IDENTIFICATION OF POTENTIAL NEURAMINIDASE INHIBITORS USING ENSEMBLE-BASED VIRTUAL SCREENING

LIM KOK KEONG

UNIVERSITI SAINS MALAYSIA

2013

IDENTIFICATION OF POTENTIAL NEURAMINIDASE INHIBITORS USING ENSEMBLE-BASED VIRTUAL SCREENING

by

LIM KOK KEONG

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

November 2013

PENGENALPASTIAN PERENCAT NEURAMINIDASE YANG BERPOTENSI MENGGUNAKAN SARINGAN MAYA BERASASKAN ENSEMBEL

oleh

LIM KOK KEONG

Tesis yang diserah untuk
memenuhi keperluan bagi
Ijazah Sarjana Sains

November 2013

ACKNOLEDGEMENTS

First and foremost I offer my sincerest gratitude to my supervisor, Prof. Habibah A. Wahab for supporting me throughout this research project with their patience and knowledge whilst allowing me the room to work in my own way. I attribute the level of my degree to their encouragements, comments, contribution and ideas; without them this research project would not have been completed or written. One simply could not wish for better or friendlier supervisor.

I would like to thank Ministry of Higher Education for providing me the financial support (under MyBrain15 Scheme) for my Master study at University of Science Malaysia.

My special thanks to Dr. Choi Sy Bing, Dr. Belal and Mr. Muhammad Yusuf for providing many valuable information toward my research project. Thus enables my research project to be success. Besides, I would like to thank Prof. Chan Kit Lam for providing me plant crude extract for my research project. Apart from that, I would also like thanks to my batch mates and friends, especially to Nur Kusaira, Lee Guan Sheng, Teh Ban Hong and Hanim for their kind supports and encouragements.

Last but not least, I would like to express my gratification to beloved my parents for continues support and sacrifices toward me throughout all my days at University of Science Malaysia.

TABLE OF CONTENTS

CHAPTER 3 – RESULTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Acceptor
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
3D	Three-dimensional
4-MU	4-methylumbelliferone
$CaCl_2$	Calcium chloride
cRNA	complementary RNA
D	Donor
DANA	2-deoxy-2,3-dehydro-N-acetylneuraminic acid
ER	Endoplasmic reticulum
FDA	Food and Drug Administration
FEB	Free energy of binding
Glu	Glutamic acid
Gly	Glycine
HA	Hemagglutinin
His	Histidine
IC_{50}	Half maximal inhibitory concentration
Ile	Isoleucine
Lys	Lysine
M1	Matrix protein
M1NS	Molecular Dynamic representative snapshot of 1 ns (from mutant-type structure-PDB ID:3CKZ)

M4NS	Molecular Dynamic representative snapshot of 4 ns (from mutant-type structure-PDB ID:3CKZ)
M9NS	Molecular Dynamic representative snapshot of 9 ns (from mutant-type structure-PDB ID:3CKZ)
M2	Ion channel protein
MD	Molecular Dynamic
MES	2-Morpholinoethanesulfonic acid
mRNA	messenger RNA
MUNANA	2’2’-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid sodium salt hydrate
N1	Neuraminidase subtype-1
N2	Neuraminidase subtype-2
NA	Neuraminidase
NADI	Natural Product Discovery System
NANA	α-D-N-acetylneuraminic acid
NaOH	Sodium hydroxide
NCI	National Cancer Institute
NEP	Nuclear export protein
NIH	National Institutes of Health
NLS	Nuclear localization signals
NMR	Nuclear magnetic resonance
NP	Nucleoprotein
NPC	nuclear pore complex
NS1	Non-structural basic protein 1
NS2	Non-structural basic protein 2
PA	Polymerase acidic protein

PB1	Polymerase basic protein 1
PB1-F2	Polymerase basic protein 1-F2
PB2	Polymerase basic protein 2
PDB	Protein Data Bank
Pro	Proline
RdRp	RNA dependent RNA polymerase
RFU	Relative Fluorescence Units
RMSD	Root Mean Squared Deviation
RNP	Ribo-nucleoproteins
SA	Sialic acid
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine
vRNA	Viral Ribonucleic acid
vRNP	viral ribo-nucleoprotein
W1NSC0	Molecular Dynamic representative snapshot of 1 ns cluster 0 (from wild-type structure-PDB ID:2HU4)
W1NSC2	Molecular Dynamic representative snapshot of 1 ns cluster 2 (from wild-type structure-PDB ID:2HU4)
W6NS	Molecular Dynamic representative snapshot of 6 ns (from wild-type structure-PDB ID:2HU4)
W9NS	Molecular Dynamic representative snapshot of 9 ns (from wild-type structure-PDB ID:2HU4)

LIST OF SYMBOLS

#	Number
%	Percentage
μgmL^{-1}	Microgram per milliliter
Å	Angstrom
g	Gram
$kcalmol^{-1}$	Kilocalorie per mole
M	Molarity
mg	Milligram
mL	Milliliter
mM	Millimolar
µM	Micromolar
nm	Nanometer
ns	Nano second
ΔG	Free energy of binding
µL	Microliter

PENGENALPASTIAN PERENCAT NEURAMINIDASE YANG BERPOTENSI MENGGUNAKAN SARINGAN MAYA BERASASKAN ENSEMBEL

ABSTRAK

Sehingga kini, wabak virus Influenza A telah menyebabkan impak yang serius dalam kesihatan manusia. Ia telah muncul sebagai ancaman wabak di seluruh dunia pada abad ke-21 dan menjejaskan populasi manusia yang besar setiap tahun. Pada masa ini, Oseltamivir (Tamiflu) dan Zanamivir (Relenza) digunakan sebagai pilihan ubat yang penting untuk penyakit wabak influenza. Walaubagaimanapun, rintangan virus influenza kepada ubat-ubatan ini telah dilaporkan kebelakangan ini. Jadi, pencarian perencat influenza yang baru adalah sangat penting untuk mengatasi penularan wabak influenza ini. Projek ini adalah mengenai penemuan perencat baru yang berpotensi untuk merencatkan virus Influenza melalui kaedah penskrinan virtual berasaskan pendokkan ensembel. Sebagai reseptor yang membebaskan sel daripada virus, neuraminidase (NA) telah menjadi target reseptor yang terkenal untuk direncat dan kajian ini tertumpu kepada NA subjenis-1. Variasi konformasi NA daripada Protein Data Bank (PDB) dan struktur simulasi dinamik molekul (MD) telah digunakan dalam kajian ini. Dengan bantuan perisian komputer, penskrinan virtual berasaskan pendokkan ensembel telah dijalankan. NA telah disaringkan terhadap pangkalan data daripada Institut Kanser Kebangsaan US atau “National Cancer Institute” (NCI) dan Sistem Penemuan Hasil Semulajadi “Natural Product Discovery System” (NADI) untuk penemuan sebatian yang berpotensi sebagai perencat NA. Dari hasil pendokkan, 20 sebatian daripada Pangkalan Data NCI telah dipilih. Untuk Pangkalan Data NADI, terdapat 40 sebatian telah dipilih dan mereka telah berkelompok kepada 7 tumbuh-tumbuhan. Semua sebatian (NCI dan NADI)

mampu mengikat kepada semua 13 struktur ensembel. Ini telah menunjukkan activiti-activiti yang berkemungkinan untuk anti-neuraminidase. Kemudiannya, sebatian ini tertakluk kepada penilaian perencatan aktiviti melalui MUNANA assay. Dari hasil assay, ia telah menunjukkan bahawa kebanyakan sebatian NCI dan kesemua ekstrak tumbuhan menunjukkan perencatan aktiviti NA untuk kedua-dua jenis reseptor NA samada jenis normal ataupun jenis mutan. Secara khususnya, NSC5069, NSC114449, NSC343344 dan NSC373427 dari Pangkalan Data NCI dan Jambu Batu daripada Pangkalan Data NADI telah menunjukkan perencatan aktiviti NA yang lebih baik dengan nilai IC_{50} dari 118.0 μgmL^{-1} sehingga 250.0 μgmL^{-1} untuk kedua-dua jenis reseptor NA samada jenis normal ataupun jenis mutan. Walau bagaimanapun, kesan potensi perencatan mereka didapati lebih rendah berbanding dengan DANA [4 μgmL^{-1}]. Oleh itu, sebatian ini dianggap sebagai perencat yang lemah untuk NA subtype-1. Walaupun perencatan aktiviti NA untuk sebatian ini tidak sebaik DANA, sebatian ini masih mampu menunjukkan perencatan aktiviti NA. Oleh itu, penerokaan Pangkalan Data NADI dan NCI melalui penskrinan virtual berasaskan pendokkan ensembel telah menunjukkan kesan perencatan yang memuaskan pada neuraminidase subjenis -1.

IDENTIFICATION OF POTENTIAL NEURAMINIDASE INHIBITORS USING ENSEMBLE-BASED VIRTUAL SCREENING

ABSTRACT

To date, influenza A virus cause a serious impact in human health. It has emerged as a worldwide pandemic threat in the 21st century where large human populations were affected annually. At present, Oseltamivir (Tamiflu) and Zanamivir (Relenza) have become important treatments for influenza infectious disease. Unfortunately, the resistance of influenza viruses to these drugs has been reported recently. So, it is important to discover new anti-influenza inhibitors to overcome the on-going and potential influenza outbreak. This project is about the discovery of the potential inhibitor for influenza infectious disease via ensemble-based virtual screening. As a receptor destroying enzyme, neuraminidase has been widely used as a drug target for drug discovery. Thus, this study was focused on Neuraminidase subtype-1. Variation of neuraminidase conformations from Protein Data Bank (PDB) and molecular dynamics (MD) simulation structures were used in this study. With the aid of computational resource, ensemble-based virtual screening was performed. Neuraminidase was screened against the National Cancer Institute (NCI) Database and the Natural Product Discovery System (NADI) Database to discover the potential compounds as the neuraminidase inhibitors. From docking results, 20 compounds from NCI Database were selected. For NADI Database, there were 40 compounds have been selected and they were clustered into 7 plants. All these compounds (NCI and NADI) were able to bind to all 13 ensemble structures. This has exhibited the probable anti-neuraminidase activity. These compounds were then subjected to inhibitory activity evaluation via MUNANA assay. From the assay results, they showed that most of the NCI compounds and all plant extracts have

exhibited NA inhibitory activity on both wild-type and mutant-type of NA. Particularly, NSC5069, NSC114449, NSC343344 and NSC373427 from NCI Database and Jambu Batu from NADI Database showed a better NA inhibitory activity with IC_{50} value ranging from 118.0 μgmL^{-1} to 250.0 μgmL^{-1} on both wild-type and mutant-type NA. However, their inhibition potency were found lesser compared to DANA [4 μgmL^{-1}]. Thus, these compounds were considered as weak inhibitors toward NA subtype-1. Although the NA inhibitory activities of these compounds were not as good as DANA, these compounds have exhibited NA inhibitory activities. Hence, this suggested that exploration of NADI and NCI Database through ensemble-based virtual screening had demonstrated promising inhibitory effects on neuraminidase subtype-1.

CHAPTER 1

INTRODUCTION

1.1 General background

Influenza which is also known as flu is an infection disease of respiratory tract (Ge et al., 2010). Every year, millions of peoples are affected (Wang et al., 2006). Influenza infection is characterized as acute febrile illness. It causes variable degree of systemic symptoms such as fever, cough, sore throat, sputum production, nasal obstruction and sneezing. For complicated case, bronchitis and pneumonia are observed. Sometime, aggravation of asthma and chronic obstructive pulmonary disease will happen. The severity of influenza infection can range from mild to severe. For serious case, it can lead to death (Johnston, 2001, CDC, 2011a).

To date, influenza infectious diseases remain as a life threatening illness due to its ability to spread rapidly and undergo antigenic drift and antigenic shift (Lin et al., 2004). Influenza virus generally attacks human respiratory tract and it doesn't have to travel very far into body to take root. This is why until today, influenza virus exists and brings serious impact to human health. Because of the ability of antigenic drift and antigenic shift of influenza virus, periodic epidemics and occasionally pandemics happen causing significant morbidity and mortality in human population and pose serious impact in global economy (Webster et al., 1992, Cox and Subbarao, 2000, Roberts, 2008).

Influenza viruses are seems to be change and evolve in two different ways: antigenic drift and antigenic shift. Antigenic drift occur gradually by randomly accumulates point mutations within the viral glycoproteins genome (in the HA, the NA, or both) in response to immune pressure (Murphy and Webster, 1996). Point

mutations results in the minor changes to the viral glycoproteins and this cause virus may not recognized by the host immunity. So, the newly forms viruses cannot be recognize and defend well by antibodies induced by the previous infectious strains (Olsen, 2002). Antigenic shift is a sudden and more dramatic form of genetic and antigenic change in influenza virus. This can occurs through genetic reassortment between two or more different virus strains, mixing together to produce a new strain of influenza virus (Olsen, 2002). As influenza A viruses have a broad range of susceptible host to infect, reassortment and recombination of genetic materials from different animal species are commonly detected. The resulted changes will thus create novel subtypes that have not been present in human viruses for a long time. Consequently, the introduction of a novel strain into human population is usually a pandemic or a worldwide epidemic that will result in hundreds of thousands or millions of influenza-related deaths (Scholtissek, 1994, Treanor, 2004).

In the 21st century, influenza pandemic turns up four times, in 1918, 1957, 1968 and 2009 (Table 1.1). Each of these pandemic was caused by the emergence of new strain of influenza virus. Influenza pandemic occur when influenza virus spread rapidly throughout the world and affecting a large proportion of human population. It usually happens with the emergence of a novel strain of influenza virus and caused serious impact to human health (Michaelis et al., 2009).

Historically, 1918 pandemic (Spanish flu) caused the most devastating impact to human population where about 25% of the world's population affected. It is one of the worst natural disasters in modern history. Estimated, Spanish flu killed over 50 million peoples globally and it was caused by influenza virus H1N1 subtype. Usually influenza kills the weaker member of the human population, but Spanish flu affected healthy adults quickly. It seriously affected human in the age between 20 to

Table 1.1 Summaries of influenza pandemics

Year	Name	Virus subtype	Global Deaths
1918	Spanish flu	H1N1	20-50 million
1957	Asian flu	H2N2	2-4 million
1968	Hong Kong flu	H3N2	1-2 million
2009	Swine flu	H1N1	In progress

Adapted from (Michaelis et al., 2009, Hudson, 2009).

40 years and half of the influenza death was made up of this group (Johnson and Mueller, 2002, Taubenberger and Morens, 2006). In 1957, there was another pandemic occurred which called Asian flu. This 1957 pandemic was caused by influenza virus H2N2 subtype resulted in 2-4 million peoples died. Although the proportion of people affected was high, Asian flu pandemic was relatively mild compared to Spanish flu pandemic (Taubenberger and Morens, 2010). During early of 1963, third pandemic happened in Hong Kong thus was then known as Hong Kong flu. Hong Kong flu was caused by H3N2 subtype of influenza virus which replaced previous the H2N2 subtype virus. It is estimated that 1-2 million peoples died from this pandemic (Taubenberger and Morens, 2010). Influenza viruses are undergoing antigenic drift and antigenic shift from time to time and during 2009 resulting in another pandemic. Influenza virus H1N1 subtype appeared again after the 1918 pandemic. This time, the pandemic was known as swine flu or 2009 pandemic influenza virus which was actually derived from four different strains of influenza virus result of re-assortment of North American swine influenza, North American avian influenza, human influenza and a swine influenza typically found in Asia and Europe. It was entirely new combination that was not seen before (Ge et al., 2010, Taubenberger and Morens, 2010). The first case was detected in Mexico during March-April 2009 (Perez-Padilla et al., 2009). Since March 19, 2010, World Health Organization (WHO) reported there were a million of cases detected and from that number at least 16,813 deaths were documented (Taubenberger and Morens, 2010). Although swine flu was relatively mild compared to Spanish flu, pandemic 2009 is still a threat (Padlan, 2010). Therefore, efforts to improve understanding of the pathogenicity and transmissibility of influenza are important and develope new

and improved antiviral and vaccine are crucial for controlling the future influenza outbreak (Ge et al., 2010).

Although vaccines play important role in preventing the influenza, but the use still have significant drawbacks as they provide limited scope of control. Furthermore, the virus mutates rapidly in order to escape from human immune system (Johnston, 2001). Antiviral is another option for influenza treatment. The number of antiviral has increase in the past few years (Table 1.2) where antiviral agents targeting M2 ion channels and neuraminidase (NA) of influenza have been introduced. M2 ion channels inhibitors were the first drugs available for influenza treatment. Amantadine and Rimantadine are two Adamantane-based M2 ion channel inhibitors. Both of these drugs are effective only for influenza A virus because influenza B and C virus does not have M2 ion channel protein. Currently, Amantadine and Rimantadine have not been widely used because of their side effects and rapid emergence of drug resistance strains (Johnston, 2001, Ge et al., 2010). Zanamivir and Oseltamivir are other antiviral drugs but they are targeting on NA. Both these drugs are effective against all types of influenza viruses. Although NA inhibitors are important option for influenza treatment, the increasing use of these drugs will raise the chance of the emergence of drug resistance influenza strain (Johnston, 2001, Ge et al., 2010).

Recently, the resistance of influenza virus to NA inhibitors has been reported. The resistance of influenza virus to Zanamivir has been observed in an immune-compromised child (Gubareva et al., 2000). During late January 2008, there are unexpected spreading of Oseltamivir resistance influenza strain in Europe country. The resistance occured because of the substitution of amino acid H275Y (H274Y in N2 numbering) in NA of these viruses (Lackenby et al., 2008). There were also a

Table 1.2 Anti-influenza drugs

Drug	Originator/Licensee	Year (country) of first launch
Amantadine hydrochloride (Symmetrel®)	Endo	1964
Rimantadine hydrochloride (Flumadine®)	Forest	1987 (France)
Zanamivir (Relenza®)	Biota/GlaxoSmithKline	1999 (Australia)
Oseltamivir (Tamiflu®)	Gilead/Roche	1999 (Switzerland)

Adopted from (Ge et al., 2010)

large portion of Oseltamivir-resistance influenza virus circulating in Norway and ready to be transmitted between peoples (Lackenby et al., 2008). Up until 8 December 2009, WHO reported there are 109 Oseltamivir-resistance viruses have been detected worldwide (WHO, 2009). This provokes the need for discovery and development of new and novel drug to fight the on-going and future influenza outbreak.

1.2 Biology of influenza virus

1.2.1 Classification of influenza virus

Influenza virus is a spherical shape enveloped virus that belongs to Orthomyxoviridae family (Kawaoka, 2006, Cheung and Poon, 2007, Bouvier and Palese, 2008). Influenza virus is categorized into three serologically distinct types: type A, type B and type C based on the antigenic characteristics of the core proteins such as nucleoproteins and matrix proteins (Fields et al., 1996, Hampson and Mackenzie, 2006, Cheung and Poon, 2007, Ferraris and Lina, 2008). The different types of influenza viruses are summarized in Table 1.3. Influenza A viruses are the most common virulent pathogens that caused major pandemics around the world since 1918 (Cox and Subbarao, 1999, Kilbourne, 2006, Lewis, 2006). Influenza B viruses are less common and it produces less severe cases such as seasonal epidemics (Osterhaus et al., 2000, Hay et al., 2001). Influenza C cases are seldom reported but it can cause local epidemics (Matsuzaki et al., 2002). Influenza C virus can attack upper and lower respiratory system of their host. Bronchitis and pneumonia are usually observed when lower respiratory are infected. (Katagiri et al., 1983, Moriuchi et al., 1991, Matsuzaki et al., 2002).

Table 1.3 Type of influenza virus

Type of influenza virus	Number of gene segment	Surface glycoprotein	Host range
Influenza A virus	8	Hemagglutinin & Neuraminidase	Wide (humans, pigs, horses, whales, seals and birds)
Influenza B virus	8	Hemagglutinin & Neuraminidase	Humans and seals
Influenza C virus	7	Hemagglutinin-esterase-fusion (HEF)	Mainly humans, but also found in swine

Adapted from (Lamb and Krug, 2001, Cheung and Poon, 2007)

These three serologically distinct type influenza viruses have a similarity where they infect human causing the disease in human population. Influenza A viruses are likely to have more widely host range where they can infects birds (swine, chicken, duck, turkey and geese), humans and many other species of mammals (pigs, horses, seals). Influenza viruses type B and C have only limited host range where they infect human and some mammal species (Sun, 2009).

The major differences among influenza virus are the glycoprotein found on the envelope of virion and the genetic material of influenza viruses (Figure 1.1). Influenza A and influenza B virus contain 2 enveloped glycoproteins namely Hemagglutinin (HA) and NA while Influenza C virus contains only one envelope glycoprotein called Hemagglutiin-esterase fusion protein. Influenza virus can also be distinguished by their genetic materials where influenza A and B virus contain eight RNA segment but influenza C virus contains only seven RNA segments (Cox et al., 2010). Among them, influenza A viruses tend to become most of the public attention because of high mortality rate happen in human population. Therefore our studies mainly focused on the influenza A viruses because of its virulence, widely host range and pandemic treat.

1.2.2 Genome and structure of influenza A virus

Influenza A virus is an enveloped virus and it is roughly spherical in shape which consists of an outer lipid bilayer membrane (Figure 1.2). Lipid bilayer membrane of influenza virus is actually derived from the plasma membrane of the infected host cell. M1, Matrix protein found underneath the lipid bilayer membrane of influenza virus forms a shell that gives strength and rigidity to the lipid envelope. On the surface of the lipid envelope are glycoproteins and M2 ion channel proteins.

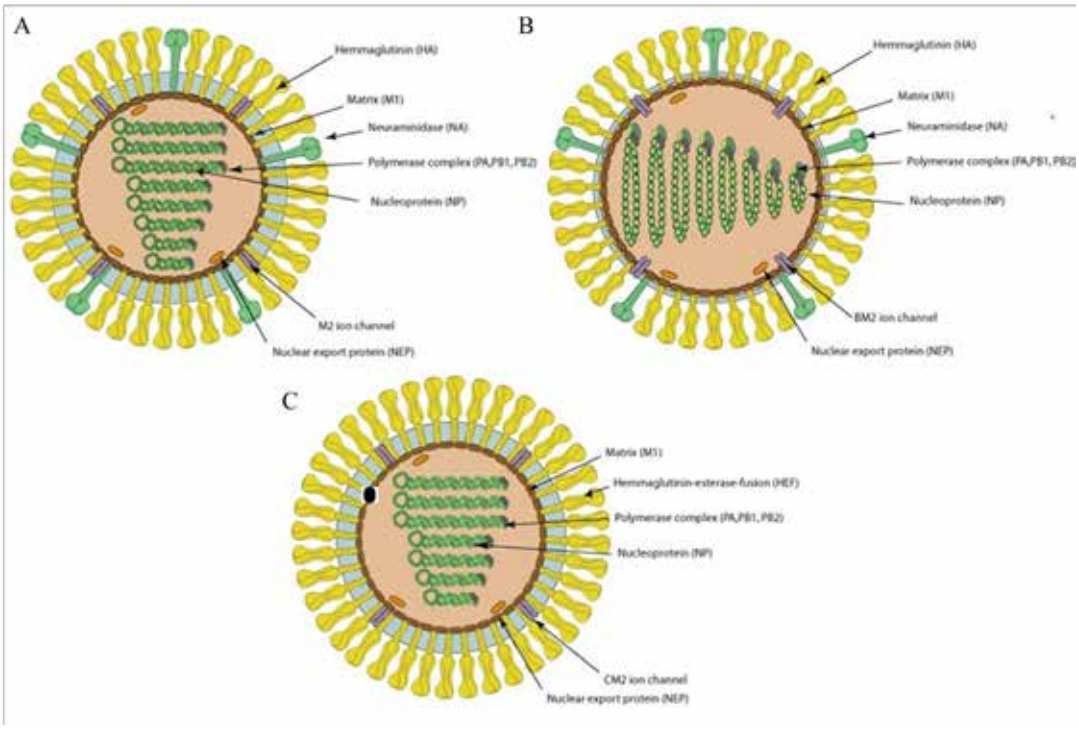

Figure 1.1 Structural of three different types of influenza viruses: A, B, and C. Adopted from (SIB, n.d.)

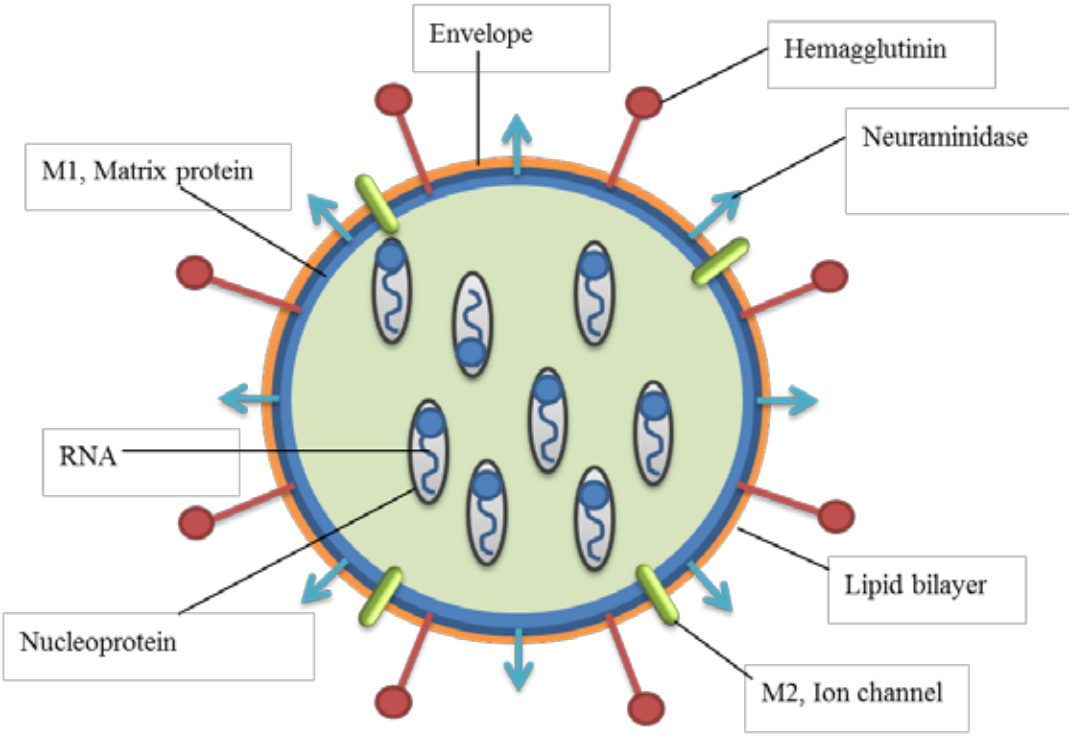

Figure 1.2 The structure of influenza A virus

The glycoproteins look like a spike protrudes from the lipid bilayer membrane of virus. There are two different types glycoproteins found on the surface of the envelope of influenza virus: HA and NA. Influenza A virus are categorised into different subtypes based on the antigenic properties of the viral envelope surface glycoprotein HA and NA. There are 16 known types of HA (H1-H16) and 9 known types of NA (N1-N9) (Webster et al., 1992). Different subtype of influenza A virus encodes for different HA and NA glycoprotein and are name based on the combinations of these two viral envelope surface glycoprotein (Russell et al., 2006a, Palese and Shaw, 2007, Bouvier and Palese, 2008).

The genetic materials of influenza virus are embedded within the interior of the virion. There are total 8 negative-stranded RNA segments, covered by nucleoprotein (NP). These 8 RNAs carry all the important information needed for making a new influenza particle. Different RNA encodes for different protein (Table 1.4) (McGeoch et al., 1976, Bouvier and Palese, 2008, Racaniello, 2009b). Each of viral RNA wrapped into viral ribo-nucleoprotein (vRNP) complex (Figure 1.3) by nucleoproteins (NP) and RNA-dependent RNA polymerase complex: PB1, PB2 and PA (Cheung and Poon, 2007, Chan, 2009).

1.2.2.1 Neuraminidase

Neuraminidase (NA), also known as sialidase, is the major envelope glycoprotein of influenza virus. It is a glycoside hydrolase (EC3.2.1.18) that has essential role during viral replication and infection. NA has a shape that looks like mushroom where it has a long thin stalk with a spherical head on the top. Generally, NA is constituted of a stalk region, a globular head, a cytoplasmic tail and a transmembrane domain. The cytoplasmic tail is anchored in the lipid bilayer of the

Table 1.4 Genetic materials of Influenza A virus and encoded proteins

RNA Segments	Encoded Proteins
1	Polymerase basic protein 2, PB2
2	Polymerase basic protein 1, PB1; Polymerase basic protein 1-F2, PB1-F2
3	Polymerase acidic protein, PA
4	Hemagglutinin, HA
5	Nucleoprotein, NP
6	Neuraminidase, NA
7	Matrix protein, M1; Ion channel protein, M2
8	Non-structural basic protein 1, NS1; Non-structural basic protein 2, NS2

Adapted from (Nicholson et al., 1998, Cheung and Poon, 2007, Michaelis et al., 2009)

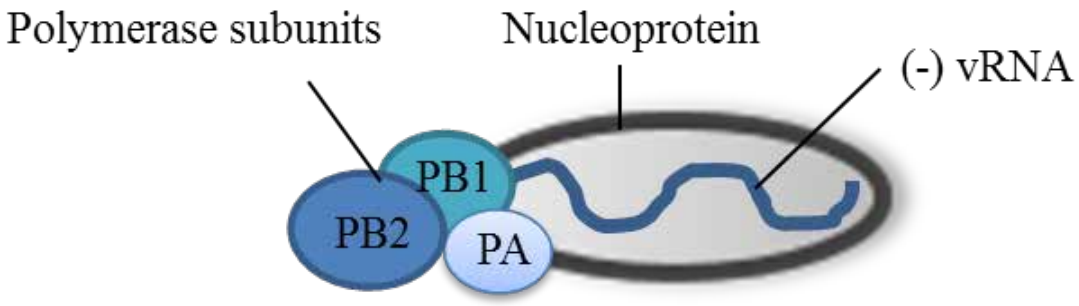

Figure 1.3 The structure of viral Ribo-nucleoprotein complex (vRNP complex). Adapted from (Agustin and Paul, 2002)

viral membrane that makes the rest of NA structure projected out. (Laver and Valentine, 1969, Grienke et al., 2012).

NA play critical role in facilitating the release of the newly form progeny virion to infect other host cell and prevent them from self-aggregation (Li et al., 2010). NA destroys receptors recognised by hemaglutinin by catalyses the cleavage of the terminal of α-ketosidic linked sialic acid residue from the infected cell during the final step of the replication cycle. The cleavage enables progeny virions to bud out from the infected cell and spread the infection to neighbouring cell. (Palese et al., 1974). Besides, NA also assists the mobility of the influenza virus within the host respiratory tract. This is done by decoy receptors found on the mucins, cilia, and cellular glycocalyx of host respiratory tract and allows virus to go through host ciliated epithelium to reach to the host cells for infection. (Matrosovich et al., 2004, Gong et al., 2007).

In nature, there are 9 NA subtypes have been identified in the influenza virus family (Swayne and Halvorson, 2003). They can be classified into 2 phylogenetic groups (Figure 1.4): Group 1 is consisting of N1, N4, N5, and N8 and Group 2 include N2, N3, N6, N7 and N9 (Russell et al., 2006a). Recently, a comparative study of crystal structure of NA showed that, Group 1 NA contains an extra cavity near the binding site while this cavity is lacked in Group 2. This is because 150-loop of Group 2 NA retains a close conformation, so it leads to the missing of this cavity. (Russell et al., 2006a, Xu et al., 2008).

The crystal structure of head domain of NA and its complex: sialic acid and other inhibitors have been identified and solved (Varghese et al., 1992). It is actually composed of four identical subunits forming a mushroom-like shape homotetramer (Figure 1.5). Analysis of these structures showed that NA has a highly conserved

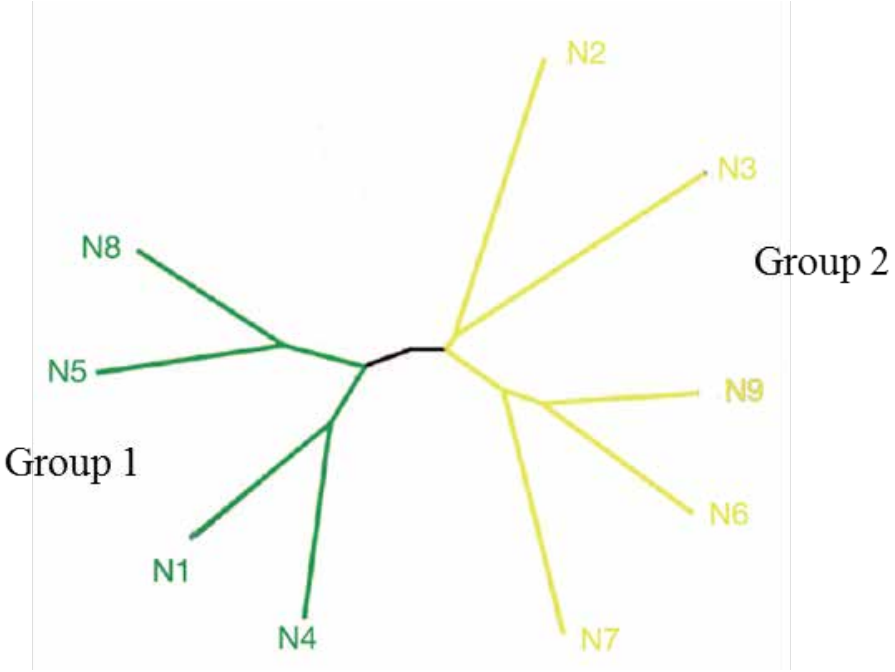

Figure 1.4 Phylogenetic tree of neuraminidase of influenza A virus. Nine subtypes of NA have been identified in nature and are classified into 2 groups: Group1 and Group 2. Adapted from (Russell et al., 2006a)

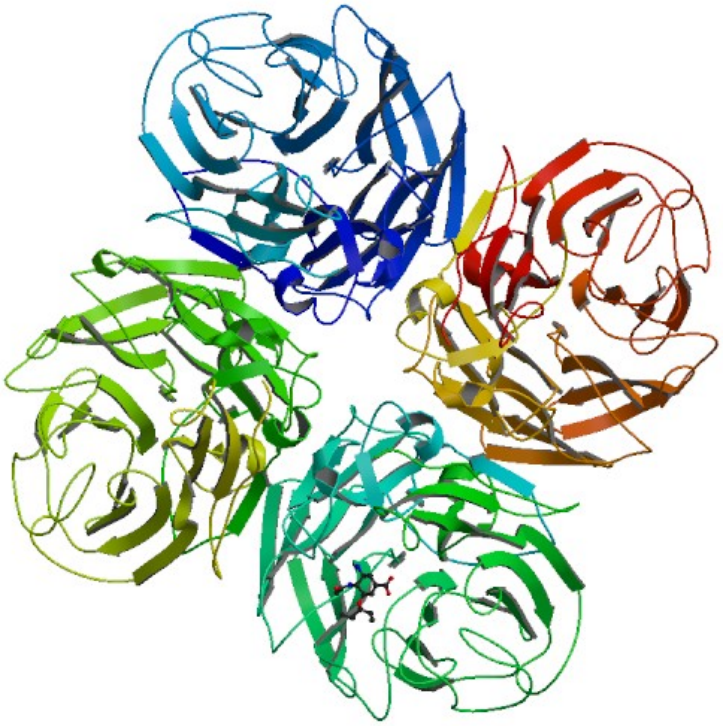

Figure 1.5 Three-dimensional structure of an N1 neuraminidase tetramer [PDB entry: 2HU0]. Neuraminidase is making of four identical polypeptides. Adopted from (Russell et al., 2006a)

binding site across all serotypes of influenza A and B viruses (Russell et al., 2006a). NA was an established target in structure-based drug design (Grienke et al., 2012). Due to this property, NA becomes an attractive target for drug discovery and development. The use of NA crystal structure in drug discovery and development has resulted in the discovery of two potent NA inhibitors; Zanamivir and Oseltamivir (von Itzstein et al., 1993, Kim et al., 1997).

1.2.2.2 Hemagglutinin

Hemagglutinin (HA) is a viral envelope glycoprotein. It is a triangular rod-shape molecule that protrudes from the influenza virus envelope. HA have two specific roles during the replication cycle of influenza virus. First, it provides initial contact of influenza virus to the host cell by attaching the virus to the sialic acid receptor found on the surface of the host cells (Russell et al., 2006b). Then, it mediates the entry of influenza virus into the cytoplasm of host cell by triggering membrane fusion process of virus and host cell (Chizmadzhev, 2004, Sollner, 2004, von Itzstein, 2007). During 1981, Wilson and colleagues have published the first HA structure (Wilson et al., 1981). Since then, there is continuously appearance of other HA crystal structures resolved and deposited in Protein Data Bank (PDB). Both the HA and fusion activity of influenza virus have been targeted in drug discovery and development research, but until today there is no successful drug come out from this approach (Gong et al., 2007).

1.2.2.3 M2 and other viral proteins

M2 protein is an integral membrane protein that is distributed on the cell membrane of influenza virus (Lamb et al., 1985). It acts as proton channel that

mediates the pH regulation activity of viral particles trapped in endosomes. M2 protein allows the acidification of the interior of the viral particle inside the endosome which then enables the release of the vRNP into the host cell (Martin and Helenius, 1991a, Bui et al., 1996). The ion channel activity of M2 protein is also important for stabilizing the native conformation of the newly formed HA during intracellular transport of HA for viral assembly. This process is done by maintaining a high pH in the Golgi vesicles (Takeuchi and Lamb, 1994). A number of M2 proteins have been determined and it is reported to exist as homotetramer. There are two known M2 inhibitors: Amantadine and Rimantadine. Both are bound to the transmembrane region of M2 protein to inhibit the H^+ ion from entering into the viral particles (Ge et al., 2010).

Other interesting proteins such as nucleoprotein, viral polymerase and non-structural protein 1, NS1 have been targeted for drug discovery and development (Das et al., 2010). Nucleoprotein is the molecule that wraps viral RNA. It is a part of vRNP where numerous copies of nucleoproteins packaged together with viral RNA and viral polymerase to form a vRNP complex. Nucleoproteins are believed to take part in transportation of vRNPs and viral replication. Viral polymerase exists as a heterotrimer and it is composed of PA, PB1 and PB2. It is important during replication cycle for viral transcription and replication. Non-structural protein 1 (NS1) is a multifunctional protein that exists as an oligomer and it is normally found in the nucleus. NS1 protein is important for regulation of viral and cellular protein expression (Cheung and Poon, 2007). Until today, there is not any small molecule against these targets are reported (Das et al., 2010). The rest of viral proteins are less common thus are not discussed here.

1.2.3 Life cycle of influenza virus

Influenza virus cannot reproduce outside the cell; the reproduction must take place within a living host cell. Therefore, Influenza virus requires a host cell for reproduction. The life cycle of influenza A virus can be divided into multiple steps, which is the attachment and the entry of the virus into the host cell, transport of ribo-nucleoproteins into host nucleus, synthesis of viral RNA, synthesis of viral proteins, and assembly of new particles and release of particles from the host cell (Figure 1.6).

1.2.3.1 Attachment of influenza virus into the host cell

Influenza virus begins the life cycle by attaching to the sialic acid containing receptors found on the host cell surface (Figure 1.8). The attachment involved the binding of influenza outer membrane glycoprotein HA to the sialic acid receptors of host (Figure 1.7) (Mochalova et al., 2003). HAs recognise the sialic acid moiety receptors and then bind to them (Bouvier and Palese, 2008).

1.2.3.2 Virus entry

Once virus attaches to the host cell, the bound virus is taken up by the host cell via receptor-mediated endocytosis in an endosome. (Palese and Shaw, 2007, Luo, 2012). Inside the host cell, the fusion and un-coating of influenza virus involved the acidification of the endosome environment to a low pH (pH5-6). A pH drop inside the endosome causes the viral HA protein to undergo structural/conformation changes which then result in the fusion of viral and endosomal membrane. The acidic environment of the endosome is not only important for the fusion of viral and endosome membrane but also important to enable the opening of M2 ion channel proteins. The opening of M2 ion channel protein acidified the viral core environment

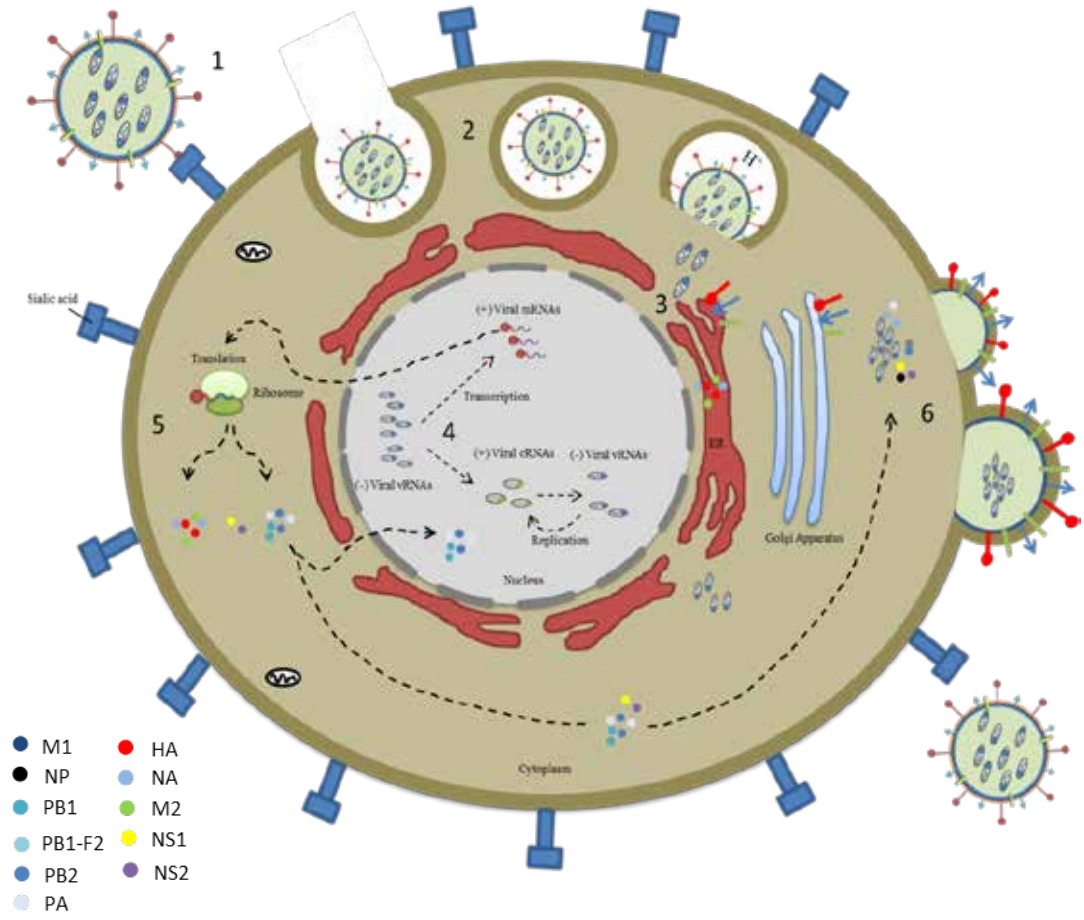

Figure 1.6 The influenza virus replication cycle. 1. Attachment of virus to the host cell, 2. Entry of virus into the host cell, 3. Transport of ribo-nucleoproteins into host nucleus, 4. Synthesis of viral RNAs, 5. Synthesis of viral proteins, and 6. Assembly of new particles and release of particles from the host cell

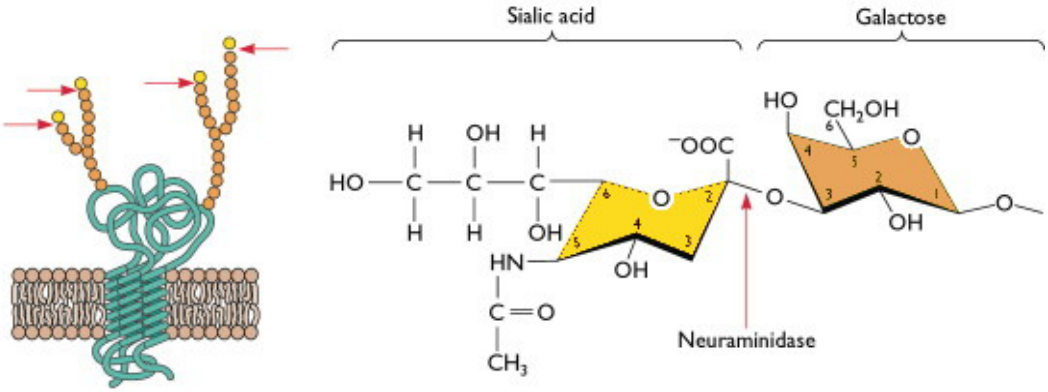

Figure 1.7 Sialic acid on the host cell surface. On the left is the sialic acid found at the terminal position of glycans attached to the cell surface. The spheres (orange) are sugars and the sialic acid (yellow) is found at the last sugar in the chain that attached to protein (cyan). On the right is the chemical structure of SA-galactose linkage. Adopted from (Racaniello, 2009a)

by allowing the influx of H^+ ions into the virion. This process disrupts protein-protein interaction and result the releases of the viral genome segments or Ribo-nucleoproteins (RNPs) from M1 matrix protein to the host cell's cytoplasm (Stegmann et al., 1987, Martin and Helenius, 1991b, Stegmann, 2000, Palese and Shaw, 2007).

1.2.3.3 Transport of Ribo-nucleoproteins into the host nucleus

After the fusion and un-coating of influenza virus, the cycle is continued by the transcription and replication of influenza virus. Transcription and replication of influenza virus occur in the nucleus of the host cell. Therefore, to enable the transcription and replication of influenza virus to be taking place, the viral RNPs must be transported into the host cell nucleus. The size of viral RNPs is relatively large (10-20nm) and they cannot enter the host cell's nucleus through passive diffusion process. They must therefore transport into the nucleus of host cell via active transport mechanism through the host cell nuclear pore complex (NPC) with the aid of viral protein nuclear localization signals (NLS) (Figure 1.6). (O'Neill et al., 1995, Cros and Palese, 2003, Fodor and Smith, 2004, Cros et al., 2005, Palese and Shaw, 2007).

1.2.3.4 Synthesis of viral RNA

Influenza A virus possesses 8 negative-stranded RNA segments. These viral RNAs segments are serve as the templates for the synthesis of messenger RNA (mRNA) and complementary RNA (cRNA). Newly synthesized cRNAs are used as template for the further synthesis of the new viral RNA (Cheung and Poon, 2007, Resa-Infante et al., 2011). In general, the synthesis of cRNA occurs in the early

process before the viral RNA synthesis take place (Resa-Infante et al., 2011). Viral RNA polymerase subunits (PB1, PB2 and PA) and nucleoprotein (NP) are important during the transcription and replication of influenza virus. They enter host cell nucleus as a part of viral RNP complex. The first viral mRNA generated is transported out of the host cell nucleus into host cell cytoplasm for the translation process into protein. In the cytoplasm, the mRNAs will translate into their respectively proteins according to the RNA segments of influenza virus. The newly synthesized viral polymerase proteins and nucleoprotein will then be transported back to the nucleus for further catalyses the mRNA transcription and viral RNA/cRNA replication. This process will keep repeating to produce viral genome. During the late infection process, the process of synthesis of viral RNA is increased (Krug, 1981, Braam et al., 1983, Kawakami and Ishihama, 1983, Huang et al., 1990, Cros and Palese, 2003, Fodor and Smith, 2004, Neumann et al., 2004, Deng et al., 2005, Amorim and Digard, 2006, Engelhardt and Fodor, 2006, Boulo et al., 2007). The newly synthesized viral RNA will form viral RNP complex by binding to RNA polymerase subunits and nucleoprotein and later exported into the cytoplasm to prepare for the construction of new virion (Braam et al., 1983, Cros and Palese, 2003, Boulo et al., 2007).

Unlike other organism, the transcription of influenza virus involved cap-snatching mechanism (Figure 1.8). This is because viral RNAs contain only a poly (A) tail which lacks of 5'-cap. In order to initiate the transcription, influenza virus need to grab 5'-capped RNA fragments from host cell mRNA (Krug, 1981, Hagen et al., 1994, Samji, 2009, Resa-Infante et al., 2011). RNA dependent RNA polymerase (RdRp) constitutes PB1, PB2 and PA proteins are important for the influenza virus transcription. During the transcription process, PB2 binds to 5'-7-methylguanosine