HETEROLOGOUS EXPRESSION OF AVIAN INFLUENZA A (H5N1) NEURAMINIDASE IN *Kluyveromyces lactis* AND *Escherichia coli*

NOOR ZAFIAH TAUFIK

UNIVERSITI SAINS MALAYSIA

2016

HETEROLOGOUS EXPRESSION OF AVIAN INFLUENZA A (H5N1) NEURAMINIDASE IN *Kluyveromyces lactis* AND *Escherichia coli*

by

NOOR ZAFIAH TAUFIK

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

September 2016

ACKNOWLEDGEMENT

Alhamdulillah, thank you Allah for giving me the opportunity to finally write these words. To my supervisor Professor Dr. Mohd Razip Samian, thank you for opening the door and giving me the chance to do this project under your guidance, and most importantly thank you for not giving up on me. To Prof. Nazalan Najimudin, and Prof. Dr. Habibah A. Wahab, thank you for allowing me to use your lab equipment and material during this project.

A wise man used to say, "Family don't end with blood" and 414 lab members are the living proofs of it. Thank you to Kak Qiss, Kak Kem, Kak Aini, and Abang Chai for making the bonding possible. Without the home set up of the lab, the aquarium will not exist and there'll be no precious party memories to remember.

To my fellow labmates, Hani, Sya, Hadi, Faisal, Syafiq, Fui Ling, Suhaimi, Chek, Luan, Khansa, Deva, Pow, Xuan Yi, Haida, Mardani, Suba, Ashraf, Yifen, Chee Wah, Kee Shin, and Peter, thank you for all the good, the bad and the ugly things we did together. Special thanks to Matt and Fifi for keeping up with my "volcanic" behaviour. And to Afiqah Bakhari, thank you for the warmest welcome I ever had on the day I first step my foot in the lab.

I would also like to express my deepest appreciation to my family; Opah, Chu and Paksu, Adik, Omar, Adam and Amira, Hafiz, Nina, and my beloved cousins, Aan, Luqman, Arif, Aimi and Amni. Thank you for your unconditional love and support throughout this arduous journey.

I'm also grateful for the financial support from Ministry of Science, Technology and Innovation (MOSTI) through National Science Fellowship (NSF).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

%	Percentage
×g	Relative centrifugal force (RCF)
µF	Microfarad
µg	Microgram
µL	Microliter
4-MU	4-methylumbelliferone
APS	Ammonium persulfate
BSA	Bovine serum albumin
BSL-3	Biosafety level three
CDC	Centers for Disease Control and Prevention
CFU	Colony forming unit
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
FDA	Food and Drug Administration
g	Gram
gDNA	Genomic DNA
GndCl	Guanidine chloride
HA	Haemagglutinin
IBs	Inclusion bodies
IPTG	Isopropyl β-D-thiogalactopyranoside
kb	Kilobase pair
kDa	Kilo Dalton

L	Liter
LB	Luria-Bertani
LIC	Ligation independent cloning
M	Molar
mA	Milliampere
Mb	Megabase pair
MBP	Maltose binding protein
MCS	Multiple cloning site
min	Minute
mL	Milliliter
mM	Millimolar
mRNA	Messenger RNA
ms	Millisecond
MUNANA	2'-(4-methylumbelliferyl)-α-D-*N*-acetylneuraminic acid
NA	Neuraminidase
NAHD	Neuraminidase head domain
ng	nanogram
nm	Nanometer
^{o}C	Degree Celcius
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pmole	Picomole
RE	Restriction enzyme
Ref	Refolding buffer

RFU	Relative fluorescence unit
RNA	Ribonucleic acid
rNA	Recombinant neuraminidase
rpm	Rotation per minute
SA	Sialic acid
SDS	Sodium dodecyl sulphate
sec	Second
Seq.	Sequence
Sol	Solubilization buffer
TE	Tris-EDTA
TEMED	N, N, N, N-tetramethyl-ethylenediamine
U	Unit
UV	Ultra violet
V	Volt
v/v	Volume per volume
vRNA	Viral ribonucleic acid
vRNPs	Viral ribonucleoprotein complexes
w/v	Weight per volume
WHO	World Health Organization
YCB	Yeast carbon base
α-mf	α-mating factor
Ω	Ohm

PENGEKSPRESAN NEURAMINIDASE VIRUS SELESEMA BURUNG A (H5N1) SECARA HETEROLOGUS DALAM *Kluyveromyces lactis* DAN *Escherichia coli*

ABSTRAK

Pengekspresan neuraminidase (NA) virus H5N1 di dalam sistem pengekspresan mikrob yang ringkas dengan menggunakan *Kluyveromyces lactis* dan *Escherichia coli* sebagai perumah telah dikaji. Untuk penghasilan protein NA dalam *K. lactis*, dua jenis fragmen DNA NA terdiri daripada 1) jujukan penuh (1347 bp) dan 2) jujukan domain kepala (residu 63-449, 1161 bp) NA daripada A/Chicken/Malaysia/5858/2004 (H5N1) telah diamplikasi dan diklon ke dalam vektor pengekspresan pKLAC2 lalu menghasilkan konstruk plasmid yang dinamakan pKLAC2-NA and pKLAC2-NAHD. Kedua-dua plasmid kemudiannya diluruskan dan diintegrasikan ke bahagian promoter *LAC4* di dalam genom *K. lactis* melalui rekombinasi homologus. Walaubagaimanapun, analisis protein yang dijalankan sebaik sahaja pengekspresan protein diinduksi tidak menunjukkan sebarang kehadiran protein rekombinan NA (rNA) terembes daripada mana-mana transforman yis. Oleh itu, jujukan penuh gen NA turut diamplikasi dan diklon ke dalam vektor pengekspresan pET-32 Xa/LIC dan seterusnya digunakan untuk mentransfomasikan *E. coli* Rosetta-gami 2(DE3). Analisis protein yang dijalankan selepas induksi pengekspresan protein menunjukkan bahawa protein rNA bersaiz ~ 62 kDa telah diekspres dalam bentuk jasad rangkuman (IBs). Seterusnya, untuk mendapatkan semula protein rNA yang bio-aktif, jasad rangkuman tersebut telah dipencil, dilarut dan dilipat semula menggunakan kaedah pencairan "pulsatile". Pengaruh pH, suhu dan bahan kimia tambahan turut dikaji. Sekurang-kurangnya 27% protein terlarut berjaya dilipat semula, dan nilai aktiviti enzim yang rendah telah direkodkan.

HETEROLOGOUS EXPRESSION OF AVIAN INFLUENZA A (H5N1) NEURAMINIDASE IN *Kluyveromyces lactis* AND *Escherichia coli*

ABSTRACT

The expression of H5N1 virus neuraminidase (NA) in simple microbial expression system utilizing *Kluyveromyces lactis* and *Escherichia coli* as the host was studied. For NA protein production in *K. lactis*, two types of NA DNA fragment composed of 1) full length sequence (1347 bp) and 2) head domain sequence (residue 63-449, 1161 bp) from A/Chicken/Malaysia/5858/2004 (H5N1) were amplified and cloned into pKLAC2 expression vector generating plasmid constructs named pKLAC2-NA and pKLAC2-NAHD. Both plasmid were then linearized and integrated into the *LAC4* promoter region of *K. lactis* genome through homologous recombination. However, protein analysis carried out upon the induction of protein expression did not exhibit the presence of secreted recombinant NA (rNA) protein by any of the yeast transformants. Hence, the full length sequence of NA gene was also amplified and cloned into pET-32 Xa/LIC expression vector and subsequently used to transform *E. coli* Rosetta-gami 2(DE3). Protein analysis performed after the protein expression induction shows that ~ 62 kDa size rNA protein was expressed in the form of inclusion bodies (IBs). Next, to recover the bioactive rNA protein, the IBs were isolated, solubilized and refolded using pulsatile dilution method. The effects of pH, temperature, and chemicals additives were also investigated. At least 27% of the solubilized IBs were successfully refolded, and low amount of enzyme activity was recorded.

CHAPTER 1

1.0 Introduction

For centuries, influenza virus specifically human influenza virus has become one of the most successful and threatening virus in human population (Adams, 2006). After three influenza pandemics occurred in 20th century, H1N1 'swine flu' arrived as the first influenza pandemic for 21st century (Trifonov *et al.*, 2009). Even though H1N1 virus had always been the major player in the event of influenza pandemics, another pandemic threat comes from highly pathogenic avian influenza (HPAI) A H5N1 virus. According to World Health Organization (WHO) (2015), ever since the first outbreak in 1997 and followed by 2004-2005 epidemic, this virus had constantly causing outbreaks around the globe.

As the name implied, avian influenza A H5N1 mainly infected birds causing millions of loss in the poultry industries. However, cases of human infection had also been reported which mainly due to direct contact with the infected animals (Subbarao *et al.*, 1998; Ku and Chan, 1999). A H5N1 is yet to obtain the ability to easily transmit within human population but recent finding which linked avian influenza virus with 1918 H1N1 virus (Taubenberger, 2006; Qi *et al.*, 2012; Worobey *et al.*, 2014) suggested that the possibility is definitely there. With possible cases of human to human infection reported (Wang *et al.*, 2008; Hill *et al.*, 2009) and the emergence of H7N9 recently, many fear that the event might be sooner than later (Horby *et al.*, 2013). As the virus will continually circulate and mutate, no one can ever predict when the new reasserted H5N1 virus will emerge (Martin *et al.*, 2006). Therefore, developing strategies and preparedness for the prevention of the influenza in the future are worthy steps to take.

Lesson from 2009 H1N1 pandemic showed that many aspect of the contingency plan need to be improved especially the availability and the accessibility of vaccines and antivirals (Morens and Taubenberger, 2012). Due to the unavailability of universal vaccine, influenza infection treatment had normally relied on the antiviral such as M2 ion channel inhibitors (adamantanes) and neuraminidase (NA) inhibitor (oseltamivir, zanamivir and peramivir) [Food and Drug Administration (FDA), 2015]. However, with the ability of influenza viruses to continually mutate, having a long lasting drug is nearly impossible. Adamantanes for example are no longer recommended as at least half of the influenza A subtypes had gained resistance toward the drug [Centers for Disease Control and Prevention (CDC), 2015; Dong *et al*., 2015). To date, only NA inhibitor remains effective albeit the emergence of a few resistance cases reported towards oseltamivir and zanamivir (Samson *et al.*, 2013).

NA inhibitors were designed to block the activity of NA protein of influenza virus which is important for the released of new viral progeny, thus preventing the spread of the viruses in the host body (Shtyrya *et al.*, 2009). When the crystal structure of A (H5N1) NA was discovered and solved in 2006, it was revealed that group-1 NA (N1, N4, N5 and N8) and group-2 NA (N2, N3, N6, N7 and N9) are structurally different. In group-1 NA, a cavity was found adjacent to the active site close to ligand binding hence suggesting possibilities to develop new drug by exploiting this particular region (Russell *et al.*, 2006).

Drug discovery however is a very demanding process where large amount of the protein in pure and active form is required. Due to low number of NA protein on the virus, obtaining the desired amount could be the limiting factor for this process. Most method currently used to produce the protein in large scale are costly

and laborious, and may also restricted to limited number of laboratories equipped with at least biosafety level three (BSL-3) facilities (Yongkiettrakul *et al.*, 2009). Therefore, in this study two types of microbial host, *Kluyveromyces lactis* (yeast) and *Escherichia coli* (bacteria) were utilized to express the influenza A (H5N1) NA.

1.1 Objectives

The main objective of this study is to produce influenza A (H5N1) NA using *K. lactis* and *E. coli* as hosts. The following steps are employed:

1) Transformation and expression of avian influenza A (H5N1) NA in *K. lactis.*
2) Optimization of avian influenza A (H5N1) NA expression in *E. coli.*
3) Screening of refolding condition to recover bioactive recombinant NA produced in *E. coli.*

CHAPTER 2

2.0 Literature Review

2.1 Influenza: epidemiology and virology

Listed alongside Ebola, Hantavirus, and Marburg virus as one of the most deadly virus, influenza earned its reputation by claiming millions of human life throughout the history (Schwar, 2014). The oldest record of influenza was probably written by Hippocrates in 412 BC who was reporting a major epidemic called "Cough of Perinthus" in the seventh chapter of the sixth book of Epidemic as a part of an ancient Greek medical collection known as Corpus Hippocraticum (Adams, 2006; Pappas *et al.*, 2008; Wohlbold and Krammer, 2014). The name influenza however is not that ancient. The word "influenza" appeared for the first time in 1357 AD originating from an Italian word meaning influence. The first influenza pandemic was recorded in 1580 that began in Europe and later found its way to Asia and Africa (Adams, 2006).

Following the timeline, series of influenza epidemic and pandemic had occurred but none could be compared to the 1918-1919 influenza pandemic so called "Spanish Flu" caused by H1N1 virus killing ~ 50 million people or perhaps more (Patterson and Pyle, 1991; Johnson and Mueller, 2002; Crosby, 2003). Other than the unusual mortality rate, many aspect of the pandemic such as the origin of the strain, the vulnerability of young adults (20-40 years old) towards the infection, and the characteristics of three distinct influenza epidemic waves remain obscure (Shanks and Brundage, 2012; Morens and Taubenberger, 2012; Worobey *et al.*, 2014). Later, 20th century witnessed another pandemics; Asian Flu (H2N2, 1957-1958) and Hong Kong Flu (H3N2, 1968-1969) resulting 1 to 4 million of death (van Riel *et al.*, 2010;

Ligon, 2005). Studies showed that the latter pandemic was indeed related to the 1918-1919 influenza (Taubenberger and Morens, 2006).

Due to its glorious history, the arrival of the first 21st century influenza pandemic the 2009 H1N1 was not a surprise. Commonly known as "Swine Flu", the new strain appeared to be a combination of a previous triple re-assortment of bird, swine and human flu viruses with a Eurasian pig flu virus (Trifonov *et al.*, 2009; Dawood *et al.*, 2012). Albeit years of efforts and preparedness, more than 200 000 people were estimated to lost their lives globally and in the United State alone, the death penalty was in the range of 8000 to 18 000 lives (Shrestha *et al.*, 2011; Dawood *et al.*, 2012). This number might be statistically low compared to the older pandemics but the 2009 H1N1 virus showed two outstanding and worrisome features; able to cause major out-of-season epidemics in temperate countries and severely infected the young and healthy people which had been seen in the 1918-1919 influenza pandemic hence raising the anxiety for the emergence of more notorious influenza strain in the future (Shindo and Briand, 2012).

2.1.1 Viral type and structure

There are three types of influenza virus; influenza A, B and C representing three of the five genera of Orthomyxoviridae family (Palese and Shaw, 2007). Even though all of the three types infected humans, influenza B and C almost exclusively replicated in human while influenza A has the widest range of host including non-human such as poultry, pigs, sea mammals, horses and bats (Hay *et al.*, 2001; Taubenberger and Morens, 2010; Wohlbold and Krammer, 2014). These viruses normally evolved through antigenic drift (mutation within the virus) or antigenic shift (reassortment of two or more strains). While influenza B continually undergoes antigenic drift causing severe epidemics, influenza A with its capability to evolve

both ways caused greater impact in terms of annual epidemics and pandemics (Hay *et al.*, 2001).

Based on its diverse genetics, influenza A virus subtypes are further classified based on the sequence and the antigenicity divergence of the virus' surface glycoprotein, haemaglutinin (HA) and neuraminidase (NA) (Palese and Shaw, 2007). Currently there are 18 HA subtypes (H1-18) and 11 NA subtypes (N1-N11) had been identified in which only a certain type of HAs (H1, H2 and H3) and NAs (N1 and N2) were found to circulated among humans. For human cases associated with poultry outbreaks, H5, H6, H7, H9, H10, N3, N7, N8, and N9 are the usual suspects (Wohlbold and Krammer, 2014).

The structure and organization of all influenza are generally similar but influenza A and B are more alike compared to influenza C. The virus particle of influenza A and B are commonly in spherical shape with ~ 100 nm in diameter but can also exceed 300 nm in length when in filamentous shape, whereas for influenza C, the viruses can exist in a long cordlike structure with ~ 500 µm in size on the infected cell surface (Bouvier and Palese, 2008). These enveloped viruses also share the same genetic material which is negative sense single-stranded RNA albeit different in number, seven pieces for influenza C and eight for influenza A and B. In total the genome size of influenza A and B is ~ 13, 600 and ~ 14, 600 bp respectively and as influenza C has one less RNA segment the size is smaller ~ 10, 000 bp (Klenk, 2007; Swiss Institute of Bioinformatics, 2010). Through mRNA splicing, the eight pieces of influenza A and B vRNA encode ten or eleven proteins needed for the survival of the virus. These proteins are 1) basic polymerase subunit; PB1 and PB2 (some strain of influenza A produce an accessory protein PB1-F2), 2) acidic polymerase subunit PA, 3) surface glycoprotein; HA and neuraminidase NA, 4)

nucleoprotein NP (RNA binding protein), 5) matrix protein; M1, M2 for influenza A, and NB and BM2 for influenza B, and 6) the interferon antagonist protein; non-structural protein 1 (NS1) and nuclear export protein (NEP/NS2). The main difference between influenza C and influenza A and B genomic composition is that, there is only one glycoprotein present on influenza C viral surface which is haemagglutinin-esterase-fusion (HEF) protein replacing the HA and NA proteins on the surface of influenza A and B (Bouvier and Palese, 2008).

As shown in Figure 2.1, despite the variations of the envelope composition, the virion organization of all influenza is quite alike (Bouvier and Palese, 2008). The viral core contains NEP protein and helical viral ribonucleoprotein complexes (vRNPs) composed of NP coated RNA segments and heterotrimeric RNA-dependent RNA polymerase (consist of PB1, PB2 and PA) (Nayak *et al.*, 2009). Surrounding the core is the lipid-containing viral envelope made by M1 protein and the variations of envelope proteins (Influenza A: M2, HA, and NA; influenza B: BM2, NB, HA and NA; influenza C: CM2 and HEF) (Klenk, 2007; Bouvier and Palese, 2008).

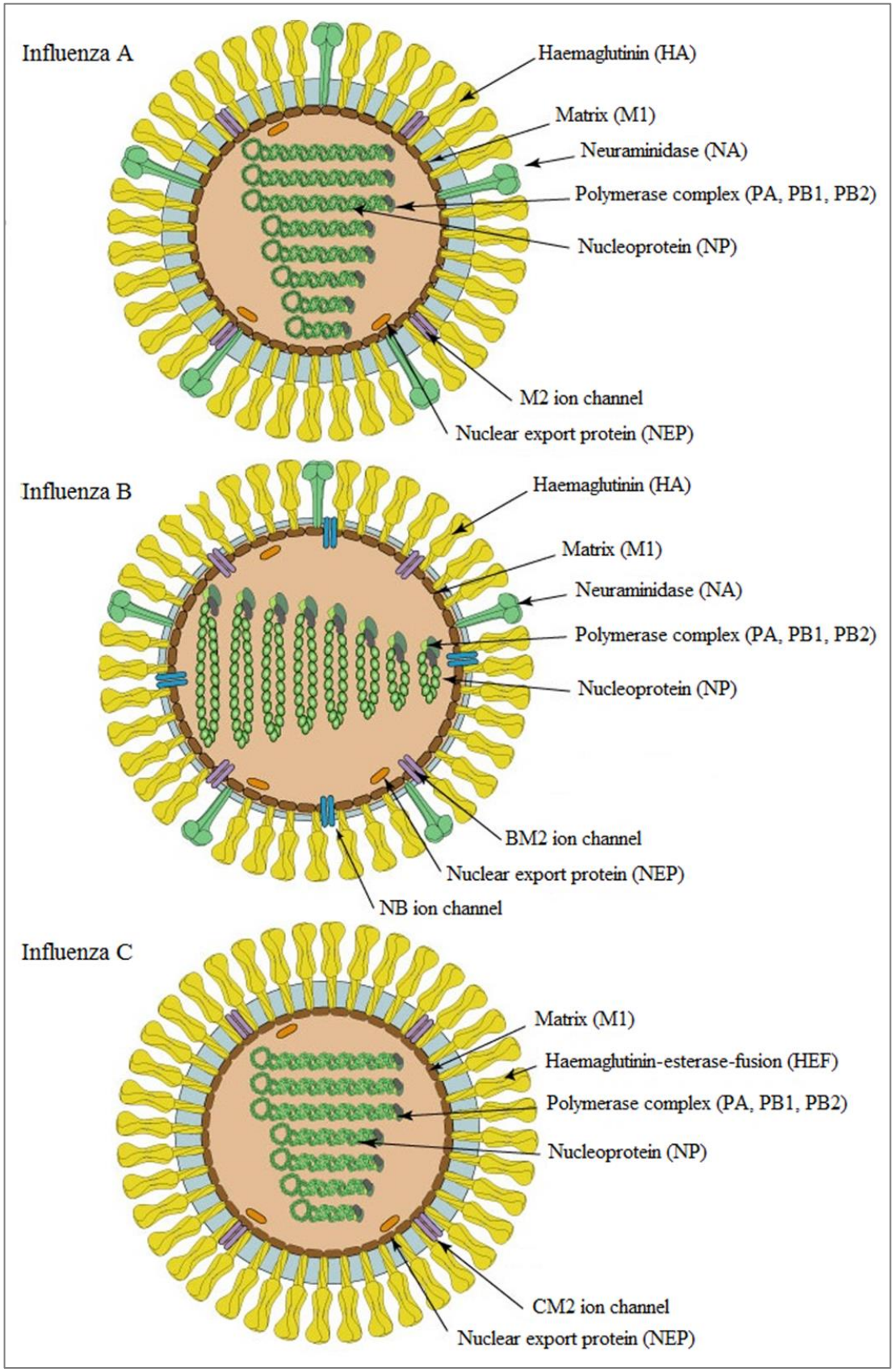

Figure 2.1: Different structure of influenza A, B and C. Adapted from Swiss Institute of Bioinformatics (2010).

2.1.2 Virus replication cycle

The life cycle of influenza virus was basically a rotation of four stages; 1) host cell invasion; 2) nucleus take over and host cell resources' exploitation (vRNPs enter the nucleus, transcript and replicate its own genome); 3) departing of the vRNPs from the nucleus; and 4) final assembly and budding at the host cell plasma membrane getting ready for the next cell invasion (Samji, 2009). The first stage began with attachment process involving HA, the most abundant surface glycoproteins. Approximately 80% of the influenza virus surface glycoproteins are made of HA. Crystal structure of HA solved by Wilson *et al*., (1981) showed that this molecule is a trimer comprised of two distinct regions; a triple-stranded coiled-coil of α-helics as a stem extended from the membrane and antiparallel β-sheet at the top of the stem as the globular head where N-acetylneuraminic (sialic) acid (SA) receptor binding site is located. For H1 subtype HA, surrounding the receptor binding site are five predicted antigenic sites. The homotrimer will bind to SA found on the host cell's membrane surface (Skehel and Wiley, 2000). Influenza HA preferentially binds to two types of galactose-linked sialic acid; SA α2,6-Gal for human viruses and SA α2,3-Gal for avian virus. In human, SA α2,6-Gal is dominantly expressed on the upper respiratory tract while the lower part expressed both types of SA. This phenomenon explained why H5N1 infection tends to cause pneumonia and the main reason for human to be more susceptible to H1N1 virus compare to H5N1 virus (Skehel and Wiley, 2000; Voet and Voet, 2004).

Following the attachment, the virus was endocytosed allowing the virus to enter the host cell in the form of endosome. The low pH of the endosome of around five to six will trigger the conformational change of HA exposing the fusion peptide that mediates the viral and endosomal membranes fusion hence internalizing the

vRNPs into the cytoplasm (Sieczkarski and Whittaker, 2005; Stegmann, 2000). vRNPs were then directed into the host cell nucleus though viral proteins nuclear localization signals (NLSs) for the synthesis of both viral RNA, the capped and polyadenylated messenger RNA (mRNA) to be used as the template for viral protein translation and vRNA segments to form the genomes of progeny virus (Boulo *et al.*, 2007; Bouvier and Palese, 2008).

By hijacking the host's transcription machinery the viral mRNA was synthesized using a process called "cap snatching" process, and once polyadenylated and capped, the viral mRNA will be transported out of the nucleus and translated. As for the negative sense vRNPs, the RNA complex appeared to be exported from the nucleus by the viral proteins M1 and NEP/NS2 (Cros and Palese, 2003). The three envelope proteins, HA, NA and M2 synthesized on the membrane-bound ribosomes will undergo post-translational modification prior to the viral assembly. Then the vRNA was encased in the capsid protein, and together with new matrix protein, the protein complex was transported to sites at the cell surface where envelope HA and NA components have been incorporated into the cell membrane. Progeny virions were formed and released by budding which probably initiated by an accumulation of M1 matrix protein at the cytoplasmic side of the lipid bilayer. The progeny virions were released from the host cell surface by cleaving of the terminal sialic acid residues by NA protein (Bouvier and Palese, 2008).

2.2 Avian influenza A H5N1

Among influenza A virus, other than H1N1 virus, another influenza threat came from avian influenza virus particularly a highly pathogenic avian influenza (HPAI) A H5N1 virus. First identified in Italy in the early 1900s, H5N1 was known as "avian flu" or "bird flu". This virus had caused a large scale of death in poultry

worldwide ever since, yet it was very rare to find the infection of H5N1 in human. At least, not until the first report of chicken-to-human transmission case in 1997 in Hong Kong (Subbarao *et al.*, 1998; Ku and Chan, 1999). Following the first outbreak and the 2004-2005 epidemics, H5N1 had continually caused outbreaks around the globe and occasionally accompanied by transmission to human with 50% fatality rate (Neumann *et al.*, 2010). Despite predominantly targeting animals, the possibility for human to be the main target is undeniable as the notion was supported by the possible case of human-to-human transmission (Wang *et al.*, 2008; Hill *et al.*, 2009) and findings that linked the avian influenza virus with the notorious 1918 H1N1 virus (Taubenberger, 2006; Qi *et al.*, 2012; Worobey *et al.*, 2014).

2.2.1 The evolution of human avian influenza A (H5N1)

For decades avian influenza had only infected bird causing outbreaks mostly in Europe and America (Adams, 2006; Petsko, 2005). However, in 1996 when an HPAI H5N1 virus was isolated from geese following the virus outbreak in Guangdong Province in China, the virus demography was no longer the same. It was proven that this virus is the HA donor for the latter reassortant H5N1 virus in the 1997 Hong Kong outbreak (de Jong and Hien, 2006). Though the spreading of the virus was managed to be controlled by killing millions of chicken and the number of infection was relatively low, the death rate deserved an attention as 33% of the infected human died (6 out of 18) (Chan, 2002).

By the end of 2003 and in the early 2004, as H5N1 virus strain started making its way out of China, outbreaks began to occur in a few other countries, mostly South East Asia (Petsko, 2005). The 2004 strains which evolved from 1999-2002 were also known as Z genotype or “Asian lineage HPAI A (H5N1), and can further be divided into two antigenic clades; clade 1 and clade 2. Human and bird

isolates from Thailand, Vietnam and Cambodia, and bird isolates from Malaysia and Laos belongs to clade 1 (Webster and Govorkova, 2006; Pal, 2014). In February of the same year, the virus was found to have gained the ability to infect non-human mammals when the strain was detected in pigs in Vietnam (Petsko, 2005; Choi *et al.*, 2005). The second clades which responsible for 2005 -2006 human infection were first identified from China, Indonesia, Japan and South Korea birds' isolate (Pal, 2014). The viral strain continued to spread outside Asia when it hits Turkey and Iraq in Jan 2006 followed by other Middle East countries such as Azerbaijan and Northern Africa. From 2007 onward most of human case infection was reported mainly in Indonesia and Egypt. To date over 800 cases were reported worldwide (from 2003) with near 60% of fatality (WHO, 2015).

2.2.2 A H5N1 in Malaysia

The first outbreak of HPAI A (H5N1) subtype in Malaysia was reported in August 2004 in a village in the state of Kelantan (Alexander, 2007). The virus was transmitted from smuggled fighting cocks from neighbouring country to the local village chickens, and weeks after, eight more outbreaks were reported. No human infection or death was reported and the disease was brought under control by depopulation, poultry and birds quarantine/clinical surveillance and movement restriction of birds and their products (Tee *et al.*, 2009).

However, in February 2007 another outbreak of HPAI A (H5N1) occurred in Malaysia in other regions over wider areas including several villages in Kuala Lumpur, and along the industrial area of western coast of Peninsular Malaysia in the state of Perak and Pulau Pinang. If 2004 HPAI A (H5N1) strain was revealed to be highly similar to strains isolated from Vietnam and Thailand, the 2006 strains phylogenetic analysis showed that the latter strain was closer to the isolates from

China and Indonesia. The disease was again brought down by control and preventative efforts, yet another outbreak re-occurred in another village in the state of Selangor in June 2007 and finally resolved several month later (Tee *et al.*, 2009).

2.2.3 The pandemic potential

Ever since the outbreak in 1997, H5N1 virus had continually undergone various adaptations and multiple genetic re-assortments and became more virulence with wider host range and location (Le and Nguyen, 2014). Irrefutably, a new strain of HPAI A H5N1 virus is expected and for HPAI A H5N1 virus to be the main character in the next pandemic, all it has to obtain is the ability to easily transmit in human population unnoticed or hardly recognized by human immunity system and causing severe disease (Moscona, 2005). To do that, a very efficient attachment and replication system is needed and HA protein plays a very important role as it is on the front line of the virus invasion process by recognizing the cell glycan receptor for attachment. The human receptors, SA α2,6-Gal (SA α2,3-Gal expressed by avian host) is the first obstacle and with a few changes of amino acid in the receptor-binding pocket of the H5N1 HA protein there is possibility for the changes of receptor specificity (Naeve *et al.*, 1984). For replication process, PB2 protein is one of the main components for RNA replication. Glu-to-Lys mutation at location 627 of PB2 was found related to the adaptability of the virus in mammal (Gambotto *et al.*, 2008). Evidently these idea might actually work as some scientists had experimentally proven that with certain mutations this virus can evolve to a very threatening strain and even capable of transmitting through air which is very alarming (Herfst *et al.*, 2012; Imai *et al.*, 2012; Wilker *et al.*, 2013).

Though we are lucky that the virus has yet to obtain the ability to be easily transmitted among humans, the possibility of the event might be closer as 2013 witnessed the most recent influenza outbreak in China caused by a novel avian influenza A H7N9 virus which infected more than 100 of people in that year alone with 20% - 30% of fatality rate (Horby *et al.*, 2013). So far, the virus has been hypothesized to be evolved at least from four origins (Liu *et al.*, 2013). Quoted as "…unusually dangerous virus to humans.", H7N9 showed unusual prevalence of older males, easily transmitted to human from poultry (in comparison with H5N1) and harder to spotted as it didn't kill the poultry thus allowing it to spread silently hence wrecking the nerves among researchers (Shadbolt, 2013; Arima *et al.*, 2013; Devitt, 2013).

2.2.4 Pathogenesis of avian influenza A (H5N1) in human

Once being infected by the influenza virus, after two or more days of exposure the person will normally experience an influenza-like illness such as fever, cough, shortness of breath, sore throat, muscle aches, diarrhea, and conjunctivitis (CDC, 2008; Gambotto *et al.*, 2008) And if not well treated, it may progress to pneumonia and in worse scenario, may lead to acute respiratory distress syndrome (ARDS), multi-organ failure and eventually death (Maines *et al.*, 2011; Thanh *et al.*, 2008).

The alveolar epithelial cells has always been the main target for influenza A replication, which is why diffused alveolar damage (DAD) and haemorrhage in the lung were often observed on the infected patients (Uiprasertkul *et al.*, 2007; Korteweg and Gu, 2008; Chan *et al.*, 2013). However, compared to seasonal influenza infection and 2009 H1N1 which mostly affected the upper respiratory tract, H5N1 preferred the lower region and in some rare and severe cases the virus spread

beyond the lung by infecting trachea, brain, intestine, lymphoid tissues and even penetrate the placental barrier (Uiprasertkul *et al.*, 2005; Gu *et al.*, 2007; Korteweg and Gu, 2008; Gao *et al.*, 2010; van den Brand *et al.*, 2014).

The cell and organ damage due to H5N1 infection were thought to be associated with viral replication through either cytolytic (degeneration of cells) or apoptotic (programmed cell death) mechanisms (Korteweg and Gu, 2008; Alberts *et al.*, 2002; van den Brand *et al.*, 2014). This theory though not fully proven was supported by the evidences of active viral replication in many part of the respiratory tract isolated from the infected humans, in mammalian model such as ferret and mice, and also in *in vitro* models of human respiratory epithelial cells (Korteweg and Gu, 2008; Chan *et al.*, 2013; Belser and Tumpey, 2013).

Other than viral replication, extensive studies on the role of dysregulation of cytokines (cells signalling molecules which covered a wide range of peptide, glucopeptides, and protein responsible for mediating the immune system) deduced that it could be one of the key mechanisms involved in the pathogenesis of H5N1 influenza (Yu *et al.*, 2011; Hui *et al.*, 2009; Mandal, 2013). Besides that, even though the exact roles remains unclear, it was believed that other factor such as up-regulation of tumor necrosis factor-related apoptasis-inducing ligand (TRAIL) and reduced cytotoxicity of $CD8^+$ lymphocytes are also taking part in the pathogenesis of this virus (Zhou *et al.*, 2006; Hsieh and Chang, 2006).

2.2.5 Prevention and treatment

Following the outbreak of influenza A H5N1 in 2003 which triggered the surge for pandemic preparedness, WHO issued a checklist to guide the development of national influenza pandemic preparedness plans in 2005 (Berera and Zambon,

2013). The list basically addressed the need to develop a good communication plan in handling the pandemic, surveillance system to monitor the antivirals usage and or sale, health and research facilities, and increasing the awareness among the population regarding the pandemics and actions to take when it strike (WHO, 2005). In 2007, WHO also outlined another guideline to help national authorities to stop the development of the pandemic when it first detected (Berera and Zambon, 2013). Based on early mathematical modelling studies, rapid containment strategies and protocols were laid out (WHO, 2007).

Therefore, when the imminent threat of A (H1N1) pdm2009 was issued by WHO in April 24, 2009, the years of plans and strategies were rapidly implemented across the globe. Even though the plans manage to lower the number of death in 2009 pandemic as compared to the previous one, many insisted that we can and should do better in the future and if there is one thing we can learn from this genius virus it is that survival requires non-stop progression, and so must we. As we learned from the pandemic post mortem, the availability of the treatment is useless if it cannot be delivered on time (Fedson, 2013). Therefore, the revision of every aspects of the plans such as better trained clinician, additional experienced technicians, more superior surveillance system and even better vaccines and antivirals in terms of effectiveness, accessibility and rapid production are required (Fineberg, 2014; Berera and Zambon, 2013).

In the fight against influenza pandemic, many believe that the ultimate way to win the war is by equipping ourselves with the ultimate vaccine and or the capacity to produce the vaccines rapidly (Fedson, 2013). However the continual mutation of the virus makes the idea almost impossible to materialize and despite the promising potential shown by several developing “pre-pandemic vaccines” the

question remains whether it will actually work against the actual pandemic virus as no can ever know for certain which virus will be responsible (Osterhaus, 2007). Nevertheless, by June 2008 at least three H5N1 vaccines for human were licensed; Panvax by CSL Limited approved by Australia (June 2008), Prepandrix produced by GlaxoSmithKline approved by European Union (May 2008), and Sanofi Pasteur's vaccine approved by United States in April 2007 (Roos, 2008). In 2013 despite containing the controversial oil-in-water emulsion adjuvant AS03, FDA approved GlaxoSmithKline's vaccine Q-Pan to be included in the US pandemic emergency vaccine stockpile (La Vigne, 2013).

In the 2009 pandemic, as the vaccines for pdm2009 H1N1 were only approved by FDA by September 15 2009, antivirals were the first defence against the pandemic where they were stockpiled by most of countries (Berera and Zambon, 2013). Influenza treatment has always highly relied on antiviral drugs which mostly designed by targeting viral proteins essential for the virus replication cycle. Among the designed drugs are viral attachment and fusion inhibitor, M2 blockers, viral RNA-dependent RNA polymerase (vRdRP) inhibitors (to interrupt transcription and viral genome replication process), NP inhibitor, and NA inhibitor (de Clercq, 2006; Loregian *et al.*, 2014).

Historically, adamantane derivatives, amantadine and rimantadine were the first two drugs available for influenza treatment designed specifically active against influenza A virus. These M2 ion channel blockers function by interrupting the viral uncoating process (de Clercq and Neyts, 2007). However, they only worked on influenza A virus due to the exclusive presence of M2 ion channel on this virus type, and were always associated with rapid emergence of drug-resistance viral strain (von Itzstein, 2007; Bright *et al.*, 2005; He *et al.*, 2008). Up to 2013, nearly half of the

influenza A viruses had conferred the resistance toward adamantanes with H1, H3, H5, H7 H9, and H17 subtype bearing the high-level resistance, hence restricting the usage of the antiviral (Dong *et al.*, 2015; CDC, 2015)

As of for today, only NA inhibitors (oseltamivir, zanamivir, and peramivir) were proven affective, licenced, and primarily recommended against the circulating influenza viruses (CDC, 2015). These antivirals however were not without limitation. Oseltamivir for example was not to be used by high-risk patients (patients less than one year or diagnosed with liver or renal failure) and the uptake has to be within 24-48 hours after the onset of infection. Zanamivir and peramivir were only approved for patient older than five and 18 years old respectively (FDA, 2015). And most recently, a polymerase inhibitor was approved known as favipiravir but only in Japan with a very strict usage (Ison, 2015).

2.3 Neuraminidase: the drug target

The envelope of influenza A virus is generally composed of a lipid bilayer and integral membrane proteins including the two major glycoproteins, HA and NA protruding from the membrane covering the surface of the virus like spikes (Colman, 1994; Goodsell, 2009; Reddy *et al.*, 2015). The nomenclature of the influenza A strains are based on these two glycoproteins which are also the key determinant of the virus virulence and pathogenesis (Korteweg and Gu, 2008; Zhou *et al.*, 2009; Gamblin and Skehel, 2010). HA plays the starring role by recognizing and binding to the sialic acid existed on the host cell during the attachment process and the subsequent membrane fusion. NA major role on the other hand is to cleave the α-ketosidic linkage between the sialic acid and an adjacent sugar residue that anchoring the newly assembled virion on the host cell surface so that it will not get stuck (Wagner *et al.*, 2002; Goodsell, 2009; Shtyrya *et al.*, 2009).

For decades, both HA and NA had received significant amount of attention as drug target, but only a few of NA inhibitors were proven to be affective as a broad-spectrum inhibitor while most of the candidates were either strains dependent, cytotoxic, or resistance prone (Eyer and Hruska, 2013; Edinger *et al.*, 2014). Despite the lack of successful antiviral or vaccine, given the continuous threat, future efforts should not be deter but continued and progress should be made to further understand the virus and improve the technology (Littler and Oberg, 2005; Atkins *et al.*, 2012; Saha *et al.*, 2013).

2.3.1 Neuraminidase structure

Influenza NA was first discovered by George Hirst and soon after became the target of influenza therapeutics approach. But the limited technology at that time was not able to fully unfold the protein structure (Hirst, 1942; Wohlbold and Krammer, 2014). Admittedly, it took nearly 40 years for the first NA to be purified and crystallized, and the N2 NA was found to be an association of four identical monomers where each monomer was made of six four-stranded antiparallel β-sheets arranged in the formation propeller blade (Varghese *et al.*, 1983) as illustrated in Figure 2.2A. Each of the nascent polypeptide consists of four parts; N-terminal cytoplasmic domain, transmembrane domain, stalk region and the globular head domain where the catalytic site is located (refer Figure 2.2B) (Air, 2012; Chen *et al.*, 2012). The total size of NA protein is around 450-480 amino acids depends on the NA types, portraying the low sequence similarity yet they do have a stable topologies and a conserved nine amino acids sequence “ILRTQESEC” as part of the enzymatic active site (Chen *et al.*, 2012; Doyle *et al.*, 2013).

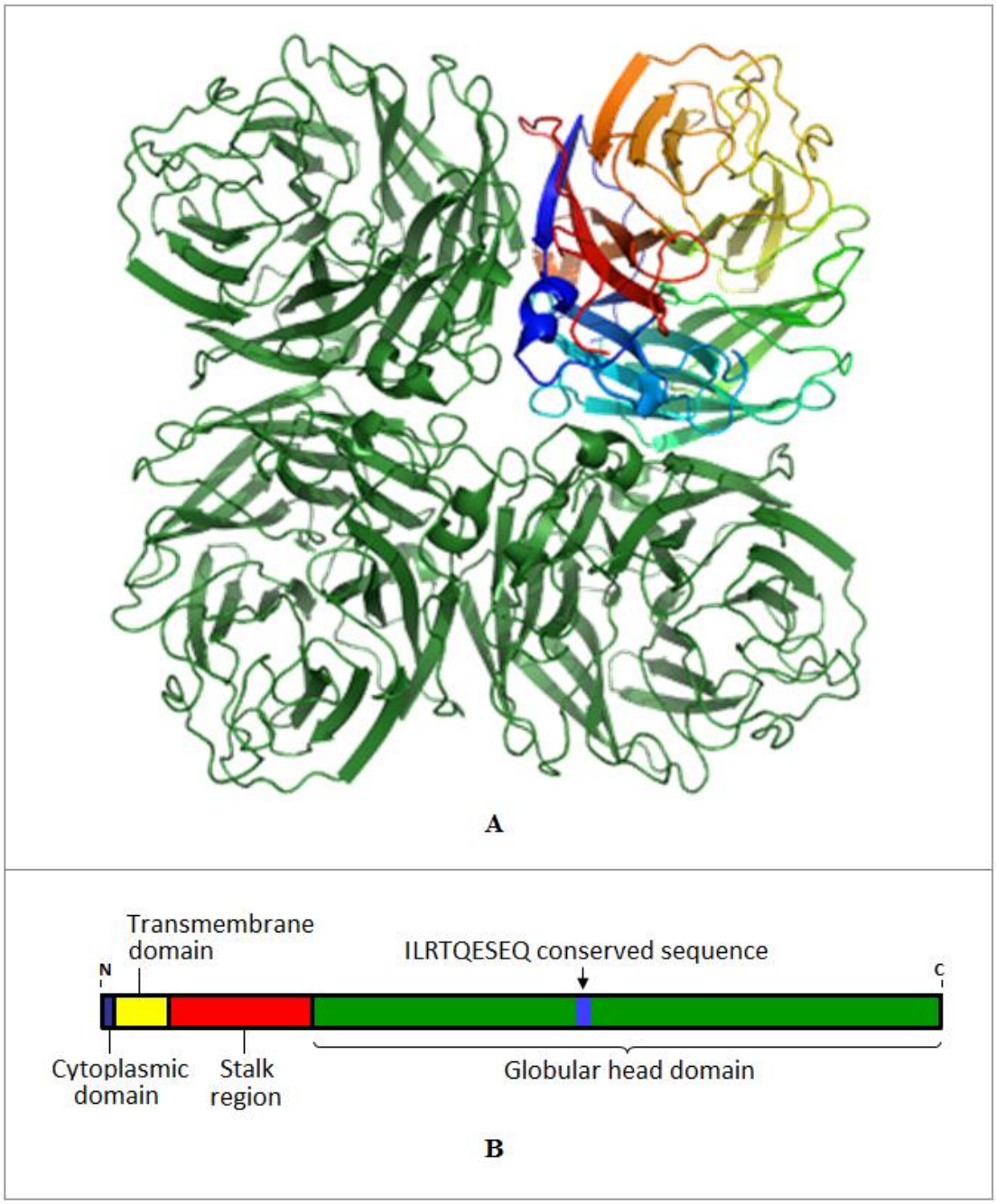

Figure 2.2: Structure of influenza A virus NA.

(A) 3D crystal structure of 1918 influenza A H1N1 NA globular head domain shown as tetramer displaying conserved 6-bladed propeller structure coloured in green with one of the subunits coloured differently from blue to red along the N to C chain terminus for each blade. Each blade is made up of four anti-parallel beta sheets stabilized by disulphide bonds and connected by loops of varying length. Picture was adapted from Royuela (2013).

(B) Schematic diagram of NA protein showing cytoplasmic, transmembrane, stalk region and globular head domains, and universally conserved ILRTQESEC sequence (arrow pointed) adapted from Wohlbold *et al*., (2014).

For avian influenza A (H5N1) NA specifically, the crystal structure was solved in 2006 showing that NA of influenza A are phylogenetically divided into two distinct groups, group-1 containing N1, N4, N5 and N8 subtypes, and group-2 contain the rest of NA subtypes (N2, N3, N5, N7 and N9). Superposition of the active site of three group-1 NA, (N1, N4 and N8) produced a higly similar active site region whereas superposition of the active of N1 (group-1) and N9 (group-2) in contrast showed a markedly difference in the 150-loop region (Russell *et al.*, 2006). The latest isolated NA (N10 and N11) however, showed little resemblance to any group (Wohlbold and Krammer, 2014).

2.3.2 Neuraminidase inhibitor

The SA binding site of NA is comprised of the following amino acids, Arg118, Asp151, Arg152, Arg224, Glu276, Arg292, Arg371 and Tyr406 which are conserved among influenza A and most NA inhibitors were designed to interrupt the binding by mimicking the SA composition (Mitrasinovic, 2010). The analogue of SA, 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (DANA) (Figure 2.3A) was developed in 1960s as inhibitor for many bacterial and viral NA (Meindl and Tuppy, 1969). With the successful configuration of NA structure in 1983, the first effective NA inhibitor known as zanamivir (Figure 2.3B) which is a derivative of DANA was discovered, and marketed under the trade name Relenza (von Itzstein *et al.*, 1993; Wohlbold and Krammer, 2014). Another NA inhibitor, oseltamivir (Figure 2.3C; marketed as Tamiflu) was also designed based on the same structure but with a different approach and received more attention due to its ease of administration and systemic availability as it was designed as capsule compared to the inhalant Relenza (de Clercq and Neyts, 2007; Air, 2012).

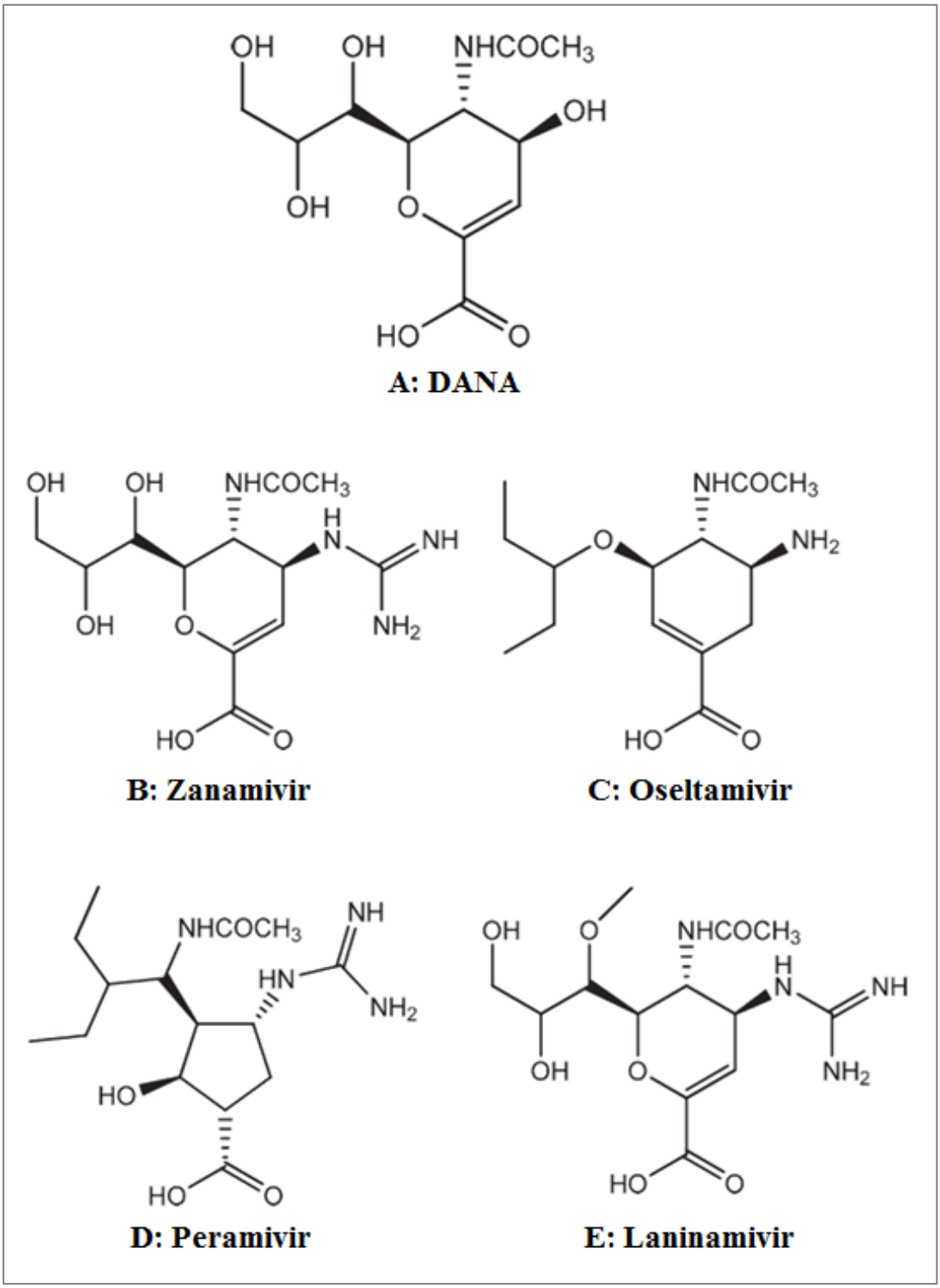

Figure 2.3: Chemical structure of the NA inhibitors. Adapted from McKimm-Breshkin (2012).

One of the reasons why the search for new inhibitor was never stopped is that from time to time the influenza virus will most definitely become resistance towards the current inhibitor as shown by the worldwide spread of oseltamivir-resistant seasonal A (H1N1) during the 2007-2009 influenza seasons. Clinical and *in vitro* reduced susceptibility mostly towards oseltamivir had been reported for H5N1 and H3N2. And most of the resistance strains carried NA substitution mutations such as H274Y, N294S, E119G and I122R (Samson *et al.*, 2013). Therefore, along with advances in technology particularly in drug discovery such rational drug design and X-ray crystallography, dozens of new or modified NA inhibitors had been suggested and tested (Eyer and Hruska, 2013).

One NA inhibitor candidate developed by structure-based design is an injectable drug known as peramivir (Figure 2.3D) developed by Biocryst. A cyclopentane derivative was finally approved and commercially marketed as Rapibav. Because it is injectable, peramivir is very applicable for patient who cannot be treated by pills or with limited lung capacity (Kohno *et al.*, 2011; FDA, 2015). Another promising candidate is laninamivir (code name, R-125489) (Figure 2.3E), another SA analogue that demonstrated a long-acting inhibitory activity and was approved for usage by Japan (Kubo *et al.*, 2010; Samson *et al.*, 2013). Groups of researcher had also studied the possibility to develop the inhibitor by using the natural compound. This approach was still under development and with numbers of possible compounds, many trial and error are on the way (Shan *et al.*, 2012; Ikram *et al.*, 2015).

2.3.3 Neuraminidase production

When the NA was discovered and set as the drug target, the urgency to extract the active protein to study the protein began (Wohlbold and Krammer, 2014). Conventionally, NA was prepared by proteolytic cleavage or detergent-treatment of the viral envelop followed by subsequent purifications to obtain the catalytic head domain. Two proteases commonly used for the extraction procedure are pronase and bromelain (Seto *et al.*, 1966; Aitken and Hannoun, 1980; Cabezas *et al.*, 1982; McKimm-Breschkin *et al.*, 1991; Franca de Barros *et al.*, 2003). Pronase refers to a protease cocktail (contain at least ten proteolytic components) secreted by the bacteria *Streptomyces griseus* and bromelain is another protease mixture produced by pineapple mostly from fruit and stem parts (Sweeney and Walker, 1993; Kelly, 1996). As both of the proteins have a broad range of protein specificity, this method tends to complicate the downstream purification of NA due to cleavage of other protein such as HA. Therefore, a new approach to increase the protein specificity was applied by genetically engineered a thrombin cleavage site at the stalk region of NA (Wanitchang *et al.*, 2010).

As the demand for the protein will never cease to stop mainly for the inhibitor screening purposes, new and perhaps simpler production methods were indeed needed. As an alternative, the production of NA protein by using recombinant technology had been considered and implemented. Various type of protein expression system had been explored such as insect cells with baculovirus expression system (Johansson *et al.*, 1995; Schmidt *et al.*, 2011; Margine *et al.*, 2013), CHO-K1 cells (Tanimoto *et al.*, 2004), human embryonic kidney cells (Nivitchanyong *et al.*, 2011), and the yeast such as *Pichia pastoris (Yongkiettrakul et al., 2009)*and *Kluyveromycis lactis* (Tsai *et al*., 2011) .