CLONING AND OVEREXPRESSION OF INFLUENZA A H1N1 NS1 PROTEIN IN *Escherichia coli*

ONG LIH LIH

A dissertation submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Science (Biotechnology)

Faculty of Biosciences and Bioengineering
Universiti Teknologi Malaysia

DEC 2010

To my dearest parents and siblings

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Chan Giek Far, for initiating this research project and for her continuous guidance during the study. From this project, I have increased my own independence, creative thinking and confidence. In addition, she has given advice, patient teaching, read and critically commented on my draft of this thesis.

A thank you also goes out to all members of Faculty of Biosciences and Bioengineering for guiding and giving me a warmth helping hands along throughout the project. Thanks for their moral support, interest and constructive suggestion of this project. There are no words to describe how much I appreciate their kindness. Last but not least, I would like to thank my family and friends for all of their love and support.

You were there for me even when I was in difficulty and faced problems during my research. These are deeply appreciated. Thank you again to all of you. To you I owe this thesis.

ABSTRACT

Influenza virus is globally pathogenic and it is usually associated with zoonotic respiratory disease. It possesses a lipid-bounded segmented genome which encodes at least one biochemically-distinct protein. Its subtype A can be classified according to antigenic differences. NS1 protein is defined as nonstructural protein in the virus. It is a known multifunctional virulence factor. It can only be detected in the infected cell. In this study, the NS1A gene was successfully cloned into the *Bam*HI/*Sac*I cleaved-pET-32c(+) vector and subsequently electro-transformed into *E. coli* BL21(DE3) expressing host. There were three positive clones confirmed to contain the gene of interest by sequencing. Protein expression in soluble and insoluble fractions was observed in *E. coli* BL21(DE3). The clone 104 was selected for subsequent analysis. Better NS1A protein expression was found at 37°C by 5mM lactose induction. Purification of the NS1A recombinant protein from the inclusion bodies fraction was attempted by Ni-NTA affinity chromatography and ion exchange chromatography. The physical condition along the purification column and the biological properties of the protein itself may perhaps result in the loss of protein and its corresponding immunogenicity. Ammonium sulfate at 20% saturation was attempted to sufficiently concentrate and partially purify the NS1A recombinant protein. The ammonium sulfate precipitated NS1A recombinant protein has shown significant immno-response to the polyclonal antibody in Western blot. The 37kDa NS1 protein was detected to react with the H1N1 NS polyclonal antibody.

ABSTRAK

Influenza virus merupakan patogen dalam dunia global dan ia biasanya berkaitan dengan penyakit pernafasan zoonosis. Ia mempunyai genom tersegmentasi yang diselaputi oleh lipid dan setiapnya mengekodkan sekurang-kurangnya satu protein yang berbeza dalam sifat biokimia. Influenza A boleh diklasifikasikan mengikut antigennya. NS1 protein ditakrifkan sebagai protein nonstruktural virus yang merupakan faktor virulensi yang multifungsi. Protein ini hanya boleh dikesan dalam sel yang dijangkiti. Dalam kajian ini, gen NS1A telah berjaya diklonkan ke dalam vektor pET-32c(+) melalui potongan enzim pada *Bam*HI/*Sac*I. Plasmid rekombinan kemudian ditransformasikan melalui denyutan elektrik ke dalam sel bakteria perumah *E. coli* BL21(DE3). Terdapat tiga klon telah dikenalpastikan sebagai klon positif yang mengandungi gen NS1A. Pengekspresan protein dalam fraksi larut dan tidak larut diperhatikan dalam *E. coli* BL21(DE3). Klon 104 telah dipilih untuk analisis selanjutnya. Pengekspresan NS1A protein didapati berkesan pada suhu 37 °C di bawah 5mM laktosa induksi. Protein rekombinan NS1 dari fraksi tidak larut cuba ditulenkan berdasarkan kromatografi afiniti Ni-NTA dan kromatografi penukaran ion. Namun, keadaan fizikal sepanjang kolum penulenan dan sifat biologi daripada protein itu sendiri mungkin menyebabkan kehilangan protein dan aktiviti keimunan. Pemendakan garam dengan 20% amonium sulfat adalah cukup untuk memendakan lalu separa menulenkan protein NS1A. Protein rekombinan yang dimendakkan dengan amonium sulfat telah menonjolkan keputusan positif dalam kaedah pemblotan Western dan protein NS1A bersaiz 37 kDa telah dikesan bertindak balas dengan antibodi poliklonal NS H1N1.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS/ ABBREVIATIONS/ NOTATIONS/ TERMINALOGY

AEC	-	Anion Exchange Chromatography
APS	-	Ammonium persulfate
BSA	-	bovine serum albumin
bp	-	base pairs
$CuSO_4$	-	copper sulfate
CV	-	column volume
dH_2O	-	distilled water
DNA	-	deoxyribonucleic acid
dNTPs	-	deoxynucleotide triphosphates
dsRNA	-	double stranded RNA
E. coli	-	*Escherichia coli*
ED	-	effector domain
ELISA	-	Enzyme-linked immunosorbent assay
EtBr	-	Ethidium bromide
g	-	gram
HA	-	hemagglutinin
HCl	-	hydrochloric acid
His	-	histidine
IB	-	inclusion bodies
IFN	-	interferon
IMAC	-	Immobilized Metal Ion Affinity Chromatography
IPTG	-	isopropyl-β-D-thiogalatoside

KCl	-	Potassium chloride
kDa	-	kilo dalton
KH_2PO_4	-	Potassium dihydrogen phosphate
kV	-	kilo volt
L	-	liter
LB	-	Luria-Bertani
m	-	mille
mmol/L; mM	-	milli molar
NA	-	neuraminidase
NaCl	-	Sodium chloride
NaH_2PO_4	-	Sodium phosphate
NAOH	-	Sodium hydroxide
NEP	-	nuclear export protein
NeuAc	-	N-acetylneuraminic acid
NLS	-	nuclear localization sequence/signal
nm	-	nano meter
No.	-	nombor
OD	-	optical density
ORF	-	overlapping open reading frame
PBS	-	phosphate buffer saline
PCR	-	polymerase chain reaction
PI3K	-	phosphatidylinositol 3-kinase
PKR	-	protein kinase R
RBD	-	dsRNA-binding domain
RNA	-	ribonucleic acids
RNP	-	ribonucleoprotein
RT-PCR	-	Reverse transcriptase- polymerase chain reaction
SDS-PAGE	-	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEAR	-	Southeast Asia region
TAE	-	Tris-Acetate electrophoresis buffer

TEMED	-	N, N, N′, N′-Tetramethyl ethylenediamine
TM	-	transmembrane
Tris	-	2-hydroxymethyl-2-methyl-1,3-propanediol
Trp	-	tryptophan
μ	-	macro
vRNP	-	viral ribonucleoprotein
WHO	-	World Health Organization

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Influenza is generally defined as a type of infection caused by influenza virus. This virus is highly pathogenic and it is usually considered as the causative agent of zoonotic respiratory disease (reviewed by Pushko, 2009). In addition, it is well documented that an avian as well as a vertebrate including human can be the intermediate host of the influenza virus (reviewed by Cox, 1998; Gibbs *et al*., 2009). This virus is often detected as a reassortant from more than one parent (reviewed by Hampson and Mackenzie, 2006; Gibbs *et al*., 2009) and its infection is usually associated with cellular alteration, apoptosis and host mortality (Schultz-Cherry *et al*., 2001).

The origin of the influenza virus always raises intriguing questions for the world. Recently, the influenza virus H1N1 outbreak is of great concern to the world. It is believed that the influenza A strain may share circulation among the genetically-distinct hosts (Guan *et al*., 1996; Rappole and Hubálek, 2006). It is not surprising that the influenza virus subtype was detected seasonally showing co-circulation with the earlier pandemic strains (reviewed Cox, 1998; reviewed by Hampson and Mackenzie, 2006; Nelson *et al*., 2008). In 1918, the respiratory infection in swine was firstly detected

(reviewed by Nichols and W. LeDuc, 2009) and has caused approximately over 40 million deaths in the world (reviewed by Nicholson *et al*., 2003 and Pushko, 2009).

Until recently, the isolation of influenza viruses has been done extensively from a spectrum of fowls and other mammalian species including human (reviewed by Hampson and Mackenzie, 2006). In addition, researchers nowadays are able to visualize the viral genome in three-dimensional structure by using advanced technology. There are many scientists concerned with the innate properties of influenza virus and most studies are related to gene regulation, gene expression, ecology and serology of the influenza virus. The cloning and expression experiments have generally resulted in a better understanding of the properties of viral proteins for antigenicity analysis and vaccine study. Subsequent advances in genetic engineering as well as protein engineering are broadly utilized for rapid virus detection.

The influenza viruses, which can be classified into types A, B and C, are included into the family of *Orthomyxoviridae* (Pringle, 1996; Bouvier and Palese, 2008). The influenza A has shown identical pathogenic potential with influenza B and it was extensively characterized as pandemics as well as epidemic threat (reviewed by Pushko, 2009) with a high transmission rate (Gibbs *et al*., 2009; reviewed by Nichols and W. LeDuc, 2009). The influenza C characterizes an occasional spread and it is less harmful to the human health as compared to influenza A or B (Matsuzaki *et al*., 2004) Nevertheless, evidence has shown that the influenza C might be latent in the swine (Matsuzaki *et al*., 2004) as well as the newborn (reviewed by Hampson and Mackenzie, 2006).

Influenza A virions are normally found in spherical shape with 80 to 120nm in diameter (Donatelli *et al*. 2003; reviewed by Pushko, 2009). However, its size may reach 300nm in length for filamentous form (Suri, 2007). This progeny virus particle is unconquered. This virion is known to govern its genetically-distinct proteins either in

extracellular or intracellular activity. Its membrane-bounded proteins, detected on or in the coated virion, involved not only in the viral replication but also ribonucleoprotein (RNP) assembly, thereby suggesting they are deployed to help in viral regulation during viral infection when the influenza virus is resisting the ongoing host immune response (reviewed by Cox, 1998; Bouvier and Palese, 2008).

In influenza A virus, the NS1A protein is encoded by the shortest viral RNA segment. This protein is specifically assembled by at least 230 amino acids. It is a multifunctional protein, involving significantly in the protein-RNA (Qiu and Krug, 1994) and protein-protein interaction (Xia *et al*., 2009). The NS1A protein plays an important role not only in the antiviral response but also in the post transcriptional activity in its host (Lin *et al*., 2007). Further, Zohari *et al.* (2008) by studying the phylogenetic relationship of NS1A gene isolated from genetically distinct infected cells, demonstrated that the NS1A protein could undergo evolutionary divergence occasionally.

1.2 Research Objectives

This research presented here focused on three main objectives. First, it aimed to clone the targeted gene, NS1A gene. Next, overexpression of the recombinant protein was attempted in the *E.coli* strain. Then, the partially purified recombinant protein was further determined through immunodetection. The specific objectives of this study were:

i. To clone influenza A NS1 gene in pET-32c(+) vector

ii. To over express influenza A NS1 recombinant protein in *E. coli* BL21(DE3)

iii. To partially purify influenza A NS1 recombinant protein

iv. To determine the immunogenicity of influenza A NS1 recombinant protein

1.3 Research Scope

The research was divided into four main parts which were cloning, overexpression, purification and immunogenicity analysis. The NS1 recombinant protein of influenza A H1N1 was successfully cloned into pET-32c(+) vector and this led to subsequent expression of the recombinant protein in *E. coli* BL21 (DE3) strain. In this project, a series of protein separation and purification process were used to purify NS1A recombinant protein. Furthermore, the protein immunoblotting was briefly performed to detect the targeted NS1 protein from the separated protein.

1.4 Problem Statement

The influenza virus spreads globally in the biosphere and it may lead to critical causalities during an outbreak. Since the Spanish influenza in 1918, the record has indicated significantly its potential circulation around the world and successively caused the Asian influenza in 1957, Hong Kong influenza in 1968 (reviewed by Cox, 1998 and Pushko, 2009) and the recent pandemic influenza 2009. In an update to influenza situation in Southeast Asia region (SEAR) by World Health Organization (WHO), up to 5th of August 2010, the pandemic H1N1 2009 has caused severe outbreaks, killing 3% of the population in Southeast Asia alone. The epidemiological summary has indicated that Ukraine which is the second largest country of Eastern Europe and India which possesses the largest population in Southern Asia, were detected to be still active in influenza A H1N1 virus.

Accumulated studies have revealed that the NS1A protein could give rise to a higher virulence (reviewed by Cox, 1998; Nicholson *et al*., 2003). Yet, most of the researchers have widely studied the viral surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA), which are considered to involve significantly in the virus

classification and antigenic shift accompanied with genomic reassortment (reviewed by Cox, 1998; Wagner *et al*., 2001).

Besides that, this nonstructural protein is not synthesized within the virion itself (reviewed by Cox, 1998) but it expresses abundantly in the nucleus of the newly-infected cell (Li *et al*., 1998). Investigation done by Birch and his colleagues (1997) has revealed that both the healthy cell and the vaccinated cell are able to inactivate or attenuate the NS1A protein. In addition, the expression of the NS1 protein was attempted in the prokaryotic bacteria, yeast and mammalian cells for protein interaction study. The characteristics of expressing vector (Ma *et al*., 2009), codon-tRNA correlation (Gouy and Gautier, 1982; Ikemura, 1985) as well as the toxicity of NS1 protein (Ward *et al*., 1994) could affect the protein expression.

Nowadays, the cloning and expression work is ubiquitous in biotechnology field. Nevertheless, the cloning and expression of NS1 protein is still new, elementary and not very in-depth exploration particularly in the Asian region. However, it can be considered as the fundamental source for future research. This nonstructural protein can be further investigated by studying the protein characterization, their innate properties, the protein binding mechanism, protein topology and others. Much remained to be learnt about this protein.