EXPRESSION AND PURIFICATION OF HEMAGGLUTININ AND NEURAMINIDASE RECOMBINANT PROTEIN OF AVIAN INFLUENZA A VIRUS USING IMMOBILISED METAL AFFINITY CHROMATOGRAPHY (IMAC)

by

GOH WEI CHIANG

Dissertation submitted in partial fulfillment of the requirement for degree of

Bachelor of Health Science (Honours) (Biomedicine)

2016

ACKNOWLEDGEMENT

I would like to thank all those who have directly or indirectly contributed to my thesis writing.

I would like to express my gratitude towards my supervisor, Professor Shaharum Shamsuddin for his guidance throughout my final year project. Besides that, I would also like to thank Mr. Wong Weng Kin and Puan Nurul Aini for their patient in guiding me while I was completing my project.

I would like to thank my parents and brothers who gave their full support to me. Moreover, I would like to thank all my friends for their support.

In addition, I would like to thank all the staffs in the School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian for helping me throughout the year.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND UNITS

%	Percentage
~	Approximately
°C	Degree celcius
µl	Microlitre
µg	Microgram
µg/ml	Microgram over millilitre
bp	Base pair
g	Gram
kb	Kilo base
kDa	Kilo Dalton
L	Litre
M	Molar
mg	Milligram
mg/ml	Milligram over millilitre
ml	Millilitre
mM	Millimolar
nm	Nanometer
ng	Nanogram
rpm	Rotation per minute
U	Unit
v/v	Volume to volume
w/v	Weight to volume
V	volt
α	alpha

LIST OF ABBREVIATIONS

dH_2O	Distilled water
ddH2O	Deionized distilled water
DNA	Deoxyribonucleic acid
HA	Hemagglutinin
HPAI	Highly pathogenic avian influenza
IMAC	Immobilized Metal Affinity Chromatography
MW	Molecular weight
NA	Neuraminidase
NCBI	National Centre for Biotechnology Information
OD	Optical density
RNA	Ribonucleic acid
UK	United Kingdom
US	United States
USA	United States of America
UV	Ultraviolet
WHO	World Health Organization
CDC	Centers for Disease Control and Prevention
TEMED	N,N,N',N'-Tetramethyl-ethylenediamine
TAE	Tris acetate EDTA
SDS	Sodium dodecyl sulphate
PAGE	Polyacrylamide gel electrophoresis
Ni-NTA	Nickel-Nitrilotriacetic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
E.coli	*Escherichia coli*
BSA	Bovine Serum Albumin

LIST OF APPENDICES

ABSTRAK

Virus selsema burung Influenza A tergolong dalam keluarga *Orthomyxoviridae*. Virus ini mempunyai lapisan luar sepanjang 120nm diameter. Ia mempunyai 8 bahagian 'negative-sense RNA'. Terdapat dua protein utama yang memainkan peranan penting dalam jangkitan virus ini iaitu hemagglutinin (HA) dan neuraminidase (NA). Virus selsema burung Influenza A berupaya untuk menjalani peralihan antigen. Jujukan DNA dalam virus diulangi oleh 'RNA polymerase' yang cenderung untuk melakukan kesilapan. Semua ini menyebabkan kemunculan strain selsema burung Influenza A yang mempunyai hemagglutinin (HA) dan neuraminidase (NA) yang berbeza. Akibatnya, 3-5 juta kes selsema burung dan sebanyak 300,000-500,000 kematian direkodkan di seluruh dunia. Oleh itu, kawasan yang terpelihara dalam HA dan NA glikoprotein penting untuk penghasilan vaksin atau alat pengesan virus. Dalam kajian ini, protein yang terdapat dalam kawasan terperlihara HA dan NA glikoprotein selsema burung Influenza A dihasilkan. Sintetik HA dan NA gen dihantar untuk penjujukan dan dikenalpastikan menggunakan BLAST daripada NCBI. Protein-protein tersebut akan ditulenkan dengan menggunakan 'Immobilized Metal Affinity Chromatography (IMAC).

ABSTRACT

Avian influenza A virus is a member of *Orthomyxoviridae* family. It has an envelope of 120nm in diameter. It has 8 segmented negative-sense RNA. There are two major proteins that play important roles in virus infectivity which are hemagglutinin (HA) and neuraminidase (NA). Influenza A virus are able to undergo antigenic drift and antigenic shift. The gene sequence in the virus is replicated by error-prone RNA polymerase. All these leads to the emergence of new subtype of influenza A virus that expressed different HA and NA glycoprotein. Consequently, the burden of influenza epidemics is approximately 3-5 million cases and results in 300,000-500,000 deaths globally. Thus, conserved region in HA and NA glycoprotein are important as it can be used for future development of universal influenza vaccine (UIV) or detection kit. In this study, the gene that encodes for the conserved regions in HA and NA glycoprotein of avian influenza A virus were expressed using bacterial expressing system. The recombinant sequence were firstly verified and confirmed the sequence from NCBI genebank. pET-47b(+) plasmid that contains synthetic HA and NA gene were transformed into BL21 *E.coli* strain. The transformed *E.coli* was induced with IPTG to produce desired HA and NA recombinant proteins. The His-tagged NA and HA protein were purified using immobilized metal affinity chromatography (IMAC). Recombinant protein present in the eluted fraction was further purified by size exclusion method. Purified recombinant protein was dialyzed with PBS to allow refolding back to its native conformation state. In conclusion, recombinant HA and NA were successfully expressed and purified.

CHAPTER 1: INTRODUCTION

1.1 Viruses

1.1.1 General characteristic of virus

Viruses are small obligate intracellular parasite that requires a host to replicate. Without a host, the virus cannot initiate the cellular machinery to produce the essential components of the virus. An infectious virus particle is known as a virion. (Lodish et al., 2000) stated that the simplest virus can encode for 4 proteins while the most complex virus can encode about 100-200 proteins. Virus will use host cell's components such as ribosomes, tRNA and translation factors to synthesis viral protein. During the late stage of infection, virus will trigger the host cellular machinery to generate a large amount of viral mRNA and proteins instead of the normal cellular macromolecule.

A simple virus can assemble by itself when mixed in a solution. Example of a simple virus is the tobacco mosaic virus (TMV). TMV consists of a single RNA molecule and one type of protein. However, more complex virus cannot assemble spontaneously. It occurs in stages where subviral particles will be developed first then become a complete virion (Lodish et al. 2000).

Most of the viruses have a narrow host range. Viruses that only infect bacteria are known as bacteriophage. Viruses that infect animal cells are known as animal viruses whereas viruses that infect plant cells are known as plant viruses. There are some viruses that can infect plant and insect as well. An example is the potato yellow dwarf virus that is able to infect leafhopper insect and potato plant as well. Within the infected animal or plant, there

are specific cells that will be infected. These specificities are determined by receptors that are present on the cell surface which allow the virus to adhere (Lodish et al. 2000).

1.1.2 Components of virus

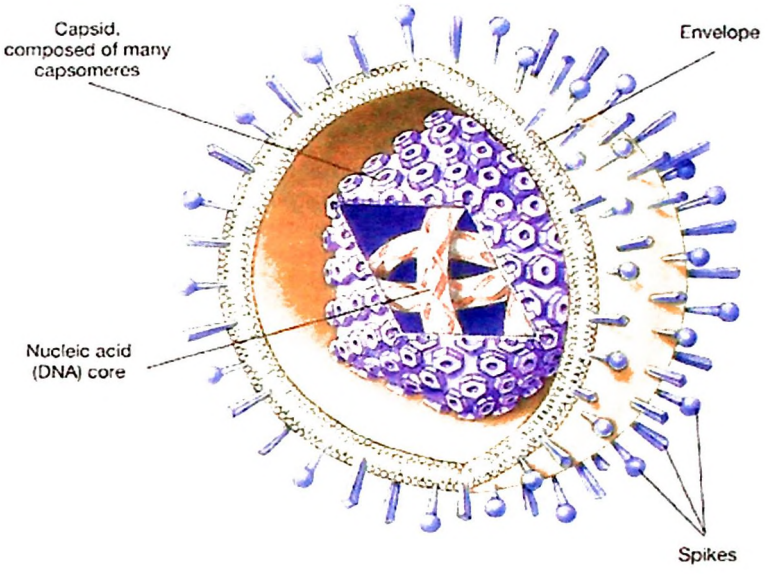

Figure 1.1: Basic components in virus (Black, 2006)

Figure 1.1 shows the basic components in a typical virus. A typical virus has envelope protein, envelope, viral genome, nucleocapsid and viral tegument. The virus that has envelope is known as enveloped virus whereas virus without envelope is known as naked virus. Viral envelope membrane is made up of lipid bilayer. These viruses normally acquire theenvelope from the host cell's membrane. This occurs when the virion buds off from the infected host cells. It will take a portion from the host cell membrane to protect itself from the host immune system and modified it by adding viral glycoprotein (Gelderblom, 1996).

1.1.3 Classification of virus

Virus can be classified into groups based on its distinctive properties. Examples are size, morphology, chemical composition, structure of genome and mode of replication (Gelderblom, 1996). Chemical composition, the structure of viral genome and mode of replication can be classified using Baltimore classification. Additional properties such as the presence of envelope and antigenic and biological properties of virus also can be used to classify virus.

The size and shape of a virus play an important role in virus classification. Virus can exist in helical nucleocapsid or icosahedral nucleocapsid. An example of helical nucleocapsid virus is the tobacco mosaic virus as shown in Figure 1.2. Helical nucleocapsid is formed when identical capsid proteins (protomers) coats around helical nucleic acid filament. This arrangement confers the genetic structure to be rigid or flexible. Other parameters such as length, width, pitch of helix and number of protomers per helical turn can further subdivide the helical nucleocapsid viruses (Gelderblom, 1996).

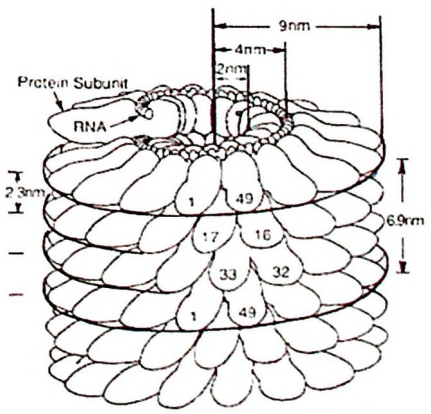

Figure 1.2: Helical structure of the rigid tobacco mosaic virus rod (Gelderblom,1996).

An example of icosahedral nucleocapsid virus is the adenovirus as shown in Figure 1.3. The building subunit of icosahedral nucleocapsid is known as capsomere. The number and arrangement of capsomere are important for virus identification. Most of the spherical viruses do have icosahedral nucleocapsid. Icosahedron structure consists of 20 equilateral triangular faces and 12 vertices (Gelderblom 1996).

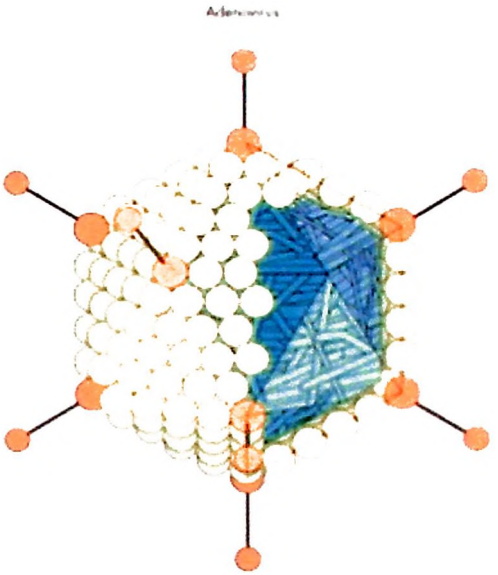

Figure 1.3: Adenovirus structure (Murray and Rosenthal, 2005).

Baltimore classification of viruses, 1971 is based on the genetic contents and mode of replication of viruses. Under this classification, viruses are divided into seven classes which are double stranded DNA viruses, single stranded DNA viruses, double stranded RNA viruses, positive sense ssRNA viruses, negative sense ssRNA viruses, RNA reverse transcribing viruses and DNA reverse transcribing viruses.

Virus genome can exist as deoxyribonucleotide (DNA) or ribonucleotide (RNA) but not both. 70% of all viruses are RNA virus and the remaining 30% are DNA viruses. Single stranded RNA can be either sense strand (plus strand) or antisense strand (minus strand). Examples of RNA virus and DNA virus are SARS and chickenpox respectively (Gelderblom, 1996).

The viral genome can either exist in single stranded or double stranded but not both. This applies as well to the linear and circular structure of viral genome. The virus can exist as one nucleic acid molecule (monopartite genome) or several nucleic acid fragments (multipartite genome) (Gelderblom, 1996).

1.1.4 Replication of virus

Virus requires host cells to propagate. This is because the host cells able to provide energy, low molecular weight precursor for viral protein and nucleic acid synthesis and cellular machinery system (Gelderblom, 1996).

In viral replication cycle, the first step is viral attachment. This is done through the interaction with complementary receptors present on the virion and targeted host cell. After the attachment, the virus will enter into the host cell. This step is known as penetration. The virus can enter into the cell via endocytosis or fusion between the virus envelopes and host cell plasma membrane. Fusion process involves interaction between viral fusion protein and second cellular receptor. The next steps are uncoating. Uncoating is defined as the separation of viral nucleic acid from outer structural components of the virion. This is important as the viral genome can be released into the host cell to perform its function (Brooks et al., 2012).

The genetic material can exist as free nucleic acid or nucleocapsid. The viral genome will be transcribed to produce mRNAs. After mRNAs are produced, the virus will use the host cellular machinery to translate the mRNA, thus producing viral proteins. Viral proteins can be divided into 3 categories which are enzyme protein for viral replication, inhibitory factor that stop host DNA, RNA and protein synthesis and structural protein for new virion development (Lodish et al., 2000).

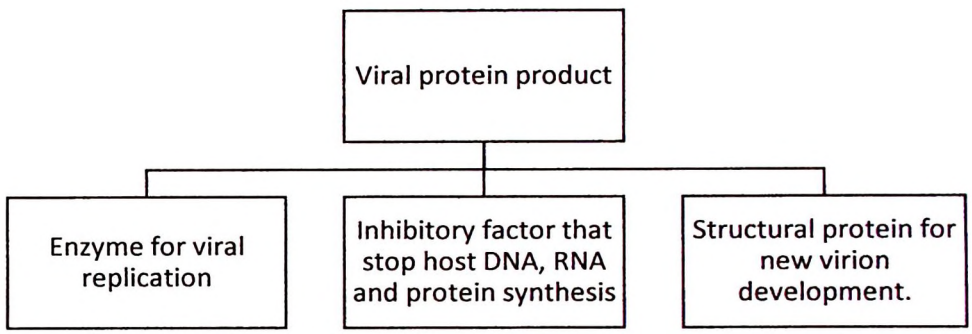

Table 1.1: 3 categories of viral protein products (Lodish et al. 2000)

There are some virus that carries RNA polymerase to produce mRNA without using the host polymerase. This type of virus is single-stranded RNA virus. Example of virus that has RNA dependent RNA polymerase is the polio virus. The newly synthesis viral genome and viral polypeptide will assemble together forming progeny virus. For non-enveloped virus, it will accumulate in the infected cells and the host cells will be lysed to release all the accumulated virus particles. For enveloped virus particle, it undergoes the budding process from the host cells. Viral glycoprotein will then inserted into the host membrane and the virus will bud off at this modified membrane regions. Enveloped-virus are non-infectious until it acquire envelopes (Brooks et al., 2012).

Viral growth cycle can be divided into lytic or lysogenic replication. Lytic replication refers to new virion production in the host cells which eventually leads to the lysis of host cell to release virion. An example of virus that undergoes lytic replication is T4 bacteriophage. Lysogenic replication refers to the viral genome that intergrate into the host's genome. The viral genome remains dormant and replicates as part of the cell's genome from generation to generation. The intergrated viral genome is known as prophage. However, after a certain period of time, the prophage will be activated and it will enter lytic cycle. This type of virus is known as temperate phage and an example is the Human Immunodeficiency Virus (HIV)

Since the virus particle has the capability to infect cells with foreign genetic material, it is normally used as a transporter in gene therapy. Theoretically, the virus will be weakened to prevent any disease development and the desired genetic material will be inserted into targeted cell's genome. The desired genetic material will be expressed to produce the desired protein. An example of a disease that can be treated using gene therapy is cystic fibrosis.

1.2 Avian Influenza A virus

1.2.1 General characteristic of Influenza A virus

Influenza virus can be divided into 3 types (A, B, C) based on antigenicity of nucleoprotein and matrix protein. Influenza A, B and C virus are members of *Orthomyxoviridae* family. It is an enveloped virus and about 120nm in diameter. It has 8 segmented negative-sense RNA which encode for at least 10 proteins that makeup the entire virus structure. Among all viral proteins, there are 2 major proteins that plays important part in virus infectivity which are hemagglutinin (HA) and neuraminidase (NA) (Chee Wei et al., 2014). These HA and NA glycoprotein able to combine together forming different subtype of influenza A virus. Figure 1.4 shows the influenza virus structure. The virus contain different types of viral proteins such as Polymerase A (PA), Polymerase B1 (PB1), Polymerase B2 (PB2), Hemagglutinin (HA), Neuraminidase (NA), Nucleoprotein(NP), Matrix protein (M1 and M2) and Non-structural protein (NS1 and NS2) (Kamps et al., 2009).

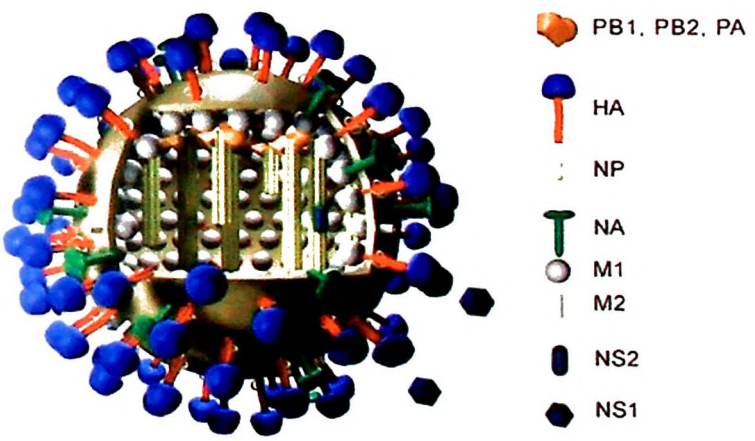

Figure 1.4: Structure of influenza virus (Kamps et al., 2009)

Polymerase B1 protein (PB1), Polymerase B2 protein (PB2) and Polymerase A protein (PA) combined together and form RNA dependent RNA polymerase enzyme. This RNA dependent RNA polymerase enzyme is responsible for the replication of viral RNA genome. Besides, it also exhibits endonuclease activity. Nonstructural protein, NS1 and NS2 function by regulating the synthesis of viral protein in the infected host cell. Matrix protein (M) are divided into M1 and M2 protein. The function of M1 is to build the virus's matrix. M2 protein can only be found in influenza A virus where it act as ion pump channel to manipulate the pH in the endosome (Kamps et al., 2006).

Avian influenza A viruses can be classified into 2 groups which are highly pathogenic avian influenza (HPAI) and low pathogenic avian influenza (LPAI). Influenza A virus can be categorized into one of these groups based on molecular characteristics of the virus and the ability of the virus to cause morbidity and mortality in animal testing such as chickens in a laboratory setting (CDC, 2015). Highly pathogenic avian influenza (HPAI) are only restricted to the H5, H7 and H10 subtypes. This is because these virus subtype are able to undergo mutation to become HPAI which results in severe respiratory problems and death of poultry animals. However, not all subtype of H5, H7 and H10 cause severe diseases (Behzadian et al., 2013).

Highly pathogenic avian viruses can survive in different environments. Based on Kamps et al., 2006, HPAI can survive in low temperature environment for a long duration. Besides, it can survive about four days at 22°C and exceed 30 days living in 0°C environment. For instance, H5N1 influenza A virus isolated in 2004 was able to survive at 37°C for 6 days compared to H5N1 influenza A virus isolated in 1997 outbreak which can only survive for

just 2 days. Thus, these indicates that the HPAI virus had shown the capability to survive in different kinds of environment which renders the effort to eliminate them more difficult.

1.2.2 Source of avian influenza A virus

Influenza A viruses can be found in the wild birds of the world especially in waterfowl. Webster et al. 2006 stated that virus in its natural host will remain in evolutionary stasis. During these periods, the virus show minimal evolution at the amino acid level. Besides, influenza A viruses are able to infect wide range of animals such as chicken, turkey, whale, seal and human. All subtype of influenza A virus can be found in birds except subtype H17N10 and H18N11. These two subtype can only be found in bats (CDC, 2014).

1.2.3 Epidemiology

Influenza infection occur throughout the worldwide. Every year, the burden of influenza epidemics is approximately 3-5 million cases and results in 300,000-500,000 deaths globally (Laursen and Wilson, 2013). Every year, the attack rate of influenza A virus in adult is approximately 5-10% and 20-30% in children. The outbreak of influenza is dependent on the geographical area. In temperate region, the influenza outbreak normally occur during winter season. In tropical area, there is no distinct pattern to indicate that the influenza occur in seasonal. However, the spreading of influenza is throughout the year and highest during rainy season. These influenza A viruses are able to infect the respiratory tract

of humans which leads to severe respiratory diseases. Examples are pneumonia, respiratory failure and in severe situation death may occur. For example, 60% of patients infected with HPAI H5N1 virus died from their illness (WHO, 2016).

1.2.4 Transmission of influenza A virus

Infected patient can shed influenza virus about 2 days before the onset of influenza like illness. The transmission of influenza A virus is due to the exposure of infected live or dead animal poultry. Sneezing and coughing from infected patients will release contaminated water droplets into the surrounding environment. Surrounding people will inhale these droplets into their respiratory system and become ill (WHO, 2016). Besides, human can also be infected due to the contaminated environment with influenza A virus where there is chances for direct skin-to-skin contact or indirect contact with respiratory secretions. For examples are live bird market and hospital (Kamps et al., 2006).

1.2.5 Pandemic threat of avian influenza A virus

Avian influenza A virus is a segmented RNA virus. This virus has RNA dependent RNA polymerase that tend to make error during synthesizing high amount of viral mRNA (Staneková and Varečková, 2010). Viral mRNA is important for the viral genome and viral protein production. Error in synthesizing mRNA strand will results in new mRNA strand that encodes for different protein structure and antigenicity. Thus, the new generation of virion will be different from the original virus. This is known as antigenic drift. In addition, there is no proofreading mechanism during the synthesis of mRNA in RNA virus (Michel Tibayrenc, 2007).

Besides, the genome of avian influenza A virus exist in 8 segment. Thus, this virus can undergo genetic reassortment to change its genetic segment with other viruses which results in novel influenza A virus. This process is known as antigenic shift (William Robert Fleischmann, 1996)

Permanent and small changes in the antigenicity of influenza A viruses are known as antigenic drift. Antigenic drift is defined as the small changes in the influenza genetic materials after replication of the virus at some point. This leads to the development of new a subtype influenza virus that is responsible for the occurrence of influenza epidemics in the world. The annual influenza vaccine that is given to people consists of 2 influenza A strain (influenza A H1N1 virus, influenza A H3N2 virus) and 1 influenza B strain (Laursen and Wilson, 2013). The human body are only able to produce antibody specifically against those 3 influenza strain. Antigenic drift results in the production of HA or NA glycoprotein where the structure and antigenicity is different from the original. The antibody that was produced

by the body cannot recognize these changes and allow the new virus to evade the host immunity system and cause diseases. CDC, 2015 state that the influenza vaccine need to be checked regularly in order to keep pace with the changing influenza virus.

Besides antigenic drift, antigenic shift causes a more devastating situation which leads to pandemic outbreak. Antigenic shift is defined as major genetic changes in the antigenicity of influenza virus which allow the virus to jump from one species to another species. Table 1.2 shows the occurrence of devastating pandemic throughout human history:

Table 1.2: Occurrence of influenza pandemic in human history (Kamps et al., 2006)

Year	Type of influenza A virus	Result of pandemic	Death toll
1889	H3N2	Moderate	unknown
1918	H1N1(Spanish flu)	Devastating	50-100 million
1957	H2N2(Asian flu)	Moderate	1 million
1968	H3N2(Hong Kong flu)	Moderate	1 million

Antigenic shift can occur in 3 ways. One of the ways is two different strains of influenza A virus such as the bird strain and the human strain present in the same mixing host such as swine. Both influenza strain virus infect and coexist together in the same cell. During this period, the genetic reassortment between the bird strain virus and the human strain virus leads to the development of a new subtype of influenza virus that can move efficiently from bird to human. Besides, there are some virus that can directly jump from bird to human. Inside the human, the bird strain virus undergoes adaptive mutation which allows it to move from one human to another. Finally, the bird strain of influenza virus will infect intermediate host and then to human. During these process, there are no genetic changes occur (Kamps et al., 2006).

Kamps et al., 2006 stated that novel epidemic influenza A strains emerged every 1 to 2 years due to selective point mutations in two surface glycoproteins which are hemagglutinin (HA) and neuraminidase (NA). These novel variants are able to evade the human host defense mechanism. Thus, there is no long-lasting immunity against influenza virus despite having a natural infection or vaccination. Antigenic drift and antigenic shift will results in an unforeseeable variability of avian influenza A virus which can cause epidemic and pandemic outbreak.

Besides, frequent emergence of novel influenza A virus and lack of herd immunity in the population remains a threat to human health (Laursen and Wilson, 2013). An example of avian influenza A virus, H5N1 has become the main concern for future pandemic outbreak. H5N1 viruses have started to develop resistance to antiviral drugs such as Amantidine and Rimantidine. Furthermore, this virus is able to spread worldwide very fast and acquire genes from another virus. This will make the virus become more adapt in the mammalian host and more efficient in human to human transmission (Behzadian et al., 2013).

From 22nd February 2016 to 6th March 2016, data about influenza infection had been collected from 96 laboratories throughout the world. 47 202 cases were confirmed positive for influenza virus infection. 74.2% of the total cases were due to influenza A virus whereas the remaining 25.8% were due to influenza B virus. Among the influenza A virus infection, about 15851(87.3%) cases were due to influenza A(H1N1) virus and the remaining 2300 (12.7%) were due to the influenza A(H3N2) virus (WHO,2016).

1.2.6 Clinical symptom of influenza infection

Patient infected with influenza virus will experience flu like symptoms. Examples are like high fever (> 38°C), cough, sore throat, runny nose, headaches and malaise. Some people may experience nausea, vomiting and diarrhea. Normally, patient can recover in just few days up to two weeks (CDC, 2015a).

However, influenza virus can also cause severe complications in infected patient. Examples of complications are pneumonia, bronchitis and ear infections. They may require hospitalization for treatment. Death may occur when high risk groups are infected with influenza virus. Examples are elderly, very young individuals and people with underlying medical conditions such as diabetes, pulmonary disease and cardiac disease (Harper et al., 2009). For instance, H5N1 influenza A virus can cause mortality rate of 60% in infected patient (CDC, 2015).

1.2.7 Prevention and Treatment

The reservoir of influenza A virus is poultry animals. Therefore, controlling poultry animals will prevent the spreading of virus. Poultry owner must be well informed about avian influenza control and follow recommended biosecurity and infection control practices by the government. Examples are use of appropriate personal protective equipments such as gloves when handling animals and regular checkup on the animals. Besides, case where highly pathogenic avian influenza outbreak, stamping out or slaughter the animals in the particular farm is the only options.(Swayne and Suarez, 2000).

Antiviral drugs such as oseltamivir, peramivir, and zanamivir can be used to treat patients infected with influenza A virus infection. However, highly pathogenic avian influenza (HPAI) such as Asian H5N1 virus have shown resistance against the current drugs available (CDC,2015).

Vaccine is needed to elicit humoral and cellular mediated immune response against the influenza virus. Nowadays, trivalent and quadrivalent vaccine are available in the market. Most of the vaccines are trivalent which made of 2 influenzas A strain and 1 influenza B strain. In quadrivalent vaccine, it consists of 2 influenzas A strain and 2 influenzas B strain. Chemically inactivated of detergent split-virion subunit vaccine, live attenuated vaccine and trivalent inactivated vaccine are the vaccines approved for influenza vaccine. All these vaccines are produced by using embryonated chicken egg as substrate. The difference between live attenuated vaccine and trivalent inactivated vaccine is that live attenuated vaccine consist of living virus but weaken whereas trivalent inactivated vaccine consist of dead influenza virus. Dead influenza virus can elicit good humoral immune response but poor

cellular mediated immunity. Weaken living virus able to elicit good humoral and cellular mediated immunity(Laursen and Wilson, 2013).

However, there are a few drawback of these vaccines. The vaccine production is time-consuming as it is grown in the embryonated chicken egg. The vaccines may also trigger allergic reactions in people who are sensitive to the vaccine and shortage of egg supply for vaccine production during the influenza outbreak. Furthermore, these vaccine are only able to induce short term immunity that maybe not completely effective (Behzadian et al., 2013).

1.2.8 Neuraminidase (NA)

Neuraminidase (NA) is viral glycoprotein encoded by 6th RNA segment that can be found on the surface of virus (Shtyrya et al., 2009). Center for Disease Control and Prevention (CDC 2015) stated that influenza A virus has 11 subtype of NA (N1 through N11).

The gene that encodes for NA protein is about 1413 nucleotide long and the size of NA protein is about 470 amino acid. However, the size of NA protein varies from one subtype to another and even among the same subtype influenza virus. Neuraminidase exist in non-covalently linked homotetramer which resemble mushroom structure. Molecular weight of NA monomer is about 60kDa and molecular weight for tetramer is about 240kDa. Generally, one virus particle has about 50 tetramers (Shtyrya et al., 2009). NA structure consists of short hydrophilic amino terminal tail, hydrophobic transmembrane domain, stalk and globular head domain which contain active site. The short hydrophilic amino terminal tail only contain 6 amino acids which are MNPNQK. This region is highly conserved in majority of influenza A virus. Besides, globular head domain has the most abundant of conserved amino acids. The function of stalk is to allow influenza virus to bud off from the host cell (Matthew J. Sylte and David L.Suarez, 2009)

Neuraminidase is an exosialidase enzyme that cleave α-ketosidic linkage which is located between sialic acid and sugar residue. NA will cleave sialic acid during virus transportation from cytoplasm to the cell surface. This is important to facilitate release of virion from the infected cell surface. In cases where NA is inactivated, sialic acid cannot be removed. This results in hemagglutinin (HA) of the virus bind to the sialic acid on the infected cell surface. This lead to aggregation of virion at the cell surface which prevent the

budding of virion from infected cells. Besides, NA able to cleave neuraminic acid residue which are present in the respiratory tract mucin. This will facilitate the attachment of influenza virus to the respiratory tract cells (Shtyrya et al., 2009).

Affinity of NA protein towards α-2,3 and α-2,6 sialic acid linked glycoprotein are different. Swine and human influenza virus has greater affinity towards the α-2,6 sialic acid whereas avian influenza virus tend to bind to α-2,3 sialic acid (Sylte & Suarez, 2009)

Active site of NA which is located at the globular head is strictly conserved among the NA subtype. Example of antiviral drugs are Oseltamivir and Zanamivir which are similar to transition state of the hydrolysis reaction. These drug prevent release of virion from infected cells to healthy cells. By this way, the influenza virion only confined in the infected cells ((Matthew J. Sylte & David L.Suarez, 2009; Shtyrya et al., 2009).

1.2.9 Hemagglutinin (HA)

There are 18 HA subtype which can be classified into 2 major phylogenetic groups. Group 1 consists of H1, H2, H5, H6, H8, H9, H11, H12, H13, H16, H17 and H18 whereas group 2 consists of H3, H4, H7, H10, H14 and H15 (Mallajosyula et al., 2015). There are 2 functions of HA protein. One of the function is to allow the influenza virus to bind to host cell. This is done by binding of HA to sialic acid cell surface receptor of host cell. Besides, HA also enable fusion between the host and virus membrane which allow release of the virus into the host cell (Laursen and Wilson, 2013). HA protein are able to bind to α-2,3 or α-2,6 sialic acid linked glycoprotein. However, it is different for different strain of influenza. HA protein from avian influenza tends to bind to α-2,3 sialic acid whereas HA protein from human influenza tends to bind to α-2,6 sialic acid (Sylte & Suarez, 2009).

Hemagglutinin glycoprotein was synthesized by the infected host cell in a single polypeptide chain. The produced hemagglutinin exist in trimeric spike (HA0). Each trimeric HA0 consist of head and stem region. The head region consist of small indentation on the surface which act as receptor binding site to bind influenza virus to targeted host cells. The stem region exist in helical structure and facilitated fusion as it houses fusion machinery. Hemagglutinin is made up of combination between glycan and protein. The glycan ensures that the hemagglutinin fold in correct conformation. Besides, since glycan was synthesized by the host, thus it will be recognized by the host immunity system as 'self-structure'. Therefore, glycan can help the influenza virus to avoid detection by the host humoral immunity system(Laursen and Wilson, 2013).

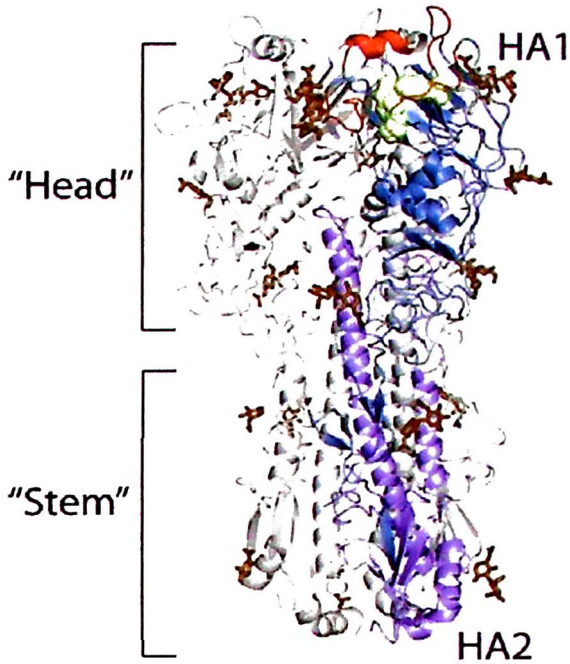

Figure 1.5: Crystal structure of hemagglutinin (Laursen and Wilson, 2013)

Highly pathogenic avian influenza (HPAI) such as H5N1 contain polybasic amino acid in hemagglutinin .This polybasic amino acid act as cleavage site in HA0. This cleavage site is recognize by host protease furin and cleaved HA0 into HA1 and HA2 subunits. The polybasic amino acid motif in H5N1 virus of (A/chicken/Malaysia/5858) is RERRRKKR. HA1 subunit important as it facilitate contact with the cell membrane. Then, HA2 subunit will fuse viral envelope and endosome membrane together which will eventually release the virus content into the cell. Low pathogenic avian influenza may acquire this polybasic amino acid cleavage site and mutate to become high pathogenic avian influenza. These polybasic cleavage site was found to be highly conserved among H5N1 viruses which makes it a candidate for development of universal influenza vaccine (Chee Wei et al., 2014).

1.2.10 Detection methods of avian influenza A virus

The presence of avian influenza A virus in the environment need to be monitor continuously. By detecting these viruses, we will be able to improve the management of patient's condition and control the spreading of the virus. Besides, we also can monitor the mutation in the virus from time to time. Thus, in order to detect influenza A virus, sophisticated detection methods are needed. There are 3 conventional methods to detect influenza virus which are fluorescent antibody (FA) staining, reverse-transcriptase polymerase chain reaction (RT-PCR) and virus isolation method.

In fluorescent antibody (FA) staining, the suspected specimen was stained with monoclonal antibody which are specific to influenza A virus. This antibody was conjugated with fluorochrome which will emit fluorescence in positive samples. However, in order to perform this test, the laboratory must be well equipped with fluorescent microscope and professional trained personnel to interpret the results (Grody et al., 2009). The time taken to complete the test is normally less than 1 hour. This test is not a point-of-care (POC) test, thus it is subjected to other factors that may affect the time taken to complete the test. The result of tests may vary as the interpretation is subjective and depends on the personnel decision (Kamps et al., 2006). Nowadays, there are fluorescent antibodies that specific against influenza A and influenza B virus (Loeffelholz et al., 2008).

Reverse-transcriptase polymerase chain reaction (RT-PCR) function by producing cDNA from RNA genetic material. Then, targeted gene region was amplified by using primer to produce high amount of amplicon. It is able to detect influenza A and influenza B virus based on the conserved regions in nucleoprotein and matrix protein gene. In addition, it is able to differentiate influenza A subtype(Kamps et al., 2006). However, there are few drawbacks. In order to ensure high specificity, selection of primer, probes, optimization of the process and interpretation of result must be carefully done. Therefore, only few laboratories are able to perform the test as there are a lot of strict regulation that needs to be adhered and require BSL 2 laboratory. Besides, this process has a high chance of producing false positive result due to cross-contamination. In addition, the cost for performing this test is high which will burden the patient and takes about 3 hours to produce result (Loeffelholz et al., 2008). Furthermore, the sensitivity of the test can decrease when the RNA samples degrade which require immediate processing of the sample (Kamps et al., 2006).

Virus isolation method requires culture of virus in the embryonated egg. The specimen are allowed to incubate at 38°C for 10 days. The presence of cytopathic effect was then observed to determine type of influenza virus. Most human influenza viruses are unable to be isolated in embryonated chicken eggs, thus cell cultures such as Madin–Darby Canine Kidney (MDCK), is needed for the isolation (Kawaoka and Neumann, 2012). The function of culture is to increase the amount of virus which is more sensitive than fluorescent antibody (FA) staining but less sensitive than RT-PCR.(Kamps et al., 2006) However, in order to culture virus, biosafety level three (BSL3) condition are required. Examples like wearing respirators, decontamination of waste materials and showering before leaving the laboratory are required. Thus, only few laboratory are able to adhere to the requirements.

1.3 Aim of the Study

This study aims to produce recombinant proteins that consists of the highly conserved, hydrophilic and immunogenic region of the influenza A virus hemagglutinin (HA) and neuraminidase (NA). The expressed proteins will be purified for various future purposes such as the development of rapid detection test kit or a potential vaccine candidate. Thus, in order to achieve this aim, a few objectives were targeted:

General Objective:

1. To purify hemagglutinin (HA) and neuraminidase (NA) recombinant protein of avian influenza A virus in BL21 *Escherichia coli*

Specific Objective:

1. To verify the HA and NA gene cloned in pET 47b(+) vector.
2. To express HA and NA recombinant protein in BL21 *E.coli* host cell.
3. To purify HA and NA recombinant protein using Immobilized Metal Affinity Chromatography (IMAC).