CONSTRUCTION AND CLONING OF SYNTHETIC M2e-NP GENE FROM ASIAN INFLUENZA A VIRUSES

by

TEH YEW SENG

Dissertation submitted in partial fulfillment of the requirements for the degree of Bachelor of Health Sciences (Biomedicine)

May 2013

ACKNOWLEDGEMENT

First and foremost, I would like to express highest gratitude of appreciation to my supervisor, Associate Professor Shaharum Shamsuddin for his continuous motivation, insight, guidance, patience and critical comments until the completion of all my lab works and thesis writing. My appreciation also extended to Dr Rapeah Suppian for the knowledge sharing on assembly PCR which guided me through the challenges for carrying out this research.

I would like to give a big thank you to final year project course coordinator, Dr. See Too Wei Cun for all the arrangement, invaluable advice and encouragement throughout this final year project. I would also like to express my sincere thanks to Dr Few Ling Ling, Dr Lim Boon Huat and Encik Nor Azmi Zainal for the endless spiritual support and continuous motivation in my undergraduate life.

Not forget to express my sincere thanks to postgraduate seniors, Chang Ching Ching, Sam Kuan Chee Sian and Tee Chee Wei for providing useful input and constructive comments which help me complete my lab works and broaden my horizon in molecular biology field.

I would like to thank my best friend Tan Yew Seng and Wong Yeng Wei as well for their continued emotional support throughout these four years. It is such a comfort to know that no matter what, you are always with me and ready to listen to me.

Lastly, I would like to thanks my mother as well as my sister for their spiritual support and love which provide me strength until this stage. My journey towards earning this degree is not possible without the support and motivation from you all. I am very grateful for meeting all of you here which make me have a fruitful experience in my undergraduate life.

TABLE OF CONTENT

LIST OF TABLE

LIST OF FIGURE

LIST OF PLATE

LIST OF SYMBOL, ABBREVIATION AND ACRONYM

~	Approximately
%	Percentage
µM/µL	Micromolar per microliter
°C	Degree Celsius
µM	Micromolar
∞	Infinity
aa	Amino acid
bp	Base pair
$CaCl_2$	Calcium chloride
CTL	Cytotoxic T lymphocyte
dH_2O	Distilled water
DMF	Dimethyl formamide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
et al	and others
EtBr	Ethidium bromide
g	Gram
g	Gravity
h	Hour
HA	Hemagglutinin
kb	Kilobase
L	Liter
LB	Luria-Bertani
M1	Matrix 1 protein
M2	Matrix 2 protein
M2e	Extracellular domain of Matrix 2 protein
M2e-NP	Extracellular domain of Matrix 2 protein with Nucleoprotein
mg/mL	Milligram per milliliter
$MgCl_2$	Magnesium chloride
min	Minute
mL	Milliliter
mM	Millimolar
NA	Neuraminidase
NaCl	Sodium chloride
NEP	Nuclear export protein
nm	Nanometer
NP	Nucleoprotein
NS1	Non structural protein 1
OD	Optical density
PA	Polymerase Acidic Protein
PB1	Polymerase Basic Protein 1
PB2	Polymerase Basic Protein 2
PCR	Polymerase Chain Reaction

Pfu DNA Polymerase	*Pyrococcus furiosus* DNA Polymerase
psi	Pounds per square inch
rpm	Revolutions per minute
s	Second
TAE	Tris-acetate EDTA
Taq DNA Polymerase	*Thermus aquaticus* DNA polymerase
U	Unit
U/µL	Unit per microliter
UV	Ultraviolet
w/v	Weight per volume
X	Times
β	Beta
µL	Microliter

ABSTRAK

Kemunculan influenza A novel yang mampu merebak dengan cepat di dalam kalangan komuniti merupakan satu ancaman kepada kesihatan bagi warga seluruh dunia pada zaman ini. Vaksin konvensional influenza umumnya mensasarkan dua antigen permukaan utama iaitu "Hemagglutinin (HA)" dan "Neuraminidase (NA)" yang sentiasa bermutasi melalui mekanisma "antigenic drift" dan "antigenic shift". Hal sedemikian telah menjadi satu cabaran yang besar kepada ahli virologi untuk pengenalpastian strain dan juga penyediaan vaksin yang efektif ketika wabak berlaku. Oleh itu, satu vaksin umum yang berupaya untuk memberi perlindugan immunisasi terhadap semua jenis influenza A virus amat diperlukan. Kebelakangan ini dua jenis protein iaitu "nucleoprotein (NP)" dan "extracellular domain of ion channel matrix protein (M2e)" telah dilaporkan berpotensi sebagai calon-calon untuk pembangunan vaksin umum disebabkan jujukan acid amino mereka adalah terpelihara bagi semua virus influenza A. Di dalam kajian ini, kita berharap dapat membina, mengamplifikasikan serta mengklonkan gen sintetik M2e-NP yang berpotensi untuk dibangunkan sebagai vaksin umum. Kajian ini terdapat 2 bahagian. Di dalam bahagian pertama, analisis "*in silico*" telah dijalankan terhadap 24 jenis virus influenza A daripada negara Asia untuk mengenal pasti jujukan acid amino yang terpelihara serta immunogenik. Di dalam bahagian 2 pula, sebanyak sebelas oligonukleotida yang panjangnya di antara 32 hingga 38 bp dihimpunkan serta diamplifikasikan dengan menggunakan *Taq* DNA Polymerase melalui tindak balas PCR himpunan. Produk PCR kemudian diligasikan ke dalam pCR®2.1-TOPO® vektor diikuti dengan proses transformasi ke dalam sel kompeten *E. coli* DH5α. Kemudian, beberapa koloni putih dipilih dan disaringkan dengan kaedah PCR koloni. Vektor rekombinan diekstrak daripada koloni positif. Kemudian pencernaan

vector rekombinan dengan enzim pembatasan *Hind*III dan *Xho*I bagi memastikan selitan sintetik M2e-NP terdapat di dalam vector rekombinan yang diekstrak sebelum dihantar untuk semakan jujukan. Keputusan diperolehi di dalam kajian ini menunjukkan kita telah berjaya menghimpunkan serta mengamplifikasikan gen sintetik M2e-NP. Produk DNA yang diperolehi adalah bersaiz lebih kurang 250 bp. Keputusan semakan jujukan DNA menunjukkan semua oligonucleotida dihimpun degan betul meskipun dengan sedikit mutasi bes. Sebagai kesimpulan, tindak balas himpunan PCR menyediakan satu cara yang baru untuk menghasilkan gen sintetik daripada kombinasi species yang merupakan mikroorganism yang berbahaya serta membolehkan penukaran bes kodon jujukan asid amino yang dapat meningkatkan pengekspresian gen.

ABSTRACT

The sudden emergence and explosive spreading of novel influenza A viruses in the community represents a major threat to global public health. Conventional influenza vaccines are mainly targeting two surface glycoproteins namely hemagglutinin (HA) and neuraminidase (NA) that undergo antigenic drift and antigenic shift continuously. This poses extraordinary challenges to the virologist for strain identification and effective vaccine preparation during a pandemic. Hence, there is a pressing need to develop an effective universal vaccine that can confer immediate protection against different variants and subtypes of influenza A viruses. In recent years, nucleoprotein (NP) and extracellular domain of ion channel matrix protein (M2e) have been widely targeted as possible universal vaccine candidates since these two proteins are highly conserved among all influenza A viruses. In this study, we aimed to construct, amplify and clone synthetic M2e-NP gene that can serve as universal influenza vaccine in future. This study comprises of two major parts. First part involves *in silico* analysis of M2e and NP amino acid sequence from 24 subtypes and variants of Asian influenza A viruses in order to identify conserved and immunogenic peptide sequences. In second part, 11 newly synthesized oligonucleotides with length ranged from 32-38 bp were assembled and amplified using *Taq* DNA polymerase in assembly PCR. Subsequently, PCR products were ligated into the pCR®2.1-TOPO® vector followed by transformation into the *E. coli* DH5α competent cells. Few white colonies were selected and screened via colony PCR. Then, recombinant plasmids that harboured desired insert were extracted from positive clone and subjected to restriction enzyme digestion using *Hind*III and *Xho*I to verify the presence of the insert prior to sequencing. In the present study, we had successfully constructed and cloned synthetic M2e-NP gene. Target band with size

of approximately 250 bp was seen on 2% agarose gel in assembly PCR. DNA sequencing results show colonies carrying the correct assembled M2e-NP gene albeit with minor base mutations. In conclusion, assembly PCR had offered an alternative way to synthesize gene from combination of various strains of highly infectious microorganism without any risk of exposure meanwhile it allow codon optimization for high protein expression in future work.

CHAPTER 1

INTRODUCTION

1.1 History of Influenza Outbreak

1.1.1 Influenza Pandemics in the 20th Century

There have been three major influenza pandemics occurred in the 20th century. These 3 major pandemics refer to Spanish flu (1918-1920), Asian flu (1957) and Hong Kong flu (1968). The Spanish flu in 1918-1920 caused more than 50 million deaths globally and infected as many as one third of the world's population making it the worst influenza pandemic recorded (Johnson and Mueller, 2002; Taubenberger and Morens, 2006; Pada and Tambyah, 2011). Fatalities cases rates were 5-20 times higher than any other influenza pandemic or epidemic (Taubenberger and Morens, 2006). Clinical symptoms were largely confined to the respiratory tract, and many patients died with typical bacterial pneumonia resulting from secondary bacterial infection. The first identification of cases occurred in the northern hemisphere and subsequently spread globally in three waves.

The first wave, primarily concentrated in Europe, Asia, and North America, began in the spring of 1918. The second and third waves followed in the fall and winter of 1918-1919 were exceedingly more contagious with a 10-fold increase in the death rate among young and healthy adults compared to the first wave (Potter, 2001; Taubenberger and Morens, 2006). Phylogenetic analysis of the influenza RNA isolated from a formalin-fixed, paraffin-embedded lung tissue autopsy of a victim in the pandemic suggested the virus was an A/H1N1 subtype, and had an avian-like origin (Reid *et al.*, 1999; Reid *et al.*, 2000; Reid *et al.*, 2004). Following the 1918 pandemic, A/H1N1 viruses continued to circulate in the global population, causing

seasonal influenza epidemics until 1957 (Logan and MacKay, 1951; Viboud *et al.*, 2006).

Humanity succumbed to another worldwide influenza outbreak nearly forty years after the great pandemic of 1918. In the Asian flu pandemic in 1957, a new virus subtype, A/H2N2, emerged as a reassortant virus carrying the HA, NA and PB1 of an avian virus, and other segments of the circulating A/H1N1 (Scholtissek *et al.*, 1978a; Kawaoka *et al.*, 1989). Acquisition of the novel avian HA and NA had allowed the A/H2N2 virus to escape human immune response and produced a pandemic, displacing the A/H1N1 virus (Cox and Subbarao, 2000). The alert of a rising pandemic first emerged from Hong Kong news reports citing an epidemic of 250,000 people. The infecting H2N2 strain caused disease in over 25% of the USA population, localized attack rates as high as 40-60%, and disproportionately affected pregnant women and those with chronic underlying conditions (Henderson *et al.*, 2009; Pada and Tambyah, 2011)

In 1968, a genetic reassortment event occurred in which the circulating A/H2N2 acquired another avian HA and PB1, resulting in a new virus subtype A/H3N2 (Scholtissek *et al.*, 1978a; Kawaoka *et al.*, 1989; Lindstrom *et al.*, 2004). The pandemic was termed as Hong Kong flu since the A/H3N2 virus was first isolated in Hong Kong (Cockburn *et al.*, 1969). Widespread disease with increased excess mortality was observed in the United States during the winter of 1968–1969, but in some other countries, including the United Kingdom, an epidemic did not occur until the winter of 1969–1970. Attack rates were the highest (40%) among 10- to 14- year-old children. Total influenza-associated excess mortality for this pandemic was estimated at 33,800 in the United States (Cox and Subbarao, 2000).

In 1977, a Russian influenza A/H1N1 strain nearly identical to the human A/H1N1 strains that had circulated widely in the 1950s resurfaced into the population (Nakajima *et al.*, 1978; Scholtissek *et al.*, 1978b). Though not considered a true pandemic, the reemergence of influenza A H1N1 viruses in 1977 has had a significant influence on the epidemiology of influenza in recent years. The first outbreak of disease was recorded in Tianjin, China in May 1977. They spread to other parts of Asia and reached Russia by November 1977, then spread to Europe, North America, and the Southern Hemisphere. Attack rates of over 50% were observed among schoolchildren, but illness occurred almost exclusively among persons younger than 20 years of age as the older populations had immunity from previous exposure to nearly identical virus (Cox and Subbarao, 2000).

1.1.2 2009 H1N1 Pandemic

Although A/H1N1 viruses reappeared in 1977 and continued to circulate among humans, the seasonal epidemics of influenza A virus from 1968 to 2009 were dominated by A/H3N2 virus variants generated by antigenic drift (Smith *et al.*, 2004; Taubenberger and Morens, 2006). Then in April 2009, a novel swine-origin influenza A/H1N1 virus was transmitted to humans in Mexico and rapidly spread to the rest of the world through human-to-human transmission, and generating the first influenza pandemic of the 21st century (Garten *et al.*, 2009; Smith *et al.*, 2009; Girard *et al.*, 2010). The virus was found to be antigenically unrelated to human seasonal influenza viruses but genetically related to viruses known to circulate in pigs. Molecular studies of the new A (H1N1) 2009 pandemic virus genome showed that it was arose as a result of reassortment between two swine influenza viruses, North American triple reassortant A/H1N2, and Eurasian avian like swine A/H1N1

(Babakir-Mina *et al.*, 2009; Garten *et al.*, 2009; Smith *et al.*, 2009). Due to this reassortment, the pandemic A/H1N1 2009 virus carried HA, NP and NS from the classical swine lineage, PB2 and PA from the North American avian lineage, PB1 from the human lineage, and NA and M from the Eurasian swine lineage (Neumann *et al.*, 2009). Figure 2.1 showing the mechanism for genesis of A/H1N1 2009 pandemic virus.

1.1.3 H5N1 Avian Influenza Outbreak

In general, avian influenza viruses such as A/H5N1 do not replicate efficiently in humans. High doses of avian viruses are required to produce a quantifiable level of replication in human volunteers (Beare and Webster, 1991). This is due to the different receptor-binding specificity between the HA in human and avian influenza viruses. Human influenza strains preferentially bind to receptors with a terminal α 2, 6-linked SA, whereas the avian influenza strains preferentially recognise receptors with a terminal α 2, 3-linked SA (Rogers and Paulson, 1983). However, α 2, 3-linked SA receptors are found in the human lower respiratory tract. Direct exposure to avian influenza virus, such as coming in close contact with secretions and excreta from infected birds or contaminated poultry products, will allow the virus to penetrate into the lower respiratory tract and replicate predominantly (Shinya *et al.*, 2006). In humans, A/H5N1 virus is capable of causing severe and often fatal diseases characterized by viral pneumonia and multi-organ failure (de Jong *et al.*, 2006).

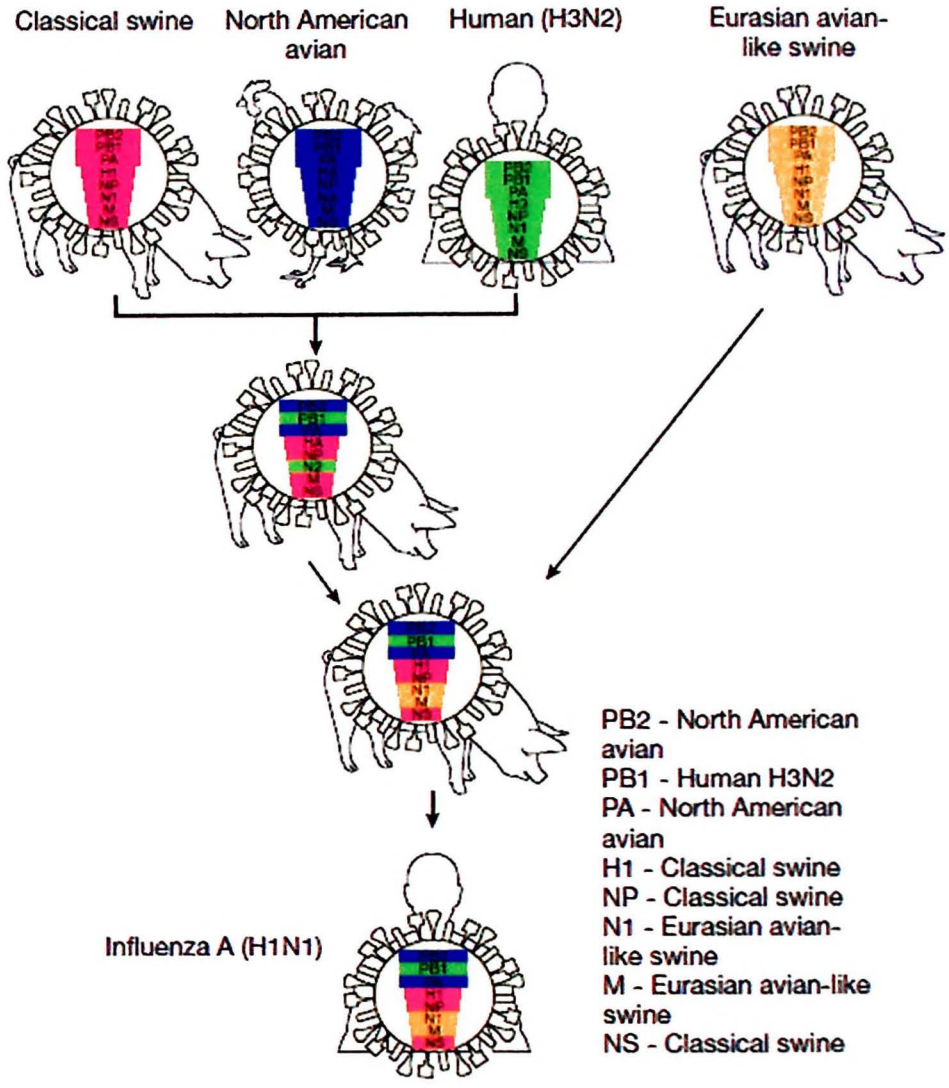

Figure 1.1 Mechanism for genesis of swine-origin H1N1 influenza viruses. In the late 1990s, reassortment between human H3N2, North American avian, and classical swine viruses resulted in triple reassortant H3N2 and H1N2 swine viruses that have since circulated in North American pig populations. A triple reassortant swine virus reassorted with a Eurasian avian-like swine virus, resulting in the pandemic A/H1N1 2009 viruses that are now circulating in humans (Neumann *et al.*, 2009).

1.2 Nomenclature and Host of Influenza

Influenza virus is belonging to viral family *Orthomyxoviridae* that includes negative sense, single strand RNA viruses. Viruses in the family *Orthomyxoviridae* have a genome consisting of 6-8 single stranded negative sense RNA segments. This family of RNA viruses consists of five genera, namely *Influenzavirus A*, *Influenzavirus B*, *Influenzavirus C*, *Thogatovirus* and *Isavirus*. The first three genera each have one species which are Influenza A virus, Influenza B virus and Influenza C virus respectively. Influenza A, B and C viruses were distinguished based on the antigenic differences in their internal proteins such as nucleocapsid (NP) and matrix (M) protein. Influenza A viruses are further divided by the antigenicity of their HA (H1-17) and NA (N1-9) glycoproteins. The genome of influenza A and B viruses consists of 8 negative stranded segments, while influenza C has 7 segments. All three types can infect humans, with infections by type A generally being the most severe and type C being the least (Horimoto and Kawaoka, 2005).

Influenza viruses are named according to the standard nomenclature established by the World Health Organization (WHO). According to this standard an influenza virus name begins with the influenza virus type (A, B, or C), followed by the species from which the virus is isolated (omitted if human), the geographical location of the isolation, the isolate identifier, the year of isolation and follow by the hemagluttinin (H) and neuraminidase (N) subtype for the case of influenza A virus. For example, the virus of H1N1 subtype isolated from humans in Puerto Rico in 1934 is: A/Puerto Rico/8/34 (H1N1). In addition to humans, various other animals serve as influenza virus hosts. Common hosts include various types of birds (avian), hosts, pigs, horses, ferrets and others (Trampuz *et al.*, 2004).

1.3 Structure of Influenza A Virus

1.3.1 Morphology of Influenza A Virus

Influenza A viruses are small, pleomorphic particles. Spherical virions range from 80 to 120 nm in diameter while filamentous particles can measure up to several micrometers in length. Fresh isolates are generally filamentous which change into almost entirely spherical morphology following several passages (Choppin *et al.*, 1960). Under electron microscope, the virion shows two distinct layers which consist an evenly spaced outer zone covering a densely packed genetic material (Hoyle *et al.*, 1962). Figure 1.2 showing the photomicrograph of influenza virions.

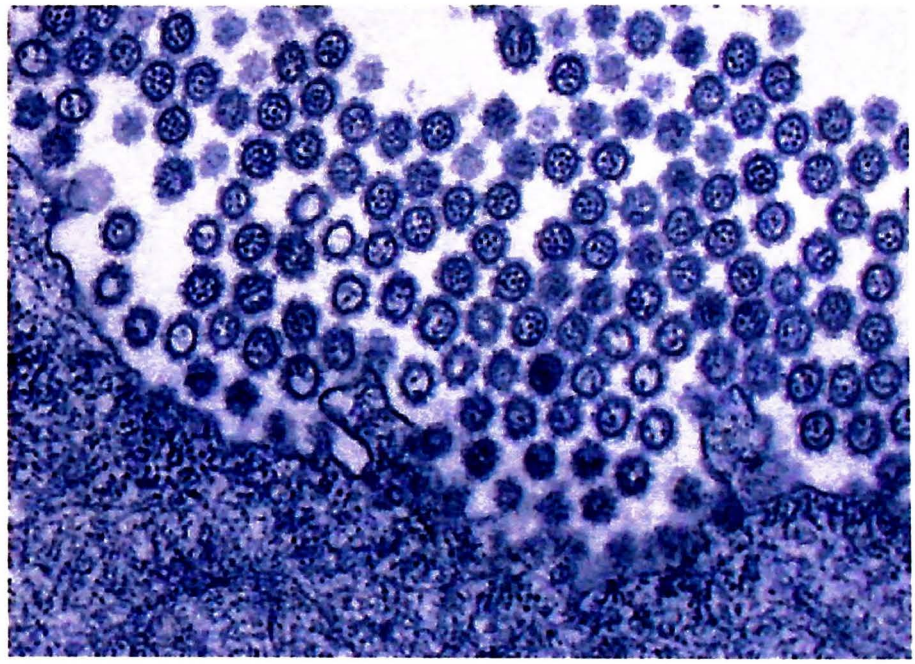

Figure 1.2 Photomicrograph of influenza virions (Noda *et al.*, 2006)

1.3.2 Viral Genome and Encoded Proteins

The genome of influenza A virus consists of eight single-stranded negative-sense RNA segments, ranging in size from 890 to 2341 bases that encode for 12 viral proteins (Lamb, 1989). A schematic diagram of the influenza virus genome showing the eight RNA segments and locations of the viral proteins is illustrated in Figure 1.3. These proteins fall into 4 major categories which including surface glycoproteins, matrix proteins, ribonucleoproteins and non-structural proteins.

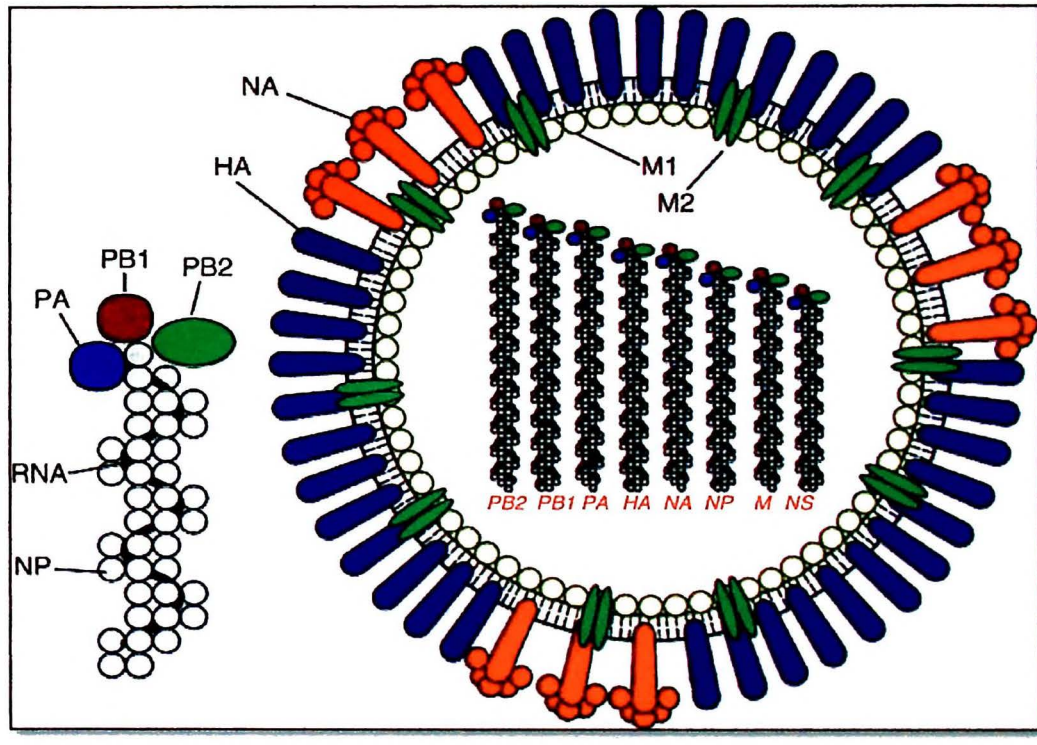

Figure 1.3 Schematic diagram of the influenza A virus genome. The two surface glycoproteins HA and NA, and the M2 ion channel protein are embedded in the viral envelope. The RNP complex on the left consists of vRNA, and PB2, PB1 and PA associated with NP. The genome is not drawn to scale (Cox *et al.*, 2004).

1.3.2 (a) Surface Glycoproteins

Segment 4 and 6 encode for the surface glycoproteins of influenza virus which including hemagglutinin (HA; 566aa) and neuraminidase (NA; 469aa). HA is responsible for receptor binding and membrane fusion during viral entry into the cell. Each HA polypeptide has three sub-domains: receptor binding, vestigial esterase and fusion domain (Gamblin *et al.*, 2004). Biological activation of HA requires three posttranslational modifications: (a) cleavage of the amino terminal signal peptide; (b) glycosylation and palmitoylation; and (c) proteolytic cleavage of precursor HA into two disulfide-linked subunits, HA1 and HA2. Newly synthesized HA is cleaved to remove the amino-terminal hydrophobic sequence of 14 to 18 amino acids (aa), which are the signal sequence for transport to the cell membrane. After the HA is cleaved into two disulfide-linked chains, HA1 and HA2, carbohydrate side chains are added subsequently (Mitnaul *et al.*, 2000; Skehel and Wiley, 2000). The fully processed HA consists of HA1 (typically) of about 324 aa plus variable carbohydrate, and HA2 (typically) of about 222 aa plus variable carbohydrate and 3 palmitate residues.

The hemaglutinin (HA) protein exists as a trimeric protein, composed of three copies of the hemaglutinin monomer. Each HA molecule consists of a globular head on a stalk. The head is made up exclusively of HA1 and contains the receptor-binding cavity as well as most of the antigenic sites of the molecule. The stalk consists of all of HA2 and part of HA1. The carboxy-terminal region of HA2 contains the hydrophobic transmembrane sequence and a terminal cytoplasmic anchor sequence where palmitate is attached (Webster *et al.*, 1992). It is the most important viral antigen against which neutralising antibodies are directed (de Jong *et al.*, 2006; Stevens *et al.*, 2006). On the other hand, NA that exists as a tetrameric

protein play role in enzymatic cleavage of sialic acids from cell surfaces to allow the release of newly synthesized progeny virions. Thus, it is thought to be important in the final stages of release of new virus particles from infected cells, prevent new virus particles agglutinating, thus increasing the number of free virus particle and hence spread of the virus from original site of infection (de Jong *et al.*, 2006).

1.3.2 (b) Matrix Proteins

Segment 7 encodes two products, matrix protein M1 (252aa), and M2 (97aa), which share overlapping reading frame due to alternate splicing. M1 is the most abundant component of the virus, and forms a shell beneath the lipid bilayer (Compans *et al.*, 1970). It interacts with the RNP complexes, HA, and NA, and facilitates the export of progeny RNPs out of the nucleus into the cytoplasm (Nayak and Hui, 2002). M2 resides in the viral envelope and functions as an ion channel. It is crucial for the dissociation of RNP complexes from M1 in the early phase of the infectious cycle (Bui and Helenius, 1996).

1.3.2 (c) Proteins of Ribonucleoprotein Complex

The single stranded RNA of influenza A virus is quickly degraded within a cell if not complexed with other proteins. The vRNP complex consists of one of the eight viral RNA segments complexed with multiple nucleoproteins (NP), and one of each of PA, PB1 and PB2. Genome segment 5 encodes the nucleoprotein (NP; 498aa). The nucleoproteins provide a valley that the vRNA strand binds to and stabilizes the RNA. On average there is one NP protein per 24 bases of viral RNA (Portela and Digard, 2002). It is also required for the efficient elongation of influenza replicative products (Honda *et al.*, 1988). On the other hand, RNA-dependent RNA

polymerase complex which including the polymerase basic protein 2 (PB2; 759aa), polymerase basic protein 1 (PB1; 757aa), and polymerase acidic protein (PA; 716aa) are encoded by segment 1-3 (Honda *et al.*, 1990). The PB2 subunit of the influenza polymerase binds to cap-1 structures of host pre-mRNAs, and is involved in transcription initiation. As for PB1 polymerase, it responsible for elongation of the primed nascent viral mRNA and also as elongation protein for template RNA and vRNA synthesis. Recently, a novel protein known as PB1-F2 (87aa) has been identified to be expressed from an alternate open reading frame (+1 ORF) of PB1 gene. This viral protein localizes to the mitochondria and play role in influenza virus pathogenesis through induction of apoptosis (Webster *et al.*, 1992; Li *et al.*, 2001). However the precise function of PA is not well established. There are some evidences suggests that it involves in both transcription and replication, and contains the domain for endonuclease activity (Dias *et al.*, 2009).

1.3.2 (d) Non-structural Proteins

Genome segment 8 also encodes two proteins due to alternate splicing. They are non-structural protein 1 (NS1; 230aa), and nuclear export protein (NEP, formerly knownas NS2; 121aa). The non-structural proteins 1 (NS1) is known to operate as a counter-measure to the host defense mechanisms against viral infection thus ensuring effective replication of the influenza virus within the host cell (Ludwig *et al.*, 2002). NEP mediates the nuclear export of progeny RNP of the cell by acting as an adaptor between the RNP complexes and nuclear export machinery (O'Neill *et al.*, 1998).

1.4 Replication Cycle of Influenza A Virus

The replication cycle of influenza virus consists of 6 main steps. These including attachment and endocytosis, fusion and uncoating, transportation of viral ribonucleoprotein, transcription and replication, translation and assembly and budding. A schematic diagram illustrating the complete replication cycle of Influenza A virus is shown in Figure 1.4.

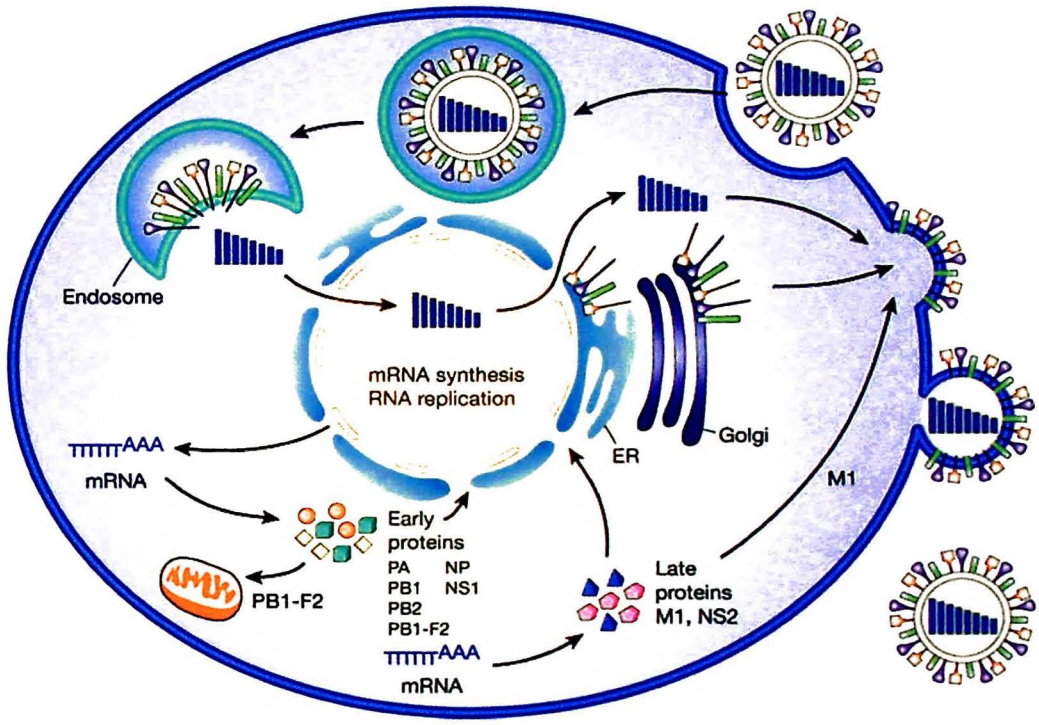

Figure 1.4 Replication cycle of influenza A virus. First, influenza A virus via its HA binds to SA receptors on the cell surface. This triggers endocytosis. Low pH conditions of the endosome induce membrane fusion and particle uncoating which allows the release of RNPs into the cytoplasm. vRNPs are transported to the nucleus where all transcription and replication take place. Messenger RNAs are exported to the cytoplasm for translation. Early viral proteins, that is, those required for replication and transcription, are transported back to the nucleus. Late in the infection cycle, the M1 and NS2 proteins facilitate the nuclear export of newly synthesized vRNPs. PB1-F2 associates with mitochondria. The assembly and budding of progeny virions occurs at the plasma membrane. Figure was taken from (Neumann *et al.*, 2009).

1.4.1 Attachment and Penetration (Endocytosis)

The influenza viral life cycle begins with the binding of a virion to a host cell. The binding of human influenza A viruses to epithelial cells is mediated by interaction of HA with cell surface receptors containing terminal α-2,6-linked SA (Rogers and Paulson, 1983). Following attachment, virions are endocytosed into cellular compartments. Several endocytic mechanisms including clathrin-coated pits as well as clathrin-caveolin independent pathways have been described for influenza virus entry (Matlin *et al.*, 1981; Rust *et al.*, 2004).

1.4.2 Fusion and Uncoating

Once internalized, infusion of protons (H+) into the virion via the M2 ion channel stimulates fusion of the viral and endosomal membranes. This is because acidification causes a structural change in the viral hemagluttinin proteins and frees the fusion peptide of the hemagluttinins HA2 subunit to interact with the endosomal membrane. This in turn facilitating the release of RNPs from M1 which allow M1-free RNPs to be released into the cell cytoplasm for the next step (Bui *et al.*, 2000).

1.4.3 Transportation of Viral Ribonucleoprotein into Host Nucleus

Once the ribonucleoproteins have been released into the cytoplasm of the cell they are then transported into the host nucleus. The transport process begins with the recognition of the nuclear localization signal (NLS) located on each nucleoprotein of the vRNP by proteins known as karyopherins. A trimeric complex of the vRNP and two karopherin proteins forms and then binds with the nuclear pore and the host cell nuclear membrane. The vRNP are then transported and released into the cell nucleus (Martin and Helenius, 1991; Kemler *et al.*, 1994; Whittaker *et al.*, 2000).

1.4.4 Transcription, Replication and Translation

Once inside the nucleus, influenza viral ribonucleic acids (vRNA) are then transcribed and replicated. The viral negative-strand RNA (vRNA) serves as a template for the synthesis of both capped, polyadenylated viral messenger RNA (mRNA) and full length positive-strand RNA, "complementary RNA" (cRNA). The cRNA in turn serves as a template for the synthesis of new vRNA molecules. (Shapiro and Krug, 1988; Mullin *et al.*, 2004; Vreede *et al.*, 2004; Vreede and Brownlee, 2007). Host cellular machinery is used to translate viral proteins from viral mRNA in the cytoplasm. Viral mRNAs are then transported out into the cytoplasm for viral protein synthesis. Nucleoprotein (NP) and NS1 protein are preferentially synthesized at the early stage of viral infection (Hatada *et al.*, 1989). HA, NA and M2 proteins are transported to the cell surface while PB2, PB1, PA, NP, NS1, NS2 and M1 are actively transported to the nucleus of cells using a nuclear localization signal through nuclear pore complexes (Portela *et al.*, 1999).

1.4.5 Assembly and Budding

Progeny RNPs assembled inside the nucleus are exported into the cytoplasm with the help of M1 and NEP. Interaction of M1 with progeny RNPs and the cytoplasmic domains of HA and HA facilitates the movement of RNPs to the assembly site in the plasma membrane. Hemagluttinin proteins are then assembled into trimers while NA and M2 are assembled into tetramers and all are exported to the Golgi apparatus. The final assembly steps include association of the HA, NA and M2 into lipid layer along the plasma membrane. The packaging of the eight vRNP occurs by selective incorporation or random incorporation. Upon completion of

budding the influenza virion is released by way of neuraminidase enzymatic release of sialic acid (Chen *et al.*, 1999; Leser and Lamb, 2005; Boulo *et al.*, 2007).

1.5 Pathogenesis of Influenza A Virus

In mammalian hosts, influenza virus is mainly a respiratory pathogen with disease severity ranging from asymptomatic to severe systemic fatal infection. The virus primarily infects and replicates in the ciliated columnar trachea-bronchial epithelium and then spreads to the lower respiratory tract or sometimes to systemic organs. Pathogenesis of influenza virus can mainly be ascribed to its ability to shut off host cell protein synthesis, induce apoptosis and regulate several cellular transcription systems in order to avoid host immune system. The PB2 polymerase protein binds to host cell mRNA cap structures in the nucleus and cleaves them to prime viral mRNA synthesis resulting in degradation of the decapped cellular mRNAs. Nuclear export of cellular pre-mRNAs that escape "cap snatching" is blocked by NS1 protein by inhibiting their splicing. Productive replication of influenza virus in respiratory epithelium results in apoptotic cell death (Lu *et al.*, 1994; Brydon *et al.*, 2005). Besides, NS1 protein plays an important role in regulating cellular transcription systems. The NS1 protein is an antagonist of IFN α/β production and it also inhibits maturation and cytotoxic T cell stimulation of dendritic cells and thereby retards the development of an adaptive immune response against the virus (Fernandez-Sesma *et al.*, 2006). One major determinant of influenza virus pathogenesis is the cleavability of precursor hemagglutinin protein. The proteolytic cleavage of precursor hemagglutinin protein into two disulfide linked subunits, HA1 and HA2, is a prerequisite for successful fusion of viral and host

endosomal membranes and therefore is necessary for viral infection (Steinhauer, 1999).

1.6 Host Immune Response to Influenza A Virus

Similar to other virus, both innate and adaptive immunity system are involved in the immune response that against influenza A virus infection. The innate immune system comprises physical barriers, soluble chemical factors and cellular components. The epithelial surfaces form the first physical barrier to any bacterial or viral infections. Mucins produced by respiratory epithelial cells trap bacteria or virus particles while movement of broncho-pulmonary cilia clears them out of the airway.

Pathogens that successfully breach the physical barrier are neutralized by soluble antimicrobial factors. For example, human alpha defensins are short cationic antimicrobial peptides of neutrophils that can inhibit influenza virus replication. Complement present in human serum can neutralize influenza virus and *in vivo* experiments show that in association with natural antibody (IgM), complement can provide protective immunity in influenza naive hosts (Jayasekera *et al.*, 2007; Salvatore *et al.*, 2007). Besides, collectins are important soluble PRRs of the innate immune system. Among the nine members that have been identified so far, mannan-binding lectin (MBL) and surfactant proteins A and D (SP-A and SP-D) are the most studied for their role in host defense against infectious agents. They bind to a variety of carbohydrate residues on pathogen surfaces and thereby activate cells to respond by way of agglutination, complement activation, opsonization and phagocytosis. During influenza infection, human MBL can bind to both HA and NA on virus and thereby neutralize and contain the viral spread (Rase *et al.*, 1999; Gupta and Surolia, 2007). Cells such as macrophages and natural killer cells constitute the cellular

compartment of innate immune system. Influenza infected macrophages show an increased transcription level for an array of anti-viral cytokines including IL-1 beta, IL-6, TNF-alpha and IFN α/β. They also produce chemokines such as RANTES, monocyte chemotactic protein-1 (MCP-1) and macrophage inflammatory protein-1 alpha (MIP-1α) which further attract more mononuclear cells to the site of infection to aid in viral clearance. As for Natural killer (NK) cells, they are a population of large granular lymphocytes that are particularly kill the infected cell via lysis method. They are found abundantly in blood, liver, spleen, lymph nodes and other non-lymphoid organs such as pregnant uterus. They circulate in the body as "resting" cells containing a pool of constitutively expressed high levels of IFN γ, perforin and granzyme B mRNA transcripts. Following activation, they produce an array of cytokines and cytolytic mediators (perforins, granzymes). Activated NK cells kill infected cells either by creating pores on membranes with perforins and granzymes or by activating apoptotic signalling in the infected cells (Stetson *et al.*, 2003; Fehniger *et al.*, 2007).

The adaptive immunity is an antigen-specific, cellular defense system that requires activation by the innate immunity to eliminate a microbial infection. The adaptive immune system predominantly comprises T and B lymphocytes. Once the influenza virus infects the upper respiratory tract, the infection can be controlled either by influenza-specific CD8+ T cells, or release of antibodies promoted by influenza-specific CD4+ T cells. CD8+ T cells induce apoptosis of cells on which they recognize influenza viral antigens presented by major histocompatibility complex (MHC) class I molecules (Legge and Braciale, 2003). CD4+ T lymphocytes are often referred to as helper T (Th) cells as their primary function is to assist B lymphocytes in producing antibodies against microbial infections. Influenza virus

infection activates DCs to produce IL-12 which in turn drives CD4+ T cells towards Th1 phenotype. During influenza virus infection, CD4+ T cells provide protection in a B cell dependent manner, although they themselves possess some cytotoxic (perforin) activity too. Neutralizing antibodies produced by B cells during an influenza infection have an important protective role and are therefore the main targets of vaccine induced immunity. Development of humoral immunity to an influenza infection consists of an early rise in IgM antibody titer followed by their affinity maturation and immunoglobulin class-switching to IgG, IgA and IgE antibodies (Baumgarth *et al.*, 1999; Brown *et al.*, 2006). A study carried out by Mozdzanowska *et al.* (2000) did show that mice that are deficient in B cell are 50 – 100 times more susceptible to a lethal influenza infection and they show a higher mortality rate when challenged with a pathogenic PR8 strain of influenza. Hence, this clearly shows the significant contributions of B cells and anti-viral antibodies in recovery from an influenza infection.

1.7 Current Influenza Vaccine

The ideal measure to control influenza infections is to prevent infection by the application of influenza vaccine especially from strains that have become antiviral resistant. Antibodies have been demonstrated to be the major mediator of protection in immunocompetent hosts and current influenza vaccines are designed mainly to elicit these humoral responses (Gerhard, 2001). The surface glycoproteins HA and NA are the primary target of human immune response. Antibodies against the HA will neutralise the virus infectivity, and antibodies against the NA can reduce the severity and spread of disease (Nichol, 2003). Influenza strains circulating throughout the world are continuously being monitored by scientists and each year

the WHO meets and recommends the specific strains that should be included in that season's vaccine. There are currently two types of influenza vaccines licensed for human use in the U.S., the inactivated influenza vaccine and the live attenuated, cold adapted influenza vaccine (LAIV). Both of these vaccines are grown in eggs and are trivalent, containing A/H1N1, A/H3N2 and influenza B strains without any adjuvant. Both A/H1N1 and A/H3N2 in the inactivated influenza vaccine are high growth reassorted viruses that contain the HA and NA from selected wild type circulating strains, and the six internal genes from a laboratory-adapted parent strain A/Puerto Rico/8/34 that confers the virus to grow up to high yield in embryonated chicken eggs (Hoffmann *et al.*, 2002; Plotkin and Minor, 2010). There are four different types of inactivated influenza vaccines which including split, subunit, whole inactivated influenza vaccines and virosomal influenza vaccines. The inactivated vaccine is administered intramuscularly and is 70-90% effective in healthy individuals (Subbarao, 1999; Minne *et al.*, 2007). Inactivated vaccines do elicit strong humoral responses against the HA and NA proteins, they have been shown to be poor inducers of cytotoxic T lymphocytes (Gerhard, 2001). It is believed by many that this lacking of CTL response makes this vaccine ineffective at protecting against heterologous strains of influenza, especially when compared to the LAIV (Belshe *et al.*, 2000; Belshe *et al.*, 2007).

In contrast to the inactivated influenza vaccines which are administered by intramuscular or subcutaneous injection, LAIV is administered intranasally. LAIV is a 6/2 reassortant which contains the genes encoding the 6 internal segments from the attenuated donor strain (PB1, PB2, PA, M, NP and NS) and the 2 surface proteins (HA and NA) of the wild-type virus. The donor strain is cold-adapted by genetic modification which only allows it to grow in human nasal cavities but its temperature

sensitive causes it not able to stay in internal organs such as the lungs which have temperature more than 37°C. The underlying idea of vaccination with LAIV via the upper respiratory tract (nose) is to induce a secretory and systemic immune response that more closely resembles the immune response observed after natural infection. In other words, LAIVs induce a broad mucosal and systemic immune response. It has also been shown to be better at stimulating CTL responses, including IFN-γ and TNF-α responses, better at generating mucosal IgA responses, and to have a better capacity to interact with and be presented by antigen presenting cells (Gorse *et al.*, 1995; Muszkat *et al.*, 2003).

1.8 Universal Influenza Vaccine as Future Vaccine

Immunity to influenza that is mediated by antibody responses to HA and NA is very effective for antigenically matched strains, but can be ineffective against different subtypes or even within subtypes that have significant drift. Another surface protein in influenza, M2, has been the subject of recent vaccine strategies. M2 has a short ectodomain (M2e) that is exposed to antibodies as well as a long cytoplasmic tail (Zebedee and Lamb, 1988). However, due to the small size and low immunogenicity of M2e, therefore many studies have focused on M2e peptide fusion constructs using a variety of carrier molecules: hepatitis B virus core (De Filette *et al.*, 2006), keyhole limpet hemocyanin (Tompkins *et al.*, 2007), liposome (Ernst *et al.*, 2006) and flagellin (Huleatt *et al.*, 2008). Studies show that M2 vaccines based on M2e fusion carriers or DNA – recombinant vector combination could provide cross protection against lethal infection with different strains (Fan *et al.*, 2004; Tompkins *et al.*, 2007). These studies showed that M2e can be one of the candidates for universal vaccine development in future. Another internal protein known as

nucleoprotein (NP) was also highly conserved among all Influenza A virus. NP is an antigen with >90% protein sequence homology among influenza A isolates and containing dominant Cytotoxic T lymphocytes (CTL) target epitopes (Epstein *et al.*, 2005). One study did by Lo *et al.* (2008) on mice in order to evaluate the role of CTL in eliciting cross-protective immunity against all the subtypes and variants of influenza A virus. In their study, a live-attenuated H3N2 (A/Alaska/6/77) on a cold adapted H2N2 (A/Ann Arbor/6/60) background was compared to DNA prime-adenovirus boost vaccines expressing PR8 nucleoprotein with M2 (A/NP+M2). CTL responses was measured by IFN- ELISpot and the result showed that A/NP+M2 having a better response. On the other hand, they also challenge the mice with A/Vietnam/1203/04 (clade 1), similarly they found that A/NP+M2 protected much better than live-attenuated vaccine.

1.9 Problem Statements

The sudden emergence and explosive spreading of novel influenza A viruses in the community represents a major threat to global public health. Conventional influenza vaccines are mainly targeting two surface glycoproteins namely hemagglutinin (HA) and neuraminidase (NA) that undergo antigenic drift and antigenic shift continuously. Antigenic drift occur is due to the low fidelity RNA-dependent RNA polymerase, influenza A viruses have a high mutation rate ranging from $1x10^{-3}$ to $8x10^{-3}$ substitutions per site per year. Mutations that alter amino acids found in antigenic regions of HA and NA likely confer selective advantages via evasion of pre-existing immunity. On the other hand, antigenic shift occurred is due to the reassortment of segmented genome between different strains which allow influenza A escape from host immunity. Three biggest pandemics occur in 20^{th} century is due to this antigenic shift. Currently at least 6 months are needed to prepare and distribute the Influenza vaccine once a potential pandemic strain had been identified. Therefore, this poses extraordinary challenges to the virologist for strain identification and effective vaccine preparation during a pandemic. Hence, there is a pressing need to develop an effective universal vaccine that can confer immediate protection against different variants and subtypes of influenza A viruses. In recent years, nucleoprotein (NP) and extracellular domain of ion channel matrix protein (M2e) have been widely targeted as possible universal vaccine candidates since these two proteins are highly conserved among all influenza A viruses and some previous animal studies did show that combination of these two proteins in vaccine formulation greatly improve the immunogenicity of the vaccine.

1.10 Objectives

1.10.1 General Objective

In this study, we aimed to construct, amplify and clone synthetic M2e-NP gene that can serve as universal influenza vaccine in future.

1.10.2 Specific Objectives

1) To identify highly conserved regions and immunogenic epitopes for M2e and NP amino acid sequences by using *in silico* method.
2) To construct and amplify synthetic M2e-NP gene via assembly PCR.
3) To clone synthetic M2e-NP gene into pCR®2.1-TOPO® vector and confirm the sequence by sequencing.

CHAPTER 2

MATERIALS AND METHODS

Flow chart for methodology in this study was shown in Figure 2.1.

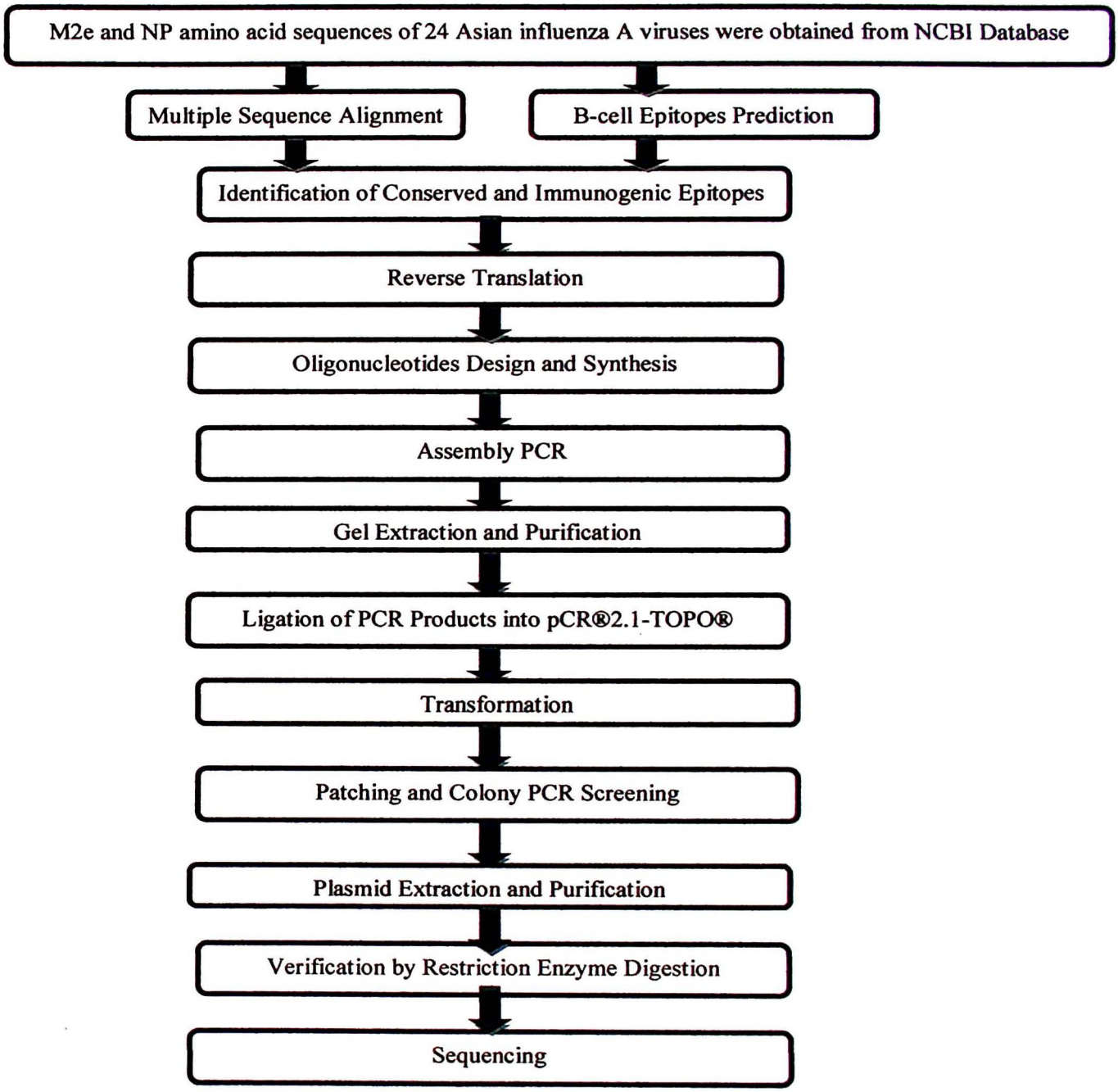

Figure 2.1 Flow chart of methodology in this study.