VIRTUAL HIGH THROUGHPUT SCREENING FOR AVIAN INFLUENZA (H5N1) NEURAMINIDASE INHIBITORS

by

NURUL IZZA BINTI ISMAIL

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

July 2011

ACKNOWLEDGEMENTS

بسم الله الرحمن الرحيم

First and foremost, I would like to express my sincere gratitude to Allah SWT for His gracious, compassion and countless blessing to us throughout all of our lives.

I would like to thank my supervisors, Assoc. Prof. Mohamed Razip Samian and Assoc. Prof. Habibah A Wahab for their consistent support, guidance and advice during the past years. Besides my supervisor, my sincere thanks also goes to Dr Hassan Hadi, for his encouragement and inspiring comments.

I am grateful for the financial support provided by Universiti Sains Malaysia and Ministry of Higher Education during my studies. I also would like to express my thanks to my fellow labmates: Nani, Sue, Imtiaz, Belal, Ita, Hamdah, Sy Bing, Hafiz, Wai Keat, Yee Siew, Faezah and Nurul for their help and discussion in a friendly environment for me. A special thank to all who have helped in giving support in my work.

Finally, my special gratitude is dedicated to my lovely parents and family. To my beloved parents, thank you for raised me with love. To my dear husband, thank you for being by my side with your love and endless support since 10 years back. To my lovely siblings, Nazuha, Nazira, Roharty and Ikhwan, thank you for everything I could ever ask for. My sincere love for all from the bottom of my heart.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF SYMBOLS AND ABBREVIATIONS

$(\sigma_{ij}/r_{ij})^{12}$	Van der Waals repulsion force
$(\sigma_{ij}/r_{ij})^{6}$	Van der Waals attraction force
2D	Two dimensional
3D	Three dimensional
Å	Angstrom
b	Observed bond length before stretching
b_o	Unstrained or reference bong length after stretching
cDNA	complementary DNA
DANA	2-Deoxy-2,3-didehydro-N-acetylneuraminic acid
DOF	Degree of freedom
dpf	Dock parameter file
dsRNA	Double stranded RNA
E_{bend}	angle bending energy
E_{elec}	Electrostatic energy (non-bonded interaction)
$E_{stretch}$	Bond stretching energy
E_{tors}	Tortional or rotational energy
E_{vdw}	van der Waals energy (non-bonded energy)
GA	Genetic algorithm
gpf	Grid parameter file
HA	Hemagglutinin
k_b	Parameter that controls the stiffness of the bond spring
K_i	Inhibition constant

k_θ	The angle-bending force constant, which controls the stiffness of the angle spring
LGA	Lamarckian genetic algorithm
LS	Local search
M1	Matrix protein
M2	M2 ion channel protein
MC	Monte Carlo
MD	Molecular dynamic
mRNA	Messenger RNA
n	A number of minimum points the bond rotates with full cycle (360)
NA	Neuraminidase
NMR	Nuclear magnetic resonance
PDB	Brookhaven Protein Data Bank
q_i	Atomic charges of interaction atom i
q_j	Atomic charges of interaction atom j
r_{ij}	Distance between the atoms i and j
RMSD	Root mean square deviation
RNA	Ribonucleic acid
SA	Simulated annealing
V_n	Tortional barrier
ΔG	Free energy of binding
ΔG_{conf}	Free energy of binding for deviations from covalent geometry
ΔG_{elec}	Free energy of binding for electrostatics
ΔG_{hbond}	Free energy of binding for hydrogen bonding

ΔG_{sol}	Free energy of binding for desolvation upon binding and the hydrophobic effect (solvent entropy changes at solute-solvent interfaces)
ΔG_{tor}	Free energy of binding for restriction of internal rotors and global rotation and translation
ΔG_{vdw}	Free energy of binding for van der Waals (repulsion and dispersion)
ε_0	Dielectric constant
ε_{ij}	Well depth of the energy minimum
θ	Observed value for the bond angle after angle-bending
Σ	Sum
ω	Observed torsion angle

PENYARINGAN MAYA BERDAYA TINGGI BAGI PERENCAT NEURAMINIDASE SELESEMA BURUNG (H5N1)

ABSTRAK

H5N1 adalah satu daripada sub-jenis virus influenza yang boleh menyebabkan ancaman pandemik bukan sahaja kepada burung, tetapi juga kepada kebanyakan spesis lain termasuk manusia. Neuraminise N1 adalah enzim glikoprotein yang terdapat di permukaan partikel virus influenza. Di dalam kajian ini, penyaringan maya berdaya tinggi dijalankan bagi mengenalpasti molekul kimia yang dapat menyahaktifkan fungsi neuraminidase N1. Dengan menggunakan struktur molekul kimia daripada pangkalan data NADI-CHEM, sejumlah 2498 struktur molekul kimia yang boleh diabstrak daripada tumbuhan semulajadi di Malaysia didokkan kepada enam konformasi neuraminidase N1 yang berbeza menggunakan perisian AutoDock 3.0.5. Enam konformasi neuraminidase yang berbeza diperolehi dengan melakukan simulasi dinamik molekul terhadap struktur monomer dan tetramer neuraminidase. Keputusan yang diperolehi disusun mengikut tenaga bebas pengikatan terendah (BFE). Bagi proses penyaringan yang dilakukan terhadap konformasi-konformasi monomer neuraminidase, lima ligan menduduki kedudukan 100 teratas tenaga bebas pengikatan terendah. Ligan-ligan tersebut adalah drimene, faradiol, ochrolifuanine A, beta-amyrin dan cycloartenol. Bagi proses penyaringan yang dilakukan kepada konformasi-konformasi tetramer neuraminidase, empat ligan iaitu drimene, ochrolifuanine A, asid jacoumarik dan asid ifflaionik menduduki kedudukan 100 teratas tenaga bebas pengikatan terendah. Kesemua ligan yang dipilih mempunyai tenaga bebas pengikatan yang lebih rendah daripada oseltamivir (-9.36 kcal/mol). Lima ligan iaitu, drimene, faradiol, beta-amyrin, ochrolifuanine A, dan asid ifflaionik mengikat kepada konformasi-konformasi monomer dan tetramer

dengan bacaan Ki antara 0.1 nM dan 10 nM. Semua ligan yang dianalisa tersebut berada pada kedudukan poket aktif neuraminidase. Malah, drimene dan ochrolifuanine A berupaya untuk mengikat pada konformasi kawasan aktif '*150-loop*' 'terbuka' dan '*150-loop*' 'tertutup' dengan tenaga bebas pengikatan yang lebih rendah berbanding tenaga bebas pengikatan oseltamivir. Daripada analisa mendalam yang dijalankan didapati semua ligan tersebut mempunyai potensi untuk dijadikan sebagai penghalang fungsi neuraminidase.

VIRTUAL HIGH THROUGHPUT SCREENING FOR AVIAN INFLUENZA (H5N1) NEURAMINIDASE INHIBITORS

ABSTRACT

Avian influenza H5N1 virus which caused virulent influenza in birds is a pandemic threat causing viral disease in many species including humans. Neuraminidase N1, an antigenic glycoprotein enzyme found on the surface of the influenza particle, was chosen as a drug target for virtual high-throughput drug screening of small molecules that is capable of blocking its activity. A total of 2498 compounds derived from Malaysian natural plants in NADI-CHEM Structure Database were docked to six different neuraminidase conformations derived from molecular dynamics simulation of neuraminidase monomer and tetramer using AutoDock 3.0.5. The results were ranked according to the lowest binding free energy (BFE). From this screening process, five ligands exist in all top 100 BFE ranking in the molecular docking with different conformations of the neuraminidase monomer. The ligands were drimene, faradiol, ochrolifuanine A, beta-amyrin and cycloartenol. In molecular docking with different conformations of the neuraminidase tetramer, drimene and ochrolifuanine A in addition to jacoumaric acid and ifflaionic acid scored in 100 BFE ranking, and were chosen for detailed binding studies. Interestingly, all selected ligands from molecular docking with N1 monomer and tetramer conformations scored more favourable binding free energy compared to oseltamivir (-9.36 kcal/mol). Five ligands, namely drimene, faradiol, beta-amyrin, ochrolifuanine A, and ifflaionic acid bind to N1 monomer and tetramer conformations with acceptable *Ki* values, ranging from 0.1 nM to 10 nM. All the selected ligands bind significantly to the neuraminidase binding pockets. Drimene

and ochrolifuanine A are capable of binding to both 'open' and 'closed' 150-loop active site conformation with binding free energy better than oseltamivir. The results indicate that these ligands have potential to be explored as new neuraminidase inhibitor.

CHAPTER ONE

INTRODUCTION

The first highly pathogenic H5N1 virus case was detected on 1996, isolated from a farmed goose in Guangdong Province, China. In 1997, similar outbreaks were reported in poultry at farms and live animal markets in Hong Kong. . That same year, also witness the first human infections with avian influenza H5N1 case in Hong Kong. 18 cases (6 fatal) were reported in Hong Kong. Altogether, 18 cases (6 fatal) are reported in the first known instance of human infection with this virus (World Health Organization, 2008).

After several years, in 2003, two human cases of avian influenza H5N1 infection (one fatal) are confirmed in a Hong Kong family. H5N1 infections on human then are reported in several Asian countries like Hong Kong, China, Vietnam, Thailand, Cambodia and Indonesia. In early 2006, the first case in Middle East is detected in Turkey. The viruses are similar to those currently circulating in birds but it rapidly end in the coming week. Other cases of H5N1 are reported in Iraq, Azerbaijan, and Egypt. On January 2007, Nigeria, a West Africa country, confirms its first human case (World Health Organization, 2008). Until now, the numbers of reported cases continue to increase and there are no signs to stop (Poland et al., 2007).

From the increasing number of avian influenza cases from time to time and with the highly fatality, this pandemic should be stopped. Many researches and

pharmaceutical company came with efforts to discover effective drugs to stop the pandemic (Kim et al., 1999, Babu et al., 2000, Du et al., 2007).

At present there are four drugs available worldwide for avian flu treatment namely amantadine, rimantadine, oseltamivir, and zanamivir. However, none of the four drugs have been shown to effectively prevent serious influenza-related complications such as bacterial or viral pneumonia. In addition, reports proved that in some cases the viruses were resistance to the drugs (Le et al., 2005, Gubareva, 2004, McKimm-Breschkin, 2000, Wang et al., 2007). Studies have shown that all four drugs can reduce the duration of symptoms by one day only if taken within 2 days of the onset of the illness. According to the report published by the National Institute of Health, these drugs generally cause side effects such as nausea, vomiting, wheezing or serious breathing problems, headache and diarrhea. Thus, there is a need for the discovery of a new drug to address this problem. With the availability of the structure of neuraminidase N1, a virtual high throughput screening (vHTS) program can be initiated (Fanning et al., 2000, Kim et al., 1997, Russell, et al., 2006, Wei et al., 2006, Wang et al., 2007).

The cost of developing drug in the conventional manner is prohibitive with the advent of cheap computing power, we can use *in silico* high throughput drug screening techniques to identify new hit compound. Using this technique, the screening of potential ligands will be done using molecular docking simulation. A library of ligand, the NADI-CHEM database will be used for this purpose (Natural Based Drug Discovery Intelligent-Chemical Structure Database, 2007).

The objectives of this research are:

1. To identify hit compounds that have potential to inhibit neuraminidase N1. The small chemical compounds that used in this study come from local natural products and will hopefully generate a high value hit compounds that can react directly to numerous comformations of neuraminidase N1.

2. To characterize the molecular interaction of the hit ligands. Small ligands which have significant interactions to neuraminidase N1 will be identified and will be studied further as drug lead compounds.

CHAPTER TWO

LITERATURE REVIEW

2.1 Avian Influenza Virus

2.1.1 Influenza virus classification

Group A influenza virus can be classified into a subtype on the basis of two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). HA and NA are two surface glycoproteins on the surface of avian influenza virus. Altogether, there are sixteen subtypes of hemagglutinin and nine subtypes of neuraminidase (Chen, et al., 2007).

This formation of various subtypes for avian influenza A is caused by phenomenon called antigenic drift and antigenic shift. Antigenic drift happens all the time, but antigenic shift happens only occasionally. Antigenic drift happens when point of mutations or gradual genetic change occurring in the two known genes that contain the genetic material to produce glycoproteins, HA and NA. Antigenic drift produces new virus strains that may not be recognized by antibodies. Further, it may lead to loss of immunity. This type of antigenic drift occurs in all types of influenza group A, B, and C (Carrat and Flahault, 2007; Itzstein, 2007).

Antigenic shift refers to an abrupt change in HA and NA combination to produce new influenza A virus subtype. It happens through a process called genetic

reassortment. During the rearrangement, strains are enabled to jump from one species to another.

Two different strains of influenza A can combine and form a new subtype. A new global pandemic may occur if a new subtype of influenza A virus is introduced into human population. Antigenic shift can occur in three ways; (a) from wild waterfowl to humans through genetic change in an intermediate host such as chicken or pig, and a human virus infects the intermediate host at the same time (reassortment); (b) directly from waterfowl to humans without genetic change; or (c) from waterfowl to an intermediate host without genetic change (Figure 2.1) (Goldrick, 2007). Thus far, antigenic shift has been reported to occur only in influenza group A.

Influenza viruses required host cells to supply viral enzymes in their replication process. This was first suggested more than 25 years ago, when it was shown that influenza A virus required a functioning nucleus in which to replicate, in contrast to other RNA viruses that replicate in the cytoplasm of their host cells (Voyles, 2002).

Both HA and NA are involved in avian influenza replication process. HA is involved in the attachment and fusion protein. HA binds to the terminal sialic acid on the viral receptor. HA then mediates the fusion of the virion and vesicle membrane process. Neuraminidase cleaves sialic acid linkages and releases the newly formed progenies from the host cell.

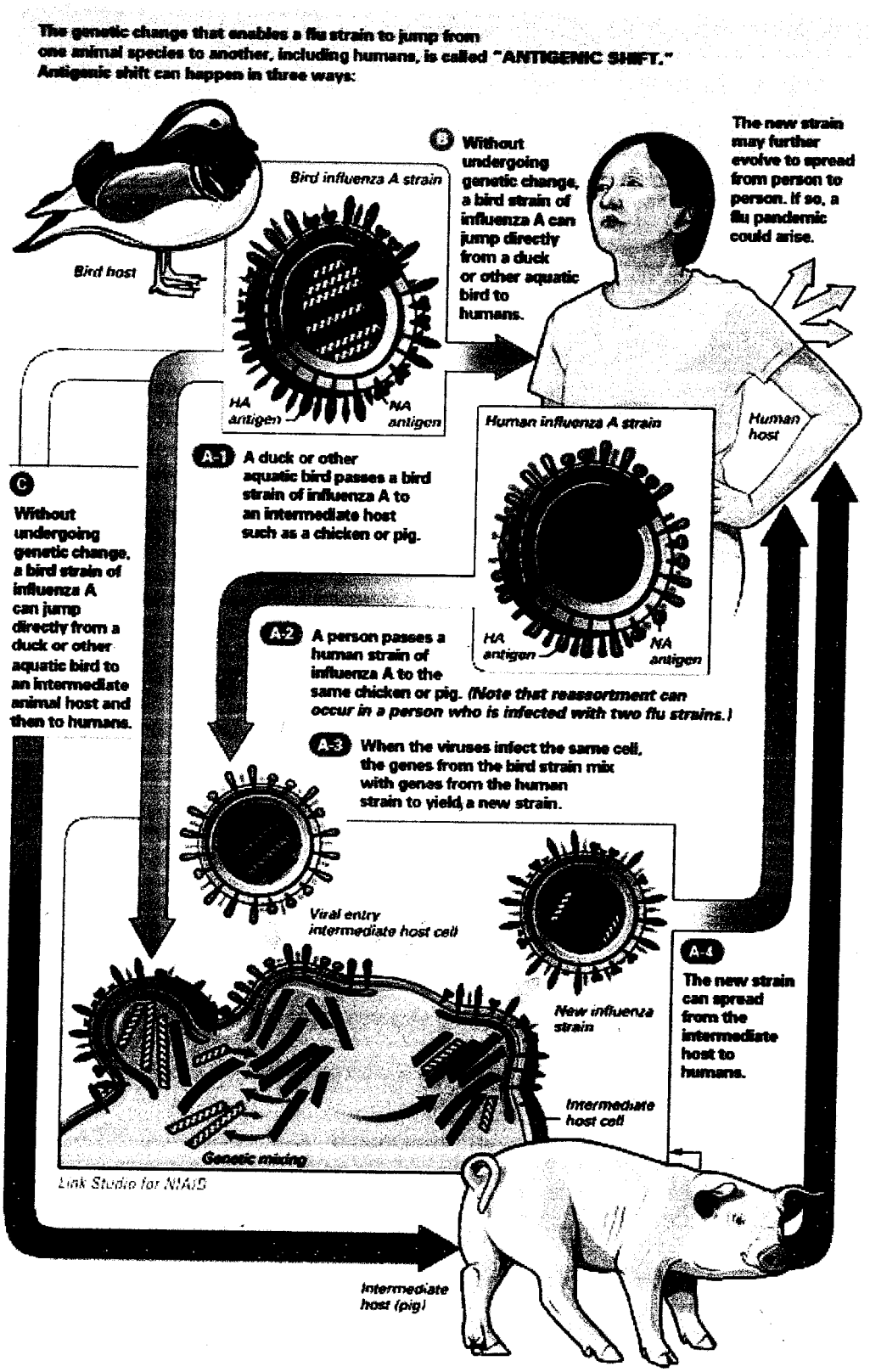

Figure 2.1 The process of avian influenza antigenic shift (Source: National Institute of Allergy and Infectious Diseases accessed on July 9, 2008)

2.1.2 Avian Influenza H5N1

Avian influenza or also known as bird flu is an infectious disease of birds and mammals caused by RNA virus (Figure 2.2). H5N1 virus is an influenza A virus that occurs mainly in birds, is highly contagious among birds, and can be deadly to them. To date, animal infected H5N1 virus cases have been reported in seventy countries according to the World Health Organization (WHO, 2008). During late 2003 and early 2004, H5N1 affected more than 100 million birds that either died from the disease or were killed in order to try to control the outbreaks.

Chillingly, there have been reports of human fatilities caused by this virus. About 258 human infections (as of November 13, 2006), have been reported in Cambodia, China, Indonesia, Thailand, Turkey and Vietnam. H5N1 infection may follow an unusually aggressive clinical course, with rapid deterioration and high fatality. This is demonstrated in the current situation in Asia, where the mortality rate is approximately 50%. The impact of H5N1 pandemic is catastrophic as shown by the US congressional Budget Office that has estimated that up to 2 million of the US population might die, with up to 40% of all workers ill for as long as 3 or more weeks should the pandemic occur. Financially, this is translated into a total cost of US$683 billion covering the deaths, clinically ill, and outpatient treatment cost. Before the virus mutate into a form that infect human aggressively, it is necessary to develop drugs to stem the progression of isolated incidences into a full-blown epidemic (Levi, et al., 2007).

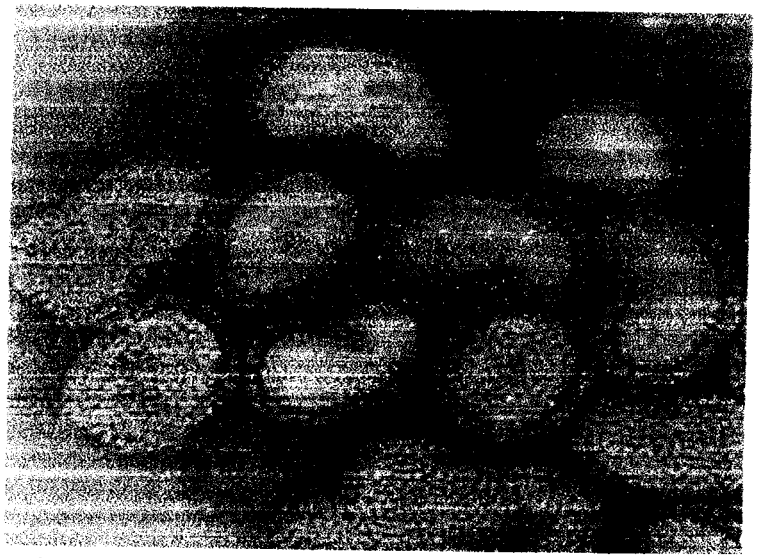

Figure 2.2 Influenza A Virus (Source: International Committe on Taxonomy of Viruses accessed on March 8, 2011)

2.1.3 Pathogenicity

Further classification of the avian influenza virus is according to the basis of specific molecular genetics and pathogenesis criteria that need specific testing. Low pathogenic (LPAI) virus is usually associated with mild disease in poultry and high pathogenic (HPAI) virus refers to the strain causing severe illness and high mortality in poultry. The "North American H5N1 is referred to as the Low Pathogenic H5N1, whereas the "Asian" H5N1 is known as the High pathogenic H5N1 because of its high rates of mortality and causing worldwide concern (Spackman, et al., 2007).

From 2003 to 2006, high pathogenic A (H5N1) influenza viruses have had a devastating impact on domestic or wild birds in many parts of South East Asia, Europe, the Middle East and Africa. The HPAI subtype A (H5N1) has been the cause of large scale death in poultry and the subsequent infection and death of over 140 humans (Hurt, et al., 2007). While the virus is not easily transmitted from human to human, there might be further mutation or genetic reassortment that can produce a strain which has the potential to cause a pandemic.

2.1.4 Avian influenza viral protein

Influenza virus is an orthomyxovirus and the structure is well known (Figure 2.3). Avian influenza A virus is an enveloped virus. It contains of eight segments of negative sense single stranded RNA that encode 11 proteins. These proteins are polymerase-7mG CAP binding, polymerase-elongation, polymerase, hemagglutinin,

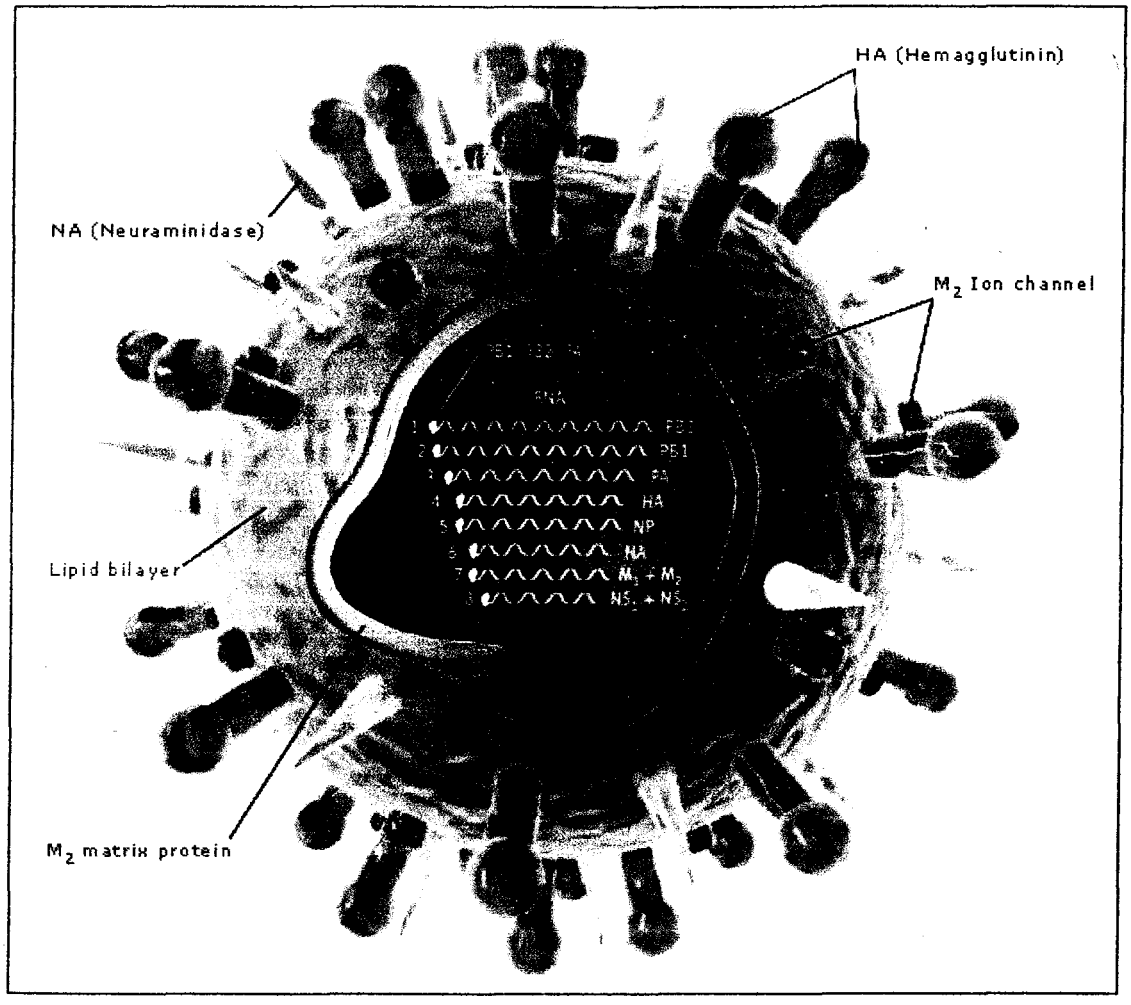

Figure 2.3 Schematic structure of influenza virus (Source: International Federation of Pharmaceutical Manufacturers & Associations accessed on January 24, 2008)

nucleoprotein, neuraminidase, matrix proteins, and non-structural proteins (Chen, et al., 2006). Table 2.1 presents the gene assignments in influenza A virus (Voyles, 2002). The nucleocapsids of the virus are helical. This virus is a pleomorphic virus, meaning its shape varies. The lipid envelope contains the M2 ion channel protein and the viral glycoproteins, hemagglutinin and neuraminidase. Within the envelope, the HA and NA proteins are relatively abundant, whereas the M2 protein is present in relatively few copies (Basler, 2007).

Beneath the viral membrane is the viral matrix protein, M1. The nucleoprotein (NP) associates with the eight genomic RNA segments to form the ribonucleoprotein complexes (RNPs) lie inside the virus. United with this structure is the viral polymerase; PA, PB1 and PB2, and also viral nuclear export protein (NS2).

2.1.5 Avian influenza replication process

The replication process of this orthomyxovirus differs from other type of viruses. The genome is transcribed to mRNAs and replicated to form new genomes. Influenza viruses bind through hemagglutinin onto sialic acid sugars on the surfaces of epithelial cells (Itzstein, 2007; Oxford, 2000; Skehel and Wiley, 2000). The virus enters the host cell by endocytosis, and further appears in endosome. Following endocytosis of the viral particle, the endosome undergoes acidification (Basler, 2007; DeTulleo and Kirchhausen, 1998; Yamashiro and Maxfield, 2004). Because of the acidic environment in the endosome, modification at the hemagglutinin spikes

Table 2.1 Gene assignments for influenza A virus segments

Segment Number	Protein	Name Function
1	PB2	Polymerase-7mG CAP binding
2	PB1	Polymerase elongation
3	PA	Polymerase
4	HA	Hemagglutinin
5	NP	Nucleoprotein
6	NA	Neuraminidase
7	M1, M2	Matrix (membrane) proteins
8	NS1, NS2	Non-structural proteins

happened, creating a hole, and the influenza virus releases viral RNA (vRNA) molecules, viral proteins and RNA-dependent RNA transcriptase into the cytoplasm. vRNA and proteins are then transported into host cell nucleus via the nuclear pores. Three virus polymerase proteins are involved in the both transcription and translation process, namely PB1, PB2, and PA. The vRNA produces various kinds of viral messenger RNAs (vmRNA) which travel out through the nuclear pores. The vmRNA carries genetic information that is used in manufacture of protein.

Some of the newly synthesized vmRNAs have roles in the synthesis of nucleoprotein that travels back into the nucleus. Other vmRNAs direct the production of viral proteins and transmembrane viral proteins. Other viral proteins have multiple actions in the host cell, including degrading cellular mRNA and using the released nucleotides for vRNA synthesis and inhibiting the translation of host-cell mRNAs.

In the nucleus, the negative sense vRNA produces full-length positive sense copies of itself. This process is catalyzed by RNA-dependent RNA transcriptase. These are then used to create further copies of the negative sense vRNA. These new negative sense vRNAs become associated with nucleoproteins and some viral proteins that have migrated into the nucleus. Such newly formed nucleocapsids and their associated M proteins exit the nucleus via nuclear pores. Around the new nucleocapsid, the viral proteins are collected beneath the cell membrane, while above the cell membrane; hemagglutinin and neuraminidase have coated the host cell surface (Skehel and Wiley, 2000).

With all these viral elements now in place, the newly formed particles can begin to take shape and to bud from the cell surface. The cell membrane that envelopes the emerging nucleocapsid and matrix protein becomes the viral envelope (complete with projecting spikes) and the virus particle is then released once neuraminidase cleaved the sialic acid residues from host cells (Fanning, et al., 2000). After the release of the new progeny influenza viruses, the host cell dies.

2.1.6 Influenza virus inhibitors

At this moment, vaccination is the important way for preventing avian flu. But, in some condition, vaccination is inadequate and effective antiviral agents would be of the utmost importance. Antigenic drift of influenza virus may occur between any influenza pandemic seasons. This will cause the currently available vaccine to be less protective. Furthermore, once a pandemic occurs, the production of influenza vaccine cannot accommodate the needs of the population. Therefore, antiviral still play a major role as a treatment of influenza infection.

There are several known anti-influenza inhibitors nowadays (Figure 2.4) (Clark, 2004; Clercq and Neyts, 2007). Amantadine and rimantadine are known as M2 channel inhibitors. IMP dehydrogenase inhibitors like ribavarin and viramidine can be used to block the RNA replication and transcription steps which require repeated cycles of (-)RNA $\leftrightarrow$ (+)RNA polymerization reactions occur in the nucleus (Clercq and Neyts, 2007). While, neuraminidase as a glycoprotein enzyme is

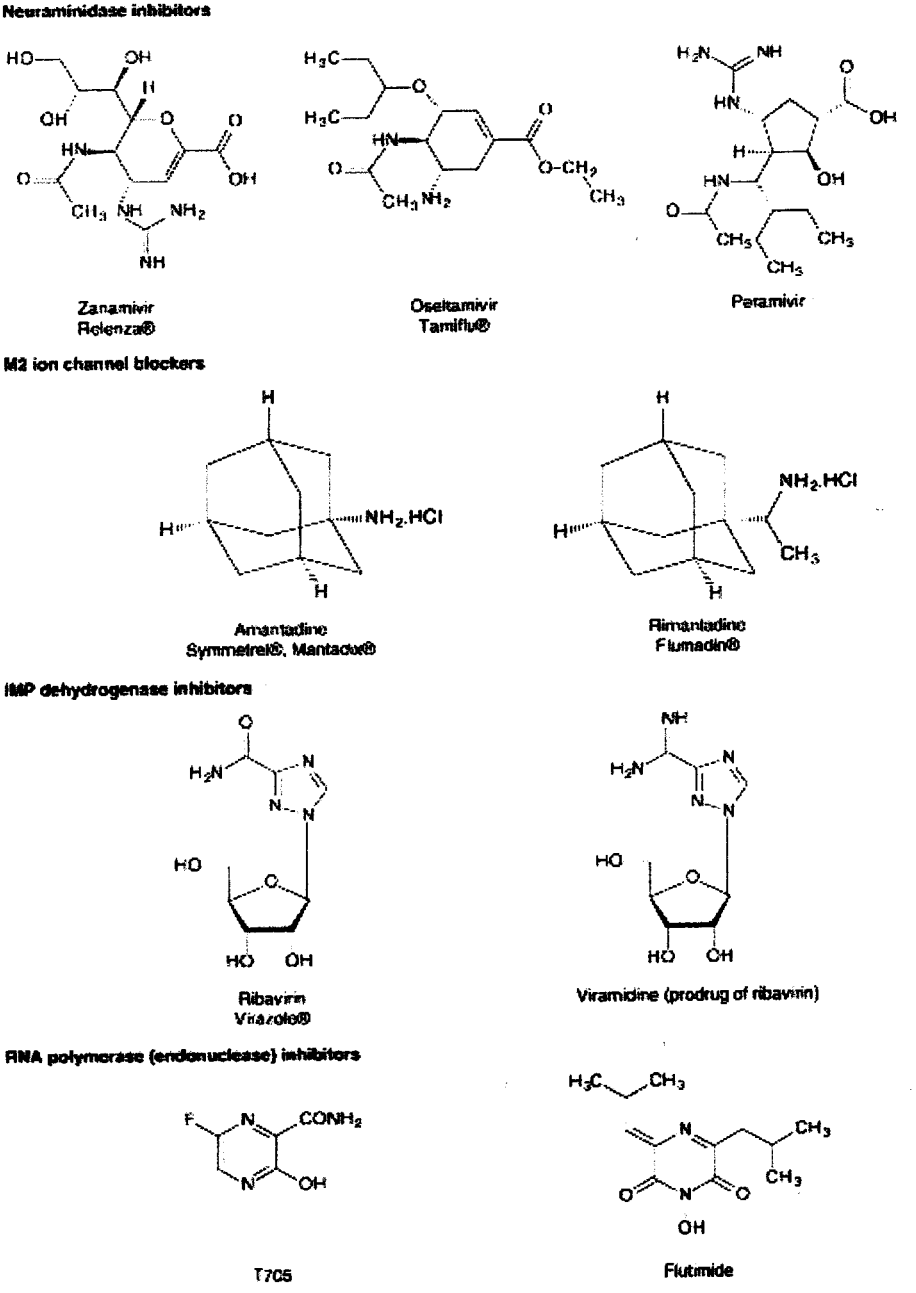

Figure 2.4 Major classes of anti-influenza virus compounds

inhibited by oseltamivir, zanamivir and peramivir (Boltz, et al., 2007; Gubareva, 2004; Hayden, 2001; Kim, et al., 1999; Leneva, et al., 2000; Sun, et al., 2006; Wei, et al., 2006).

The experimental inhibition parameters for known inhibitors of influenza virus neuraminidase are shown in Table 2.2 (Taylor, 2003). Sialic acid is a weak inhibitor for neuraminidase with a K_i value of 1mM. FANA was the best inhibitor of neuraminidase for some years with a K_i = 0.8 μM (Taylor, 2003). The most effective inhibitors that have been developed are oseltamivir and zanamivir with K*i* values lower than 1.3 nM (Taylor, 2003).

However, there are many weaknesses of the inhibitors reported. None of the inhibitors for M2 channel ion and neuraminidase have been shown to effectively prevent serious influenza-related complications such as bacterial or viral pneumonia. Studies have shown that some of the drugs can reduce the duration of flu symptoms by 1 day only if taken within 2 days of the onset of the illness (Smith, et al., 1980).

According to the report published by National Institute of Health, these drugs generally, cause fewer side effects. In some cases, oseltamivir caused nausea and vomiting. The common side effects of zanamivir are wheezing or serious breathing problems, headache and diarrhea (National Institute of Health, 2008).

Table 2.2 Inhibition parameters for known inhibitors of influenza virus neuraminidase

Inhibitors	Influenza A virus	
	K_i	IC_{50}
Sialic acid	1 mM	-
DANA	4 μM	0.015 mM
FANA	0.8 μM	-
Zanamivir	0.06-1.3 nM	0.3-2.3 nM
Oseltamivir	0.10-1.3 nM	0.01-2.2 nM

2.2 Neuraminidase as a Drug Target

There are many ways to inhibit influenza virus function. M2 ion channel, IMP dehydrogenase, RNA polymerase, hemagglutinin and neuraminidase were known as anti-viral targets. One of the most chosen targets among researches is neuraminidase (Fanning, et al., 2000; Hayden, 2001; Kim, et al., 1997).

Neuraminidase has been chosen as a drug target because it has a broader antiviral spectrum, better tolerance, less potential for emergence of resistance, close correspondence of the conserved residues of the active site from all influenza neuraminidases, and *in vitro* studies demonstrated both HA and NA susceptibility to NA inhibitors (Zhang and Xu, 2006). Besides, the existing inhibitors, such as oseltamivir and zanamivir, were developed based on different structures of neuraminidase (Du, et al., 2007).

2.2.1 The neuraminidase active site

Neuraminidase can be categorized into two groups structurally; group-1 includes N1, N4, N5, and N8 and group-2 includes N2, N3, N6, N7, and N9 (Landon, et al., 2008). The active site for N1 and all group-1 neuraminidases is different from group-2 especially at the 150-loop (residues 147-Arg152) region. The conformation of the 150-loop is such that the Cα position of the group-1 specific Val149 is about 7Å distant from the equivalent isoleucine residue in group-2. The hydrophobic side chain at Val149 is pointed away from the active site. There is a

large cavity adjacent in N1 active site region. The 150-cavity is about 10 Å long and 5 Å wide and deep (Russell, et al., 2006).

In vitro, N1 can bind with an inhibitor in either the 'open' or 'closed' conformation of the 150-loop, depending on the soaking conditions. Incubating N1 crystals in 20 μM oseltamivir for 150 minutes showed binding of the inhibitor with the 150-loop in the open cavity conformation (Collins, et al., 2008; Russell, et al., 2006). While, at a higher concentration of inhibitor, the 150-loop changes its conformation to close-cavity conformation. The main change in the conformation is the Glu119 and Asp151 residues are now oriented facing the bound inhibitor to form a 'closed' conformation.

This 150-loop's motion is coupled to the motion of the neighbouring 430-loop. The 430-loop comprises residues Arg430-Thr439. The 430-loop helps to expand the active site's cavity. These coupled motions significantly expanded the active site cavity, increasing its solvent-accessible surface area as compared with both open and closed crystal structures (Amaro, et al., 2007; Landon, et al., 2008).

A study showed that for both the apo and holo molecular dynamic simulations of neuraminidase, the greatest structural diversity is found in the 150-loop and 430-loop areas, although the entire binding-site region was used in the clustering. This indicates that the 150-loop and 430-loop areas are particularly flexible (Landon, et al., 2008). The two most variable regions in the binding site are highlighted in Figure 2.5.

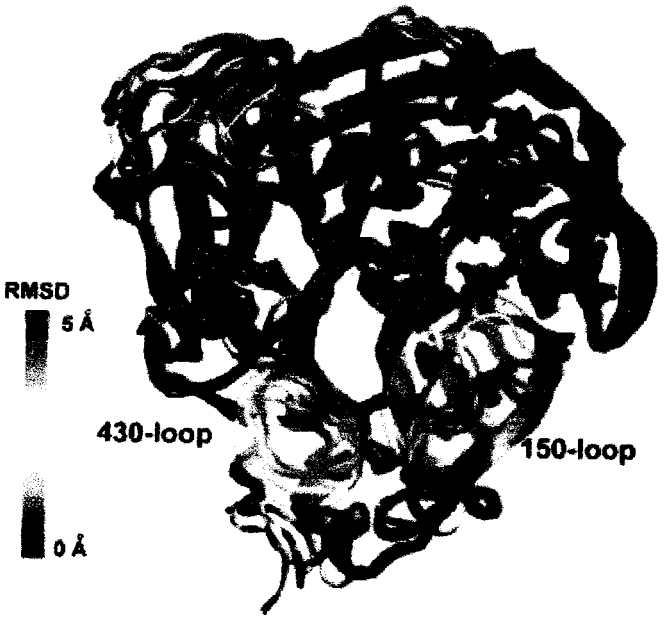

Figure 2.5 The location of 150-loop and 430-loop region in neuraminidase 2HU4. (Landon et al., 2008)

The neuraminidase active site can be divided into three major binding pockets as shown in Figure 2.6 (Lew, et al., 2000). The polar residues in the N1 active site Glu276, Glu277, Arg292, Asn294, and hydrophobic Ala246 make up Pocket 1. Pocket 2 consists of the highly conserved amino acids residues Ile222, Arg224, and Ala246. The third binding pocket is the largest and is made up from hydrophobic and hydrophilic residues; Glu119, Asp151, Arg152, Trp178, Ser179, Ile222, and Glu227 (Lew, et al., 2000; Wang, et al., 2007).

2.2.2 3-Dimentional structure of neuraminidase

The first step in running a screening process is to get a suitable protein as a target. The ideal receptor is one that is closely linked to human disease and binds a small molecule in order to carry out a function. The target receptor should have a well-defined binding pocket. Other designed small molecule can compete with the natural small molecule in order to modulate the function of the receptor at a required level of potency. The receptor should be essential and its elimination should lead to the pathogen's death. Finally, the target receptor should be able to be inhibited by binding a small molecule.

Crystal structures are the most common source of structural information for drug design. A crystal structure should be evaluated for the resolution of the diffracted amplitudes or often simply called resolution, reliability or R factors, coordinate error, temperature factors and chemical correctness. Typically, crystal structures determined with data extending to beyond 2.5 Å are acceptable for drug

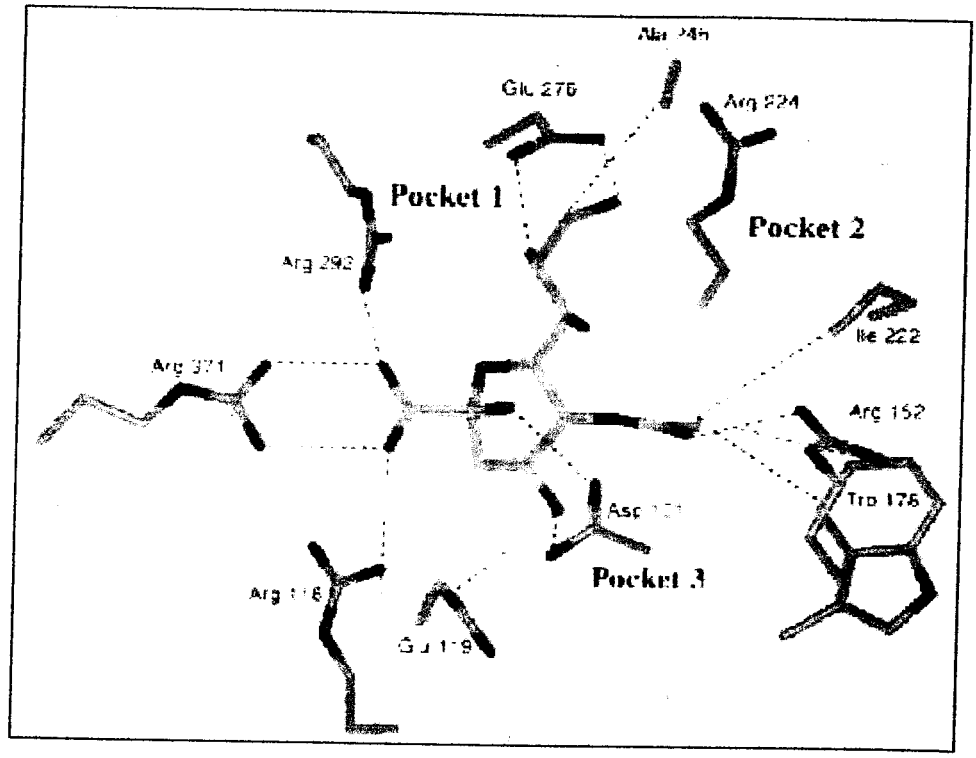

Figure 2.6 Three major binding pockets of neuraminidase active site.

design purposes since they have a high data to parameter ratio, and the placement of residues in the electro-density map is unambiguous (Anderson, 2003). The R_{free} should be below 28% and R factor should be well below 25%. Coordinate error in Protein Data Bank is calculated using Luzzati method and should be in the range of 0.2 Å – 0.3 Å. Temperature factors of atoms in the region of interest should be no greater than the average temperature factor for the molecule. Finally, the molecule should be refined to be consistent with all rules of stereochemical "correctness". After a careful observation in the Brookhaven Protein Data Bank, the N1 in complex with oseltamivir (PDB code: 2hu4) was chosen as a drug target (Figure 2.7). It was found in 2006 through x-ray diffraction. The neuraminidase structure contains eight chains with 388 residues each.

The chosen X-ray crystal structure of neuraminidase N1 contains eight chains; they are chain A, B, C, D, E, F, and G. All of them are similar in term of the size and amino acids sequences. Thus, it is sufficient for the purpose of this study to select one chain. The weight of the protein is 42,111.6 with 385 amino acid residues (Figure 2.8).

2.3 Introduction to Drug Discovery and Molecular Modelling

Drug discovery and development is an expensive process due to the high research and development costs and extensive clinical test. A study in year 2003

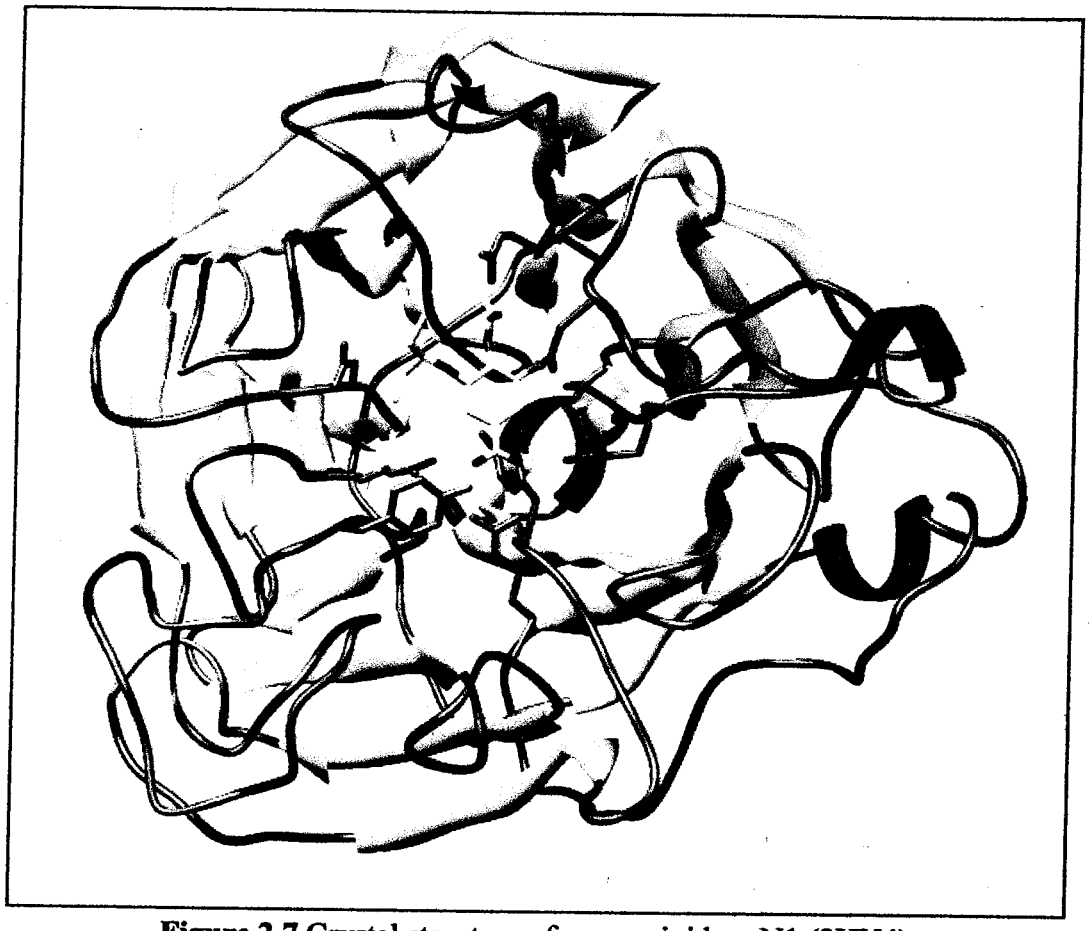

Figure 2.7 Crystal structure of neuraminidase N1 (2HU4)