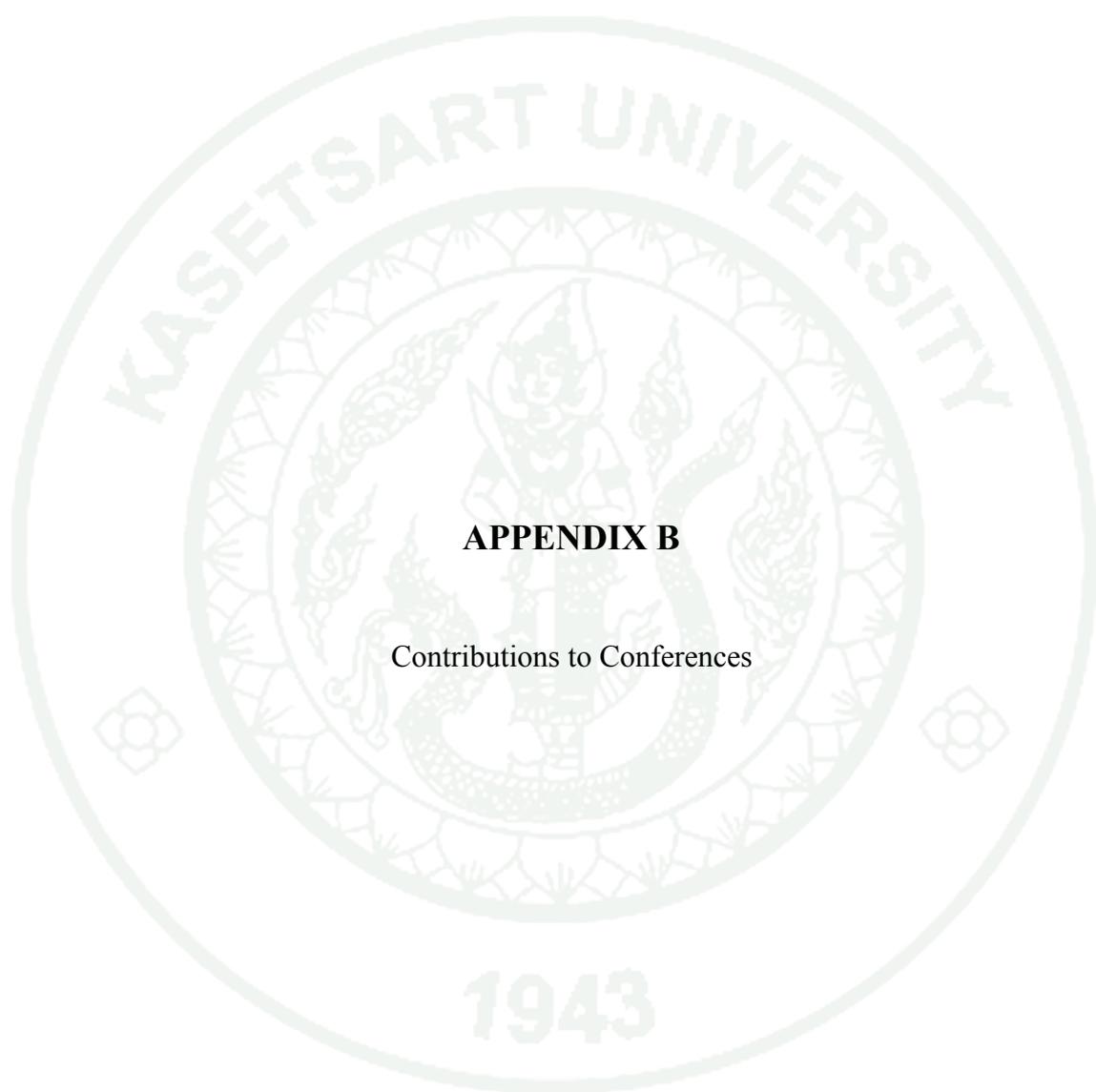
Val = valine

1943

### The pHW2000 Expression Vectors



**Figure 5A** Map and restriction endonuclease sites for pHW2000



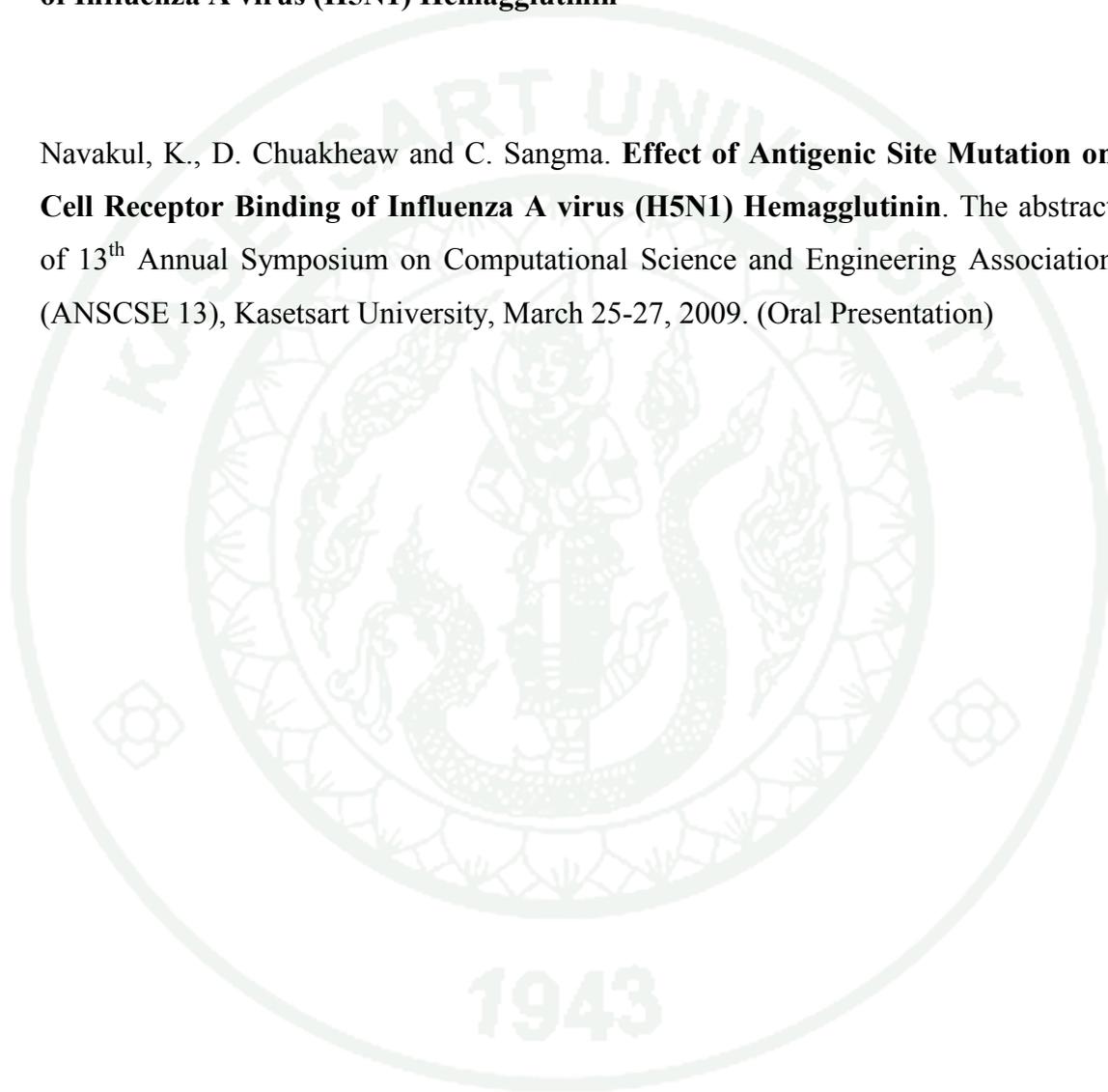
**APPENDIX B**

Contributions to Conferences

### Oral Presentation to Conferences

**- Oral presentation: Effect of Antigenic Site Mutation on Cell Receptor Binding of Influenza A virus (H5N1) Hemagglutinin**

Navakul, K., D. Chuakheaw and C. Sangma. **Effect of Antigenic Site Mutation on Cell Receptor Binding of Influenza A virus (H5N1) Hemagglutinin**. The abstract of 13<sup>th</sup> Annual Symposium on Computational Science and Engineering Association (ANSCSE 13), Kasetsart University, March 25-27, 2009. (Oral Presentation)



## Effect of Antigenic Site Mutation on Cell Receptor Binding of Influenza A virus (H5N1) Hemagglutinin

Krongkaew Navakul<sup>1</sup>, Daungmanee Chuakheaw<sup>1</sup> and Chak Sangma<sup>1\*</sup>

<sup>1</sup>*Cheminformatics Research Unit, Department of Chemistry, Faculty of Science, Kasetsart University, Jatuchak, Bangkok, Thailand 10900*

*\*E-mail: fscicsm@ku.ac.th, Tel: +66-2-562-5555 ext 2192*

Hemagglutinin (HA) is an antigenic glycoprotein found on the surface of the avian influenza, responsible for binding the viruses to sialosacharides on the host cell surface. It facilitates the release of progeny viruses from infected cells. Previous study it was found that mutations along the loop 130 (H5 numbering) were important for host cell selectivity of HA (Auewarakul *et al.*, 2007). In other HA subtype, beyond residue 137 this loop is also an antigenic site. Interestingly, in swine influenza A virus (H9), the antigenic site is absent and, coincidentally, human receptor can bind with the H9 HA. In this study we try to study whether or not this deletion can also cause the change in host type selectivity in avian influenza. We used molecular dynamics simulations to build H5 HA model from the sequences where amino acid were removed or changed between residues 137 to 140. The result showed more selectivity toward human cell receptor after the deletion of this loop.

### REFERENCES

1. Auewarakul, P.; Suptawiwat, O.; Kongchanagul, A.; Sangma, C.; Suzuki, Y.; Ungchusak, K.; Louisirirochanakul, S.; Lersamran, H.; Pooruk, P.; Thitithanyanont, A.; Eiamudomkan, A.; Pittayawonganon, C.; Guo, C-T.; Hiramatsu, H.; Jampangern, W.; Chunsutthiwat, S.; Puthavathana, P. *J. Virol.* **2007**, 9950-9955, 81.
2. Ha, Y.; Stevens, D.J.; Skehel, J.J.; Wiley, D.C; *Proc. Natl. Acad. Sci. USA* **2001**, 11181-11186, 98.

### Poster Presentation to Conferences

- **Poster presentation: Cell Receptor Binding Mechanism of Influenza A virus (H1N1) 2009 Hemagglutinin.**

**Krongkaew Navakul**, Daungmanee Chuakheaw, Sissades Tongsimma and Chak Sangma. The abstract of The Pure and Applied Chemistry International Conference 2010 (PACCON 2010), Sunee Grand Hotel and Convention Center, Ubon Ratchathani, Thailand, January 21 - January 23, 2010.





## Cell Receptor Binding Mechanism of Influenza A virus (H1N1) 2009 Hemagglutinin.

K. Navakul<sup>1</sup>, D. Chuakheaw<sup>1</sup>, S. Tongsim<sup>2</sup> and C. Sangma<sup>1\*</sup>

<sup>1</sup>Department of Chemistry, Faculty of Science, Kasetsart University, Jatuchak, Bangkok, Thailand 10900

<sup>2</sup> National Center for Genetic Engineering and Biotechnology, Thailand Science Park, Phahonyothin Road, Klong 1, Klong Luang, Pathumthani, Thailand 12120  
E-Mail : fscicsm@ku.ac.th, Tel. +66-2-562-5555 ext 2192



### INTRODUCTION

The outbreak in 2009 of a new strain of the influenza A (H1N1) virus infection in humans has raised concerns of the risk of a global flu epidemic. The new influenza A (H1N1) virus contains combination of gene segments of swine, avian, human, influenza virus. Based on genetic characterization, the hemagglutinin (HA) gene is similar to that of the swine influenza virus but there is no proposed binding mechanism reported. In this work homology modeling was used to build and study the binding mechanism of the virus HA.

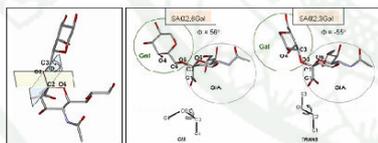
### METHODOLOGY

Influenza virus attaches to the host cell by binding of the hemagglutinin to cell receptor on the host cell surface. The HA of human influenza virus prefers to bind SA $\alpha$ 2,6Gal-terminated receptor, whereas avian influenza viruses prefers SA $\alpha$ 2,3Gal-terminated receptor.

Starting structures were generated using the crystal structures from H1N1 HA A/swine/Iowa/15/30 (Swine 1930) (pdb code 1rv0 and 1rv1) with 84 % sequence identity as templates for the simulations of A/California/10/2009(H1N1) (Swine 2009) HA binding to trisaccharide SA $\alpha$ 2,3Gal and SA $\alpha$ 2,6Gal respectively. Avian flu H5N1 and A/Puerto Rico/8/34 (H1N1) (Puerto 34) HA simulations were used as controlled systems for avian and human virus.

Complex structures were solvated using TIP5P water model and all bonds involving hydrogen were constrained by SHAKE algorithm. The complexes were minimized by heating them from 0 to 300 K with 10 kcal/mol harmonic restraints via a 100-ps NVT equilibration. The 3 ns productive runs at 300K using SANDER in AMBER9. Torsion angles define in Fig 1 were monitored and compared among the simulated systems [1,2]. Free energy were calculated using Sietraj program [3].

The receptor bound conformation were investigated by torsion angle measurement (Fig.1) and compared (see Fig. 2).



### RESULTS & DISCUSSION

In Swine 2009 HA, the receptors had the binding patterns similar to human virus (Puerto 34) according to the bound conformations.

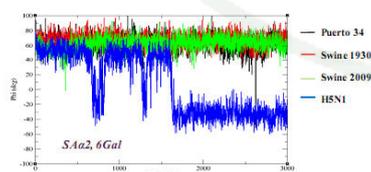
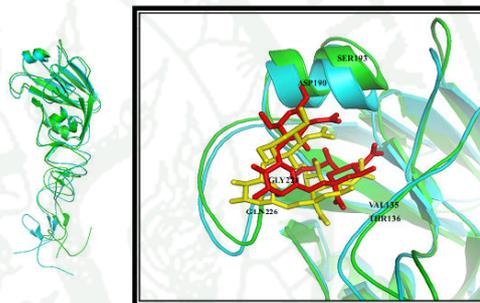


Fig2. An example of torsion angle comparison. In this picture, the time each receptor analog spent having a particular angles for SA $\alpha$ 2, 6Gal is shown.

Table1. The relative binding free energy and torsion angle monitoring of avian and human cell receptor analogs complex to 1930 H1N1, 2009 H1N1 and avian flu H5N1.

Protein	Avian ( $\alpha$ 2, 3)		Human ( $\alpha$ 2, 6)	
	Torsion angle ( $\phi$ )	$\Delta G_{\text{bind}}$ (kcal/mol)	Torsion angle ( $\phi$ )	$\Delta G_{\text{bind}}$ (kcal/mol)
Swine 1930	trans	-5.99	cis	-7.92
Swine 2009	trans	-5.61	cis	-8.44
H5N1	trans	-7.92	trans	-6.18



Structural comparison of Swine 1930 (green ribbon) and A/California/10/2009(H1N1) (blue ribbon) binding with SA $\alpha$ 2, 6-Gal analogs (red and yellow sticks).

### CONCLUSIONS

Receptors bound to HA from influenza A virus (H1N1) 2009 in the same manner as when they bound to human virus. Selectivity to human host was predicted as Swine 2009 > Swine 1930 > H5N1 while selectivity to avian host was predicted as H5N1 > Swine 1930 > Swine 2009.

### REFERENCES

1. Auewarakul, P. et al., 2007. An Avian Influenza H5N1 Virus That Binds to a Human-Type Receptor. *J. Virol.* 81 (18) : 950-9955.
2. Jongkon N, Mokmak W, Chuakheaw D, Shaw PJ, Tongsim S, Sangma C. Prediction of avian influenza A binding preference to human receptor using conformational analysis of receptor bound to hemagglutinin. *BMC Genomics* 2009, 10 (Suppl 3) : S24.
3. Naim et al., 2007, 52. Åqvist et al., 1994.
4. Gatherer, D. *Journal of Clinical Virology*, 2009, 45, 174.
5. Garten, R. J.; Davis, C. T., et al. *J. Science*. 2009, 325, 197.

### Acknowledgements

Thailand Graduate Institute of Science and Technology(TGIST)  
National Center for Genetic Engineering and Biotechnology (BIOTEC)  
Kasetsart University Research and Development Institute (KURDI)

## CURRICULUM VITAE

**NAME** : Miss Krongkaew Navakul

**BIRTH DATE** : January 20, 1985

**BIRTH PLACE** : Chainat, Bangkok, Thailand

**NATIONALITY** : Thai

<b>EDUCATION</b>	<b>: <u>YEAR</u></b>	<b><u>INSTITUTION</u></b>	<b><u>DEGREE/DIPLOMA</u></b>
	2003-2006	Chiangmai University	B.Sc. (Chemistry)
	2007-2009	Kasetsart University	M.Sc. (Chemistry)

**SCHOLARSHIPS** :

- 1) The Scholarship for Science Talented Students (full),  
Comission on Higher Education (2007-2008)
- 2) Commission on Higher Education (National Center of  
Excellence for Petroleum, Petrochemicals, and Advanced  
Materials, NCE-PPAM) (2007-2008)
- 3) Thailand Graduate Institute of Science and Technology  
(TGIST) (2008-2010)
- 4) National Center for Genetic Engineering and  
Biotechnology (BIOTEC) (2008-2010)

### PRESENTATION AND PUBLICATIONS

Navakul, K., D. Chuakheaw and C. Sangma. **Effect of Antigenic Site Mutation on Cell Receptor Binding of Influenza A virus (H5N1) Hemagglutinin**. The abstract of 13<sup>th</sup> Annual Symposium on Computational Science and Engineering Association (ANSCSE 13), Kasetsart University, March 25-27, 2009 (Oral Presentation).

Navakul, K., D. Chuakheaw, S. Tongsimma and C. Sangma. **Cell Receptor Binding Mechanism of Influenza A virus (H1N1) 2009 Hemagglutinin**. The abstract of The Pure and Applied Chemistry International Conference 2010 (PACCON 2010), Sunee Grand Hotel and Convention Center, Ubon Ratchathani, Thailand, January 21-23, 2010 (Outstanding Poster Award).

