CHAPTER 2

LITTERATURES REVIEW

2.1 Avian influenza virus

Influenza A virus cause an infectious disease of poultry called “avian influenza”. The virus was firstly recognized in 1878 in Italy (Ligon, 2005). The virus was then isolated from chickens in 1902. In 1933, the first human influenza virus was isolated (Lamb and Takeda, 2001). Influenza viruses are classified into Orthomyxoviridae family (orthos means “standard, correct” and myxo means “mucus” in Greek). Orthomyxovirus particles are roughly spherical or filamentous with a diameter of approximately 100 nm. Curently, the Orthomyxoviridae family is divided into five genera; 1) influenza virus A, 2) influenza virus B, 3) influenza virus C, 4) ogotovirus, 5) isavirus. Influenza virus genome is single-stranded RNA with negative sense (that is the viral RNA is complementary to the messenger RNA) (Alexander, 2005 and 2007). These viruses are polymorphic particles covered with a lipid bilayer and several glycoproteins such as haemagglutinin (HA) and neuraminidase (NA) (Perez *et al.*, 2005).

A viral particle contain eight RNA segments (Figure 2.1). This segments encode 11 proteins: polymerase A protein (PA), polymerase B1 protein (PB1), polymerase B2 protein (PB2), Nucleoprotein (NP), haemagglutinin (HA), neuraminidase (NA), matrix protein (M1), M2, non-structural 1 protein (NS1), NS2 and polymerase B1-F2 protein (PB1-F2) (Swayne, 2006; Webster *et al.*, 1992). Influenza viruses are divided into types A, B, and C based on antigenicity of M1 and NP proteins (Chen and Deng, 2009; Suarez and Schultz-Cherry, 2000). Influenza type A viruses infect humans, horses, mink, pigs, marine mammals and birds. While influenza type B and C viruses infect only humans (Olsen *et al.*, 2006; Tamura *et al.*, 2005). Moreover, influenza type A viruses are further subtyped based on antigenicity of haemagglutinin (HA) and neuraminidase (NA) (Capua and Alexander, 2001; Swayne, 2008). At present, 16 haemagglutinin subtypes (H1-H16)

and 9 neuraminidase subtypes (N1-N9) are founded (Fouchier *et al.*, 2005). Each virus contains a combination of one H and one N antigen. Only H1, H2, H3, N1, N2 are found in humans (Yuen and Wong, 2005). In contrast, H1, H3, N1, N2 and N7 are found in pigs and H7N7 and H3N8 found in horses (Horimoto and Kawaoka, 2001).

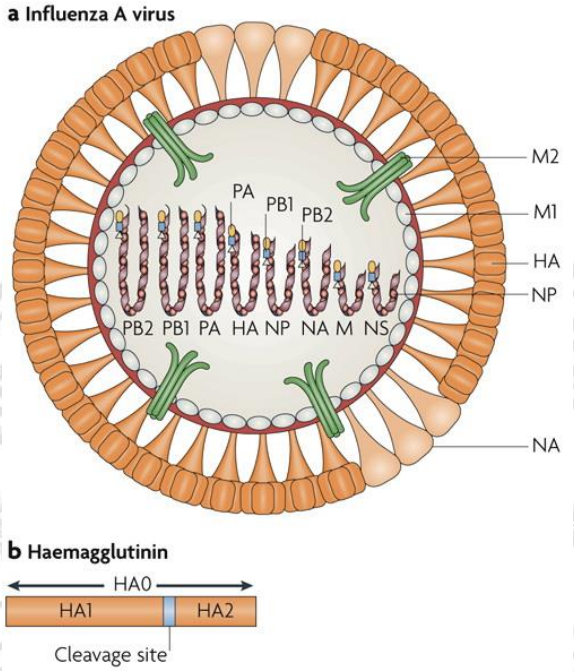

Figure 2.1 Schematic diagram of influenza A virus (Subbarao and Joseph, 2007).

On the basic of the the severity of clinical outcomes, avian influenza viruses are classified into highly pathogenic avian influenza (HPAI) or low pathogenic avian influenza (LPAI) (The World Organisation for Animal Health, 2009). HPAI kill at least 75% of susceptible chickens within 10 days post-inoculation (Swayne, 2008) while LPAI kill less than 75% (The World Organisation for Animal Health, 2009). Only H5 and H7 subtypes can cause HPAI (Capua and Alexander, 2001; Alexander, 2007a). However, H5 and H7 subtypes can be found as LPAI virus (Swayne, 2008).

2.1.1 Influenza virus proteins

2.1.1.1 Haemagglutinin (HA)

The haemagglutinin (HA) protein which binds and agglutinated a red blood cell, is a Type I integral membrane glycoprotein that is part of the viral lipid bilayer

surface architecture forming spikes-like structures. It can bind to host cell receptors by interacting with sialic acid moieties allowing a fusion steps between the host cell and the virion envelope occurred (Wilson *et al.,* 1981; Wilson and Cox, 1990). The HA protein is encoded by segment four of the viral genome with a protein size of 60 KDa prior to the post-translational modification. Next to the signal sequence of a newly synthesized HA is cleaved to remove an amino-terminal hydrophobic sequence of 14 to 18 amino acids, which are the signal sequence for transport to the cell membrane. The carbohydrates containing alpha-2-6-linked or alpha-2-3-linked sialic acid moieties are added, whose position and number vary with the virus strain. Palmitic acid is added to cysteine residues close to the HA carboxy terminus. In the final processing step, the HA was cleaved by host enzyme into HA1 and HA2 (Shehata *et al.,* 2012a). The virus-cell fusion is mediated by the free amino terminus of HA2 which required this cleavage for infectivity (Lewis 2006; Stevens *et al.,* 2006). The three-dimensional structure of a complete HA trimer has been determined by the fully processed HA consists of HA1 (typically) of about 324 amino acids plus variable carbohydrate, and HA2 (typically) of about 222 amino acids plus variable carbohydrate and 3 palmitate residues. Each HA molecule consists of a globular head on a stalk. The head is contains the receptor-binding cavity as well as most of the antigenic sites (epitopes) of the molecule which is made up exclusively of HA1. HA2 covers most of the stalk. It is coboxy-terminal which is hydrophobic consists of transmembrane region and cytoplasmic anchor site (Webster *et al.,* 1992). The HA mediates of virus is binding to a sialic acid residues on glycoproteins or glycolipids on the cell surface to entry the virus into host cells which partly accounts for the host specificity of the various influenza A virus subtypes. A great variability due to mutation in RNA by amino acid substitutions at several sites on the HA1 molecule is involved the epitopes in receptor-binding (Brown, 2000; Horváth *et al.,* 1998). The specificity and affinities of influenza viruses from different hosts can differ as the 2-3 or 2-6 linkage of N-acetyl neuraminyl-galactose (SA 2,3 Gal; SA 2,6 Gal). These binding specificities correspond to the types of sialic acid linkages within these hosts and therefore avian strains preferentially binds to sialic acids linked to galactose by an α-2,3 linkage, whereas human strains preferentially binds to sialic acids with an α-2,6 linkage (Wood *et al.,* 1993). Moreover, the influenza virus HA is subject to a very high rate of mutation due to the error-prone viral RNA polymerase activity. Apart from random mutation, the host

immune response also forced the amino acid substitution mutation. Because HA is one of main target for host immune system. Several recent reports demonstrated that HA and HA1 fragment containing the majority of antigenic determinants are responsible for generation of vaccines and virus-neutralizing antibodies based on conserved antigens provide broader protection (Kreijtz *et al.*, 2007; Shoji *et al.*, 2008; Tonegawa *et al.*, 2003). Therefore, the HA protein is the main target of the T - cell mediated immune responses (Horváth *et al.*, 1998) and protective humoral immunity by neutralizing antibody (De Jong and Hien, 2006; Stevens *et al.*, 2006).

2.1.1.2 Neuraminidase (NA)

NA is a second major surface antigen of the virion and also a type II integral glycoprotein of the membrane (Mitnaul *et al.*, 2000), which is involved in budding of new virions from infected cells. NA cleaves terminal sialic acid from glycolipids or glycoprotein. Thus, it is functions to be important in the final stages of release of new virus particles from infected cells or host cell receptors, prevent the new virus particles agglutinating, thus increasing the number of free virus particle and hence permit progeny virions to escape from original site of infection in which they arose (De Jong and Hien, 2006). NA is highly mutate capable similar to HA. This is partly due to avoid host immune system. NA have been identified to nine subtypes in nature; they are not serologically cross-reactive. The several subtypes which different variants are known (Brown, 2000; Webster *et al.*, 1992).

2.1.1.3 The others protein

The Apart from HA and NA, the virus also has six segments encoding polypeptide that are essential and their respective functions are summarized in Table 2.1.

Table 2.1 Another protein of avian influenza segments and their encoded polypeptides

Segment	Protein name	Encoded polypeptide	Length (nt)	Functions
PA	Polymerase acidic	PA	2233	Component of RNA polymerase, cap recognition
PB1	Polymerase basic1	PB1, PB1-F2	2341	Component of RNA polymerase, endonuclease activity, elongation
PB2	Polymerase basic2	PB2	2341	Component of RNA polymerase, cap recognition
M	Matrix	M1, M2	1027	M1: interaction with vRNPs and surface glycoprotein, nuclear export, budding M2: ion channel activity, assembly
NP	Nucleoprotein	NP	1565	RNA binding, RNA synthesis, RNA nuclear import
NS	Nonstructural	NS1, NS2 (NEP)	890	NS1: multi-functional, viral IFN antagonist NS2/NEP: nuclear export of vRNPs

2.1.2 Influenza virus receptors and replication

Influenza virus particles bind to neuraminic acid (sialic acid, SA) containing cellular receptors on the surface of host cells via the viral HA glycoprotein (the alpha-2-3-linked and/or alpha-2-6-linked sialic acid carbohydrates) to initiate infection and replication (Figure 2.2). Virions penetrate into the cell membrane by a clathrin-dependent (Connor *et al.,* 1994) and independent endocytosis (Sieczkarsk and Whittaker, 2002). On binding at the cell surface, the haemagglutinin are endocytosed and the cellular protease reacts with haemagglutinin and reveal the fusion between the endosomal membrane and the viral envelope. The virus are further activated by the maturation of the endosome. During the endocytosis the matrix protein 2 ion channels allows influx of protons leading to an acidic environment inside of the virion particle and un-coating the RNP complex from the viral genome (Matlin *et al.,* 1981). The virion genome are then cytoplasmic release which originated by the pH-dependent fusion of haemagglutinin to the endosomal membrane (Medina and García-Sastre, 2011). The genetic information is then actively transported into the nucleus through nuclear pores of the cell where the viral RNA-dependent RNA polymerase initially transcribes a series of RNA molecules into the negative-sense RNA molecules (mRNA) that are necessary templates for packaging co-assembly with the viral proteins (Martin and Heleniust, 1991). The viral proteins from mRNA (complementary, positive sense RNA segments) are transported to the cytoplasm and translated to create the viral protein by host cellular machinery (Lamb and Choppin, 1976) which used in the synthesis of the negative sense small viral RNA (svRNAs) involved in regulating the switch between the replication and transcription processes of the viral messenger RNA in the cytoplasm (Medina and García-Sastre, 2011; Perez *et al.,* 2010). The M1 and NP proteins return to the nucleus and interact with the negative sense vRNA after the translation processes of the mRNAs has started (Medina and García-Sastre, 2011; Chenavas *et al.,* 2013). The viral proteins are produced in the nucleus, can be post-translational processed in the Golgi apparatus, and directed towards the plasma membrane (HA, NA, and M2) where M1 reacts to assemble the virus particle (Medina and García-Sastre, 2011; Roberts *et al.,* 1998). The virion particle release the membrane-encapsulated from the apical plasma membrane via facilitation of the neuraminidase

enzymatic activity which cleaves the sialic acid moieties (Varghese and Colman, 1991). Finally, the virus particles are bud and released.

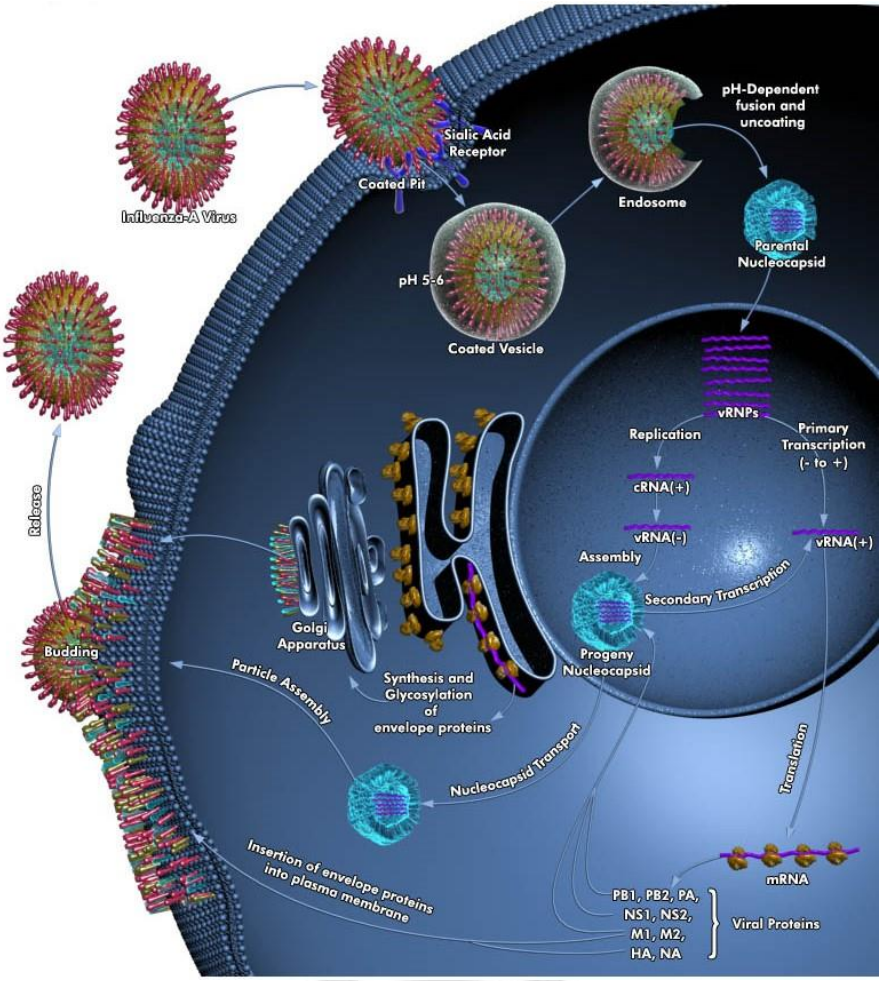

Figure 2.2 Schematic diagram of influenza virus replication (https://www.qiagen.com/)

2.1.3 Antigenic shift and drift

Antigenic changes of influenza virus can be described into 2 patterns:

1) Antigenic drift: due to the lack of a proof reading system for the RNA polymerases, point mutations in the HA and NA genes can be occurred. This allows HA and NA glycoproteins gradually and continuously undergo minor antigenic changes.

2) Antigenic shift: due to the segmented nature of the influenza viral genome, the exchange of RNA segments between two subtypes of influenza viruses infecting the same host cell may result in the emergence of a novel subtype. This exchange is called “genetic reassortment”, which can lead to major antigenic changes in the HA or NA genes. The novel subtype/strain(s) increased human infectivity and may lead to pandemics of influenza (Bouvier and Palese, 2008; Shoham, 2006; Tamura *et al.*, 2005; Webster et *al.*, 1992).

2.1.4 Human Infections by Avian Influenza H5N1

Avirulent avian influenza manifests as mild respiratory signs, depression, diarrhoea or decreased egg production, while highly pathogenic avian influenza results in systemic symptoms (depression, listlessness, ruffled feathers, cessation of egg laying, excessive lacrimation, oedema, cyanosis, diarrhoea and nervous system disorders) (Horimoto and Kawaoka, 2001). Highly pathogenic avian influenza mainly comprises H5 and H7 (Horimoto and Kawaoka, 2001).

Human influenza in adults and adolescents classically present with abrupt onset of fever, chills, myalgia, malaise, anorexia, headache, sore throat and a dry cough. Fever peaks within 24 hours and lasts 1 to 5 days (Cox and Subbarao, 1999). In children, a non-specific febrile illness, croup, bronchiolitis, bronchitis, gastrointestinal complaints, and febrile convulsions may also occur (Cox and Subbarao, 1999). The incubation period of human influenza is 1 to 4 days (Cox and Subbarao, 1999). Virus is detectable in respiratory secretions within 24 hours of the onset of illness, peaks for 24 to 72 hours and falls to low values by the fifth day (Nicholson *et al.*, 2003). This asymptomatic viral shedding may contribute to disease transmission, but it is associated with low viral titres. Viral shedding may persist for weeks in young children (Nicholson *et al.*, 2003). Human influenza virus can be recovered from children up to 6 days before and 21 days after the onset of symptoms (Liem *et al.*, 2005).

2.1.5 Epidemiology of Human Cases of Avian Influenza H5N1

Historical outbreaks demonstated that 70-100% of influenza cases had had exposure to ill poultry (Beigel *et al.*, 2005). Some reports suggested a transmission from human to human with or without illness (Ungchusak *et al.*, 2005; Katz *et al.*, 1999; Bridges *et al.*, 2000). Hence, during the pre-pandemic phase, eliciting risk factors of exposure to poultry and known H5N1 cases should be undertaken for diagnosing suspected cases.

2.1.6 Laboratory Diagnosis

Laboratory confirmation of H5N1 influenza virus may be used to detect (and initially identify) both seasonal and novel influenza A viruses in specimens from humans but it requires a high index of multiple suspicious specimens and very sensity diagnosis technique (e.g., reverse transcriptase PCR [RT-PCR]) (Kandun *et al.*, 2006; Oner *et al.*, 2006). There are several detection methods to detect influenza virus in clinical specimens for example as virus culture, detection of viral nucleic acids (RT-PCR), detection of antigen, and detection of rising titers of antibodies. Moreover, the rapid test to distinguish influenza subtypes is necessary in diagnosis, appropriate therapy, control of infection, and epidemiological study of avian influenza disease.

2.1.6.1 Virus isolation

The respiratory and enteric epithelium of avian species are the majority of avian influenza virus replication in to host cell receptor. Therefore, swabs and tissues from these systems are the preferred samples. Samples taken from avian species should include intestinal contents (faeces) or cloacal swabs and oropharyngeal swabs. Swabs should be collected into suitable sterile transport media containing isotonic phosphate-buffered saline (PBS), pH 7.0–7.4 with antibiotics. Samples can be stored at 4°C for up to 48 hours after collection. If they are unable to be cultured, storage it at -70°C is recommended.

The preferred methods which used to isolate the avian influenza viruses are the inoculation of specific antibody negative (SAN) eggs, or specific pathogen free (SPF) embryonated chicken eggs, or cell culture such as Madin Darby canine kidney (MDCK) cells or rhesus monkey kidney (LLC-MK2) cells. Positive influenza cultures from these methods may or may not exhibit cytopathic effects, thus the presence of avian influenza viruses should be identification by immunofluorescent staining with monoclonal antibodies against the nucleoprotein, haemagglutination-inhibition (HI) or RT-PCR.

2.1.6.2 Antigen detection

Viral antigen detection may be carried out by direct immunofluorescence or enzyme immunoassay (EIA) which are widely used for detection of viral antigens in clinical specimens and diagnosis of human influenza virus because they could be operated as point-of-care tests and convenient and simple to use. However, such testing is directed at conserved viral antigens (e.g., matrix protein and nucleoprotein) and detect all subtypes of influenza A viruses, whether of human or avian origin. Therefore these tests will not differentiate human from avian influenza virus subtypes. Moreover, the disadvantages are that they may be sensitivity for the diagnosis of avian influenza H5N1 virus (i.e. a negative result does not exclude H5N1 disease). They may not have been validated for different species of birds. So subtype identification is not achieved. (Beigel *et al.*, 2005; Kandun *et al.*, 2006; Oner *et al.*, 2006, Peiris *et al.*, 2004; Yuen *et al.*, 1998). Therefore, the antigen detection is mainly used for field screening of high mortality clinical cases for influenza A followed by confirmation of results using a subtype-specific diagnostic methods such as RT-PCR or virus cultures, to differentiate avian from human influenza virus.

2.1.6.3 RT-PCR

RT-PCR assays can be used to detect viruses irrespective of antigenic differences subtypes of influenza A viruses by targeted at genes (e.g., matrix gene) that are highly conserved and, separately, at the HA or NA genes to identify specific influenza A virus subtypes. Moreover, real-time RT-PCR is favored because results can be obtained

more rapidly (around 4 to 6 h), increases sensitivity and specificity by the use of probes, and enables the quantitation of the viral target gene (De Jong *et al.*, 2005).

2.1.6.4 Antibody detection

The haemagglutination inhibition (HI) assay is the method of choice to detect subtype-specific antibodies to human seasonal influenza viruses in human sera which useful for epidemiological investigations. But this technique is low sensitivity, therefore it has less value in detecting antibodies against avian influenza viruses in specimens (Beare and Webster, 1991; Kida *et al.*, 1994; Lu *et al.*, 1982, Rowe *et al.*, 1999). Therefore, the neutralization assay is the preferred method to detect H5-specific antibodies in humans which could detected antibodies against H5N1 virus within 14 or more days after the onset of symptoms in patients infected with avian flu (Katz *et al.*, 1999).

2.2 *Pichia pastoris*

Pichia pastoris is a methylotrophic yeast (Figure 2.3) that belong to the class Hemoascomycetes, family Saccharomycetoideae. The genus *Pichia* is characterized by budding cells, hat-shaped or saturn-shaped ascospores (Rose and Harrison, 1987).

Figure 2.3 *Pichia pastoris* (http://www.corbisimages.com)

P. pastoris is a useful tool for both basic laboratory research and industrial manufacture as it can express huge quantities (milligram-to-gram) of proteins. The fermentation can be readily scaled up to meet greater demands. Also, parameters influencing protein productivity and activity, such as pH, aeration and carbon source feed rate, can be controlled (Higgins and Cregg, 1998). *P. pastoris* has a eukaryotic protein synthesis pathway. It requires a simple growth medium and easy to manipulate and cultured when compared with mammalian cells. Moreover, *P. pastoris* is suitable for foreign protein expression by a several reasons, including high levels of intra- or extracellular protein expression, and the ability to manipulate post translation modifications, such as glycosylation, disulphide bond formation and proteolytic processing (Cregg *et al.*, 2000). Moreover, integration of expression vectors into host genome preventing the loss of recombinant elements during fermentation process (Romanos, 1998). Purification of secreted recombinant proteins is simple because the yeast secretes a low levels of native proteins (Cregg *et al.*, 1993). Therefore, *P. pastoris* is a system of choice for protein expression.

2.2.1 Methanol metabolism in *Pichia pastoris*

Only four genera, i.e. *Hansenula*, *Pichia*, *Candida* and *Torulopsis* of yeasts which identified to date, are able to utilize methanol as the sole carbon and energy source (Faber *et al.*, 1995). The methanol metabolic pathway is similar in all 4 genus with the unique set of enzyme (Higgings and Cregg, 1998). In *P. pastoris*, alcohol oxidase (AOX, EC 1.1.3.13) is the first enzyme in the methanol utilization pathway. For the methanol utilization pathway, the oxidation of methanol to formaldehyde and generating hydrogen peroxide in the process as shown in Figure 2.4. A by-product, hydrogen peroxide, is subsequently degraded to water and oxygen by a second peroxisomal enzyme, catalase (Cat). The formaldehyde generated by AOX follows one of two paths. A portion leaves the peroxisome and is further oxidezed by two cytoplasmic enzyme, formaldehyde dehydrogenase (Fld) and formate dehydrogenase (Fdh), to generate energy for the cell. (Lin-Cereghino *et al.*, 2005) The remaining formaldehyde is assimilated to form cellular constituents by a cyclic pathway that starts with the condensation of formaldehyde with xylulose 5-monophosphate by a third peroxisomal enzyme dihydroxyacetone synthase

(DHAS). The products of this reaction, glyceraldehyde 3-phosphate and dihydroxyacetone, leave the peroxisome and enter a cytoplasmic pathway that regenerates xylulose 5-monophosphate and produces one net molecule of glyceraldehyde-3-phosphate for every three turns of this cycle (Cereghino and Cregg, 2000).

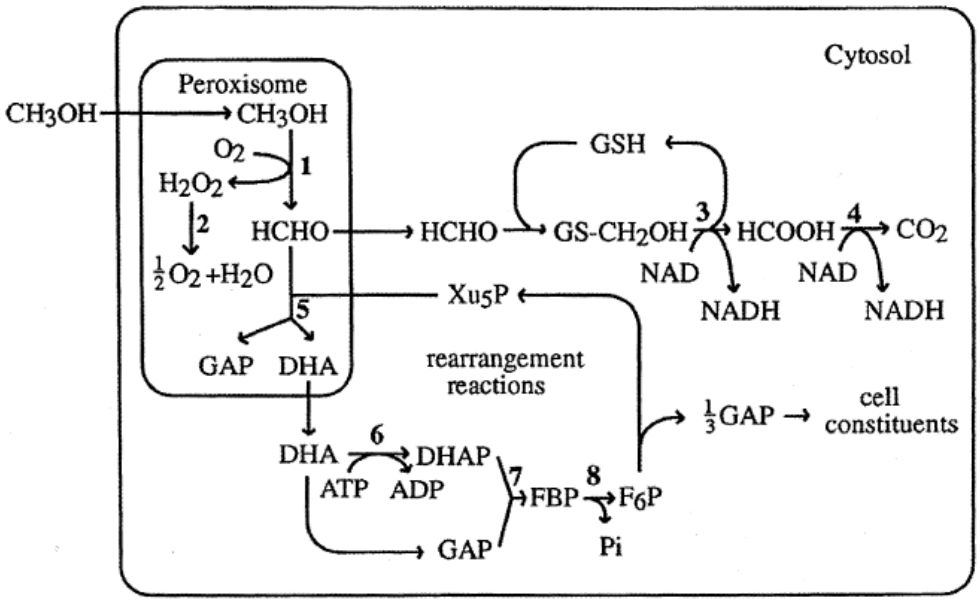

Figure 2.4 The methanol pathway in *Pichia pastoris*. 1,alcohol osidase; 2,catalase; 3,formaldehyde dehydrogenase; 4,formate dehydrogenase; 5,dihydroxyacetone synthase; 6,dihydroxyacetone kinase; 7,fructose 1,6-biphosphate aldolase; 8,fructose 1,6-biphosphatase (Cereghino and Cregg, 2000).

2.2.2 The *Pichia* expression systems

The foreign gene expression in *Pichia pastoris* contains three steps: first, insertion of the interested gene into an expression vector; second, introduction of the expression vector into yeast genome; and third, determination of the expressed protein.

2.2.2.1 Expression vector

The The plasmid vectors designed for heterologous protein expression in *Pichia pastoris* have several common features. All expression vectors have been designed as *Escherichia coli/P. pastoris* shuttle vectors (Figure 2.5), containing an origin of

replication for plasmid maintenance in *E. coli* and markers functional in one or both organisms. The vector is composed of DNA sequences containing the *P. pastoris* AOX1 promoter and a multiple cloning sites (MCS) for insertion of interested gene, and followed by the transcriptional termination sequence (for mRNAs polyadenylation) of the AOX1 gene (Cereghino and Cregg, 2000; Higgings and Creggs, 1998). For foreign proteins secretion, the vectors must contain a secretion signal of *P. pastoris* acid phosphatase gene (PHO1) or *Sacharomyces cerevisiae* α factor prepro signal sequence in the promotor.

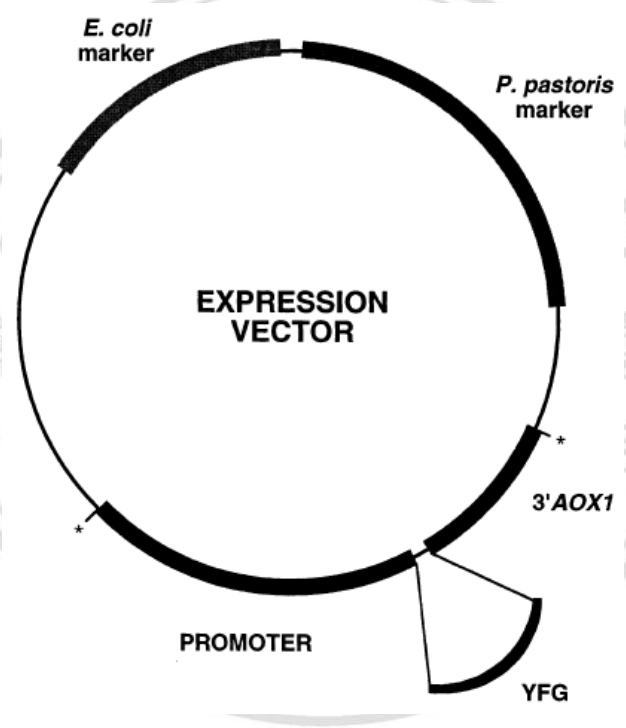

Figure 2.5 General diagram of *P. pastoris* expression vector (Cereghino and Cregg, 2000)

To obtain the best expression, it is suggested toplace the ATG of interasted gene to the position of the AOX1 ATG. This position concurs with the first restriction site in most MCS (Cereghino and Cregg, 2000). Vectors with dominant drug resistance markers that allow for enrichment of strains that receive multiple copies of foreign gene expression cassettes during transformations have been developed (Cereghino and Cregg, 2000; Higgings and Cregg, 1998).

The first generation of *P. pastoris* expression vectors (such as pHIL-D2 or pPIC9) used the functional histidene dehydrogenase gene (HIS4) as a selectable but this does not allow direct selection of multi-copy integrants and also the gene size itself is large. Therefore, this lead to the second generation of vectors which based on the ability or the *Sh ble* gene product that confer zeocin resistance in both bacteria and yeast (Figure 2.6).

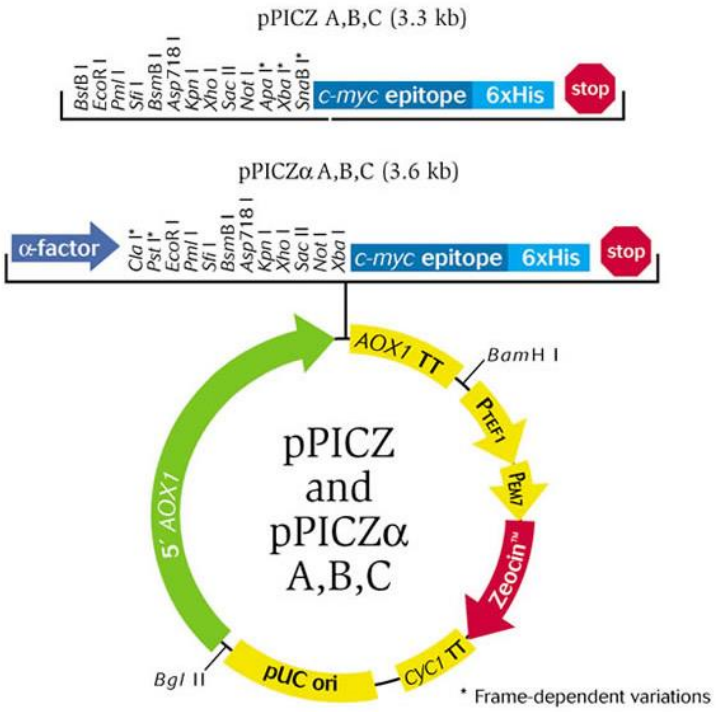

Figure 2.6 Vector diagram of pPICZ and pPICZα (Invitrogen Corp., 1998)

These zeocin-resistant vectors contain the 5'AOX1 promoter region, the AOX1 transcription terminator and the *Sh ble* gene from *Streptoalloteichus hindustanus*. These vectors also contain a multiple cloning site (MCS) with several unique restriction sites for foreign gene insertion and sequences encoding the polyhistidine (His6) and *myc* epitopes for facilitating the detection and purification of a recombination fusion protein (Higgings and Cregg, 1998). This vector allow high-level, methanol inducible expression of the gene of interest in *Pichia*, and can be used in any *P. pastoris* strain including X-33, GS115, SMD1168H, and KM71H (Invitrogen corp, 1998). Furthermore, selection of zeocin-resistant transformants at high zeocin concentration can be useful for the heterologous protein production (Higgings and Cregg, 1998).

2.2.2.2 The promoter in *Pichia pastoris*

In *Pichia pastoris*, the interested gene can be expressed under the control of inducible and constitutive promoter.

Inducible promoter

In the wild-type *P. pastoris* strain, AOX1 promoter is one of the strongest promoter and widely used for heterologous protein expression. (Cereghino and Cregg, 2000; Damasceno *et. al.*, 2004). There are two copies of the alcohol oxidase (AOX) gene in the genome of *P. pastoris*, called AOX1 and AOX2. AOX1 (alcohol oxidase 1) promoter controls the expression of the first enzyme in methanol metabolism and is responsible for a vast majority of alcohol oxidase activity in the cell (Cereghino and Cregg, 2000). With this promoter, expression of recombinant proteins is highly repressed while cultures are grown to high density in glucose or glycerol. Cultures are shifted to a methanol medium to induce rapid high-level expression (Cereghino and Cregg, 1999).

Constitutive promoter

The AOX1 promoter has been proofed successfully into expressing several genes but it is not suitable in some situations. For example, high quantities of methanol used in large scale fermentation is potential fire hazard. Therefore, promoters that are not induced by methanol for expression of certain genes are need (Daly *et al.*, 2005). The *P. pastoris* glyceraldehyde 3-phosphate dehydrogenase (GAP) gene promoter provides strong constitutive expression on glucose at a level comparable to the AOX1 promoter (Hans *et al.*, 1997; Waterham *et al.*, 1997). This gene was isolated. The GAP promoters have been tested with various carbon sources such as glucose, glycerol, oleic acid and methanol, the result shown that induction with glucose provide a highest expression levels. Furthermore, in comparison with AOX1 promoter, some expressed protein was produced at highest levels under the control of the GAP promoter (Waterham *et al.*, 1997). Therefore, GAP promoters makes the cells to be grown without the need of methanol induction and is particularly attractive for large-scale protein production.

Other promoters are the constitutive YPT1 and PEX8 promoters. The YPT1 promoters is constructed from a GTPase gene product related in the secretory pathway. Whereas, the *P. pastoris* PEX8 gene encodes a peroxisomal matrix protein that is essential for peroxisome biogenesis. These promoters provide a low but constitutive level of expression in media containing glucose, methanol, or mannitol as carbon sources. (Cereghino and Creggs, 2000).

2.2.2.3 Transformation and integration into the *Pichia pastoris* genome

Vector transformation into *Pichia pastoris cells* can be performed in several ways such as spheroplast formation, electroporation, PEG1000 and lithium chloride treatments. In general, spheroplasting and electroporation provide the highest efficiency of transformation (approximately 10^3 to 10^4 transformants per µg DNA). However, the spheroplast transformation is bothered by the risk of contamination, reduction of cell viability causing by enzyme over-digestion. Electroporation has become widely use in laboratory research as it requires fewer steps, reduced risk of contamination (Daly *et al.*, 2005). In case of people who do not have an electroporation device, a Polythylene glycol (PEG) method is another alternative method even this is not as good as spheroplasting or electroporation (Invitrogen crop., 1998).

There are two approches of integration of the expression vector into *P. pastoris* at specific locus of chromosome for genetically stability (Daly *et al.*, 2005). First, the simplest way is to insert the vector at a unique site in either the maker gene (e.g., HIS4) or the AOX1 promoter fragment and then the vector will integrate into the yeast genome as presented in Figure 2.7 (Cereghino and Cregg, 2000).

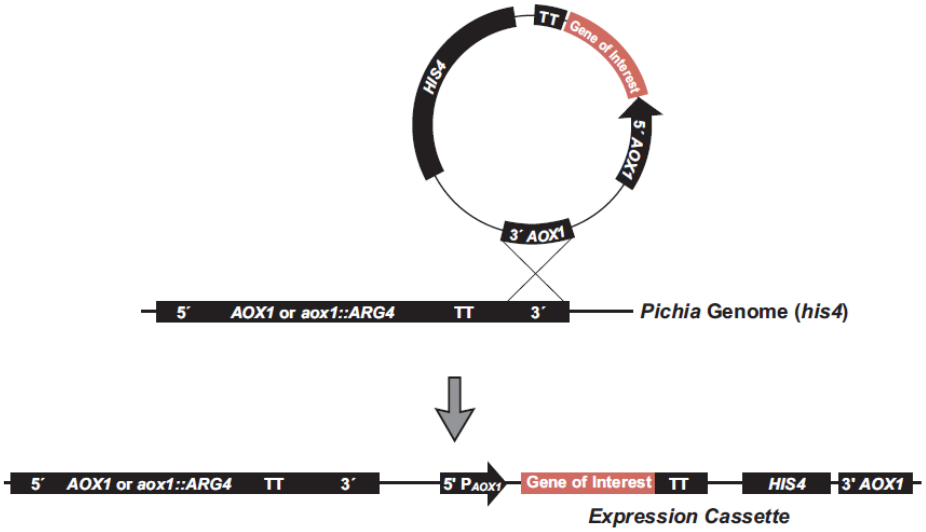

Figure 2.7 Gene insertions in *Pichia pastoris* (Invitrogen Corp., 1998).

Second approche, the expression vectors can be digested in linear form, which is flanked by 5' and 3' AOX1 sequences. This linear form will then replace the AOX1 gene of *Pichia* genome (Cereghino and Cregg, 2000) as presented in Figure 2.8.

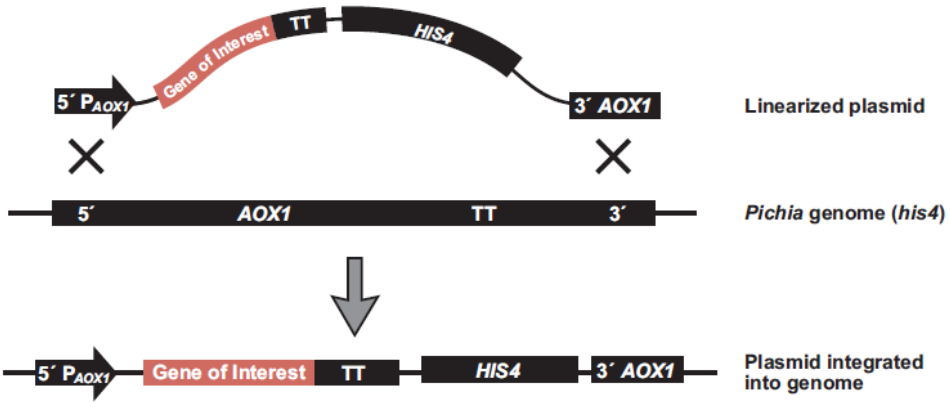

Figure 2.8 Gene replacements in *Pichia pastoris* (Invitrogen Corp., 1998).

Integration by gene insertion (either at the AOX1 or his4 loci) can result in multiple gene insertion events due to repeated recombination events at a rate of 1-10% of transformants. The number of integrated copies of the expression cassette can affect the amount of protein expressed (Daly *et al.*, 2005).

2.2.2.4 Intracellular and secretory protein expression

Heterologous protein expression in *Pichia pastoris* can be produced either secretion or intracellular expression. This choice depends on the protein to be expressed and the objective of the study.

Secretion signals

Protein secretion needs a signal sequence to target protein to the secretory pathway. Several different secretion signal sequences have been used successfully (Higgings and Cregg, 1998). In *P. pastoris* secretion systems, the most widely used signal sequence is the *S. cerevisiae* α-mating factor pre-pro leader sequence (α-MF) (Raemaeker *et al.*, 1999). The major advantages of using *P. pastoris* system are that the yeast secretes very low levels of native proteins and the vast majority of the secreted protein can be obtained from the culture medium (Invitrogen Crop., 1998).

Intracellular expression

If glycosylation or other post-translational modification processes are not desirable, intracellular expression is an alternative. However, purification of intracellular protein is much more difficult than of secreted protein because the low representation (less than 1%) of expressed protein in yeast cell compared to total intercellular protein.

An advantage of intracellular protein expression in *P. pastoris* is that an amino-terminal methionine residue is cleaved. Another advantage is that the terminal amino acid of proteins expressed in *P. pastoris* can also be acetylated by N-acetyl-transferase. Indeed, there are many proteins successfully produced in *P. pastoris* using this intracellular expression systems, such as the hepatitis B surface antigen (Daly *et al.*, 2005).

2.2.2.5 Posttranslational modification

Pichia pastoris expression system has many post-translational modification processes like higher eukaryotes for example correct folding, disulphide bond formation,

O- and N -linked glycosylation and processing of signal sequences. The production of foreign proteins from *Pichia* is limited by correct folding and disulphide bond formation. Therefore, it is difficult to optimize production of recombinant proteins unless these rate-limiting factors can be identified.

Glycosylation is one of the most common and complex process during post-translational modifications in eukaryotic cells (Eckart and Bussineau, 1996). It had been shown that the correct glycosylation patterns on recombinant proteins enhance their biological activity. Although both of *Sacharomyces cerevisiae* and *P. pastoris* are able to introduce eukaryotic post-translational modifications in protein, but the chain length of proteins expressed in *P. pastoris* have significantly shorter glycosyl chains than those expressed in *S. cerevisiae,* which making *P. pastoris* is an interesting alternative for the recombitant protein expression (Bretthauer and Castellino, 1999).

In mammals, *O*-linked oligosaccharides have a more varied sugar composition including N-acetylgalactosamine, galactose and sialic acid. In contrast, lower eukaryotes, including *P. pastoris*, add *O*-oligosaccharides solely composed of mannose (Man) residues to the hydroxyl groups of serine and threonine of secreted proteins. It is possible that *Pichia* will glycosylate heterologous proteins, even when those proteins are not normally glycosylated by the native host; and even when the protein is glycosylated in the native host, *Pichia* may not glycosylate it on the same serine and threonine residues (Cereghino and Cregg, 2000).

N-linked glycosylation glycosylation in fungi and yeasts, such as *Pichia*, is also different than in higher eukaryots. In all higher eukaryotes, it begins in the cytoplasmic side of the endoplasmic reticulum (ER) with the transfer of a lipid-linked oligosaccharide unit, $Glc_3Man_9GlcNAc_2$ (Glc=glucose; GlcNAc=N-acetylglucosamine), to asparagine at the recognition sequence Asn–Xaa–Ser/Thr (Charlwood *et al.*, 2001). This oligosaccharide core unit is subsequently trimmed to $Man_8GlcNAc_2$. In mammalian cell, golgi apparatus further trim and add oligosaccharides either $Man_{5-6}GlcNAc_2$ (high-mannose type), a mixture of several different sugars (complex type) or a combination of both (hybrid type), and are represented in Figure 2.9.

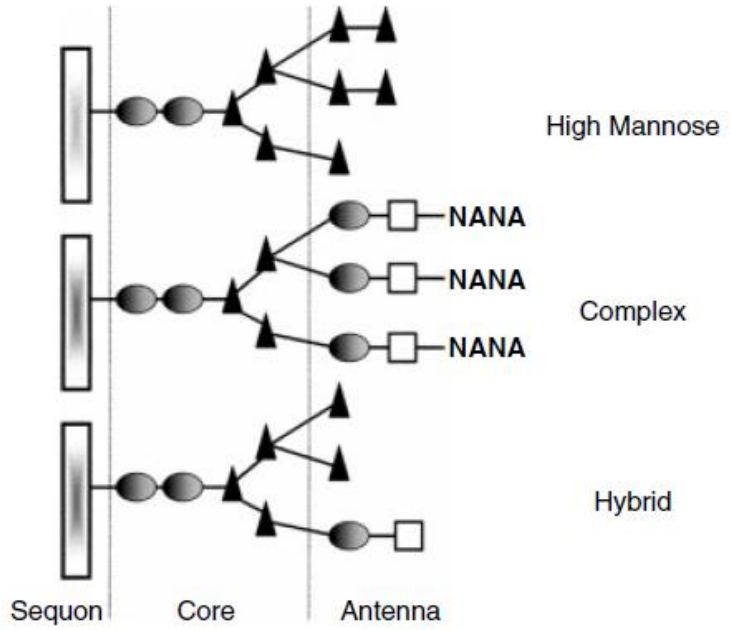

Figure 2.9 Protein glycosylation patterns. Oligosaccharide chains consist of a sequon, which is attached to the core by an Asn residue. The core is attached to the antenna, which can be high-mannose, complex or hybrid. GlcNAc,● ; mannose,▲; galactose, □ (Goochee *et al*, 1991).

2.2.2.6 Process of *Pichia pastoris* fermentation

The culture conditions are important factors for *Pichia pastoris* expression systems. This will help to emprove the productivity of the target proteins. At first stage, optimization of culture condition is employs by shake-flask, small-scale expression methods. The higher product yields for recombinant proteins will be resulted by developing a model for methanol utilization and then using this model for a lage scale protein production (Daly *et al*., 2005).

The fermentation growth is especially important for secreted proteins, as the concentration of product in the medium is roughly proportional to the concentration of cells in culture. *P. pastoris* can be cultured at extremely high cell densities (>100 g/l dry cell weight; >400 g/l wet cell weight; >500 OD600 units/ml) in the fermenter where parameters such as pH, aeration and carbon source feed rate can be controlled. The level

of transcription initiated from the AOX1 promoter can be 3-5 times greater in cells fed with methanol at growth-limiting rates compared to cells grown in excess methanol. (Cereghino *et al.*, 2002).

2.2.3 Application of *Pichia pastoris*

Nowadays, the species *Pichia pastoris* is extensively used for the expression of heterologous proteins. The protein produced from *P. pastoris* are usually properly folded and has no toxic cell wall pyrogens as *E. coli*; nor contains potentially oncogenic or viral nucleic acids as mammalian cells (Freyre *et al.*, 2000).

2.2.3.1 Therapeutics application

Pichia pastoris has been used widely for the production of therapeutically relevant macromolecules. The system has gained interest compared with *Saccharomyces cerevisiae*, and is now available in a commercial kit (Invitrogen, San Diego, CA, U.S.A.). The cattle tick Bm86 antigen has been expressed in *Pichia* for manufacturing a vaccine based on this antigen and a certified production and downstream processing protocol has been described (Canales *et al.*, 1997). HV-2, a variant of hirudin, a blood-coagulation inhibitor, has been produced to high yield (1.5 g/l) as a secreted protein in *Pichia* (Rosenfeld *et al.*, 1996) as well as a cytokine consisting of a fusion of an interleukin with a soluble form of its receptor. Several antibody fragments have been expressed at high yield in *P. pastoris* (Fischer *et al.*, 1999). Recently reported, *P. pastoris* is used to produce basic fibroblast growth factor (bFGF) which is a potent angiogenic molecule stimulates smooth muscle cell growth, wound healing, and tissue repair (Mu *et al.*, 2008). Production of recombinant N-A1 to develop a tumor marker for adenocarcinomas and as a target for antibody-directed therapeutics (Sainz-Pastor *et al.*, 2006). In addition, for development of vaccine, the haemagglutinin (HA) gene of highly pathogenic avian influenza H5N1 was cloned and expressed as His tagged protein in *P. pastoris*. After immunization with this rHA, an antigen specific immune response was observed. This indicated that the rHA could be further used as a vaccine candidate against avian influenza (Murugan *et al.*, 2013).

2.2.3.2 Enzymes production

Enzyme for blood group conversion such as erythrocytes from blood group B harbor oligosaccharide chains with terminal β-galactose residues. For conversion, β-galactosidases from different sources can be used. A *Pichia pastoris*-based production has been established for two enzymes from coffee beans (Zhu *et al.*, 1995) and from soybeans (Davis *et al*., 1996), respectively. Moreover, the methylotrophic yeast species have been developed as high-yield production systems for recombinant amylases and sugar-converting enzymes from different sources. An examples of this class of proteins produced in *P. pastoris* are α-amylases 1 and 2 which isolated from barley shown a similar structure and function to that isolates from malt extracts (Juge *et al.*, 1996). The recombinant production of the spinach phosphoribulokinase in *P. pastoris* provided the possibility to explore its function (Brandes *et al.*, 1996).

2.2.3.3 Recombinant methylotrophs as a biocatalyst

Pyruvic acid is a chemical intermediate for chemical, pharmaceutical industrials and can be used as dietary supplement. Genetically modified strains of *Pichia pastoris*, that express both hydroxyacid oxidase enzymes from spinach and endogenous catalase have been used as whole-cell biocatalysts for conversion of L-lactic acid to pyruvic acid (Gough *et al.*, 2005).

2.2.3.4 Methanol removing

Methanol is a ubiquitously present in nature and many industrial waste streams. Biological treatment with methylotrophic yeast is an alternative means to remove the methanol fraction. It is a much-reduced energy cost and complete removal of the methanol but it is slow and complex to process. It also allows the potential to generate a valuable product (Hoy *et al.*, 2006).