UNIVERSITI PUTRA MALAYSIA

THE NUCLEOTIDE SEQUENCE OF THE MATRIX (M) PROTEIN GENE OF NEWCASTLE DISEASE VIRUS (NDV) STRAIN AF 2240

SITI FATHIMAH PUTERY JEMAIN

FSAS 1999 31

THE NUCLEOTIDE SEQUENCE OF THE MATRIX (M) PROTEIN GENE OF NEWCASTLE DISEASE VIRUS (NDV) STRAIN AF 2240

By

SITI FATHIMAH PUTERY JEMAIN

Thesis Submitted in Fulfilment of the Requirements for the
Degree of Master of Science in the Faculty of
Science and Environmental Studies
Universiti Putra Malaysia

March 1999

ACKNOWLEDGEMENTS

Bismillahirrahmanirrahim, I would like to express my most sincere thanks and appreciation to my Chairperson, Associate Prof. Dr. Khatijah Yusoff, Department of Biochemistry and Microbiology, Universiti Putra Malaysia, for her guidance, support, understanding, encouragement, patience and constructive suggestions throughout the course of study. I also wish to convey my thanks to Associate Prof. Dr. Abdullah Sipat and Dr. Abd.Rahman Omar for their support and valuable suggestions during the course of this research.

I would like to convey my heartfelt thanks and gratitude to my lab-mates, Omeima, Subha, Vijay, Alan, Laili, Filza, Ainu, Wai Kean, Chiew Ling, Majid, Jeeven, Wei Hong, Chui Fung, Sing King, Siti and Priya for the friendship, support and encouragement and for making the 2 years stay, an enjoyable one.

My deepest gratitude and thanks especially go to MSMN, SRA, my parents, brothers and sisters, my beloved husband - abang Mustapha, and the 'Family' especially, mama, ibu, Azu, Nor, kak Yam, Uncle Ali and Dr. Nyonya for the continued support and prayers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS

The following abbreviations were used in the text:

bp	base pair
°C	degrees Centigrade
kDa	kilodalton
kb	kilo basepair
min	minute
s	second
v/v	volume per volume
w/v	weight per volume
µl	microlitre
NDV	Newcastle disease virus
RT-PCR	reverse transcription polymerase chain reaction
Mbp	mega base pair
M	matrix
st	strain
NDU 25828	NDV strain B1
NDU 25829	NDV strain Fontana
NDU 25830	NDV strain Hertz

NDU 25831	NDV strain Kimber
NDU 25832	NDV strain Largo
NDU 25833	NDV strain La Sota
NDU 25834	Queensland/V4
NDU 25835	Texas/GB
NDU 25836	Turkey/ND
NDU 25837	NDV strain Ulster
NDU 25838	NDV strain VGGA
VMAT_NDVA	NDV strain Australia/Victoria/32
VMAT_NDVB	NDV strain Beudette C/45

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirements for the degree of Master of Science.

THE NUCLEOTIDE SEQUENCE OF THE MATRIX (M) PROTEIN GENE OF NEWCASTLE DISEASE VIRUS (NDV) AF 2240

By

SITI FATHIMAH PUTERY JEMAIN

March 1999

Chairperson : Associate Professor Khatijah Yusoff, Ph.D.

Faculty : Science and Environmental Studies

The complete nucleotide sequence of the matrix (M) protein gene of the local Newcastle disease virus (NDV) strain AF 2240 was determined. Based on consensus primers, several segments of the M gene were reverse transcribed and amplified by the polymerase chain reaction. The RT-PCR products were cloned into a plasmid vector pCR™2.1. The DNA inserts in the clones were then cycle-sequenced. The start and polyadenylation signals have been identified. Assuming that the M gene starts and terminates at these sequences, the M gene is 1223 nucleotides long and encodes an open reading frame of 364 amino acids, corresponding to a polypeptide calculated molecular weight of 40 kDa. The M protein, is both hydrophobic and basic. Phylogenetic analysis shows that the strain has a close relationship with the North American Fontana strain. The M protein amino acid sequence also shows similarities with members of the

Rubulavirus genus such as the Simian virus 5, the human parainfluenza virus type 4 and the mumps virus. Similarities of the amino acid sequence also exist between the local strain and other members of the *Paramyxovirus* and *Morbillivirus* genus.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains.

JUJUKAN NUKLEOTIDA GEN PROTEIN MATRIKS (M) VIRUS PENYAKIT SAMPAR AYAM (NDV) STRAIN AF 2240

Oleh

SITI FATHIMAH PUTERY JEMAIN

Mac 1999

Pengerusi : Profesor Madya Khatijah Yusoff, Ph.D.

Fakulti : Sains dan Pengajian Alam Sekitar

Gen protein matriks (M) virus penyakit sampar ayam (NDV) strain AF 2240, yang merupakan strain tempatan, belum di ketahui jujukan nukleotidanya. Oleh itu kajian ini dijalankan untuk mengenalpasti jujukan nukleotida protein M tersebut. Produk RT-PCR gen M diklonkan ke dalam vektor plasmid pCR™2.1. Jujukan produk RT-PCR gen M yang telah diklonkan tadi ditentukan melalui proses penjujukan berulang. Kodon-kodon permulaan dan isyarat poliadenilat telah dikenalpasti. Dengan mengambil kira kodon-kodon ini, jujukan nukleotida protein M didapati mempunyai 1223 pasangan bes dan berupaya untuk mengkodkan protein yang mempunyai 364 asid-asid amino. Anggaran berat molekular protein M ialah lebih kurang 40 kDa. Secara keseluruhannya, jujukan asid amino protein M adalah bersifat hidrofobik dan basik. Analisis filogenetik, menunjukkan bahawa strain AF 2240 mempunyai

pertalian yang rapat dengan strain NDV daripada Amerika Utara, iaitu strain Fontana. Jujukan asid amino protein M mempunyai persamaan dengan ahli-ahli genus *Rubulavirus* yang lain, iaitu virus Simian 5, virus parainfluenza manusia jenis 4 dan virus penyakit begok dan juga ahli genus *Paramyxovirus* dan *Morbillivirus*.

CHAPTER I

INTRODUCTION

Newcastle disease or 'penyakit sampar ayam' is endemic in Malaysia and has always been a serious threat to the livestock industry. This disease is caused by the Newcastle disease virus (NDV) which belongs to the *Rubulavirus* genus within the *Paramyxoviridae* family. The disease currently has a worldwide distribution with a wide host range in which all orders of birds have been reported to be infected by NDV (Brandly, 1964). From the Office International Des Epizooties (O.I.E) homepage (http://ss.niah.affrc.go.jp/OIE/yb95/yb95b.html), it was shown that there has been 11 reported outbreaks of ND in Malaysia and frequent outbreaks also occur in Indonesia, the Philippines, India, South Korea and other developing countries from January to August 1995.

The techniques used to eradicate the virus are (1) destruction of the entire chicken population and disinfection of the affected area with viricidal agents, (2) employing quarantine procedures and (3)

vaccination of the chickens (Spradbrow, 1987). Vaccination has so far given an adequate protection against NDV, at least for chickens reared commercially. Two types of vaccines are currently being used against Newcastle disease; a live attenuated vaccine (for example, strains B1 and La Sota) usually administered in drinking water and the inactivated vaccine (for example, strain Ulster) usually administered by intra-muscular injection. In addition, researchers in UPM has used the lentogenic strain V4 (UPM) as a live attenuated vaccine and in the form of food pellet (Ideris *et al.*, 1990).

The rapid advance in recombinant DNA technology has given birth to a new generation of genetically engineered vaccine that has several advantages over the traditional vaccines. These advantages include (1) greater safety to workers and the environment because only the usage of isolated genes (cloned into specific vectors) are involved, (2) vaccine associated complications are reduced because non-essential viral components are not present in the vaccine, (3) purified proteins are more stable than virus particles and (4) cheaper production cost (Obijeski, 1985).

The local velogenic-viscerotropic NDV virus reference strain AF 2240, has been studied in UPM for the past few years. This strain has the potential to be

developed as a subunit vaccine capable of protecting poultry against a wider range of NDV isolates. Being heat resistant, its transportation and handling will be cheaper and easier as it will limit the number of essential 'cold-chains' involved in maintaining its effectiveness. Unlike the haemagglutinin-neuraminidase (HN) protein and the fusion (F) protein genes of strain AF 2240 (Tan *et al.*, 1995; Yusoff *et al.*, 1993) which have been extensively studied, very little work has been done on the matrix (M) protein gene. The M protein gene encodes the M protein, which plays an important role in the construction of the virus as well as the regulation of the viral RNA synthesis (Tanabayashi et al., 1990; Seal, 1996). It is hoped that knowledge of the M gene sequence will give a better understanding of its structure and function.

This thesis reports on the nucleotide sequence of the M gene of NDV strain AF 2240.

CHAPTER II

LITERATURE REVIEW

Newcastle Disease

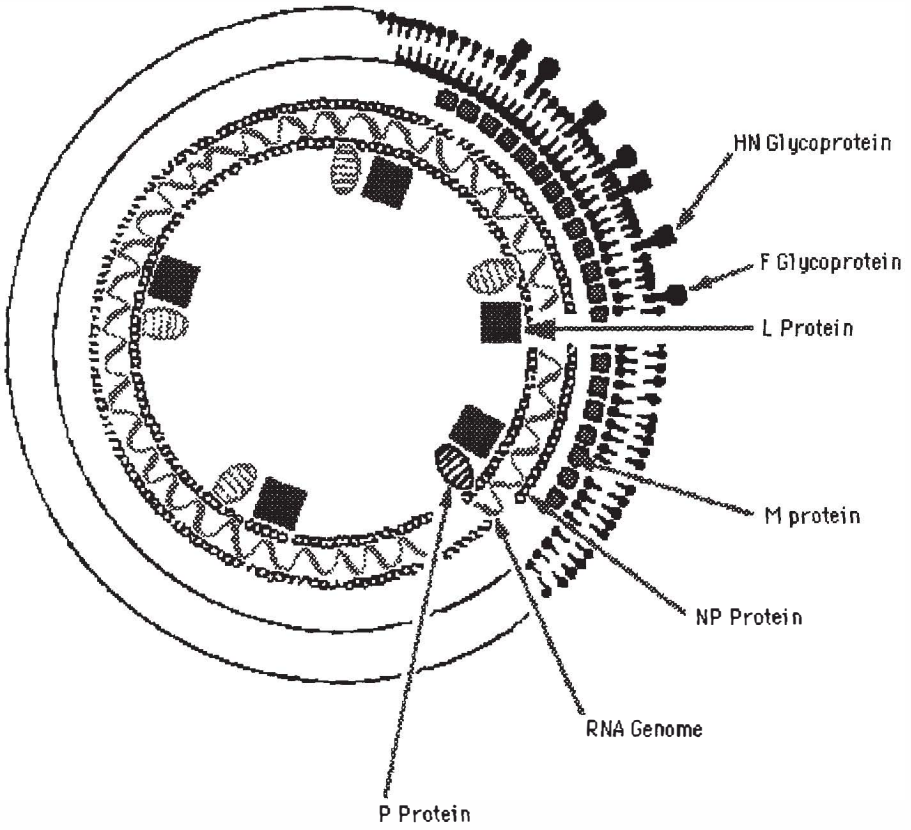
Newcastle disease which causes 100% mortality in birds (Brandly, 1964), was discovered in 1926 near Batavia (Jakarta). A similar outbreak was also reported in Newcastle-on-Tyne (hence, the name of the disease) and in Korea in that same year. By 1933, researchers in England, Philippines, Indonesia and India, through cross-immunity tests, agreed that the same virus named as the Newcastle disease virus caused these outbreaks.

Newcastle Disease Virus

The Newcastle disease virus, the human parainfluenza virus types 2, 4a and 4b, the Simian virus 5 and mumps virus are grouped under the genus *Rubulavirus* (Lamb and Kolakofsky, 1996).

The *Paramyxoviridae* family was reclassified in 1993 into two subfamilies: the *Paramyxovirinae* and the *Pneumovirinae*. The *Paramyxovirinae* contains three genera, *Parainfluenzavirus, Rubulavirus* and *Morbillivirus*. The sub-family *Pneumovirinae* contains the genus *Pneumovirus*. The new classsification is based on morphological criteria, genomic organisation, biological activities of proteins, and the sequence relationship of the encoded proteins (Lamb and Kolakofsky, 1996).

Members of this family of viruses are very pleomorphic, with a diameter of about 100 to 500 nm, and are capable of acquiring a variety of shapes ranging from circular, spherical, flattened, and often in filamentous forms. The virion is enveloped with a lipid bilayer membrane, derived from the plasma membrane of the host cell, which gives it a certain flexibility to exist in these various morphological forms suited to the pH and salinity of the surrounding environment (Waterson, 1964). The haemagglutinin-neuraminidase (HN) and fusion (F) glycoproteins protrude from the lipid bilayer membrane. They are visible as spike-like structures under the electron microscope, ranging from about 8 to 20 nm in length and spaced 6 to 10 nm apart. Below the lipid bilayer membrane is a shell of protein known as the matrix (M) protein. Protected by the lipid and protein layers, is the long helical nucleocapsid 13 to 18 nm in diameter with the "herring-bone" structure quite unique to NDV, containing the genome (Figure. 1) (Alexander, 1988; Samson, 1988).



M = matrix protein
F = fusion protein consists of 2 disulphide-linked subunits (F_1+F_2)
HN = haemagglutinin-neuraminidase protein
NP = nucleoprotein
L = large protein
P = nucleocapsid associated protein

Figure 1: Schematic Diagram of a Typical Paramyxovirus Particle.

(www.tulane.edu)

The genome of NDV is a non-segmented, single stranded linear RNA of negative polarity with a molecular weight of 5.2 to 5.7×10^6 daltons which is approximately 15 kilobase (kb) long. It codes for the following six structural proteins nucleocapsid protein (NP), nucleocapsid associated protein (P), matrix protein (M), fusion protein (F), haemagglutinin-neuraminidase (HN) protein and large (L) polymerase protein (Samson, 1988) (Table 1).

Matrix (M) Protein

Sequence Characteristics

The matrix (M) protein is a non-glycosylated, non-phosphorylated, hydrophobic and highly basic protein of 364 amino acids (Chambers *et al.*, 1986; McGinnes and Morrison, 1987; Seal, 1996). Most of its positively charged basic amino acids lie in the C-terminal portion of the molecule. It is reported that the amino acid sequence contains five cysteine residues (McGinnes and Morrison, 1987). The positions of glycine, proline and paired basic amino acid residues (R-K, K-K, K-R, R-R) of the M gene have been shown to be conserved in NDV,

Table 1: Functions of the NDV Proteins.

Protein	Approx. size (kDa)	Function
Nucleocapsid (NP)	53-56	Major structural component of nucleocapsid:complexed with genome RNA
Nucleocapsid-associated protein (P)	53-56	Associated with nucleocapsid, phosphorylated, plays a role in transcription/ replication Required for viral mRNA synthesis
Matrix or membrane (M)	38-40	Virus assembly organiser, moderates transcription
Uncleaved fusion (F_0) Larger cleaved fusion (F_1) Smaller cleaved fusion (F_2)	67 55 12	Precursor to $F_{1,2}$ Major determinant of the virulence of NDV Fusion of virus and host membranes, necessary for infection and haemolysis
Haemagglutinin-neuraminidase (HN)	72-75	Dual function: receptor binding protein responsible for haemagglutination and cleavage of neuraminic acid residues from glycoproteins/ lipids
Large (L)	180-220	RNA directed RNA polymerase, necessary for the making of + (coding) sense mRNA

(Samson, 1988)

Sendai, measles and canine distemper viruses (Bellini *et al*., 1986; Chambers *et al*., 1986; McGinnes and Morrison, 1987).

The M protein is one of the main structural elements of the virus. However, the functioning of this M protein is suggested to be mediated by its overall characteristics (basic and hydrophobic), rather than by any particular conserved sequences (Bellini *et al*., 1986; Chambers *et al*., 1986; Spriggs *et al*., 1987; Elango 1989; Limo and Yilma., 1990; Tanabayashi *et al*., 1990; Randhawa *et al*., 1996).

Functions

Firstly, the M protein seems to be a controlling factor in RNA synthesis and has a key role in virus assembly by locating nucleocapsid structures beneath those regions of the plasma membrane in which the F and HN glycoproteins are attached. This was proven when paramyxoviruses that produced defective M proteins were unable to produce virus particles (Peeples and Bratt, 1984). This was perfectly demonstrated *in vitro* when the Sendai virus nucleocapsid would not

form a complex with the viral glycoproteins unless the M protein was added. Secondly, it could be involved in moderating the virion RNA polymerase activity possibly in interactions with cellular actin (negatively charged) which drives the budding process. Thirdly, it could be involved in protein kinase activity (Blumberg *et al.*, 1984). Analysis of chimeric M proteins indicates that mutations in the amino-terminal and the carboxyl-terminal regions of the M protein all abrogate nucleocapsid binding. This suggests that the M protein conformation is important for interaction with the viral nucleocapsid (Hirano *et al.*, 1993).

Physical Location in the Virus Particle

In electron micrographs of virions (Lamb and Kolakofsky, 1996), an electron-dense layer is observed underlying the viral lipid bilayer. This is thought to represent the location of the M protein. Fractionation studies of the virus indicate that this protein is peripherally associated with membranes and is not an intrinsic membrane protein (Faaberg and Peeples, 1988). The M protein probably contains amphipathic ∝-helices that insert themselves into the inner leaflet of the lipid bilayer to coat this surface and organise its contacts with the