Molecular characterization of South African lineage II West Nile virus isolates and development of a diagnostic assay.

By

Elizabeth Magdelena Botha

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

(Microbiology)

Department of Microbiology and Plant Pathology
In the Faculty of Natural and Agricultural Science
University of Pretoria
Pretoria
South Africa

Supervisor: Prof Louis H. Nel
Co-Supervisors: Dr Wanda Markotter
Dr Janusz Paweska
Dr Maria Venter

January 2008

I declare that the thesis, which I hereby submit for the degree MSc at the University of Pretoria, South Africa, is my own work and has not been submitted by me for degree purposes at any other university.

_____

Elizabeth Botha

ACKNOWLEDGEMENTS

The student wishes to express her sincere thanks and appreciation to the following people:

1. Prof. L.H. Nel, Dr. W. Markotter, Dr. M. Venter for their guidance, support and shared knowledge throughout this study.

2. The project was done in collaboration with the NICD, Special Pathogens Unit, and in this regard special thanks to Dr. J.T. Paweska, A. Grobbelaar and P. Jansen van Vuren for use of their laboratory space and technical assistance, as well as for supplying isolates of West Nile virus and specific antisera.

3. Dr. M. Wolfaardt for her assistance in completion of the DNA sequences of the WNV H442 strain.

4. Design Biologix for donation of the BHK 21 cells.

5. Fellow students at the Department Microbiology and Plant Pathology for their interest and friendship, especially my best friend Aletta.

6. Family, loved ones and friends for their constant encouragement love and support.

7. The National Research Foundation and The Poliomyelitis Research Foundation for financial support.

SUMMARY

Molecular characterization of South African lineage II West Nile virus isolates and development of a diagnostic assay.

by

Elizabeth Magdelena Botha

Supervisor: Prof. L. H. Nel
Department of Microbiology and Plant Pathology
University of Pretoria

Co-Supervisor: Dr. W. Markotter
Department of Microbiology and Plant Pathology
University of Pretoria

Dr. J. T. Paweska
National Institute of Communicable Diseases, Special Pathogens Unit,
Sandringham

Dr. M. Venter
Dept of Medical Virology
University of Pretoria

For the degree MSc (Microbiology)

West Nile virus (WNV) belongs to the *Flaviviridae* family, a virus family of which many members are known as human pathogens. WNV has a worldwide distribution and strains that cluster in lineage II is endemic to sub-Saharan Africa. The complete nucleotide sequence of four lineage II West Nile virus strains, isolated in South Africa from patients with mild or severe WNV infections, were determined. Using a murine model, these strains had been shown to produce either highly or less neuroinvasive infections and induced similar genes to corresponding highly or less neuroinvasive lineage I strains. Nucleotide and amino acid sequence comparison between highly and less pathogenic lineage II strains demonstrated that the non-

structural genes and in particular the gene coding for the NS5 proteins were the most variable. All the lineage II strains sequenced in this study were found to possess the E-protein glycosylation site previously postulated to be associated with virulence. Comparison of the signalase cleavage sites suggested that lineage II strains may be cleaved slightly more efficiently than lineage I strains in the C-prM junction, but less efficiently between prM and E genes. Relative to the highly neuroinvasive strains sequenced in this study major deletions were found in the 3' noncoding region of 2 lineage II strains shown in previous studies to be either less- or not at all neuroinvasive. This is the first report of full genome sequences of highly neuroinvasive lineage II WNV strains.

Currently available commercial WNV ELISA kits were developed with lineage I WNV strains and are expensive to use. For these reasons the development of a potential ELISA diagnostic assay based on the South African lineage II strain, H442, was envisaged. Such assay, if reliable and efficacious would be a useful tool towards WNV surveillance. The prM and E genes were selected to be expressed as recombinant antigens because of their co-expression nature and because the envelope protein is the principal target for neutralization. After cloning of the respective genes and verification of integrity, a mammalian expression system was utilized. Different mammalian cells and transfection media were tested and BHK 21 cells with SuperFect transfection medium were found to be best. Attempted expression of proteins was tested with immunofluorescent antibody testing as well as SDS-PAGE and Western blot analysis. Expression of recombinant WNV antigens were also tested in indirect and sandwich ELISA's systems. It was however not possible to perform these two ELISA systems at a satisfactory level or clearly indicated if expression of proteins was successful.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

°C	-	Degrees Celsius
3D	-	Three dimensional
BHK	-	Baby hamster kidney cells
bp	-	Base pairs
BSL-3	-	Biosafety lab-3
C	-	Capsid protein
cDNA	-	Complementary DNA
CF	-	Complement fixation
CNS	-	Central nervous system
CO_2	-	Carbon dioxide
CSF	-	Cerebrospinal fluid
DEN	-	Dengue virus
DMEM	-	Dulbecco`s Modified Eagle's Medium
DMSO	-	Dimethylsulfoxide
DNA	-	Deoxyribonucleic acid
dNTP	-	2`deoxynucleoside-5`triphosphate
ds	-	Double stranded
E	-	Envelope protein
EDTA	-	Ethylenediaminetetra-acetic acid
ELISA	-	Enzyme-linked immunosorbent assay
ER	-	Endoplasmic reticulum
et al	-	*et alii* (and others)
FCS	-	Fetal calf serum
FDA	-	Food and drug adminstration
FITC	-	Fluorescein isothiocyanate
g	-	Gravitational Force
H_2O	-	Water
HI	-	Hemagglutination inhibition
IFA	-	Immunofluorescence assay
Ig	-	immunoglobulin
IMAC	-	imacromatography
IPTG	-	isopropyl-β-D-thiogalactopyranoside
JEV	-	Japanese encephalitis virus
kbp	-	kilo base pair
kDa	-	kilodalton
km^2	-	Square kilometre
LB	-	Luria Bertani
LD_{50}	-	50% Lethal Dose
LGT	-	Langat virus
m/v	-	Mass per volume
MIA	-	Microsphere immunoassay

min	-	Minute
ml	-	Millilitre
mM	-	Millimolar
mmol	-	Millimol
Mr	-	Molecular weight
NaAc	-	Sodium acetate
NCR	-	Non coding region
Nm	-	nano meter
NS	-	Non-structural
ORF	-	Open reading frame
PAGE	-	Polyacrylamide gel electrophoresis
PBS	-	Phosphate buffered saline
PCR	-	Polymerase chain reaction
PEG	-	Polyethylene glycol
PFU	-	Plaque forming units
prM	-	Premembrane protein
PRNT	-	Plaque reduction neutralization test
RdRp	-	RNA dependent RNA polymerase
RES	-	Reticuloendothelial system
RGD	-	Arginine-glycine-aspartic acid
RNA	-	Ribonucleic acid
rpm	-	Revolutions per minute
RT-PCR	-	Reverse transcription polymerase chain reaction
s	-	Second
SDS	-	Sodium dodecyl sulphate
SF	-	*Spodoptera Frugiperda*
ss	-	Single stranded
SVP	-	Subviral particle
TBEV	-	Tick-borne encephalitis virus
$TCID_{50}$	-	50% Tissue culture infection dose
TEMED	-	N,N,N`,N`-tetramethylethelenediamide
Tris	-	Tris(hydroxymethyl)-aminomethane
U	-	Units
UV	-	Ultra violet
V	-	Volts
v/v	-	Volume per volume
w/v	-	Weight per volume
WNV	-	West Nile Virus
X-gal	-	5`-bromo-4-chloro-3-indolyl-β-D-galactopyronoside
YFV	-	Yellow fever virus
µg	-	Microgram
µl	-	Microcli**ter**

Chapter One:

Literature Review

1.1 INTRODUCTION

West Nile virus (WNV) is a positive single stranded virus that belongs to the *Flaviviridae* family and is further classified in the *Flavivirus* genus and the Japanese encephalitis serogroup. Other genera in this family, which has over 70 members, are the *Pestivirus* and *Hepacivirus*. The *Flavivirus* genus consists of viruses associated with emerging and re-emerging human diseases such as Japanese encephalitis disease (JEV), West Nile fever, Dengue haemorrhagic fever (DEN), Yellow fever (YF) and Kyasanur forest haemorrhagic disease. (Gaunt *et al.*, 2001; Campbell *et al.*, 2002). WNV was first isolated from a febrile patient in the West Nile district of northern Uganda in 1937 (Smithburn *et al.*, 1940).

The symptoms of a WNV infection in humans include fever, rash, headaches, muscle weakness and disorientation and some cases develop encephalitis, meningoencephalitis or hepatitis (Briton, 2002). In the case of lineage II WNV strains, no direct correlation can be drawn between symptoms and pathogenesis. South African strains H442, SPU116/89 and SA93/01 had an LD_{50}dose of between 2-3 and were neuroinvasive in mice, but H442 only showed benign disease in the human patient it was isolated from. However the patient infected with SPU116/89 had fatal hepatitis and the one with SA93/01 non-fatal encephalitis. The low-neuroinvasive South African strain SA381/00 had an LD_{50} dose of 316.3 and exhibited only benign disease in the patient it was isolated from (Venter *et al*., 2005; Burt *et al*., 2002).

WNV is widely endemic in southern Africa in areas where the principal vector, *Culex univittatus* is present. Human infections only occur sporadically with large epidemics occurring when unusually high-rainfall or hot weather favored breeding of the vector (Burt *et al*., 2002). Such outbreaks occurred concurrent with epizootics in birds. Two large outbreaks were the 1974 epidemic in the Northern Cape Province, South Africa, involving tens of thousands of human cases over a 2500-km^2 area and the second occurred in the Witwatersrand-Pretoria region (Gauteng) of South Africa in 1984. Neither of these outbreaks led to human mortalities and generally all the infected patients exhibited a mild febrile illness (McIntosh, 1980; Jupp *et al*., 1986; Burt *et al*., 2002). Since then the number of human WNV infections confirmed yearly remained constant at between 5-15 cases per annum, although only a portion of suspected cases are subjected to laboratory investigation. Even with the apparent low level of

virus activity in South Africa, isolates of WNV were made from patients with severe disease, including fatal hepatitis and non-fatal encephalitis in the last few years (Burt *et al*., 2002).

WNV lineage I was introduced into the Americas in 1999. The introduction of WNV into Northern America was associated with increased frequency of neurological infections, human case-fatality rates and horse and bird deaths (Petersen and Roehrig, 2001). These observations lead to the question of whether emergence of WNV strains with increased pathogenicity occurred or whether the virulence of the virus had previously been underestimated. It also raised the question of whether there is a pathogenic difference between lineage I and II WNV strains (Burt *et al*., 2002). As the precise mechanism of pathogenesis of these viruses is not fully understood (Lindenbach and Rice, 2003) there is a need to study the pathogenesis mechanisms of WNV.

Limited epidemiological data is available for WNV in South Africa, thus the precise number of WNV cases per annum is not known and the pathogenesis of the WNV South African strains has not been well studied. According to Jupp, 2001 17.1% of humans in the Karoo area of South Africa has neutralizing antibodies against WNV, 8 % in the highveld area and 2% of humans in the KwaZulu Natal area. The few identified cases per annum suggest an active presence of WNV in South Africa, and the highly pathogenic strains isolated in the last few years indicates a presence of virulent pathogenic WNV strains in South Africa (Burt *et al*., 2002). There is a clear need for surveillance of the South African WNV situation. To perform effective surveillance, an easy to use and cost effective diagnostic assay is needed. Current available ELISA tests utilize lineage I WNV strains antigen (Prince and Hogrefe, 2005) and are expensive to use on a large scale. The answer will be to develop a safe, cost effective ELISA specific for WNV lineage II South African strains.

The literature review of this dissertation provides a brief overview of general *Flavivirus* biology with specific focus on WNV. The virulence/pathogenic determinants and possible phylogenetic relationships related to pathogenesis will be discussed. The second part of the review will discuss the properties of the membrane and envelope proteins, which are important in recognition of the host immune system as

well as an overview of previous studies producing recombinant flavivirus antigen with the focus on WNV and different diagnostic methods available.

1.2 CLASSIFICATION AND DISTRIBUTION

West Nile virus is part of the *Flaviviridae* family and *flavivirus* genus (Calisher *et al*., 1989; Lanciotti *et al*., 2002). This family is divided into 8 antigenic complexes and WNV forms part of the mosquito-borne complex of viruses. In this mosquito-borne complex it belongs to the Japanese encephalitis virus antigenic complex (Kuno *et al*., 1998). WNV formerly consisted of two distinct lineages; lineage I has a worldwide distribution from Western Africa, the Middle East, Eastern Europe, the United States and Australia and lineage II is found in Sub-Saharan Africa and Madagascar. Lineage I can be further divided into three clades; one clade represents the India isolates, another the Kunjin virus isolates from Australia and the last clade the Europe/African/Middle East and United State isolates (Lanciotti *et al.*, 2002; Lanciotti *et al.*, 1999; Burt *et al*., 2002; Beasley, 2005).

A possible new lineage (lineage 3) of WNV was identified in the Czech Republic when, in 1997, a flavivirus named Rabensburg virus strain 97-103 (RabV 97-103) was isolated from *Culex pipens* mosquitoes following floods in South Monrovia (Bakonyi *et al*., 2005). RabV 97-103 strain is antigenically related to lineage I Egypt WNV strain, Eg-101. Antigenic relationship to WNV was determined with a cross-neutralization test using Strain Eg 101. Pathogenicity characteristics in a murine model and complete nucleotide as well as putative amino acid sequences revealed that RabV 97-103 shares a 75-77 % nucleotide and 89-90 % amino acid identity with representative strains of WNV lineages I and II (Bakonyi *et al*., 2005). In 1999 another RabV strain 99-222 was isolated and has a >99 % nucleotide identity to previous RabV 97-103 strain (Bakonyi *et al*., 2005). Furthermore a unique strain have also been isolated from *Dermacentor marginatus* ticks in Eastern Europe (strain LEIV-Krnd88-190) and as in the case of RabV represents either a new lineage (Lineage 4) or a distinct member of the Japanese encephalitis virus (JEV) group based on distance analysis (Bakonyi *et al*., 2005). In 2007 an investigation of Indian isolates of WNV suggested the existence of an additional lineage (Lineage 5), which replaces lineage IC (Vijay *et al*., 2007).

In 2004 WNV was isolated from a goshawk (*Accipiter gentilis*) fledgling from a national park in south-eastern Hungary that showed symptoms of neurological disease and died as a result of this WNV infection. WNV strain isolated from this goshawk was named goshawk-Hungary/04 (Hu04). The same strain was again isolated in 2005 in the same area. These strains have the highest identity (96 % nucleotide and 99 % amino acid) to the WNV prototype lineage II strain (956D117B3(Wengler)) from Uganda (Bakonyi *et al*., 2006). Isolation of the goshawk-Hungary/04 strain was the first report of lineage II WNV outside of sub-Saharan Africa and Madagascar. This case of lineage II WNV outside southern Africa shows that WNV is capable of spreading via the bird migratory route.

1.3 VECTORS AND HOSTS

WNV is maintained in a mosquito-bird-mosquito transmission cycle primarily involving *Culex* sp. mosquitoes (Campbell *et al.*, 2002). A number of wild birds are the main reservoir hosts in endemic areas. Infected reservoir hosts (birds) develop transient high-titre viraemias that allows transmission of WNV to feeding mosquitoes (Campbell *et al.*, 2002). Humans and equines are incidental hosts; they have low viremic levels and do not play a role in the transmission cycle of WNV (Petersen and Roehrig, 2001). Other means of transmission includes transmission of WNV *via* blood transfusion or organ transplants, direct bird to human transmissions, laboratory acquired WNV infections, intrauterine transmission from infected mother to child and transmission *via* breast-feeding (Beasley, 2005; Briton, 2002).

1.4 REPLICATION CYCLE

Flaviviral RNA synthesis is semi conservative and asymmetric. The first step in the flavivirus replication cycle is the binding of the flavivirus particles to the cells at the entry point *via* interactions between the viral surface glycoprotein and cellular receptors. Virus particles will then be internalized into clathrin-coated pits *via* receptor-mediated endocytosis (Heinz and Allison, 2000; Briton, 2002; Lindenbach and Rice, 2003). The nucleocapsid is released by fusion between the virus and the host cell membranes. Fusion is introduced by low pH in the prelysosomal endoplasmic compartment and uncoating of the nucleocapsid release the viral RNA

genome into the host cytoplasm (Gollins and Porterfield, 1985; Gollins and Porterfield, 1986).

The next step in the replication cycle of flaviviruses is translation of the viral genome into a single polyprotein. The polyprotein is processed by viral serine protease, non-structural (NS) NS2B-NS3 and several host cell proteases to generate the mature viral proteins. Replication of the viral genome takes place *via* a negative strand intermediate, which serves as a template for additional positive stand genomic RNA's. Viral RNA-dependent RNA polymerase, NS5, together with other nonstructural proteins and possibly cellular proteins, are responsible for the production of these RNA copies (Lindenbach and Rice, 2003). Virus particle assembly takes place by budding into the rough endoplasmic reticulum (ER). Nascent virus particles will pass through the host secretary pathway. During this last step virion maturation will take place and virus is released by exocytosis. Released flaviviuses are small, spherical particles (50 nm) containing an electron-dense core of 30 nm and are surrounded by a lipid envelope that contains 2 viral proteins; the envelope (E) and membrane (M) protein (Chambers *et al*., 1990a; Chambers *et al*., 1990b; Heinz *et al.,* 1994; Briton, 2002; Lindenbach and Rice, 2003).

1.5 GENOME ORGANISATION

The mature WNV particle is enveloped, spherical and 50 nm in diameter. The nucleocapsid core consists of multiple copies of the capsid protein and is surrounded by a host-derived lipid bilayer. The envelope and membrane proteins are embedded in the virion membrane (Campbell *et al*., 2002).

Inside the WNV particle a positive sense single-stranded RNA genome is found. The 3' end terminates with CU_{OH} and the 5' end is capped with a type 1 cap structure (m^7GpppAmp) (Wengler *et al*., 1978; Brinton *et al*., 1986). The genome of ~11 kb is translated into a single polyprotein. The single open reading frame (ORF) is ± 10 000 amino acids in length (Lancoitti *et al*., 1999). The genome is divided into a 5' non-coding region (NCR) of ~100 nucleotides (nt) and a 3' NCR of between 100-700 nt in length. These NCR's, which flanks the ORF, contains conserved secondary structures that play important roles in genome replication and may also function as enhancers of protein translation. The individual three structural and seven non-

structural proteins, which make up the polyprotein, are derived by co- and post-translational cleavages by host cell signalases and the virus-derived NS2B-NS3 protease (Beasley, 2005; Briton, 2002; Campbell *et al.*, 2002; Nowak *et al*., 1989). The 5' end of the polyprotein consists of the three structural genes; the capsid (C), premembrane (prM) and the envelope (E) gene. They are followed by the seven non-structural genes NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5, which are situated at the 3' end of the ORF (Nowak *et al*., 1989; Rice *et al*., 1985; Lindenbach and Rice, 2003; Mcminn, 1997).

1.5.1) Non-coding region

1.5.1.1) 5' Non-coding region

The 5' NCR is capped with a type I 5' cap, $m^7GpppAmpN_2$, and is 95-132 nucleotides (nt) in length (Wengler *et al*., 1978; Brinton *et al*., 1986). The sequence is not well conserved among different flaviviruses but within members of specific antigenic serocomplexes conservation is high. The 5' NCR, however, start with a conserved AG and also contains conserved elements involved in secondary structure formation (Rice, 1996; Brinton and Dispoto, 1988). These structures are likely to influence translation of the genome. The 5' NCR also has a complementary region in the negative strand, which serves as a site of initiation for positive strand synthesis during RNA replication. The stemloop that can form near the terminus of either strand is an important determinant for genome replication (Lindenbach and Rice, 2003).

1.5.1.2) 3' Non-coding region

The 3' NCR can be variable in length (114-624 nt) (Wang *et al*., 1996; Mcminn, 1997). Although this region exhibits great variability the 3' NCR has several conserved features and secondary structures. A long stemloop structure consisting of 90-120 nt near the 3' terminus is structurally conserved between different flavivirus genomes. It is postulated that the stemloop plays a very important role in virus replication (Brinton *et al*., 1986; Lindenbach and Rice, 2003). Another feature that is conserved among flavivruses is the conserved CU sequence that terminates the 3' NCR (Rice *et al*., 1986; Rice, 1996). The conserved sequence CS2 that is found in the 3' NCR of mosquito-borne viruses is complementary to a conserved sequence in

the C gene and may be involved in circularization of the viral genome (Mcminn, 1997).

1.5.2) Coding region

The polyprotein of WNV consists of a single ORF with the proteins in the following order: capsid, premembrane, envelope, NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5 (Figure 1.1). A host signal peptidase cleaves the C-prM, prM-E, E-NS1 junctions as well as the C-terminus of the NS4A, while virus encoded serine protease is responsible for cleavage of the NS2A-NS2B, NS2B-NS3, NS3-NS4A, NS4A-NS4B and NS4B-NS5 junctions. No enzyme has been identified that cleaves the site NS1-NS2A (Lindenbach and Rice, 2003).

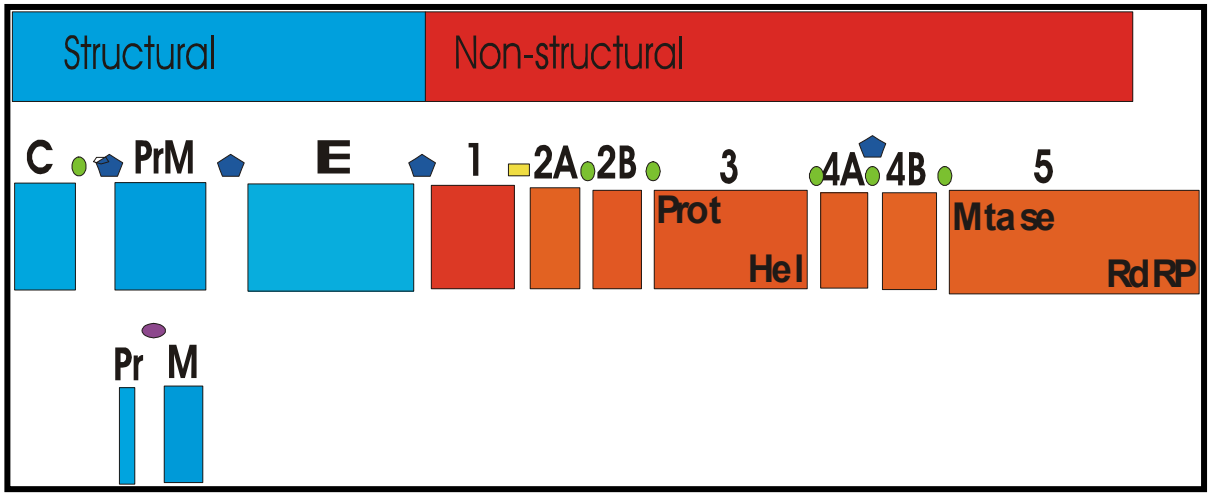


Figure 1.1: A representation of the WNV polyprotein processing. A green circle (●) indicates viral serine protease cleavage sites. Blue pentagon (⬟) indicates signal peptidase cleavage sites and purple circle (⬭) indicate cleavage with furin or furin-like enzyme. Cleavage between NS1 and NS2A (▭) are processed by unknown host enzyme. The following are also indicated on the figure: the serine protease (Prot) and helicase (Hel) on the NS3 protein and methyltransferase (Mtase) and RdRp domains on the NS5 protein.

1.5.2.1) Capsid protein (C)

The C protein is ~11 kDa in size and is rich in basic amino acids that forms the nucleocapsid encapsulating the viral RNA genome (Rice *et al.*, 1985). It is present in two forms in infected cells, the C_{arch} and C_{vir}. C_{arch} is a membrane-anchored protein that is produced when host cell signalases cleave a stretch of COOH terminal hydrophobic residues. C_{vir} is the mature form of the protein, which is produced when

this hydrophobic tail is cleaved by viral protease. It consists of charged residues at the N and C termini with an internal hydrophobic domain. This internal hydrophobic domain allows the mature C-protein to stay associated with the membranes of the endoplasmic reticulum (ER). The charged residues mediate RNA interaction (Beasley, 2005; Lindenbach and Rice, 2003).

1.5.2.2) Membrane protein (prM)

The prM protein is ~26 kDa (M is 8 kDa) and is translocated into the ER by the COOH terminal hydrophobic domain of the C-protein. Cleavage of the prM protein by signal peptidase will be delayed until this signal sequence is removed from the capsid (Lobigs, 1993; Yamshchikov and Compans, 1994). The 3' N-linked glycosylation sites and six conserved cysteine residues, which are di-sulphate linked, are present on the NH_2 terminal region of the pr part of the premembrane protein (Chambers *et al*., 1990a; Nowak and Wengler, 1987). The prM protein folds rapidly and produces a heterodimeric complex with the E-protein, shortly after synthesis. This co-synthesis is necessary for correct folding of the E-protein suggesting it may function as a chaperone. Maturation of virion particles occurs in the secretory pathway in parallel with cleavage of prM into pr and M protein by the golgi-resident furin or a furin-like enzyme of the host. The prM stabilises the E-protein and prevents it from undergoing rearrangement to the fusogenic form in the reduced pH environment of the early secretary pathway. Co-expression of prM- and E-proteins leads to the formation of subviral particles (SVPs), indicating that these proteins can associate and be released from infected cells without other viral components present (Konishi and Mason, 1993; Lorenz *et al*., 2002; Stadler *et al*., 1997). During virus release the prM-protein is cleaved by a host cell furin-like protease leaving only the smaller non-glycosylated M-protein and allowing dimerization of the E-proteins on the virion surface (Murray *et al*., 1993; Beasley, 2005; Lindenbach and Rice, 2003; Petersen and Roehrig, 2001).

1.5.2.3) Envelope protein (E)

The major virion surface protein is the envelope (E) glycoprotein (~53 kDa) and it is the most conserved structural protein. This structural protein is immunologically important, is the viral hemagglutinin and mediates virus-host cell binding and

membrane fusion. It elicits most of the virus neutralizing antibodies. The E-protein is a type I membrane protein that consists of 12 conserved cysteines that form six intramolecular disulfide bridges and contain up to three potential glycosylation motifs (Nowak and Wengler, 1987; Lindenbach and Rice, 2003). Most WNV strains encode a single glycosylation motif at residues 154-156 of the E-protein. The E protein only folds correctly if co-expressed with prM (Chambers *et al*., 1990b; Konishi and Mason, 1993; Lorenz *et al*., 2002). The 3D structure of the E protein of TBE was resolved with X-ray crystallography (Rey *et al*., 1995). There is a high amino acid sequence homology throughout the flavivirus genus and thus this structure is considered to be representative of the tertiary structure of the E protein for all flaviviruses (Rey et al., 1995; Lindenbach and Rice, 2003).

The envelope consists of the anchor domain and the ectodomain, which is divided into three domains. The central domain or domain I is discontinuous and consists of 120 residues and has an asparagine-linked glycosylation site on the E_0F_0 loop. The dimerization domain or domain II is discontinuous and consists of 180 residues. It has an extended finger like structure and a putative fusion peptide in the cd loop (Rey *et al*., 1995). This domain undergoes the most rearrangement of the three domains when exposed to acidic pH and has a hinge-like characteristic on its base part allowing projection of the cd loop toward the target cell membrane for participation in membrane fusion. Domain II is also the only domain containing flavivirus cross-reactive epitopes (Rey *et al*., 1995), while domain III and I have type and subtype-specific epitopes (Heinz, 1986; Mandl *et al*., 1989a). Domain III forms the COOH terminal end of the solubilized E protein and consists of 92 residues with putative receptor binding regions and it has an immunoglobulin-like conformation (Bork *et al*., 1994, Bhardwaj *et al*., 2001; Mandl *et al*., 2000). The residues directly following domain III do not form part of the ectodomain but form the anchor domain. This region is important for membrane anchoring of the E protein and interactions with prM. It undergoes pH-induced conformational changes and contains 2 predicted α-helical segments involved in stabilization of the prM/E interactions and trimerization of soluble E protein (Allison *et al*., 1999). This region also has two transmembrane segments that act as membrane anchors and/or signal sequences for the translocation of NS1 into the ER lumen (Mandl *et al*., 1989b).

1.5.2.4) NS1 protein

The non-structural 1 protein (NS1) is a membrane-associated glycoprotein thought to be involved in the early stages of virus replication and is ~46 kDa in size. It contains 2 N-linked glycosylation sites and 12 conserved cysteines that form disulfide bonds. NS1 is found on the cell surface and is also secreted from mammalian cells (Rice, 1996; Lindenbach and Rice, 2003; Mason, 1989). The NS1 protein is translocated into the lumen of the ER prior to cleavage at the E/NS1 junction by a cellular signalase (Chambers *et al*., 1990b). The type of proteolytic processing which occurs at the NS1-NS2A junction is still unknown but is most probably due to an ER resident host enzyme (Falgout and Markoff, 1995). As this protein consists largely of hydrophilic amino acids and has no putative transmembrane domain, the nature of this association remains controversial. One possibility is that the hydrophobic surface for peripheral association with membranes is formed by dimerization (Winkler *et al*., 1988; Winkler *et al*., 1989). The function of the extracellular forms of NS1 also still remains unknown but it was indicated that during infection strong humoral responses are produced against this protein (Falgout *et al*., 1990). It was also found that antibodies against the cell surface form could direct the complement-mediated lysis of flavivirus-infected cells (Henchal *et al*., 1988). Neither positive nor negative-strand RNA accumulates unless NS1 is supplied *in trans*, suggesting that NS1 function prior to or early in minus-strand synthesis (Lindenbach and Rice, 1999). NS1 co-localizes with vesicle packets, which is the most likely sites of RNA replication, thus NS1 is important for RNA replication (Mackenzie *et al*., 1996). Mutations at the N-linked glycosylation sites lead to dramatic defects in RNA replication and virus production (Muylaert *et al*., 1996).

1.5.2.5) NS2A protein

This protein is a medium size (~22 kDa), hydrophobic protein of which the N-terminus of NS2A is generated via the cleavage of NS1-2A by an unknown ER resident host enzyme (Falgout and Markoff, 1995). The C-terminus is generated by viral serine protease cleavage in the cytoplasm (Lindenbach and Rice, 2003; Mcminn, 1997). It is a poorly conserved membrane-associated protein whose function is unknown but this protein is most likely involved in coordinating the shift between RNA packaging and RNA replication (Khromykh *et al*., 2001).

1.5.2.6) NS2B protein

The NS2B protein is a small (~14 kDa) membrane-associated protein. It forms a complex with the NS3 protein because it is necessary as a cofactor for viral protease NS3 (Chambers *et al*., 1991) and has a conserved hydrophilic domain which is essential for its cofactor activity (Chambers *et al*., 1993). The conserved hydrophilic domain is flanked by hydrophobic regions necessary for cotranslational insertion of the NS2B-NS3 precursor into the ER membranes and for interaction between NS3 and NS2B (Clum *et al*., 1997). This protein may also be involved in modulating membrane permeability during infection (Chang *et al*., 1999).

1.5.2.7) NS3 protein

NS3 is the second largest viral protein with a size of ~70 kDa and is highly conserved. Sequence comparison and biochemical analyzes studies revealed that it is a trifunctional protein with protease, helicase and RNA triphosphatase activities (Rice, 1996; Chambers *et al*., 1991; Chambers *et al*., 1993). This protein is associated with membranes *via* it's interactions with NS2B. Significant homology is found between the NH_2 terminal region of the NS3 and serine protease. This region is subsequently thought to function in polyprotein cleavage in association with the NS2B protein (Chambers *et al.,* 1991; Chambers *et al.,* 1993; Falgout *et al*., 1993). The C-terminus also indicates homology with the family of RNA helicase proteins and it is thus postulated to be involved in RNA replication. Interactions between NS3 and NS5 would facilitate the coordination of the helicase, polymerase, and capping activities (Gorbalenya *et al*., 1989; Briton, 2002; Lindenbach and Rice, 2003; McMinn, 1997).

1.5.2.8) NS4A and NS4B proteins

Both NS4A (16 kDa) and NS4B (27 kDa) are small hydrophobic proteins and are poorly conserved membrane associated proteins of which the true functions remain unknown. These proteins may be necessary for localization and activity of the polymerase complex because they are strongly associated with the viral polymerase complex (Lindenbach and Rice, 2003; Mackenzie *et al*., 1998; Chambers *et al*., 1990b).

1.5.2.9) NS5 protein

NS5 is the largest viral protein with a size of 103 kDa, and is also the most conserved flavivirus protein. NS5 is a basic protein but does not contain long hydrophobic regions characteristic of *trans*-membrane domains (Lindenbach and Rice, 2003; Koonin, 1991). The COOH terminal portion of NS5 contains motifs characteristic of all RNA-dependent RNA polymerases (RdRps) and possesses a GDD motif as identified in other RNA-dependent RNA polymerases of positive RNA viruses (Rice *et al*., 1985; Rice *et al*., 1986). Homology between the N-terminus of NS5 and several methyltransferase proteins suggests that this domain is involved in methylation of the 5' cap (Koonin, 1993). NS5 can also be phoshorylated by an associated serine/threonine serine kinase(s) (Briton, 2002; McMinn, 1997).

1.6 MOLECULAR DETERMINANTS OF VIRULENCE

Initial replication of WNV after the bite of a mosquito is thought to occur in the skin and regional lymph nodes and produce a primary viraemia that seeds the reticuloendothelial system (RES). Virus may now seed the central nervous system (CNS) depending on the level of secondary viraemia that results from replication in the RES (Campbell *et al.*, 2002). Virus-specific as well as host–specific factors influence the level of viraemia and thus affect the clinical manifestation and disease outcome. The WNV envelope protein mediates cell attachment and neuroinvasiveness and seems to be a primary virulence factor (Campbell *et al.,* 2002; Chambers *et al*., 1998). The neuroinvasiveness of WNV refers to its ability to replicate in peripheral tissue, induce viraemia and invade the CNS. Neurovirulence also refers to the ability of the virus to initiate cytopathic infection in the CNS and cause encephalitis (Hurrelbrink and McMinn, 2003; Mcminn, 1997). The specific host factors that allow WNV entry into the CNS are unknown but most probably include factors that will promote virus entry and replication (Campbell *et al.*, 2002). Factors that increase the probability to develop meningoencephalitis include medical conditions that disrupts the cerebral endothelium e.g. hypertension and cerebrovascular disease. Increase in viraemia levels can be due to immunosuppression in patients (Campbell *et al.*, 2002). It is also known that Toll-like receptor 3 (Tlr3) facilitates WNV to penetrate the blood brain barrier and cause encephalitis. Tlr3 also enhances WNV replication in the CNS and thus induce neuronal injury by inflammation-induced cell death (Wang *et al*., 2004).

After WNV is inoculated into a murine model subcutaneously, virus replication is first detected in the draining lymph nodes (McMinn *et al.,* 1996). Plasma viraemia then develops and during viraemia extra neural tissues are infected and the release of virus from these tissues allows viraemia to persist for a few days. Virus enters the brain during the viraemic phase. Dendritic cells are the primary cellular targets of viral infection and efficient replication in these cells may be an important determinant of neuroinvasiveness, as well as free movement of virus to the bloodstream via efferent lymphatics (McMinn, 1997; McMinn and Sammels, 1997; Wu *et al*., 2000).

1.6.1) Molecular determinants of virulence of structural proteins

The structural premembrane and envelope proteins are embedded in the virus envelope and thus exposed to the host's immune system. They are also important for viral entry (Sections 1.5.2.2 and 1.5.2.3) and therefore virulence determinants are predominantly associated with these proteins. Because of the virulence importance of these proteins, their structure, function and molecular determinants of virulence will be briefly discussed below.

1.6.1.1) PreMembrane

The membrane protein is a hydrophobic integral membrane protein. Cleavage of the NH_2 terminal pr part from the COOH terminal of the premembrane takes place immediately before or during the release of virions from the infected cells resulting in mature membrane protein being produced. During virus assembly the premembrane proteins form a heterodimer with the envelope (Wengler and Wengler, 1989). This prevents the envelope from undergoing acid-catalysed conformational change during transport through the intracellular acidic compartments (Guirakhoo *et al*., 1992; Heinz *et al*., 1994). The shielding of the envelope by the premembrane is important for virulence because it prevents E-protein exposure to acidic pH and thus prevents the premature conversion of the E-protein to a fusion active form, a functional requirement for entry of the virus into the target cell (Hurrelbrink and McMinn, 2003). It was shown that if furin-mediated cleavage of pr from M is inhibited, it does not prevent release of virions from the host-infected cell but it does reduce the infectivity of these virions significantly (Guirakhoo *et al*., 1991). The prM-protein is directly

involved in preventing the premature dimer-to-trimer rearrangement of the E-protein, which is required for the fusion activity of this protein. The cleavage of the premembrane protein into its pr and M parts primes the E protein for reactivity upon exposure to the acidic conditions of the endosome (Hurrelbrink and McMinn, 2003). Only a few molecular determinants of the membrane protein that may be involved in pathogenesis were identified up to date. Attenuation studies were limited to the prM furin and glycosylation sites, as well as the prM/E signalase cleavage site. It was found that mutations in the furin cleavage site could either decrease or increase virulence in a TBE/DEN-4 chimeric virus (Pletnev *et al*., 1992). Mutations in the prM/E cleavage site of Langat (LGT) virus have also shown to attenuate neurovirulence in a mouse model (Holbrook *et al*., 2001). Mutations in the prM furin, glycosylation site as well as the cleavage site are responsible for phenotypic changes but mutations in other regions may also play a role. Further research needs to be performed in order to identify specific sites involved in virulence determination.

1.6.1.2) Envelope

The E-protein of WNV is the major glycoprotein and the principal target for neutralizing antibodies. Some of its functions are host-cell receptor binding, membrane fusion and cell entry and it appears to be an important determinant of neuroinvasiveness and neurovirulence in animal models (Lindebach and Rice, 2003; McMinn, 1997). Many molecular determinants of virulence have already been identified for other flaviviruses in the E protein. The E-protein was subdivided into 5 clusters of mutations based on the TBE E protein model (Rey *et al*., 1995). Clusters A-D are located in the ectodomain and cluster E in the anchor domain of the envelope protein. Molecular determinants are also further divided into those affecting neuroinvasiveness and neurovirulence. Figure 1.2 indicates the mutation clusters on the ectodomain of the envelope protein of TBEV (Hurrelbrink and McMinn, 2003).

The mutation clusters where identified by superimposing amino acid substitions/mutations of flaviviruses identified in previous studies onto the 3D structure of the TBEV to see if any pattern arises.

Ectodomain of TBE virus envelope protein

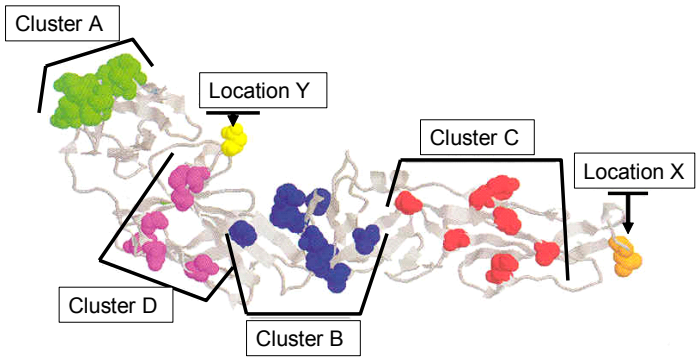


Figure 1.2: Superimposition of molecular determinants of virulence onto the three-dimensional structure of the TBEV E protein ectodomain (Rey *et al*., 1995). Four clusters of the mutations can be seen: cluster A (green), cluster B (blue), cluster C (red), and cluster D (purple), as well as two isolated mutations located within the fusion peptide (location X in orange) and the glycosylation site (location Y in yellow), respectively (Modified from Hurrelbrink and McMinn, 2003).

i. Cluster A: The Receptor binding Region

Mutations in cluster A are found on the lateral surface of domain III, the receptor binding site (Bhardwaj *et al*., 2001; Mandl *et al*., 2000). An Arginine-Glycine-Aspartic acid (RGD) motif is present in cluster A in the mosquito-borne flaviviruses (Hurrelbrink and McMinn, 2003). This RGD motif was subjected to mutagenesis studies and results indicated that mutations in this specific region can lead to attenuation of neuroinvasiveness and this may be related to the affinity of the virus for glycosaminoglycans on the host-cell surface (Hurrelbrink and McMinn, 2003; Lee and Lobigs, 2000; Lee and Lobigs, 2002). Further observations showed that mutations, which increase the net positive charge of the protein, enhance virus binding to heparin resulting in attenuation. Virus is rapidly removed from the bloodstream of infected mice, and thus virus is prevented from spreading from the extra neural sites of replication into the brain (Lee and Lobigs, 2002). It can also be that mutations in this region lead to misfolding of the E-protein which would lead to delays in virion assembly, possibly reducing viral titers and subsequently influence neuroinvasiveness (Hurrelbrink and McMinn, 2003). These mutations may be responsible for inhibition of heterodimer formation with prM and homodimer formation with E on the virion surface. Receptor binding is very important for virion infectivity. Receptor binding can be severely inhibited by any structural changes in the tertiary structure of the protein. Even minor mutations can lead to major effects on virulence.

Changes in and around residues involved in the formation of salt bridges on the lateral face of domain III were sufficient to destabilize the TBEV and cause reduction in neuroinvasiveness (Mandl *et al*., 2000). Furthermore, when the Asp residue in the RGD motif is substituted with a polar or nonpolar residue like tyrosine or asparagine respectively it causes a loss of neuroinvasiveness in mice, whereas the substitution of an alternative negatively charged residue has no effect on virulence (Hurrelbrink and McMinn, 2003). It was concluded by Hurrelbrink and McMinn, (2003) that it is possible that this and other residues in the RGD motif form structurally important salt bridges with adjacent residues in the lateral face of domain III or that charge interactions are important for the stability of receptor-ligand and/or E protein complexes.

ii. Cluster B: The Hinge Region

Cluster B is formed by mutations situated on the polar interface, linking domains I and II (Hurrelbrink and McMinn, 2003). From the 3D structure of the TBE E protein (Rey *et al*., 1995) it was suggested that this region forms part of a molecular hinge. It is further predicted to be involved in projection of the fusion peptide on the tip of domain II upwards for contact with the endosomal membrane during fusion (Rey *et al*., 1995). Studies of mutations in the hinge appear to directly disrupt low pH-mediated fusion (McMinn *et al*., 1995). It is still unclear whether mutations in the hinge cause defects in fusion by perturbing the conformational change of the protein at low pH or by disrupting the receptor-ligand complex (Hurrelbrink and McMinn, 2003; Cover *et al*., 2000). Although the exact nature of the mechanism of attenuation caused by mutations in the hinge is unclear, it is known that a delay in viral entry occurs. This affects the spread of the virus and allows reduction of neuroinvasiveness (Hurrelbrink and McMinn, 2003).

iii. Locations X and Y: The fusion Peptide and Glycosylation site

The mutation site in location X is positioned within the fusion peptide that is situated on the tip of the cd loop of domain II. This region is conserved throughout the *Flavivirus* genus (Hurrelbrink and McMinn, 2003) and is the hydrophobic region of the protein with residues involved in the initiation of fusion between viral and cellular membranes (Allison *et al*., 2001). When substituting hydrophobic residues at position

107 of TBE with hydrophilic amino acids, fusion activity was either strongly impaired or abolished (Allison *et al*., 2001).

Location Y is the conserved glycosylation site of flaviviruses (Hurrelbrink and McMinn, 2003). Mutations found in this site have been shown to reduce neurovirulence (Pletnev *et al*., 1993). DEN-2 virus selection of fusion mutants by repeated exposure to acidic pH or ammonium chloride has shown to induce mutations at the glycosylation site, in parallel with an increase in the optimal pH for fusion. The increase in the optimal pH may be linked to the proposed interaction of the E_0F_0 loop which contains the ASN-linked glycosylation site with the buried fusion peptide on the cd loop of an adjacent E protein monomer (Guirakhoo *et al.,* 1993). It is therefore possible that the N-linked glycan stabilises the dimer, thereby preventing the premature triggering of the dimer-to-trimer transition (Hurrelbrink and McMinn, 2003).

iv. Clusters C, D and E: The prM/E Interface

Mutations in these clusters have been implicated as potential determinants of virulence in a number of flaviviruses; Dengue (DEN 1-4), WNV, Yellow Fever (YF), Tick borne encephalitis (TBE) and Langat virus LGT (Lee *et al*., 1997; Chambers and Nickells, 2001; Chambers *et al*., 1998; Holbrook *et al*., 2001; Mandl *et al*., 1989b). Cluster C mutations are situated towards the distal end of domain II in *β* strands b, d, and j and in the bc loop. These mutations occur in a region of the E protein believed to be involved in trimer contacts with the proximal stem-anchor of a neighbouring monomer (Ferlenghi *et al*., 2001). It is important to take note that their distal location on domain II places them close to the predicted fusion peptide. By placing these mutations of cluster C on the 3D structure of the E protein it became clear that most of them are located on the upper surface opposed to the basal or lateral faces where contact with stem-anchor is likely to occur. It thus seems likely that these mutations in cluster C directly disrupt lateral interactions with prM and/or other E protein monomers. It is also possible that the mutations can alter the conformation of the distal tip of domain II and disturbs presentation of the fusion peptide during fusion (Hurrelbrink and McMinn, 2003). In the prM/E interface area, prM can potentially form contacts with domain I of an E-protein monomer and domain II of another E protein monomer. It was observed that many mutations along the extended finger of domain

II in the C cluster are directly adjacent to the proposed position for prM. The same situation is seen with the mutations in cluster D (Hurrelbrink and McMinn, 2003). The α-helical domain of the stem-anchor region is implicated in both the trimerization of soluble E-protein and the stabilisation of prM/E interactions and it is in this area where the mutations of cluster E resides (Allison *et al*., 1999). It may be possible for the mutations of clusters C-E to have a direct impact on the stability of the interactions at the prM/E interface and thus disrupt the trimerization of E and the stability of the glycoprotein network on the virion surface (Hurrelbrink and McMinn, 2003).

1.6.2) Molecular determinants of virulence of non-structural proteins

NS2B and NS3 form the viral encoded protease that is required for cleavage of the conserved dibasic sites at the NH_2 terminal of NS2B, NS3, NS4A and NS5 (Chambers *et al*., 1991; Ryan *et al*., 1998) as well as the COOH terminal ends of anchored-capsid and NS4A (Amberg *et al*., 1994; Nowak *et al*., 1989). The NS2B central conserved domain functions as the obligatory cofactor needed for the enzymatic activity of the NS3 protein (Falgout *et al*., 1993). The protease function of the NS2B-NS3 complex is mediated by the first part of the NH_2 terminal of NS3 (Chambers *et al*., 1990c). This also contains a Histidine-Arginine-Serine catalytic triad, typical of serine proteases. This region further also serves as a helicase and an RNA dependent nucleoside triphosphatase during RNA replication for which the motif is located at the COOH terminal two thirds of the NS3 protein (Murthy *et al*., 1999; Ryan *et al*., 1998).

The precursor replicase complex of the virus consists of the NS2A, NS3 and NS5, which are bound to a 3` RNA stem-loop structure (Khromykh *et al*., 1999). In this complex the NS2A protein binds to membrane bound NS4A, which in turn it binds with hydrophilic extensions in the lumen to dimeric NS1 to form the complete replicase complex (Khromykh *et al*., 1999). NS2A is also implicated to assist with the assembly and release of infectious particles from the infected cells (Kümmerer and Rice, 2002). The NS5 protein contains an RNA-dependent RNA polymerase (RDRP), *S*-adenosylmethionine methyl transferase (SAM) and importin-*β* binding motifs; it is thus believed to be the protein responsible for the cytoplasmic RNA replication by the replicase complex (Koonin, 1993; Bartholomeusz and Thompson, 1999).

By analysing the functions of the non-structural proteins, the possible effects that mutations can have on the virulence of the virus can be predicted. If a mutation reduce the protease activity of the NS2B-NS3 complex it will lead to reduced processing of the polyprotein and in return affect virion assembly and release, allowing the host immune system to neutralize infection.

1.6.2.1) The Viral Protease: NS2B-NS3

Mutations in and around the substrate-binding pocket seem to affect the protein function and viral replication. Specific mutations in the histidine-aspartic acid-serine catalytic triad severely inhibit protease activity of NS3 and inhibit subsequent viral replication. Mutations in the NS2B cofactor region or in the NS2B-NS3 autocatalytic cleavage site also appear to reduce replicative ability *in vitro* (Hurrelbrink and McMinn, 2003; Chambers *et al*., 1990c). Other mutations outside the substrate-binding pocket of the NS3 protease also affect viral replication (Valle and Falgout, 1998). Mutations found in a mutagenesis study (Matusan *et al.,* 2001a) were used in computer modelling of the NS2B-NS3 complex and results suggested that some of the mutations affected the stability of the complex and others disrupted the formation of functionally relevant salt bridges and/or perturbed substrate-binding specificity. Mutations other than those, which directly affect NS3 protease activity like the formation of an NS2B-NS3 complex, also have an effect on the virulence phenotype of the virus (Hurrelbrink and McMinn, 2003). In a study performed with Yellow Fever virus (Chambers *et al.,* 1993) it was indicated that mutations in the conserved central domain of NS2B eliminate its ability to associate with NS3 as well as its *trans*-cleavage activity but mutations in the flanking hydrophobic domains had little effect.

1.6.2.2) The viral replicase Complex: NS1, NS2A, NS3, NS4A and NS5

The replicase complex performs all the functions required for transcription. Some of these functions are to stabilise and transport the replicase complex to the intracellular membranes, unwinding double-stranded templates, primer extension and adding of methylated caps to RNA transcripts (Hurrelbrink and McMinn, 2003). Mutations can affect the stability of the complex or directly affect the enzymatic activities and therefore have an impact on virulence. Mutations in the methyltransferase and/or

RNA polymerase motifs of NS5 have been shown to affect the virulence of YFV vaccines strains passaged in mouse brain or Vero cells and to abolish the infectivity of clone-derived Kunjin virus (Matusan *et al*., 2001b; Holbrook *et al*., 2000; Khromykh *et al*., 1998).

Many studies have been done on the pathogenicity of wild-type and attenuated strains of flaviviruses. Studies comparing full genome RNA sequences of pathogenic and attenuated strains are summarised in Table 1.1 indicating that possible mutations leading to attenuated strains are potentially scattered throughout the genome and not restricted to constrained parts of the genome.

Table 1.1 Summary of recent studies using full genome sequences of attenuated lineage I WNV strains. The table indicates in which protein an amino acid substitution took place as well as the position of the amino acid in the protein.

Strain / GenBank accession number	**Phenotype/ virus characteristics**	**Reference**	**prM**	**E**	**NS1**	**NS2 B**	**NS4 B**	**NS5**
2003 WNV isolate Bird 1153 from Texas (AY12945)	Small plaque formation and temperature sensitive as well as mouse-attenuated phenotype. Compared to prototype WNV NY99 strain (AF196835)	Davis *et al.*, 2004	**156 V→I**	**159 V→A**			**249 E→G**	**804 A→V**
WNV Mexico strain TM171-03 (AY660002)	Attenuated in efficiency of neuroinvasiveness tested in mice. Compared to prototype WNV NY99 strain (AF196835)	Beasley *et al.*, 2004	**141 I→T**	**156 S→P**			**245 I→V**	**898 T→I**
WNV 385-99 clone 9317A (AY848695)	Loss of virulence for hamsters and a change in plaque morphology as well as persistent renal infections in hamsters. Compared to WNV stain NY385-99(AY842931)	Ding *et al.*, 2005		**167 L→F**				
WNV 385-99 clone 9317B (DQ066423)	Loss of virulence for hamsters and a change in plaque morphology as well as persistent renal infections in hamsters. Compared to WNV stain NY385-99(AY842931)	Ding *et al.*, 2005		**167 L→F**				**167 M→I**
WNV 385-99 clone 9317E (AY848696)	Loss of virulence for hamsters and a change in plaque morphology as well as persistent renal infections in hamsters. Compared to WNV stain NY385-99(AY842931)	Ding *et al.*, 2005	**21 V→M**	**167 L→F**	**183 I→T**	**99 M→I**		**390 E→G**
WNV 385-99 clone TVP9376 (AY848697)	Loss of virulence for hamsters and a change in plaque morphology as well as persistent renal infections in hamsters. Compared to WNV stain NY385-99(AY842931)	Ding *et al.*, 2005	**21 V→M**	**167 L→F**	**183 I→T**	**99 M→I**		**390 E→G**

1.7 POLYPROTEIN CLEAVAGE

Secretory proteins need transient NH_2 terminal signal sequences to initiate export across the ER. These signal sequences/signal peptides have a low degree of sequence conservation but have common structural motifs. They can be divided into three structural/function regions. Firstly a basic NH_2 terminal region (N-region), which is a major determinant of the transmembrane topology of integral membrane proteins, then a central hydrophobic region (H-region), and lastly a polar COOH terminal region (C-region). The C-region consists of small residues, like alanine in positions –1 and –3 upstream of the cleavage site and have an extended conformation (Von Heijne, 1986).

WNV polyprotein cleaves into its seperate proteins by one of two methods; 1) it can be cleaved by a cellular enzyme (signal peptidase) in the lumen of the ER or 2) by the virus serine protease (NS2B-NS3) in the cytoplasm of host cells (Rice, 1996). A third protease is required for cleavage of the COOH terminus of the NS1 protein (Falgout and Markoff, 1995). The signal peptidase cleavage between the C-prM and NS4A-NS4B proteins is regulated in such a fashion that the cellular signal peptidase cleavage in the lumen will only occur efficiently after cleavage upstream of the signal sequence mediated by the viral protease has occurred. This co-expression of the viral protease (NS2B-NS3) with the structural polyprotein region enhances the efficiency of signal peptidase cleavage at the NH_2 termini of prM and NS4B (Yamshchikov and Compans, 1993). The biological role of the sequential order of cleavage at the flavivirus C-prM junction was investigated by Lee and colleagues by introducing a mutation in the COOH terminus of the YFV prM signal sequence. The mutation in the COOH terminus uncoupled efficient signal peptidase cleavage of the prM protein indicating the pre-requisite of prior cleavage of the C protein by viral protease *in vitro* (Lee *et al*., 2000). They also made the observation that this mutation enhanced cleavage by signal peptidase but suppressed infectious virion production. A possible effect of this rapid signal peptidase cleavage of prM could be the production of a membrane-anchored form of the C protein as the predominant processing intermediate; this will be deleterious for virus replication, if the membrane-anchored C protein functioned poorly or not at all as a substrate for viral protease (Lee *et al.*, 2000).

In other studies where the entire structural polyprotein region was expressed, the signal peptidase-mediated cleavage at the NH_2 terminus of the prM did not occur efficiently, whereas at the NH_2 terminus of the envelope protein did. It was found that when prM is insufficiently produced it affects the production and lowers the secretion of prM-E heterodimers. Furthermore, when these constructs were used in vaccination studies a lack in immunogenicity was observed (Stocks and Lobigs. 1998). A mutagenesis study on Murray Valley encephalitis virus (MVE) was performed in order to identify elements in the flavivirus C-prM region that could be subjected to mutagenesis in order to overcome the controlled order of cleavage at the C-prM junction (Stocks and Lobigs, 1998). A combination of only three amino acid substitutions could override the influence of the C protein on the signal peptidase cleavage of prM (Stocks and Lobigs, 1998).

1.8 DETECTION AND DIAGNOSTIC METHODS FOR WNV

Different detection methods are available for WNV. Routinely, PCR based methods are used to detect WNV in host species e.g. mosquitoes. This test is not routinely used in diagnosis of WNV for humans due to the short period of viremia and low viral load in human hosts. It is better to use serological methods, which detects antibodies to the virus and not the virus genomic RNA (Petersen and Marfin, 2002). Serological methods used for diagnosis of WNV are: 1) Complement fixation 2) Hemagglutination inhibition 3) Plaque reduction neutralization test 4) Immunofluorescence assay 5) ELISA and 6) Microsphere immunoassay (Prince and Hogrefe, 2005; Beasley, 2005). These different detection methods of flavivruses with specific focus on WNV infection will be discussed briefly.

1.8.1) Complement fixation

The complement fixation (CF) assay utilises the ability of antigen–antibody complexes to trigger the complement cascade of the immune system (Palmer and Whaley, 1986). In this assay the absence of hemolysis indicates the presence of antibody to the antigen in question. CF antibody induced by flaviviruses infections appear two weeks after onset of disease, and the levels begin to decrease about 2 months later, reaching baseline levels in one to two years (Manath, 1995). The detection of CF antibodies is an indication of a recent infection. Cross-reactivity

among flaviviruses poses a problem and thus renders this assay non-specific for WNV and some flavivirus-infected individuals do not produce complement-fixing antibodies. Another problem with this method is that it is a highly complex assay and thus needs highly trained personnel and is labour intensive with the need for strict control of reagents for quality control (Prince and Hogrefe, 2005; Kuno, 2003).

1.8.2) Hemagglutination inhibition

Hemagglutinins of flavivirus E-proteins have the ability to bind to and agglutinate avian erythrocytes. Antibodies from infected people will block agglutination of the erythrocytes by WNV (Petersen and Marfin, 2002; Kuno, 2003). The hemagglutination inhibition test (HI) thus relies on this ability of flaviviruses. Agglutination of erythrocytes indicates an absence of WNV-specific antibodies in serum sample, and agglutination inhibition is indicative of the presence of WNV-specific antibodies in the patient serum (Shi and Wong, 2003). Problems with this method is the cross-reactivity among flavivruses that lowers the specificity and reagents cannot be stored for long periods and needs strict control for quality purposes (Prince and Hogrefe, 2005).

1.8.3) Plaque reduction neutralization test

The plaque reduction neutralization test (PRNT) is seen as the gold standard for measuring WNV-specific antibodies because it is a very specific assay (Kuno, 2003). For this assay a serial dilution of the patient's sera is prepared, which is then incubated with a WNV preparation containing a defined number of infectious units. Serum-virus mixture is then added to culture wells containing monolayer of Vero cells. After culturing for a defined number of days, the number of plaques is determined. Because WNV-specific antibodies in the serum binds to WNV envelope proteins of the virus and neutralize it, this inhibits the ability of the virus to infect the Vero cells. This can be observed in the reduction of the number of plaques (Haley *et al*., 2003). The highest serum dilution reducing plaque formation by a given level (80%-90% reduction in plaque count) is defined as the endpoint titre (Shi and Wong, 2003). The main problem of this assay is related to safety issues because of the need to work with infectious virus. It also is labour intensive and needs technical

expertise. The test is also very time consuming taking between 6-10 days to complete (Prince and Hogrefe, 2005; Martin *et al*., 2000).

1.8.4) Immunofluorescence assay

The immunofluorescence assay (IFA) depends on detection antibodies to recognise WNV antibodies in a test sample that is binded to WNV antigen fixated on a microscope slide. Briefly, a microscope slide is coated with fixed WNV-infected cells to which diluted serum or cerebrospinal fluid (CSF) is added. Secondary fluorescent-labelled anti-human immunoglobulin G (IgG) or anti-human IgM is then added and the microscope slide is examined using a fluorescent microscope (Shi and Wong, 2003). Serum specimens tested for IgM must first be diluted in sample buffer containing anti-human IgG to remove IgG in sample, which will prevent false negative IgM and also eliminates false positive IgM due to attachment of IgM rheumatoid factor to bound WNV-specific IgG (Martins *et al*., 1995; Shi and Wong, 2003). The advantage of the IFA over the other assays (CF, HI, PRNT) is that it does not require complex antigen preparation procedures and in contrast to PRNT, risk of accidental infection of the person performing the assay is eliminated by fixation of infected cells. The assay only takes one day and the sample volume required is lower. This method can also separately measure IgM and IgG antibodies as opposed to the total antibody count of the other methods and can therefore be used to measure approximate time since infection. IgG negative and IgM positive IFA results are indicative of a very recent infection, whereas an IgG positive and IgM negative IFA result are indicative of a past infection (Shi and Wong, 2003).

1.8.5) Enzyme-linked immunosorbent assay

The enzyme-linked immunosorbent assay (ELISA) is the most frequently used laboratory approach for diagnosing WNV cases. Available ELISA systems can be divided into ELISA`s using native WNV antigens and those using recombinant antigens.

1.8.5.1) ELISA using native antigens

The classic indirect ELISA for WNV IgG or IgM was designed in 1985 by Feinstein and colleagues (Prince and Hogrefe, 2005). WNV antigen from infected Vero cells was used to coat the wells after which patient serum was added and then anti-human IgG/M antibodies conjugated with alkaline phosphates. Substrate is added and the chromogenic product is measure spectrophotometrically. The indirect IgG ELISA is sensitive and specific but indirect IgM lead to false positives due to the rheumatoid factors reacting with IgG bound to the WNV antigen

In response to overcome problems encountered with the rheumatoid factors the IgM antibody capture (MAC) ELISA was next developed (Martins *et al*., 1995; Martin *et al*., 2002). In this ELISA the wells were first coated with anti-human IgM antibodies to which serum was then added allowing IgM in the serum to bind to the IgM coated on the well. Next WNV antigen was added which will bind to IgM that recognise the WNV antigen. In this method only partially purified antigen can be used so there is no need for highly purified antigen.

An IgM capture ELISA that uses unlabeled WNV antigen from infected suckling mouse brain and an enzyme-linked monoclonal antibody specific for the flavivirus group (monoclonal 6B6C-1) was developed (Martins *et al*., 1995; Martin *et al*., 2000). For results to be valid, the absorbance values obtained from infected mouse brain antigen had to be twice that of the uninfected mouse brain (Levinson and Miller, 2002). Results were expressed as a positive-to-negative (P/N) ratio. The P/N ratio was calculated by dividing the absorbance values of the patient specimen using infected mouse brain antigen by the values obtained for a known negative sample specimen using the same antigen (Prince and Hogrefe, 2005).

The indirect flavivirus IgG ELISA was modified by using monoclonal antibody reactivity (Johnson *et al*., 2000). Flavivirus monoclonal antibody (4G2) is used to coat the well. Flavivirus antigen is added and then the diluted test serum is added. Antibodies in the serum recognise the flavivirus antigen, which will then attach to the antigen. Enzyme-labelled anti-human IgG is next added and then the substrate. The advantage of this antigen-capture IgG ELISA is that highly purified flavivirus antigen is not required, since non-viral antigen does not bind to the capture monoclonal

antibodies. It also has a low level of complexity and reagents are readily available (Johnson *et al*., 2000).

An indirect IgG/M capture ELISA that utilised native WNV antigen prepared from infected Vero cells was developed (Focus Technologies, Inc (Cypress); Prince and Hogrefe, 2003b; Prince and Hogrefe, 2005). The IgM-capture ELISA is similar to the previous described ELISA with the difference that it utilizes the unlabeled 6B6C-1 monoclonal antibody and enzyme-labelled goat-anti-mouse IgG. Results were expressed as an index and not as a P/N ratio. The index is calculated by dividing the absorbance value of the patient specimen by the absorbance value of a calibrator serum. A problem with non-specific IgM reactivity mediated by rheumatoid factor and/or heterophile antibody was encountered, which lead to false positive results. A background subtraction procedure similar to the one used in the previous ELISA was employed (Prince and Hogrefe, 2005).

1.8.5.2) ELISA using recombinant antigens

Production of recombinant WNV antigen supplied a safer alternative to propagation of virus in mice and working with potentially infectious material. It was shown that it is feasible to use recombinant WNV antigens to detect WNV antibodies in a sandwich ELISA format (Wang *et al*., 2002). Recombinant WNV envelope protein produced in an *E. coli* system was used that reliably detected IgG WNV antibodies in serum from infected individuals. Three patients' sera with other flaviviral infections (Dengue, Japanese encephalitis virus and Yellow fever) were tested and none yielded positive results with this ELISA indicating no cross-reactivity.

A plasmid with the premembrane and envelope genes of Japanese encephalitis virus was constructed (Chang *et al*., 2000). Transformed mammalian COS-1 cells secreted a non-infectious subviral particle that was successful as antigens in an IgM capture ELISA. The same procedures were repeated with WNV and similar results were obtained. This WNV recombinant antigen was used in IgM capture ELISA and monoclonal-based antigen capture IgG ELISA in the place of suckling mouse brain antigen (Prince and Hogrefe, 2005; Davis *et al*., 2001). Cross reactivity was also obtained with these ELISAs as when using native antigen. It is thus still necessary to perform additional testing to identify the infecting flavivirus.

The diagnostic division of Focus Technologies also used the above-mentioned recombinant WNV antigen to develop an indirect WNV IgG ELISA kit and a WNV IgM capture ELISA kit (Focus Technologies, Inc (Cypress); Prince and Hogrefe, 2003a; Prince and Hogrefe, 2003b; Prince and Hogrefe, 2005). These kits received FDA clearance in 2003. In the IgG ELISA the wells were coated with recombinant WNV antigen and peroxidase-labelled goat anti-human IgG was used. The IgM wells was coated with rabbit anti-human IgM and used unlabelled recombinant WNV antigens and peroxidase-labelled anti-flavivirus monoclonal antibody (6B6C-1).

1.8.6) Microsphere immunoassay

The microsphere immunoassay (MIA) can simultaneously identify many different flaviviruses antibodies timeously and at a reasonable cost. This eliminates the need to identify the causative flavivrus with secondary tests when using other diagnostic methods e.g. ELISA. Screening assays for flaviviruses all have some degree of cross-reactivity problems and thus there is a need for more specific tests (Prince and Hogrefe, 2005). A MIA system using multiple polystyrene bead sets was developed (Luminex Corporation (Austin, TX); Shi and Wong, 2003). Each bead set contains distinctive proportions of red and orange fluorescent dyes that will yield a signature fluorescent pattern when analyzed by a modified flow cytometer. For detecting IgG to different antigens, each of these distinctive fluorescent bead sets can be covalently linked to a different antigen. The bead sets are then mixed together in a single reaction well, serum is then added, allowing antibody recognition of the antigens. Next goat anti-human IgG conjugated to a fluorescent reagent is added as a reporter antibody and binds to the captured IgG`s. The bead mixture will then be simultaneously analyzed for fluorescent patterns and the reporter antibody. The reporter fluorescence intensity will be directly proportional to the amount of antigen-specific IgG bound to a given bead set. Advantages of this system are its sensitivity compared to ELISA systems due to its broad diagnostic range, more available antigenic epitopes and superior fluorescence (Shi and Wong, 2003; Prince and Hogrefe, 2005). Less specimen volume is required because multiple assays are performed in a single reaction, high precision of the assay eliminates the need for replicate testing and the MIA is cost effective.

1.8.7) Nucleic acid testing (NAT)

Standard reverse transcription polymerase chain reaction assays (RT-PCR) and quantitative real-time methods e.g. Taqman and Nucleic acid sequence based amplification (NASBA) assays can be used to detect the presence of WNV genomic RNA (Lanciotti, 2003). These methods have a short turn around time, with real time PCR being more sensitive than conventional RT-PCR. Highly specialised amplification or detection equipment that is expensive is needed (Beasley, 2005). The Loop mediated isothermal amplification (LAMP) is an alternative nucleic acid amplification strategy that can overcome these drawbacks (Parida *et al*., 2004). This LAMP allows amplification of templates at a single temperature and real time detection by measuring turbidity in the assay tube (Parida *et al*., 2004; Beasley, 2005)

1.9 DIFFERENT METHODS FOR THE PRODUCTION OF WNV ANTIGEN

Different expression systems can be used to produce recombinant proteins of a desired agent. The systems discussed here are bacterial expression, insect cell expression or mammalian cell expression systems. Previous studies on flaviviruses that have used these systems to produce recombinant flavivirus proteins are discussed below.

1.9.1) Bacterial expression

Recombinant envelope, membrane and NS 1 proteins of West Nile virus were produced by transforming the bacteria *Escherichia coli* with plasmid containing gene inserts of either the envelope gene or the premembrane gene or the non-structural 1 gene (Wang *et al*., 2002). Proteins expressed during this study were evaluated in an ELISA and an immunoblot. Only the recombinant E-protein elicited an antibody response during these tests and only IgG and IgM antibodies were picked up in horses exposed to the recombinant proteins. Investigators also tested the ELISA with three patients with other flavivirus infections and found no cross-reaction (Wang *et al*., 2002). Thus the bacterial expression shows the potential to be used as an expression system for recombinant proteins used in diagnostic ELISA's. This system must first be evaluated better to prove that proteins expressed in this system will be

folded correctly and be antigenic enough to prevent cross-reaction between flavivirus when tested on a larger scale and with different strains.

1.9.2) Insect cell expression

Recombinant baculoviruses were constructed containing different combinations of gene inserts of the Japanese encephalitis virus genome. All of them contained partial or full-length sequences of the E and NS1 glycoproteins. These recombinant baculoviruses were used to infect *Spodoptera frugiperda* (SF) cells to produce recombinant protein. Immunization of mice with recombinant E glycoprotein showed 70% protection against live WNV challenge. Protection elicited by the baculovirus produced recombinant E protein was superior to recombinant E protein produced in *E. coli*. A possible reason is the lack of glycosylation of the E protein in *E.coli* systems. It was shown that the baculovirus expressed E and NS1 proteins were at least partially glycosylated (McCown *et al.*, 1990). Recombinant baculoviruses were produced which contained the full-length JEV E protein with a N-terminal 24 amino acid signal sequence derived from its adjacent prM, and a COOH terminal six-histidine tag for immobilised metal affinity chromatography (IMAC) purification (Wu *et al.*, 2003). The JEV E protein was not secreted but produced in the cytoplasm of infected insect cells with recombinant baculovirus. Mice immunised with the recombinant E protein produced in this study successfully induced neutralising antibody responses and protective immunity towards a lethal dose of JE virus (Wu *et al.*, 2003). It was also demonstrated that authentic JEV prM, E and NS1 proteins can be produced in baculoviruses systems (Matsuura *et al.*, 1989). Expressed E protein in this study was glycosylated, had the same size as wild-type E protein, reacted to anti-E monoclonal antibodies and also induced antibody responses in mice.

Recombinant baculoviruses were produced with E gene or NS1 gene or E/NS1 genes of Yellow fever virus (YFV) (Despres *et al.*, 1991). Insect (SF) cells were transfected with these recombinant viruses and it was found that the expressed E and NS1 proteins were similar to native proteins. Mice were immunized with lysates of infected cells and protection against lethal YFV encephalitis was achieved (Despres *et al.*, 1991). No significant protection was achieved with NS1 alone. Recombinant E protein also elicited a low but significant level of neutralising antibodies.

A hybrid dengue virus E protein molecule was constructed by producing a recombinant baculovirus consisting of 36 amino acids from the membrane protein, the NH_2 terminal 288 amino acids of the dengue-2 virus E protein and amino acids 289-424 of the dengue–3 virus E protein (Bielefeldt-Ohmann *et al*., 1997). It has been engineered for secretion expression by fusion to a mellitin secretion signal sequence and truncation of the hydrophobic transmembrane segment. Recombinant E protein was secreted into the culture medium. This hybrid molecule reacted with a panel of dengue virus- and flavivirus-specific Mabs that recognise linear or conformational epitopes on dengue virions (Bielefeldt-Ohmann *et al*., 1997).

The baculovirus expression system was also used to express recombinant envelope proteins for Tick-borne encephalitis virus (Marx *et al*., 2001). Two forms of the E gene were expressed; a full length E and a truncated form that has a stop codon at amino acid position 435. Both recombinant proteins had a his-tag for purification. Poor yields were found with the full-length E-protein and high yields for the truncated form of the E-protein. The truncated form was used in an ELISA as well as in an immunoblot assay to detect TBEV-specific antibodies in sera from immunised human blood donors (Marx *et al.*, 2001). This E protein exhibited the antigenic epitopes and conformation necessary for specific antigen-antibody recognition (Marx *et al.*, 2001).

Recombinant WNV particles were produced in insect cells containing the prM and E proteins or the prM/E and C proteins (Qiao *et al.*, 2004). Insect cells secreted WNV-like particles containing these recombinant proteins. These WNV-like particles had the same neutralizing activity and protected mice against challenge with WNV (Qiao *et al.*, 2004).

1.9.3) Mammalian cell expression

Tick-borne encephalitis virus envelope protein was transiently produced in mammalian (COS-1, monkey) cells in the presence or absence of the prM protein (Allison *et al*., 1995). Five different plasmid constructs where produced. Plasmids contained; 1) full length wild-type protein E gene; 2) full length wild type protein E and prM genes; 3) the same as (1) with a stop codon at amino acid 435 in the E protein; 4) the same as in (2) with the stop codon at amino acid 435 in the E protein. Plasmids

(3) and (4) thus produced a truncated form of protein E lacking the COOH terminal 62 amino acids from the membrane anchor region); 5) only the prM gene. Mammalian (COS-1) cells were transfected using CsCl-purified plasmid DNA by electroporation. In this study it was found that the formation of a heteromeric complex with prM was necessary for efficient secretion of both forms of E-protein (Allison *et al.*, 1995). Only low levels of anchor free E was secreted in the absence of the prM protein. The prM-mediated transport of E could also be obtained by co-expression of prM and E from separate constructs. Further observations was that full-length E formed stable intracellular heterodimers with prM and was secreted as a subviral particle, whereas anchor-free E was not associated with particles and formed a less stable complex with prM. This suggested that prM interacts with both the ectodomain and the anchor domain region of E. To conclude the prM gene needs to be present for the E gene to be expressed and processed correctly (Allison *et al.*, 1995).

Recombinant TBEV antigens were produced (Yoshii *et al*., 2003) and used in the development of an ELISA. ELISA results correlated with results found with an commercially available ELISA, with the exception that the developed ELISA did not show cross-reactivity with JEV as with the commercial ELISA (Yoshii *et al.*, 2003). The full length of the prM and the E genes were used as inserts to construct these recombinant proteins. Recombinant proteins retained their native form and mammalian cells released virus-like particles as shown in other expression studies (Yoshii *et al.*, 2003).

Plasmid with Japanese encephalitis virus prM and E genes were constructed (Hunt *et al*., 2001). This plasmid construct was used to transform mammalian (COS-1) cells. The cell line was modified in order to stably produce non-infectious recombinant antigen expressed as extracellular particles (Hunt *et al.*, 2001). Western blot analysis showed that extracellular particles contained the expressed envelope-protein, premembrane and membrane proteins. Recombinant proteins were analyzed in an IgM-antibody-capture ELISA and indirect IgG ELISA. Results were similar to those found with purified mouse brain antigens and purified Japanese encephalitis virus as plate-bound antigen respectively.

WNV recombinant proteins were produced in COS-1 cells by inserting the prM and E genes as a single unit into an eukaryotic expression vector (Davis *et al.,* 2001). This

construct expressed both the prM and E proteins. Transformed mammalian (COS-1) cells expressed and secreted high levels of WNV prM and E proteins into the culture medium. To concentrate these proteins the medium was treated with polyethylene glycol. The recombinant antigens were also tested in an IgM antibody-capture and indirect IgG ELISA (Davis *et al.,* 2001). It was found to be a very good and safe alternative to the use of traditional suckling-mouse brain WNV antigen.

1.10 Aims of the study

This study consisted of two parts. Part one was to investigate whether any molecular determinants of virulence between highly and less neuroinvasive lineage II WNV strains, could be identified from our analysis of full genome sequences. Part two of this study was dedicated to the generation of a recombinant WNV antigen for its application as a diagnostic reagent in an ELISA.

1.10.1) Steps in research strategy

1) To determine the complete genome DNA sequence of one non-pathogenic and three pathogenic West Nile virus lineage II strains from South Africa.
2) Perform sequence comparisons and amino acid comparisons of these genomes and compare them to other available WNV genome sequences.
3) Identify regions on the genome that may play a role in WNV virulence.
4) Cloning and expression of recombinant antigens in a mammalian expression system.
5) Testing these antigens in a diagnostic ELISA.

Chapter TWO:
Molecular characterization of WNV lineage II strains

2.1 INTRODUCTION

West Nile virus (WNV) is endemic to Africa, Asia, Europe and Australia and was introduced into the Western Hemisphere with the first outbreak occurring in 1999 in North America. Since then WNV has established itself as an important disease in the USA causing 4180 cases in 2006 alone, of which 1410 were neuroinvasive (http://www.cdc.gov/ncidod/dvbid/westnile/surv&controlCaseCount06_detailed.htm). In South Africa, WNV is widely endemic and has caused occasional epizootics, the largest of which occurred in the Northern Cape in 1974, when thousands of individuals were involved although no deaths were reported (Burt *et al.*, 2002). The number of confirmed WNV cases in recent years has been approximately 5-15 per annum; however, it is suspected that only a proportion of cases are subjected to laboratory investigation. Despite the relatively low number of annual cases, a few reports of severe disease have been made, including fatal hepatitis, non-fatal encephalitis as well as death in an ostrich chick, foal and dog (Burt *et al*., 2002).

The apparent increase in human case fatality rates, neurological infections and horse and bird deaths in the America's caused by lineage I WNV strains raised the question of whether these WNV strains have emerged with increased pathogenicity or alternatively if the severity and impact of the disease is underestimated in South Africa. The existence of two phylogenetic lineages of WNV was demonstrated, with lineage I including viruses from North Africa, Europe, Asia, America and Kunjin virus from Australia while lineage II consisted exclusively of isolates from Southern Africa and Madagascar. This supported the idea of increased virulence of lineage I while lineage II isolates were thought to be associated with endemic infection of low virulence in Africa (McIntosh *et al*., 1976; Beasley *et al*., 2002; Petesen and Roehrig, 2001; Lanciotti *et al*.,1999). Recently two unique strains, Rabensburg virus and strain LEI-Krnd88-190 have also been isolated from *Culex pipiens* mosquitoes and from *Dermacentor marginatus* ticks respectively in eastern Europe that form either two new lineages (Lineage 3 and 4) or two distinct members of the Japanese encephalitis virus (JEV) group based on distance analysis (Bakonyi *et al*., 2005). In 2007 an investigation of Indian isolates of WNV suggested the existence of an additional lineage (Lineage 5) (Vijay *et al*., 2007). In 2004 lineage II WNV were isolated from a dead goshawk (*Accipiter gentilis*) fledgling that showed central nervous symptoms from a national park in south-eastern Hungary. This WNV strain was named

goshawk-Hungary/04 (Hu04). This is first report of lineage II WNV outside of sub-Saharan Africa (Bakonyi *et al*., 2006).

The idea that lineage I WNV strains are more virulent than lineage II WNV strains has since been disputed. Firstly, it was demonstrated that the cases of severe disease in South Africa were also caused by lineage II strains (Burt *et al.*, 2002), and secondly mouse neuroinvasive experiments proved significant differences in the neuroinvasive phenotype. This correlated with genotype not with lineage suggesting that highly neuroinvasive and mild phenotypes existed in both lineages I and II (Venter *et al.*, 2005; Beasley *et al.*, 2002). The perceived virulence of WNV in recent epidemics may partly reflect the emergence and re-emergence of existing strains of WNV in geographic locations with immunologically naïve populations, or be due to high medical alertness and active surveillance programs (Burt *et al.*, 2002). Host gene expression studies further demonstrated that similar genes are induced by highly neuroinvasive lineage I and II strains (Venter *et al.*, 2005).

The WNV virion consists of a host derived lipid bilayer membrane surrounding a nucleocapsid core containing a single stranded positive sense RNA genome of approximately 11 000 nucleotides. The viral envelope and membrane proteins are embedded in the virion membrane and are associated with host range, tissue tropism, replication, assembly, and the stimulation of the B and T cell immune responses. Replication functions are associated with the non-structural (NS) proteins, which may also modulate responses to viral infection (Reviewed in Beasley, 2005). The E protein is the viral hemaglutinin that mediates virus-host cell binding, elicits most of the virus neutralizing antibodies and determines the serological specificity of the virus (Campbell *et al*., 2002; Lindebach and Rice, 2001; Petersen and Roehrig, 2001). Attenuated lineage I phenotypes that displayed reduced virulence in mice and less efficient growth characteristics in culture have been identified in Mexico. Molecular characterization of these isolates suggests that mutations that resulted in loss of E-protein glycosylation as well as mutations in the NS protein genes may be associated with these attenuations (Beasley *et al*., 2005). Sequence analysis of variants of the prototype B956 Lineage II strain obtained by molecular mutation revealed changes in the E and non-structural genes that resulted in reduced virulence in mice (Yamshchikov *et al*., 2004).

In another study, comparisons between the prototype strain (B956) and a variant of this strain that was obtained by molecular mutation (B956D117B3), revealed changes in the E and non-structural genes which resulted in reduced virulence in mice (Yamshchikov *et al*., 2004). None of the attenuated isolates could however be correlated with clinical disease in humans since they were either isolated from birds or modified in culture. Other recent studies focusing on the molecular determinants of WNV attenuation revealed that the NS4B protein may play a very important role in virulence phenotype determination. These studies all used infectious clones of the NY99 strain (Beasley *et al*., 2005; Wicker *et al*., 2006; Puig-Basagoiti *et al*., 2007; Kinney *et al*., 2006). NS4B is a small hydrophobic NS protein that is predicted to be involved in viral replication and evasion of host innate immune defences (Wicker *et al*., 2006). This protein has four cysteine residues at positions 102, 120, 227 and 237 which may be critical for protein function. Substitution of the cysteine at position 102 to serine by site-directed mutagenesis lead, to the formation of a temperature sensitive phenotype at 41ºC as well as attenuation of the neuroinvasive and neurovirulent phenotypes in mice (Wicker *et al*., 2006). An adaptive mutation of Glu to Gly at residue 249 (E249G) in the NS4B gene resulted in reduced RNA synthesis in host cells (Puig-Basagoiti *et al*., 2007). In another study, infectious clones of the NY99 strain which is highly virulent in American crows was compared in vitro with a Kenya strain (KEN-3829) that have reduced virulence in American crows. The authors demonstrated that after 72 days at 44 ºC, the KEN-3829 strain showed a 6500 fold reduction in viral RNA production compared to a 17 fold reduction of the NY99 strain. This suggested that efficient replication at high temperatures, as experienced in American crows, could be an important virulence factor that determines the pathogenic phenotype of the NY99 strain (Kinney *et al*., 2006)

To further investigate the molecular determinants of virulence of lineage II WNV strains we have determined the nucleotide and amino acid sequence of highly neuroinvasive and mild lineage II strains isolated from patients with WNV fever, meningoencephalitis and hepatitis. These isolates have previously been characterized by mice neuroinvasive experiments and gene expression analysis (Venter *et al*., 2005). These are the first full genome sequences of highly neuroinvasive lineage II WNV strains and permit comprehensive comparison to other available sequences of highly neuroinvasive and mild lineage I strains.

2.2 MATERIALS AND METHODS

2.2.1) Viruses

South African West Nile virus isolates; SPU116/89, SA93/01, SA381/00 and H442 were used in this study. These strains were chosen because all were isolated from human specimens and they have different levels of virulence. SPU 116/89 is highly neuroinvasive; the patient had necrotic hepatitis and died; whereas SA381/00 is mildly neuroinvasive and the patient only had fever, rash, myalgia and arthralgia (for full details see Table 2.1). In choosing strains with different virulence levels, sequences can be compared with other WNV full genomes of known virulence and this may help in identifying possible virulence determinants. Isolates were obtained from the National Institute for Communicable Diseases (NICD), Special Pathogens Unit (SPU), Sandringham in South Africa as freeze-dried mouse passage 2-4 and replicated by one passage in Vero cells.

2.2.2) Strain characteristics

Four lineage II WNV strains isolated from patients in South Africa with mild or severe WNV infections were selected for genome sequencing. Phenotypic pathogenicity data for these strains (H442; SPU116/89; SA93/01; SA381/00) in humans and mice are summarized in Table 2.1. Detailed clinical data for all four strains were described in Burt *et al*., (2002) while mouse neuroinvasive experiments and gene expression data for H442, SPU116/89; SA381/00 were described in Venter *et al*., (2005). Strain SA93/01 was shown to be highly neuroinvasive in a mouse model (M Venter, unpublished data) similarly as SPU116/89 and H442 strains, whereas SA381/00 was classified as being of low neuroinvasive phenotype in mice. H442 and SA381/00 caused fever, rash, myalgia and arthralgia in the patients they were isolated from. SA93/01 caused non-fatal encephalitis in two patients and SA116/89 caused fatal hepatitis (Burt *et al*., 2002).

Sequence comparisons of the four South African strains of WNV with those strains that were known to be highly or less neuroinvasive in mice, or which had been reported to be highly pathogenic or attenuated, were carried out. Lineage II strains for which both full genome sequences and neuronvirulence data in mice were available were: 1) Isolate B956D117B3 (Lanciotti *et al*., 1999) is a passaged clone of reduced

virulence (Yamshchikoc *et al*., 2004) of the prototype strain Uganda B956 which was originally associated with fever in the patient it was isolated from and neurotropic in mice (Smitburn *et al*., 1940) and 2) a Madagascar strain AnMg798 which is non-neuroinvasive (Beasley *et al*., 2002). Lineage I strains included are the highly pathogenic and neuroinvasive NY385-99 strain (Beasley *et al*., 2002), the attenuated non-neuroinvasive strain TM171-03 isolated in Mexico in 2003 (Beasley *et al*., 2004), hamster-passaged attenuated clones of NY-385-99 (Clone TYP-9376 and Clone 9317B) (Ding *et al*., 2005) and a non-neuroinvasive Kunjin virus strain MRM61C(Coia *et al*., 1988) (Table 2.1).

Table 2.1: West Nile virus strain characteristics used in this study

STRAIN	YEAR OF ISOLATION	PASSAGE LEVEL	SOURCE	LOCATION	SYNDROME	OUTCOME	LINEAGE	PATHOGENICITY IN MICE	REFERENCE
SPU116/89	1989	mouse 3	Human (*Homo Sapiens*)	South Africa	Necrotic hepatitis	Died	II	Highly neuroinvasive	Present study
SA 93/01	2001	mouse 1	Human (*Homo Sapiens*)	South Africa	Fever, rash, myalgia, Encephalitis	Survived	II	Highly neuroinvasive	Present study
SA 381/00	2000	mouse 1	Human (*Homo Sapiens*)	South Africa	Fever, rash, myalgia, arthralgia	Survived	II	Mildly neuroinvasive	Present study
H442	1958	mouse 2	Human (*Homo Sapiens*)	South Africa	Fever, rash, myalgia, arthralgia	Survived	II	Highly neuroinvasive	Present study
B956D117B3	1937	Unknown	Human (*Homo Sapiens*)	Uganda	Febrile disease	Survived	II	Less-neuroinvasive than B956	Castle *et al*., 1985
B956	1937	Unknown	Human (*Homo Sapiens*)	Uganda	Febrile disease	Survived	II	Unknown	Yamshchikov *et al*., 2004
Madagascar-AnMg798	1978	Unknown	Parrot (*Coreopsis Vasa*)	Madagascar	N/a	Died	II	Non-neuroinvasive	Keller *et al*., 2006
NY 385-99	1999	Vero 2	Human (*Homo Sapiens*)	USA	Unknown	Unknown	I	Highly neuroinvasive	Borisevich *et al.,* 2006
NY-385-99 Clone TYP-9376	2005	Hamster passage	Hamster	USA	N/a	N/a	I	Attenuated lab strain (non-neuroinvasive)	Ding *et al*., 2005
NY-385-99 Clone 9317B	2005	Hamster passage	Hamster	USA	N/a	N/a	I	Attenuated lab strain (non-neuroinvasive)	Ding *et al*., 2006
TM171-03	2003	Vero 1	Common Raven	Mexico	N/a	Died	I	Attenuated lab strain (non-neuroinvasive)	Beasley *et al*., 2004
MRM61C	1960	N/a	Mosquito (*Culex annulirostris*)	Australia	N/a	N/a	I	Non-neuroinvasive	Coia *et al*., 1988

2.2.3) Primer design

Primers used in amplification reactions and sequencing were either obtained from the literature or designed during this study (Table 2.2). Primers were designed using the lineage II Uganda strain B956D117B3 DNA sequence (GenBank accession number: M12294). DNAMAN Version 4.13 from Lynnon BioSoft was used to design the

primers. Primers were designed in order to form overlaps that span the complete genome.

Table 2.2: Primers used for the characterization of WNV lineage II strains from South Africa.

Primer Name	Sequence 5'- 3'	Position (Strain WNCFG: Accession number M12294)	Tm (°C)	Length	Reference
F10 (a)	CCTGTGGGAGCTGACAAACTT	10	57	21	Dr. M. Venter
R1008	TCCCTCCAGGAAGTCTCTGTT	1008	57	21	Dr. Palacios
R1585	AGGACTTCKCACCAACTGACATAAC	1585	57	25	Dr. Palacios
WNV E Fbac	CGCGGATCCTTCAACTGCTTAGGAATGA	967	62	28	This study
WNV E Rbac	CCCAAGCTTCTAAGCATGGACGTTGACCG	2457	65	29	This study
F2115	AAGAGGAGAACAGCAGATAAACCAT	2115	56	25	This study
R5457	GTGAARTGDGCYTCRTCCAT	5457	53	20	Briese *et al.*, 1999
F5004	GGAACDTCMGGHTCNCCHAT	5004	59	20	Briese *et al.*, 1999
R7250	GCYTYHGCYTGCMADCCDGG	7250	51	20	This study
F6918	GCTGGACAAGACCAAGAATG	6918	54	20	This study
CFD2	GTGTCTCAGCCGGCGGTGTCATCAGC	9297	67	26	Kuno *et al.*, 1998
FU1	TACAACATGATGGGAAAGAGAGAGAA	9031	56	26	Kuno *et al.*, 1998
R10962	AGATCCTGTGTTCTAGCACCA	10962	55	21	Dr. M. Venter
WNV II 1F	AGTAGTTCGCCTGTGTGAGC	1	57	20	Bakonyi *et al.*, 2006
WNV II 200R	AGCATAGCCCTCTTCAGTCC	200	55	20	Bakonyi *et al.*, 2006
WNV II 870F	CCTCGTTGCAGCTGTCATTG	870	56	20	Bakonyi *et al.*, 2006
WNV II 2491R	CTTGCCTGCCAATGTCAATG	2491	55	20	Bakonyi *et al.*, 2006
USU 3606 F	AAGAGGTGGACGGCCARRHT	3606	55	20	Bakonyi *et al.*, 2006
USU 4759 R	GTGTGCCAYAGYGTGTGGAA	4759	55	20	Bakonyi *et al.*, 2006
USU 10596 F	GWAAGCCTCYCAGAACCGTCTCGGAAG	10596	63	27	Bakonyi *et al.*, 2006
USU 11014 R	AGATCCTGTGKTCTWSYYCMCCAYCAG	11014	55	27	Bakonyi *et al.*, 2006
WNV-I	AACTCGCAGATGTGCGC	1118	55	17	This study
F5500	GAGCGAGCATCGCAGCCRGA	5500	64	20	This study
F7600	CGACRGCTATTGGCCTTTGT	7600	57	20	This study
F9600	GTCATTGGACCCKATGATGT	9600	53	20	This study
F3103	CTGGAAGCTTGAGAGGGCGG	3103	66	20	This study
F8055	ACAGACCATCAGAGGCGAGC	8055	64	20	This study
F2756	GGGATGTACAAAGCAGC	2756	57	17	This study

2.2.4) RNA Isolation

Viral RNA was extracted from 140 µl of Vero cell supernatant with the QIAamp viral RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. In brief, 140 µl supernatant was mixed with 560 µl buffer AVL containing carrier RNA. The mixture was mixed by pulse-vortexing for 15 seconds (sec). This was followed by

an incubation step of 10 minutes (min) at room temperature. Next, 96% ethanol was added to the sample and mixed with vortexing for 15 sec. Sample was then transferred to the QIAamp spin columns in a 2 ml collection tube. Columns were then centrifuged in a benchtop centrifuge at 2 000 g for 1 min. Next the columns were washed by adding 500 µl buffer AW1 to the column and centrifuging for 1 min at 2000 g. This was followed with another wash step with 500 µl Buffer AW2 and centrifugation for 3 min at 4 000 g. An additional centrifugation step for 1 min at full speed was included to ensure that all ethanol is removed. RNA was eluted by adding 60 µl nuclease-free H_2O to the spin column inserted into a clean 1.5 ml eppendorf tube and incubated at room temperature for 1 min and then centrifuged for 1 min at 2 000 g.

2.2.5) First strand cDNA synthesis

For first strand cDNA synthesis, Expand Reverse Transcriptase (Roche Diagnostics, Mannheim, Germany) was used as follows: 10 µl RNA and 0.4 µg of Random hexanucleotides (Roche Diagnostics, Mannheim, Germany) were denatured at 65ºC for 10 min and immediately cooled on ice, after which master mix was added. The master mix contained 1 X Expand Reverse Transcriptase buffer, 100 mM DTT, 200 µM of each dNTP; 20U RNase Inhibitor (Roche Diagnostics, Mannheim, Germany) and 50U of Expand Reverse Transcriptase. The reaction was then incubated at 30ºC for 10 minutes followed by 43ºC for an hour. The reaction was terminated by incubation on ice.

2.2.6) Polymerase chain reaction

Ten µl of the cDNA reaction was added to the PCR master mix. The master mix consisted of 3.75U of either Expand High Fidelity polymerase (Roche Diagnostics, Mannheim, Germany) for PCR product smaller than 2 kb or Expand Long Template PCR system polymerase (Roche Diagnostics, Mannheim, Germany) for larger products. 300 µM of each dNTP (Promega, Southampton, United Kingdom), 1 X Buffer, and 30 pmol of each primer (Table 2.2). For PCR products less than 3 kb the following cycle was used; initial denaturation at 94ºC for 2 min; followed by 40 cycles of denaturation at 94ºC for 15 sec, annealing at *ºC for 30 sec, extension at 72ºC for 2 min and a final extension step of 72ºC for 7 min. For PCR products larger than 3kb the following cycles were used: initial denaturation at 94ºC for 2 min; 10 cycles of

(*Annealing temperatures differ for al the primer pairs and each reaction were optimized separately (Table 2.3).)

denaturation at 94ºC for 10 sec, annealing at 50ºC for 30 sec, extension at 68ºC for 3 min; Followed by 35 cycles of denaturation at 94ºC for 15 sec, annealing at 50ºC for 30 sec, extension at 68ºC for 5 min plus 5 sec per cycle and a final extension step of 72ºC for 7 min.

Table 2.3: Specific annealing temperatures for PCR primer sets.

	Primer Pair	**SPU116/89**	**SA381/00**	**SA93/01**	**H442**
1	F10 and R1008	53°C	53°C	53°C	53°C
2	F10 and R1585	56°C	55°C	56°C	55°C
3	WNV EFBAC and WNV ERBAC	51°C	51°C	51°C	50°C
4	F2115 and R5457	51°C	50°C	50°C	51°C
5	F5004 and R7250	56.5°C	57°C	56.5°C	53°C
6	F6918 and CFD2	59°C	59°C	59°C	59°C
7	FU1 and R10962	56.5°C	57°C	56.5°C	57°C
8	WNV II 1F and WNV II 200R	51°C	51°C	51°C	51°C
9	WNV II 870F and WNV II 2491R	N/A	N/A	55°C	51°C
10	USU 3606F and USU 4759R	55°C	55°C	55°C	53°C
11	USU 10596F and USU 11014R	54°C	54°C	54°C	51°C

PCR product purification:

The Wizard SV gel and PCR clean-up system (Promega, SouthHampton, United Kingdom) was used to purify the PCR amplicons for sequencing according to the manufacturer's instructions. In brief 10 µl membrane binding solution was added for every 10 mg of agarose gel slice. Agarose gel was melted by incubation at 55ºC for 10 min with occasional vortexing to mix the membrane binding buffer and agarose. Next the melted gel and membrane binding solution mixture were transferred to minicolumns, incubated for 1 min at room temperature followed by centrifugation for 1 min at 10 000 g. Flowthrough was discarded and 700 µl washing buffer was added to the column followed by centrifugation for 1 min at 10 000 g. The wash step was repeated with 500 µl washing buffer and centrifugation for 5 min at 10 000 g. The minicolumn was transferred to a clean 1.5 ml eppendorf tube. DNA was eluted with 30 µl nuclease-free H_2O and incubation at room temperature for 1 min and centrifugation at 10 000 g for 1 min.

Ten µl of the PCR reaction and 2 µl of gel loading buffer (0.25% Bromophenol Blue in 40% sucrose solution) were mixed and analyzed by agarose gel electrophoresis. The bands were visualized using ethidium bromide staining (final concentration 0.5 µg/ml). A 1kb DNA ladder (Promega, Southhampton, United Kingdom) was included as a standard. Electrophoresis was carried out in 1 X Sodium boric acid electrophoresis buffer (200 mM NaOH; 728 mM Boric acid (B_3BO_3); pH 8) at 100 V in a horizontal gel electrophoresis tank.

2.2.7) DNA sequencing

PCR primers as well as sequence primers were used for sequencing the WNV full genomes (Table 2.2). DNA cycle sequencing using BigDye Terminator V3.1 kit (Applied Biosystems, Warrington, United Kingdom) was conducted as recommended by the supplier and analyzed on an ABI $P_{RISM}^{®}$ 3100/3130 genetic analyzer at the sequence facility at the Natural and Agricultural Faculty, University of Pretoria, South Africa. Briefly, 1/4 reactions with 2 µl of BigDye and 3.2 pmol primers were performed with 100 ng template for every 1 kb of PCR product. Amplification cycles were as follows; initial denaturation of 96ºC for 1 min, 25 cycles of 94ºC for 10 sec, 50ºC for 5 sec and 60ºC for 4 min. DNA sequence reactions were purified using the EDTA/NaOAc/EtOH method. (BigDye Terminator V3.1 sequencing protocol, Applied Biosystems). Briefly, 1 µl 125 mM EDTA, 1 µl 3 M sodium acetate and 25 µl 100% non-denatured ethanol (final concentration between 67-71%) were added to each 10 µl reaction. The mixture was vortexed and incubated for 15 min at room temperature. Samples were then centrifuged for 25 min at maximum speed in a bench-top centrifuge. The supernatant was removed and pellet washed by adding 100 µl 70% ethanol and centrifugation at maximum speed for 15 min. Pellets were air-dried for 20 min.

2.2.8) Sequence Analysis

Phylogenetic analysis of sequences was performed. Briefly editing and assembling of the genomes were performed with Vector NTI 9.1.0 (© Invitrogen Corporation, 2004, Carlsbad, CA, USA). ClustalW (Higgins *et al.*, 1994) multiple alignment was performed using BioEdit version 7.0.1 (Hall, 1999) and amino acid analysis was

performed in GeneDoc (Nicholas *et al.*, 1997) for windows. Amino acid changes considered to have a potential effect on the secondary structure of the proteins included substitution of hydrophilic for hydrophobic amino acids or vice versa and substitutions of cysteine, glycine, and proline residues (Nicholas *et al.*, 1997)

Comparisons are relative to the top sequence (SA381/00), which has been shown to be less neuroinvasive than the other three South African strains, with numbering referring to the sequence position of isolate SA381/00. Phylogenetic analysis was performed with MEGA version 3.1 (Kumar *et al.*, 2004) using neighbor-joining with Kimura 2 distance-parameter and a bootstrap confidence level of a 1000 replicates. P-distances matrix's were also calculated in this program. Nucleotide pairwise distance calculations and amino acid pairwise distance calculations were performed with the p-distances option. Signal cleavage predicted scores were calculated with AnalyzeSignalase 2.03 for Macintosh platform (Von Heijne, 1986).

2.3 RESULTS

2.3.1) Primer design, cDNA synthesis and PCR

Primers were obtained from previous studies or designed in this study using the prototype lineage II Uganda strain B956D117B3 sequence. For large PCR products an internal sequencing primer was designed in order to be able to sequence the complete PCR amplicon. cDNA was produced by using isolated total RNA from Vero cell supernant infected with WNV as template and random hexamers. The cDNA was then used as template in PCR reactions. PCR conditions were optimised for each primer set on all four WNV strains used in this study. PCR amplicons were analyzed on a 1% agarose gel (Figure 2.1) and purified from the gel and purified product was used in DNA sequence reactions. Complete genomes of all four WNV strains were amplified successfully using different primer pairs which overlap and spans over the entire genome.

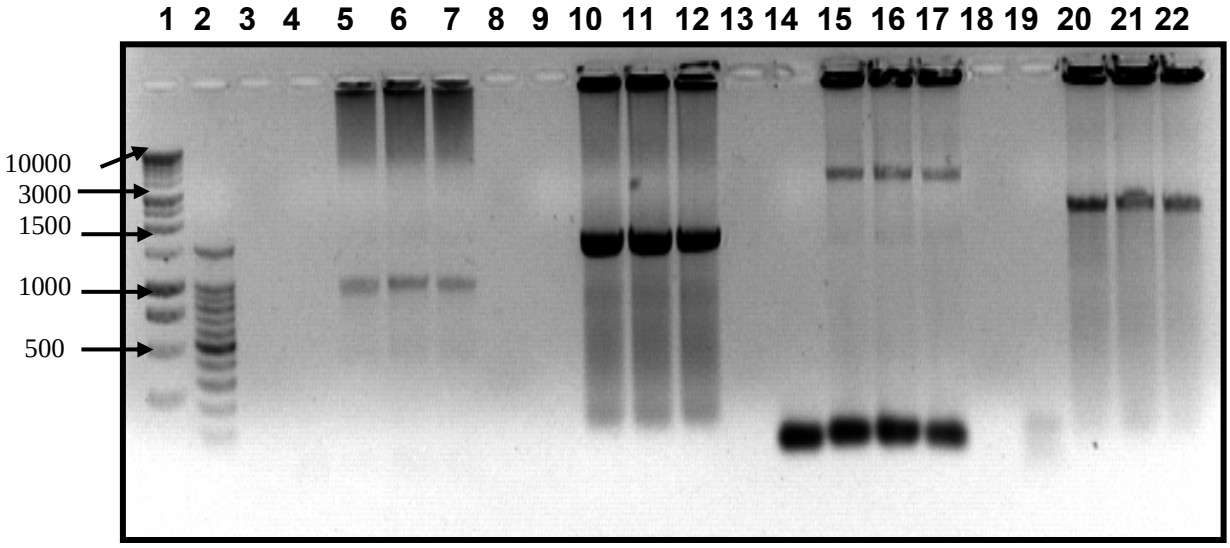


Figure 2.1: An agarose gel electrophoresis analysis of PCR reactions performed to amplify the H442 WNV lineage II strain full genome. This is a representative gel picture of only 4 of the 11 primer sets used to amplify the genome. Lane 1 is a 10Kb marker from Promega and lane 2 is a 100 nt marker (Promega). Lanes 3,8,13,18 are empty lanes; Lanes 4,9,14,19 are negative controls; Lane 5-7 is a PCR product from primers F10 and R1585 (1000 nt); Lane 10-12 PCR product from primers WNV EFBac and WNV ERBac (1500 nt); Lane 15-17 PCR product from primers F2115 and R5457 (3300 nt); Lane 20-22 PCR product from primers FU1 and R10962 (1931 nt).

2.3.2) Nucleotide sequence and phylogenetic analysis

The complete genome sequences of strains H442, SPU116/89; SA381/00 and SA93/01 were determined and deposited to the GenBank database under accession numbers EF429200, EF429197, EF429199, and EF429198 respectively. The termini of these sequences were amplified with primers published from other lineage II full genome sequences. If one assumes that the 5' and 3' termini are identical in length to other published strains these genomes were 11 052 base pair (bp) (SPU116/89; SA381/00 and SA93/01) and 11 051 bp (H442) in length. Comparison of the complete genomes of the South African strains with the prototype lineage II strain as well as lineage I strains revealed overall nucleotide P-distances of 0.0278 (97.2% similarity) between the South African strains (Table 2.5). Most differences existed in the NS5 protein gene (Table 2.4).

Table 2.4: Percentage amino acid differences for individual genes for the lineage II strains sequenced in this study (GenBank accession numbers in Appendix A1).

Strain	SA381/00	H442	SPU116/89	SA93/01	Strain	SA381/00	H442	SPU116/89	SA93/01
CAPSID					**prM**				
SA381/00					**SA381/00**				
H442	0				**H442**	0.6			
SPU116/89	0	0			**SPU116/89**	0.6	0		
SA93/01	0	0	0		**SA93/01**	0.6	0	0	
ENVELOPE					**NS1**				
SA381/00					**SA381/00**				
H442	0.8				**H442**	0.9			
SPU116/89	0.2	0.6			**SPU116/89**	0.9	0		
SA93/01	0.2	0.6	0		**SA93/01**	1.1	0.3	0.3	
NS2A/B					**NS3**				
SA381/00					**SA381/00**				
H442	0.6				**H442**	0.8			
SPU116/89	0.3	0.8			**SPU116/89**	0.6	0.5		
SA93/01	0.3	0.8	0		**SA93/01**	0.6	0.5	0	
NS4A/B					**NS5**				
SA381/00					SA381/00				
H442	0.2				H442	0.8			
SPU116/89	0.2	0.5			SPU116/89	2.1	1.5		
SA93/01	0	0.2	0.2		SA93/01	1.9	1.3	2.0	

P-distances were constructed over the full genome (Table 2.5) and for each individual gene of the WNV polyprotein (Table 2.6-2.15). Few differences existed between WNV strains from the same geographically area with less than 1% amino acid difference existing over the complete genome between the strains from South Africa and the same holds true for strains from New York (NY). The Madagascar strain AnMg798 differed with more than 3% from the other lineage II strains found in South Africa (SA381/00, SPU116/89, H442, SA93/01) and Uganda B956D117B3, suggesting that differences are related to the geographic location of isolation rather than temporal (Table 2.4 and 2.5).

Table 2.5: Percentage amino acid and nucleotide differences when comparing the entire genome of selected WNV strains. The lower-left matrix corresponds to amino acid sequences and the upper-right matrix corresponds to nucleotide sequences (GenBank accession numbers in Appendix A1).

	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	ANMg798	NY-385-99	NY-385-99 Clone TYP-9376	NY-385-99 Clone 9317B	TM171-03	MRM61C
SA381/00		***2.4***	***3.6***	***3.7***	***3.1***	***2.9***	***15.8***	***20.6***	***20.6***	***20.6***	***20.6***	***20.5***
H442	0.7		***2.7***	***3.0***	***1.8***	***1.6***	***15.6***	***20.5***	***20.6***	***20.6***	***20.5***	***20.6***
SPU116/89	0.9	0.7		***1.3***	***2.0***	***1.9***	***16.0***	***20.8***	***20.9***	***20.9***	***20.8***	***20.7***
SA93/01	0.8	0.7	0.6		***2.3***	***2.2***	***15.7***	***20.8***	***20.9***	***20.9***	***20.8***	***20.6***
B956D117B3	1.0	0.8	1.0	0.9		***0.3***	***15.7***	***20.7***	***20.7***	***20.7***	***20.7***	***20.6***
B956	0.7	0.6	0.7	0.7	0.7		***15.8***	***20.7***	***20.7***	***20.7***	***20.7***	***20.6***
AnMg798	3.4	3.3	3.5	3.4	3.5	3.4		***21.5***	***21.5***	***21.5***	***21.4***	***21.2***
NY-385-99	6.0	5.9	6.1	6.1	6.1	6	6.6		***0.1***	***0.1***	***0.4***	***11.7***
NY-385-99 Clone TYP-9376	6.1	6.0	6.2	6.2	6.2	6.1	6.7	0.1		***0***	***0.5***	***11.8***
NY-385-99 Clone 9317B	6.1	6.0	6.2	6.2	6.2	6.1	6.7	0.1	0.0		***0.5***	***11.8***
TM171-03	6.0	5.9	6.1	6.1	6.1	6	6.7	0.1	0.2	0.2		***11.8***
MRM61C	6.5	6.5	6.7	6.6	6.6	6.5	7.1	2.4	2.4	2.4	2.4	

The 5' noncoding region (NCR) had a 0.014 P-distances (98.6% similarity) between SA381/00 and the highly neuroinvasive South African WNV strains and 0.000 P-distance (100% similarity) between the other 3 highly neuroinvasive South African strains. The 3' NCR showed an overall P-distances of 0.0145 (98.55% similarity) between these strains. The lineage I strains also showed high levels of similarity in the 5'NCR and 3'NCR with 100% similarity between the NY strains, and 98.3% similarity between the NY strains and other lineage I strains in the 5'NCR and 99.3% similarity in the 3'NCR between NY strains. To note is the 100% similarity between the NY strains; the Madagascar strain (AnMg798) and the neuroinvasive strains from South Africa (H442, SPU116-89, SA93/01) in the 5'NCR. In the 3'NCR there is only an 83% similarity between these strains (Table 2.6 and 2.7). The capsid gene had the lowest levels of overall nucleotide difference of al 10 genes analysed (Table 2.8).

Table 2.6 Nucleotide sequence P-distance values for the 5' NCR of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NCR 5'	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.014											
SPU116/89	0.014	0.000										
SA93/01	0.014	0.000	0.000									
B956D117B3	0.027	0.014	0.014	0.014								
B956	0.041	0.027	0.027	0.027	0.014							
AnMg798	0.014	0.000	0.000	0.000	0.014	0.027						
NY-385-99	0.014	0.000	0.000	0.000	0.014	0.027	0.000					
NY-385-99 CLONE 9317B	0.014	0.000	0.000	0.000	0.014	0.027	0.000	0.000				
NY-385-99 CLONE TVP-9376	0.014	0.000	0.000	0.000	0.014	0.027	0.000	0.000	0.000			
TM171-03	0.041	0.027	0.027	0.027	0.041	0.054	0.027	0.027	0.027	0.027		
MRM61C	0.027	0.027	0.027	0.027	0.041	0.054	0.027	0.027	0.027	0.027	0.041	

Table 2.7 Nucleotide sequence P-distance values for the 3'NCR of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NCR 3'	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.009											
SPU116/89	0.018	0.018										
SA93/01	0.013	0.013	0.013									
B956D117B3	0.004	0.004	0.013	0.009								
B956	0.004	0.004	0.013	0.009	0.000							
AnMg798	0.139	0.139	0.139	0.148	0.139	0.139						
NY-385-99	0.184	0.184	0.170	0.179	0.179	0.179	0.188					
NY-385-99 CLONE 9317B	0.193	0.193	0.179	0.188	0.188	0.188	0.197	0.009				
NY-385-99 CLONE TVP-9376	0.193	0.193	0.179	0.188	0.188	0.188	0.197	0.009	0.000			
TM171-03	0.184	0.184	0.170	0.179	0.179	0.179	0.188	0.000	0.009	0.009		
MRM61C	0.179	0.179	0.179	0.175	0.175	0.175	0.188	0.103	0.112	0.112	0.103	

Table 2.8 Nucleotide sequence P-distance values for the Capsid gene of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

CAPSID	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.003											
SPU116/89	0.014	0.011										
SA93/01	0.022	0.019	0.014									
B956D117B3	0.014	0.011	0.011	0.014								
B956	0.011	0.008	0.008	0.011	0.003							
AnMg798	0.092	0.089	0.087	0.079	0.081	0.084						
NY-385-99	0.154	0.157	0.154	0.154	0.160	0.157	0.165					
NY-385-99 CLONE 9317B	0.154	0.157	0.154	0.154	0.160	0.157	0.165	0.000				
NY-385-99 CLONE TVP-9376	0.154	0.157	0.154	0.154	0.160	0.157	0.165	0.000	0.000			
TM171-03	0.157	0.160	0.157	0.152	0.163	0.160	0.168	0.003	0.003	0.003		
MRM61C	0.165	0.168	0.165	0.171	0.171	0.168	0.182	0.073	0.073	0.073	0.076	

Table 2.9 Nucleotide sequence P-distance values for the preMembrane gene of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

prMEMBRANE	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.028											
SPU116/89	0.034	0.016										
SA93/01	0.038	0.020	0.004									
B956D117B3	0.036	0.012	0.012	0.016								
B956	0.034	0.014	0.014	0.018	0.002							
AnMg798	0.174	0.168	0.172	0.172	0.174	0.176						
NY-385-99	0.200	0.198	0.198	0.202	0.200	0.202	0.218					
NY-385-99 CLONE 9317B	0.200	0.198	0.198	0.202	0.200	0.202	0.218	0.000				
NY-385-99 CLONE TVP-9376	0.200	0.198	0.198	0.202	0.020	0.202	0.218	0.000	0.000			
TM171-03	0.206	0.204	0.204	0.208	0.206	0.208	0.220	0.006	0.006	0.006		
MRM61C	0.202	0.214	0.212	0.216	0.212	0.212	0.224	0.112	0.112	0.112	0.114	

Table 2.10 Nucleotide sequence P-distance values for the Envelope gene of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

ENVELOPE	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.023											
SPU116/89	0.031	0.028										
SA93/01	0.038	0.036	0.013									
B956D117B3	0.030	0.023	0.015	0.023								
B956	0.028	0.021	0.014	0.023	0.003							
AnMg798	0.153	0.156	0.160	0.154	0.158	0.158						
NY-385-99	0.207	0.205	0.209	0.209	0.208	0.207	0.227					
NY-385-99 CLONE 9317B	0.207	0.205	0.209	0.209	0.208	0.207	0.227	0.001				
NY-385-99 CLONE TVP-9376	0.207	0.205	0.209	0.209	0.208	0.207	0.227	0.001	0.000			
TM171-03	0.205	0.204	0.208	0.208	0.207	0.206	0.226	0.004	0.004	0.004		
MRM61C	0.207	0.213	0.214	0.213	0.214	0.215	0.223	0.128	0.128	0.128	0.127	

Table 2.11 Nucleotide sequence P-distance values for the NS1 gene of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NS1	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.028											
SPU116/89	0.041	0.026										
SA93/01	0.043	0.029	0.009									
B956D117B3	0.036	0.017	0.020	0.024								
B956	0.034	0.015	0.018	0.022	0.002							
AnMg798	0.158	0.152	0.156	0.159	0.158	0.156						
NY-385-99	0.214	0.208	0.208	0.211	0.210	0.208	0.229					
NY-385-99 CLONE 9317B	0.215	0.209	0.209	0.212	0.211	0.209	0.230	0.001				
NY-385-99 CLONE TVP-9376	0.215	0.209	0.209	0.212	0.211	0.209	0.230	0.001	0.000			
TM171-03	0.216	0.210	0.208	0.213	0.212	0.210	0.227	0.004	0.005	0.005		
MRM61C	0.224	0.220	0.223	0.222	0.222	0.222	0.230	0.132	0.133	0.133	0.132	

Table 2.12 Nucleotide sequence P-distance values for the NS2A/B genes of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NS2A-NS2B	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.030											
SPU116/89	0.046	0.030										
SA93/01	0.047	0032	0.016									
B956D117B3	0.035	0.015	0.020	0.023								
B956	0.033	0.013	0.018	0.021	0.004							
AnMg798	0.170	0.169	0.176	0.175	0.169	0.169						
NY-385-99	0.210	0.213	0.215	0.214	0.209	0.211	0.209					
NY-385-99 CLONE 9317B	0.211	0.213	0.216	0.215	0.210	0.212	0.210	0.001				
NY-385-99 CLONE TVP-9376	0.211	0.213	0.216	0.215	0.210	0.212	0.210	0.001	0.000			
TM171-03	0.208	0.211	0.213	0.213	0.207	0.209	0.207	0.004	0.005	0.005		
MRM61C	0.191	0.198	0.196	0.194	0.195	0.195	0.201	0.116	0.117	0.117	0.117	

Table 2.13 Nucleotide sequence P-distance values for the NS3 gene of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NS3	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.023											
SPU116/89	0.042	0.027										
SA93/01	0.045	0.031	0012									
B956D117B3	0.032	0.016	0.015	0.019								
B956	0.032	0.016	0.015	0.019	0.001							
AnMg798	0.160	0.157	0.157	0.155	0.154	0.154						
NY-385-99	0.206	0.208	0.210	0.212	0.205	0.206	0.215					
NY-385-99 CLONE 9317B	0.206	0.208	0.211	0.212	0.206	0.207	0.215	0.001				
NY-385-99 CLONE TVP-9376	0.026	0.208	0.211	0.212	0.206	0.207	0.215	0.001	0.000			
TM171-03	0.205	0.206	0.208	0.210	0.204	0.205	0.216	0.004	0.004	0.004		
MRM61C	0.199	0.197	0.195	0.195	0.197	0.197	0.208	0.122	0.122	0.122	0.125	

Table 2.14 Nucleotide sequence P-distance values for the NS4A/B genes of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NS4A-NS4B	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.028											
SPU116/89	0.028	0.033										
SA93/01	0.022	0.039	0.016									
B956D117B3	0.040	0.030	0.030	0.035								
B956	0.033	0.023	0.025	0.031	0.008							
AnMg798	0.165	0.164	0.172	0.166	0.175	0.173						
NY-385-99	0.230	0.231	0.229	0.228	0.238	0.236	0.234					
NY-385-99 CLONE 9317B	0.230	0.231	0.229	0.228	0.238	0.236	0.234	0.000				
NY-385-99 CLONE TVP-9376	0.230	0.231	0.229	0.228	0.238	0.236	0.234	0.000	0.000			
TM171-03	0.232	0.233	0.231	0.229	0.239	0.238	0.236	0.004	0.004	0.004		
MRM61C	0.221	0.219	0.222	0.221	0.222	0.219	0.228	0.127	0.127	0.127	0.129	

Table 2.15 Nucleotide sequence P-distance values for the NS5 gene of selected lineage I and II WNV strains calculated by using Mega 3.1 (GenBank accession numbers in Appendix A1).

NS5	SA381/00	H442	SPU116/89	SA93/01	B956D117B3	B956	AnMg798	NY-385-99	NY-385-99 CLONE 9317B	NY-385-99 CLONE TVP-9376	TM171-03	MRM61C
SA381/00												
H442	0.022											
SPU116/89	0.040	0.031										
SA93/01	0.035	0.028	0.016									
B956D117B3	0.026	0.015	0.025	0.022								
B956	0.026	0.014	0.025	0.022	0.001							
AnMg798	0.165	0.162	0.165	0.163	0.161	0.160						
NY-385-99	0.204	0.203	0.210	0.208	0.206	0.206	0.209					
NY-385-99 CLONE 9317B	0.204	0.203	0.211	0.208	0.206	0.206	0.209	0.001				
NY-385-99 CLONE TVP-9376	0.204	0.203	0.211	0.208	0.206	0.206	0.209	0.001	0.000			
TM171-03	0.202	0.202	0.209	0.206	0.204	0.204	0.206	0.005	0.005	0.005		
MRM61C	0.212	0.212	0.216	0.212	0.211	0.211	0.210	0.111	0.111	0.111	0.110	

The P-distance Tables of all the different genes (Table 2.8-2.15) reveals a pattern of close similarity within the lineages (Less than 5% difference) and high levels of differences between lineages (more than 15% differences). The gene that showed the most difference between the lineages is the NS4A/B gene with an average of 23% differences between all the strains. (Table 2.14)

The 5'NCR had only three nucleotide (nt) differences between the lineage II sequences (Table 2.16). Noteworthy, nucleotide differences in the noncoding regions are the following deletions; The attenuated strain B956D117B3, passaged molecular clone of the prototype Uganda B956, had a 76 bp deletion in the 3'NCR from nucleotide 10 404 to 10 479. Strain H442 had a 2-nucleotide deletion at nucleotides 10 439 and 10 440. Strain AnMg798 had deletions similar to that of strain B956D117B3 at position 10 411 to 10 487, 10 501 to 10 512, 10 951, The sequence of the AnMg798 strain is incomplete in GenBank and ended at position 10866 (Keller *et al*., 2006).

Phylogenetic analysis confirmed the placing of the South African strains in lineage II (Figure 2.2). SA93/01 and SPU116/89 clustered together while H442 and SA381/00 were on separate branches within lineage II. Although the full genome phylogenetic analysis provides limited information regarding diversity the same clustering was observed when Neighbour Joining trees were constructed using the E and NS3 and NS5 genes (Appendix B). Although the Indian strains clustered with lineage I, p-distance analysis suggested that it was as distant to the lineage I strains (20% differences) as to the lineage II strains (21%-22%) relative to < 5% differences within lineage 1C and 12% differences between 1A and 1B. It was therefore termed lineage 5, as suggested by Bondre *et al.*, 2007.

.

Table 2.16: Nucleotide differences in the noncoding 5’ and 3’ regions of lineage II strains. Numbering according to WNV strain SA381/00. Black shading indicates deletions and grey shading nucleotide differences. The length of each genome is given in the last column. (Strain AnMg798 is incomplete in the GenBank database and may thus be longer than indicated) (GenBank accession numbers in Appendix A1).

Region	5`NCR				3`NCR																																	
Nucleotide Position	51	52	65	82	10403	10404	10405	10407	10411	10427	10438	10439	10440	10442	10460	10467	10473	10479	10480	10487	10487	10488	10490	10492	10495	10496	10497	10500	10501	10509	10512	10515	10519	10522	10524	10526	10527	10540
SA381/00	A	T	A	A	T	A	A	A	C	C	T	G	T	T	A	C	A	A	G	A	T	T	T	G	A	G	T	A	A	A	A	T	G	A	T	G	A	T
H442	A	C	A	A	T	A	A	A	C	C	C			T	G	C	G	A	G	A	T	T	T	G	A	G	T	A	G	G	A	T	G	A	T	G	A	T
SPU116/89	A	C	A	A	T	A	A	A	C	T	T	G	T	T	G	T	G	A	G	A	T	T	T	G	A	G	C	A	G	A	A	T	G	A	T	G	A	T
SA93/01	A	C	A	A	T	A	A	G	C	T	T	G	T	T	G	T	G	A	G	A	T	T	T	G	A	G	T	A	G	A	A	T	G	A	T	G	A	T
B956D117B3	A	C	A	G	A														G	A	T	T	T	G	A	G	T	A	G	A	A	T	G	A	T	G	A	T
B956	A	C	T	G	A	A	A	A	C	T	T	G	T	C	G	T	G	A	G	A	C	T	T	G	A	G	T	A	G	A	A	T	G	A	T	G	A	T
AnMg798		C	A	A	T	A	G	G													T	C	C	T	T	A	T	G				A	T	G	A	A	G	C

Region	3`NCR																																					
Nucleotide Position	10543	10544	10581	10582	10605	10614	10654	10669	10670	10681	10687	10700	10703	10704	10705	10728	10737	10770	10791	10796	10797	10801	10814	10815	10851	10852	10853	10854	10856/7	10858	10887	10925	10936	10939	10951	10954	10956	11017	Length of Genomes
SA381/00	A	T	G	G	G	T	G	C	T	A	G	C	T	A	A	A	A	A	T	C	T	T	A	G	C	A	A	A		T	A	C	A	G	T	C	C	T	11052
H442	A	T	G	G	G	T	G	C	T	A	G	C	T	A	A	A	T	A	T	C	T	T	A	G	C	A	A	A	A	T	G	C	A	G	T	C	C	T	11051
SPU116/89	A	T	G	G	G	C	G	C	T	A	G	C	T	A	G	A	T	A	T	C	T	T	A	G	C	A	A	A		T	G	C	A	G	T	C	C	T	11052
SA93/01	A	T	G	G	G	T	G	C	C	A	G	C	T	A	G	A	T	A	T	C	T	T	A	G	C	G	A	A		T	G	C	A	G	T	C	C	T	11502
B956D117B3	A	T	G	G	G	T	G	C	T	A	G	C	T	A	A	A	T	A	T	C	T	T	A	G	A	A	A	A		T	G	C	A	G	T	C	C	C	10962
B956	A	T	G	G	G	T	G	C	T	A	G	C	T	A	A	A	T	A	T	C	T	T	A	G	T	A	A	A		T	G	C	A	G	T	C	C	T	11038
AnMg798	G	A	A	A	T	C	A	T	T	G	A	T	C	G	A	T	T	C	C	T	C	C	G	A	A	A	G	G		C	G	A	C	T		T	T		10866

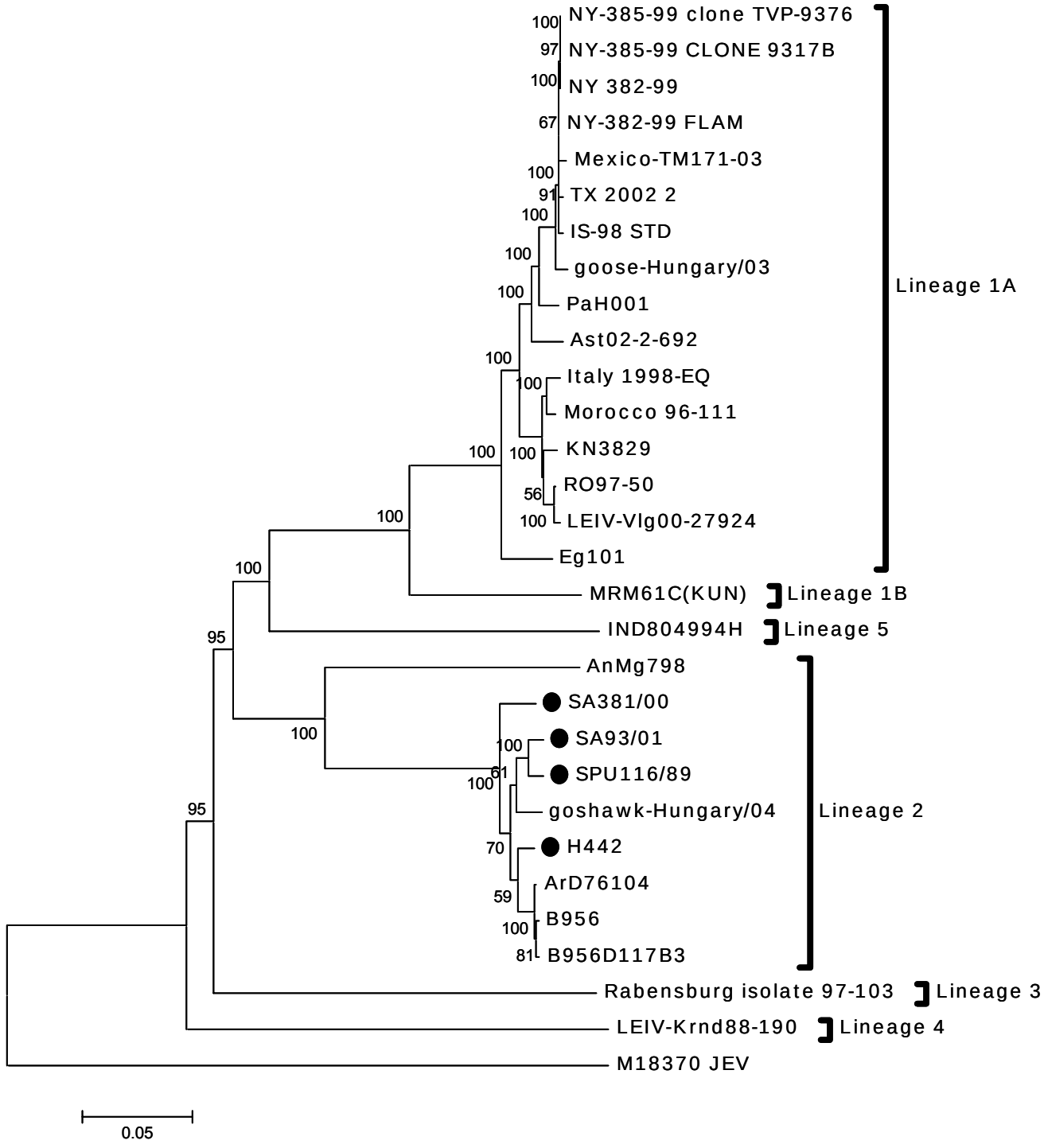

NY-385-99 clone TVP-9376
NY-385-99 CLONE 9317B
NY 382-99
NY-382-99 FLAM
Mexico-TM171-03
TX 2002 2
IS-98 STD
goose-Hungary/03
PaH001
Ast02-2-692
Italy 1998-EQ
Morocco 96-111
KN3829
RO97-50
LEIV-Vlg00-27924
Eg101
Lineage 1A
MRM61C(KUN)
Lineage 1B
IND804994H
Lineage 5
AnMg798
SA381/00
SA93/01
SPU116/89
goshawk-Hungary/04
H442
ArD76104
B956
B956D117B3
Lineage 2
Rabensburg isolate 97-103
Lineage 3
LEIV-Krnd88-190
Lineage 4
M18370 JEV
0.05

Figure 2.2: Phylogenetic analysis of full genome nucleotide sequences of lineage I and II WNV strains. The tree was constructed with the program MEGA version 3.1 using the neighbour-joining method with Kimura 2-parameter distance matrix; together with a bootstrap confidence level of 1000 replicates was used. South African strains sequenced in this study are indicated by a black dot. GenBank accession numbers for the sequences included in the tree are as follows: NY-385-99 clone 9317B (DQO66423); NY-385-99 clone TVP-9376 (AY848697); NY 385-99 (DQ211652); NY-382-99 FLAM (AF196835); IS-98 STD (AF481864); Mexico-TM171-03 (AY660002); TX 2002 2, (DQ 164205); goose-Hungary/03 (DQ118127); Eg101 (AF260968); RO97-50 (AF260969); Morocco 96-111 (AY701412); Italy 1998-Eq (AF404757); KN3829 (AY262283); LEIV-Vlg00-27924 (AY278442); PaH001 (AY268133); Ast02-2-696 (DQ411035); MRM61C (KUN) (D00246); IND804994H (DQ256376); AnMg798 (DQ176636); SA381/00 (EF429199); H442 (EF429200); goshawk-Hungary/04 (DQ116961); SPU116/89 (EF429197); SA93/01 (EF429198); B956D117B3 (M12294); B956 (AY532665); ArD76104 (DQ 318019); Rabensburg isolate 97-103 (AY765264); LEIV-Krnd88-190 (AY277251); M18370 JEV (M18370).

Neighbour-joining tree constructed with MEGA 3.1 for the 5'NCR nucleotide sequence is inconclusive due to the low bootstrap values. This can be due to the high level of conservation in this region (Appendix B1.1). The neighbour-joining tree for the capsid nucleotide sequence in Appendix B1.2 indicates that H442 and SA381/00 clusters together with a confidence level of 98. Nucleotide neighbour-joining trees for each of the individual genes prM, E, NS1, NS2A/B, NS3, NS4A/B and NS5 indicates that strains SPU116/89 and SA 93/01 sequenced in this study clusters together (Appendix B1.3 – B1.10). This could be expected, as they both are highly neuroinvasive and showed severe symptoms in patients. H442 and SA381/00 cluster at the same side of the trees. They both showed fever symptoms in patients although they are highly and less neuroinvasive respectively.

2.3.3) Amino acid differences between the highly neuroinvasive and mild South African WNV strains.

Amino acid differences between the 4 South African strains are summarised in Table 2.17. Very few amino acid differences were found in the structural proteins between the South African WNV strains. SA381/00 had only 1 difference in the prM protein at position 105 relative to the highly neuroinvasive strains (alanine [A] to valine [V]). Two differences (glycine [G] instead of alanine [A], position 54 and proline [P] rather than threonine [T], in position 70) may result in structural differences in the envelope protein of H442, which was isolated 50 years earlier than strains SPU116/89, SA93/01 and SA381/00. The attenuated lineage II Uganda strain B956D117B3 and the non-neuroinvasive Madagascar strain AnMg798 contained significant differences in the glycosylation site of the envelope proteins relative to the South African strains. B956D117B3 was found to have a deletion from residue 154 to 157, and in AnMg798, a proline replaced the serine at position 156. Both of these changes would prevent glycosylation. Further changes with potentially structural implications (substitutions of hydrophilic amino acids for proline [P] and glycine [G] residues) were found in AnMg798 at positions 156, 199 and 230.

The nonstructural proteins, with the exception of NS1, NS2A and NS2B, were the most variable viral proteins. In strain SA381/00, the least virulent of the 4 strains, a hydrophobic amino acid (alanine [A]) was found in position 160 of NS3, in contrast to the hydrophilic amino acid (serine [S]) in the other strains. Similarly, a glycine [G], which was found instead of a hydrophilic amino acid (arginine [R]) at position 298, could alter the structure of the SA381/00 NS3 protein. In the case of the highly pathogenic strain SPU116/89, threonine [T] (hydrophilic) in position 79 of NS4B contrasted with alanine [A] (hydrophobic) in the other strains. Other amino acid changes with potential structural implications were at positions 18 and 145 of the NS4A gene and 14 of the NS4B gene in strain B956D117B3, and at position 14 and 27 in the NS4B gene of strain AnMg798 (Table 2.17).

The NS5 protein was found to be the most variable, and several positions were identified where the South African strains associated with mild infections (SA381/00 and H442) and the two other lineage II strains (AnMg798, B956D117B3) associated with reduced virulence in mice had the same amino acid changes relative to strains which caused severe disease (SPU116/89 and SPU93/01). These included hydrophilic versus

hydrophobic amino acids in position 614 and hydrophobic (mild) versus hydrophilic (pathogenic) in positions 625 and 626 of the NS5 protein. SPU116/89, isolated from a patient with necrotic hepatitis, were found to have amino acid changes that affect the hydrophobicity of the NS5 protein relative to all other strains in positions 197, 623, 635, 641 and 643.

The envelope protein glycosylation motif previously identified in lineage I strains at position 154 to156 (NYS) (Beasley *et al*., 2005) was present in all 4 South African strains however due to a proline [P] substitution at position 156 was not predicted to be glycosylated in strain AnMg798 from Madagascar. The glycosylation motif is deleted completely in strains B956 and B956D117B3 (Figure 2.3).

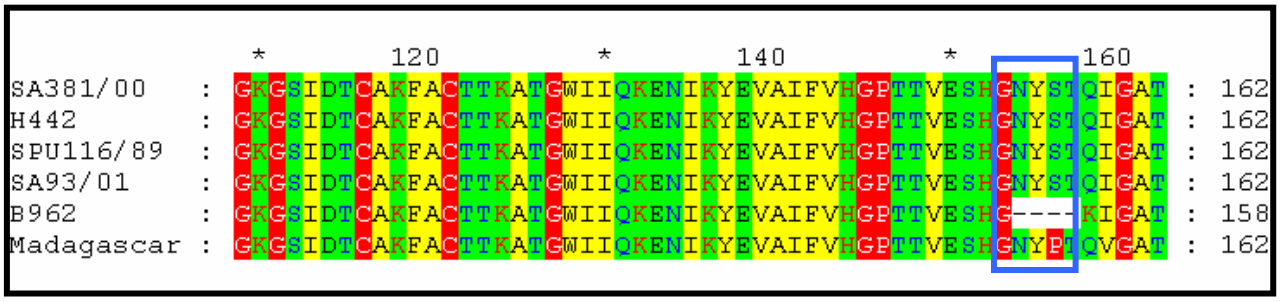


Figure 2.3: Amino acid alignment of lineage II WNV strains with the blue block indicating the glycosylation site of the envelope gene. NYS glycosylated and NYP not-glycosylation.

Table 2.17: Amino acid differences between South African lineage II WNV strains sequenced in the present study and previously published lineage II strains. Strain SA381/00 is less neuroinvasive than the highly neuroinvasive strains SA93/01, H442 and SPU116/89. Light grey shading indicates hydrophobic amino acids, hydrophilic amino acids are highlighted in dark grey and amino acids in black are structural determining amino acids. White blocks indicate the deletion of the glycosylation site in the envelope protein of the Uganda strains. Numbering according to SA381/00 for specific genes.

Protein	Cap		prM				Envelope																								NS1																					NS2A					
Amino Acid position	110	112	19	105	156	157	33	47	54	58	66	67	70	154	155	156	157	158	159	178	182	188	199	208	230	280	291	312	313	491	29	54	88	94	96	102	105	146	147	174	206	208	240	273	289	290	292	323	324	327	338	14	34	38	84	98	104
SA381/00	T	T	V	V	A	I	I	N	A	S	S	D	T	N	Y	S	T	Q	I	L	E	V	S	A	S	K	K	V	G	V	M	V	I	N	M	K	A	V	N	T	L	D	D	V	D	R	G	F	Q	N	T	L	I	M	K	R	S
H442	T	T	V	A	A	I	I	N	G	S	N	D	P	N	Y	S	T	Q	I	L	E	V	S	A	S	K	K	A	G	V	M	V	V	N	M	K	A	A	N	T	L	D	D	V	D	S	G	F	Q	N	T	L	I	M	K	R	S
SPU116/89	T	T	V	A	A	I	I	N	A	S	S	D	T	N	Y	S	T	Q	I	L	E	V	S	A	S	K	K	A	G	V	M	V	V	N	M	K	A	A	N	T	L	D	D	V	D	S	G	F	Q	N	T	L	I	M	K	K	S
SA93/01	T	T	V	A	A	I	I	N	A	S	S	D	T	N	Y	S	T	Q	I	L	E	V	S	A	S	K	K	A	G	V	I	V	V	N	M	K	A	A	N	T	L	D	D	V	D	S	G	F	Q	N	T	L	I	M	K	R	S
B956D117B3	T	T	V	A	A	I	I	N	A	S	S	D	T					K	I	L	E	V	S	E	S	K	K	A	R	V	M	V	V	N	M	K	A	A	N	T	L	D	D	V	D	S	E	F	Q	N	T	M	I	M	K	R	S
B956	T	T	V	A	A	I	L	N	A	S	S	D	T					K	I	L	E	V	S	A	S	K	K	A	R	V	M	V	V	N	M	K	A	A	N	T	L	D	D	V	D	S	G	F	Q	N	T	L	I	M	K	R	S
AnMg798	A	S	I	A	V	V	I	K	A	H	S	E	T	N	Y	P	T	Q	V	M	D	I	G	A	G	T	R	V	G	I	I	I	V	D	L	R	T	T	S	V	F	E	E	I	A	S	G	Y	T	S	L	L	V	L	R	R	N

Protein	NS2A																		NS2B			NS3																NS4A					NS4B														
Amino Acid position	117	118	123	126	129	167	177	181	183	190	191	192	201	212	214	216	217	223	26	28	41	160	210	215	230	233	253	255	298	304	333	334	347	382	421	572	597	18	55	85	98	145	14	15	19	20	21	23	24	27	30	49	79	117	150	178	184
SA381/00	A	Y	N	S	V	K	L	V	V	K	R	S	C	V	N	M	I	M	I	S	V	A	K	K	A	S	S	N	G	K	V	T	T	K	S	T	A	G	A	V	E	S	S	S	H	K	P	A	R	T	V	T	A	A	A	M	M
H442	A	Y	N	S	V	K	L	V	V	R	R	S	C	V	N	M	I	M	I	T	V	S	K	K	A	S	S	N	R	K	I	S	T	K	S	M	V	G	A	V	E	S	S	S	H	K	P	A	R	T	V	T	A	V	A	M	M
SPU116/89	A	Y	N	S	V	K	L	V	V	K	R	S	C	V	N	M	I	M	I	S	V	S	K	K	A	S	S	N	R	K	I	S	T	K	S	T	A	G	A	V	E	S	S	S	H	K	P	A	R	T	V	T	T	A	A	M	M
SA93/01	A	Y	N	S	V	K	L	V	V	K	R	S	C	V	N	M	I	M	I	S	V	S	K	K	A	S	S	N	R	K	I	S	T	K	S	T	A	G	A	V	E	S	S	S	H	K	P	A	R	T	V	T	A	A	A	M	M
B956D117B3	A	Y	N	S	V	K	L	V	V	K	R	S	C	V	N	M	I	M	I	S	A	S	K	K	A	S	S	N	R	K	I	S	T	K	S	T	A	V	V	V	E	G	G	S	H	R	P	A	R	T	V	T	A	A	V	I	M
B956	A	Y	N	S	V	K	L	V	V	K	R	S	C	V	N	M	I	M	I	S	V	S	K	K	A	S	S	N	R	K	I	S	T	K	S	T	A	G	A	I	E	G	S	S	H	K	P	A	R	T	V	T	A	A	V	I	M
AnMg798	G	F	Q	R	I	R	V	I	I	K	K	N	S	M	S	L	V	I	S	S	V	S	R	R	S	A	N	S	R	R	I	S	S	R	N	T	A	G	A	V	D	S	G	R	Y	K	L	V	K	G	I	A	A	I	A	M	V

Protein	NS4B				NS5																																																				
Amino Acid position	204	230	239	242	29	42	45	78	96	139	177	197	207	254	258	278	281	285	291	298	336	374	386	390	403	523	526	527	528	532	533	543	572	582	589	594	595	614	616	620	623	625	626	635	639	641	643	644	703	790	816	837	841	878	882	897	898
SA381/00	T	I	L	T	K	Q	R	V	R	A	R	T	V	F	V	R	S	R	S	A	T	Y	F	D	G	E	T	K	G	I	Y	R	E	V	G	D	V	T	T	V	V	M	M	D	K	G	G	K	T	V	A	K	V	V	E	T	I
H442	T	I	L	T	K	H	R	V	R	A	K	T	V	F	V	K	N	R	S	A	T	Y	F	D	G	E	T	K	P	I	Y	R	E	V	G	D	V	T	T	V	V	M	M	D	K	G	G	K	T	V	S	R	V	V	E	T	I
SPU116/89	T	I	L	T	K	H	R	V	R	A	K	I	V	F	V	K	N	R	S	A	T	H	F	D	G	E	T	K	P	I	Y	R	Q	V	G	S	L	P	T	F	G	R	R	Y	K	E	R	K	T	V	A	R	V	V	E	T	I
SA93/01	T	I	L	T	K	H	R	V	R	A	K	T	V	F	V	K	N	R	S	A	T	Y	F	D	G	E	T	K	P	I	F	P	E	G	R	D	V	P	S	V	V	R	R	D	K	G	G	N	T	V	A	R	V	V	E	A	I
B956D117B3	T	I	L	M	K	H	R	V	R	A	K	T	I	F	V	K	N	R	S	A	T	Y	F	D	G	E	T	K	P	V	Y	R	E	V	G	D	V	T	T	V	V	M	M	D	K	G	G	K	T	A	A	R	V	V	E	T	I
B956	T	I	L	M	K	H	R	V	R	A	K	T	V	F	V	K	N	R	S	A	T	Y	F	D	G	E	T	K	P	V	Y	R	E	V	G	D	V	T	T	V	V	M	M	D	K	G	G	K	T	A	A	R	V	V	E	T	I
AnMg798	S	V	F	T	R	H	K	I	K	V	K	T	V	Y	A	K	N	K	G	E	V	Y	Y	E	S	S	V	R	P	I	Y	R	E	V	G	D	V	T	T	V	V	M	M	D	R	G	G	K	V	V	A	R	I	L	G	T	T

2.3.4) Cleavage sites

In order to investigate the efficiency of polyprotein cleavage between the different WNV strains, signalase prediction algorithms were used to analyze the signal peptidase cleavage sites (Von Heijne, 1986). The results are summarized in Table 2.18. No differences were found in predicted cleavage efficiency between the highly pathogenic and less pathogenic strains. The only meaningful difference was observed in the C-prM cleavage region as indicated by the Student *t*-test probability calculated in Table 2.18, where the lineage II strains were predicted to be cleaved more efficiently than lineage I stains. Only slight differences were apparent in the PrM-E site with no differences in any other cleavage regions between lineage I and 2 strains

Table 2.18: Summary of the cleavage scores predicted for the cleavage junctions between the capsid and the prM proteins (C-prM), between the prM and the envelope (prM-E) and the envelope and NS1 proteins (E-NS1), and NS4B-NS5. Signal cleavage predicted scores were calculated with AnalyzeSignalase 2.03 (Von Heijne, 1986). The last amino acid of the first protein and the first two amino acids of the following protein are indicated. P value is the Two-tailed Student *t*-test results, indicating the probability of significance of observed differences between lineage 2 and lineage 1 strains. The arrow indicates the exact cleavage site.

	SA381/00	H442	SPU116/89	SA93/01	956D117B3(Wengler)	AnMg798	NY-385-99	NY-385-99 CLONE TYP-9376	NY-385-99 CLONE 9317B	TM171-03	MRM61C	P value
C-prM												
G	+3.69	+3.69	+3.69	+3.69	+3.69	+4.00	-0.49	-0.49	-0.49	-0.49	-1.85	0.00005
→ **A**	+9.37	+9.37	+9.37	+9.37	+9.37	+8.01	+5.93	+5.93	+5.93	+5.93	+7.37	0.00004
V	-9.14	-9.14	-9.14	-9.14	-9.14	-7.8	-9.52	-9.52	-9.52	-9.52	-10.32	0.02235
PrM-E												
Y	-10.15	-10.15	-10.15	-10.15	-10.15	-9.12	-9.12	-9.12	-9.12	-9.12	-9.45	0.00433
→ **S**	+11.27	+11.27	+11.27	+11.27	+11.27	+12.42	+12.42	+12.42	+12.42	+12.42	+11.50	0.01728
F	-5.37	-5.37	-5.37	-5.37	-5.37	-4.78	-4.78	-4.78	-4.78	-4.78	-5.27	0.01977
E-NS1												
H	-9.01	-9.01	-9.01	-9.01	-9.01	-9.71	-9.01	-9.01	-9.01	-9.01	-9.01	0.36322
→ **A**	+4.26	+4.26	+4.26	+4.26	+4.26	+4.04	+4.26	+4.26	+4.26	+4.26	+4.26	0.36322
D	-11.05	-11.05	-11.05	-11.05	-11.05	-11.15	-11.05	-11.05	-11.05	-11.05	-11.05	0.36322
NS4B-NS5												
R	-16.05	-16.05	-16.05	-16.05	-16.05	-16.05	-16.05	-16.05	-16.05	-15.83	-16.05	0.37390
→ G	-13.19	-13.19	-13.19	-13.19	-13.19	-13.19	-13.19	-13.19	-13.19	-13.08	-13.19	0.37390
G	-19.54	-19.54	-19.54	-19.54	-19.54	-19.54	-19.54	-19.54	-19.54	-19.66	-19.54	0.37390

2.4 DISCUSSION

West Nile Virus lineage II strains are endemic to Southern Africa and Madagascar (Burt *et al*., 2002). Despite the wide spread distribution of WNV in South Africa clinical cases are infrequently reported. Low levels of clinical WNV cases has lead to the hypothesis that lineage II strains are not associated with severe disease, however it has been previously reported that the few cases of severe WNV disease identified in South Africa were all due to lineage II strains (Burt *et al*., 2002). Mouse neuroinvasive experiments and gene expression analysis with some of the strains associated with severe disease in humans confirmed that all of them were neuroinvasive and have similar virulence when compared in a mouse model to lineage I New York strains (NY385/99 and NY2001Hu) (Venter *et al*., 2005). One strain that was isolated from a patient with benign illness showed less neuroinvasiveness than the others lineage I and II strains but was similar to mild lineage I strains confirming the idea that pathogenicity is associated with genotype rather than lineage (Venter *et al*., 2005).

Patients from which strain SA381/00 and H442 was isolated had fever, rash, myalgia and arthralgia; patient infected with SA93/01 had non fatal meningo-encephalitis and SPU116-89 infected patient showed symptoms of necrotic hepatitis and died (Burt *et al*., 2002). Mouse neuroinvasive experiments and gene expression data of H442, SPU116/89 and SA381/00 were described in Venter *et al*., (2005). The mouse neuroinvasive phenotype of strain SA93/01 was also determined as part of the above-mentioned study although it was not included in gene expression analysis due to financial constraints and therefore not reported before. SA93/01 was classified as highly neuroinvasive with an LD_{50}dose similar to that of SPU116/89 and H442 (between 2-3) (Venter, unpublished data), while SA381/00 had an LD_{50}dose of 316 and was classified as being of low neuroinvasive phenotype in mice.

Comparison of the four South African strains with other lineage II strains as well as with lineage I strains were aimed at identifying specific changes that may be associated with strain virulence. Alignments were therefore carried out relative to strains that were known to be highly neuroinvasive or mild in mice where possible, or that have been reported to be highly pathogenic or attenuated. Lineage II strains included in this analysis are strains for which both full genome sequences and neuronvirulence data in mice were available including isolate B956D117B3 (Lanciotti

et al., 1999) and Madagascar strain AnMg798. B956D117B3 is a passaged clone of reduced virulence (Yamshchikov *et al*., 2004) of the prototype strain B956, which was originally associated with fever in a patient and was neurotropic in mice (Smithburn *et al*., 1940). AnMg798 is not neuroinvasive (Beasley *et al*., 2002) Lineage I strains included, are the highly pathogenic NY385-99 strain which is highly neuroinvasive in mice (Beasley *et al*., 2002); the attenuated non-neuroinvasive strains TM171-03 isolate in Mexico in 2003 (Beasley *et al*., 2004); molecular attenuated characterized clones of the NY-385-99 Clone TYP-9376 and NY-385-99 Clone 9317B (Ding *et al*., 2005) and a non-neuroinvasive Kunjin virus (MRM61C).

Full genome sequencing indicated that few differences exist between the 4 South African strains at both nucleotide and amino acid levels. Phylogenetic and P-distance analysis suggest that differences between strains within a lineage are related to geographic location of isolation rather than temporal. The South African WNV strains were collected over a period of 50 years and still only demonstrate an average of 3% nucleotide (less than 1% amino acid) difference and a 21% nucleotide (more than 3% amino acid) difference with the Madagascar strain (AnMg798). The same situation was observed for lineage I strains. New York strain show a 0.1% amino acid differences between itself and the NY strain clones and 2.4% amino acid difference to Kunjin lineage I strain and an average of about 6% amino acid difference between lineage I and II WNV strains.

The envelope protein of Flaviviruses is the major envelope protein and the principal target for neutralizing antibodies. The E protein is also involved in host-cell receptor binding and membrane fusion (Lindenbach and Rice, 2003). The genetic stability observed in the surface proteins (membrane (M) and envelope (E)) of strains included in this study that were isolated over a period of 50 years suggest an absence of immune driven selection. Very few to no changes were found in the structural proteins of WNV lineage II strains. Only the H442 WNV strain, which had been isolated 50 years prior to the other strains, had two potential structural substitutions in the envelope gene (hydrophobic A to structural important G, and T to structural important P). Stability of envelope proteins which is exposed to host immune responses may have positive implications for vaccine design.

Most substitutions were found in the nonstructural proteins in particular NS3, NS4A/B and NS5. The NS3 protein is part of the protease complex, which is important for

cleavage of the polyprotein and may affect the virulence phenotype of the virus if the function is compromised (Hurrelbtink and McMinn., 202). If the polyprotein is cleaved less sufficiently, virus assembly and release are delayed which allows the host immune system to clear the infection (Hurrelbrink and McMinn, 2003). The NS3 protein of the mild strain, SA381/00 demonstrated hydrophobic and hydrophilic changes, which could lead to structural changes, thus possibly affecting the cleavage, function and thus potentially affect the virulence phenotype of the virus. SPU116/89, which is a highly pathogenic WNV strain, has a hydrophilic substitutions in the NS4A/B protein, which appear to be involved in viral replicase (Beasley, 2005). Mutations affecting replication may alter the tropism of the virus or improve or suppress it's ability to replicate in the host.

The majority of the mutations between the strains were found in the NS5 protein. This protein is thought to be responsible for cytoplasmic RNA replication by the replicase complex because it contains an RNA-dependent RNA polymerase (RDRP), *S*-adenosylmethionine methyl transferase (SAM), and importin-*β* binding motifs (Hurrelbrink and McMinn, 2003). Deletions in the NS5 protein has been shown to abolish replication (Beasley, 2005) suggesting that amino acid substitutions in this protein may effect replication efficiency and thereby the virulence phenotype of the virus. Attenuated strains with a temperature sensitive phenotype and reduced virulence in mice have been isolated in Texas that also contained mutations in the NS5 proteins (Davis *et al.*, 2004). In addition, the organ tropism of strains have also been affected by mutations in the NS5, NS2 and E-proteins (Ding *et al.*, 2005). A resent study showed that defects in both the guanine n-7 and ribose 2'-O methyltransferase (MTase) activity is lethal to the virus, but viruses defective only in 2'-O methylation are attenuated and renders protection to mice from wild-type WNV challenge. Thus N-7 methylation is important for survival of WNV (Zhou *et al*., 2007). The two strains that caused mild diseases in patients H442 and SA381/00 had hydrophilic amino acid where the other two strains had hydrophobic amino acids at position 614 and at positions 625 and 626 it's the other way around. In positions 197,623,635,641 and 643 strain SPU116/89, which was isolated from a patient with necrotic hepatitis, had amino acid changes that may affect this protein's hydrophobicity and structure. From the results of previous studies and taking these substitutions together with specific symptoms caused in patients into consideration it can be possible that these substitutions may affect tissue tropism and replication efficiency and may even possibly have an influence on pathogenicity.

The absence of a putative E-protein glycosylation site 154-156 (NYS) have previously been associated with reduced virulence in mice (Beasley *et al.*, 2004). All four South African WNV strains were glycosylated in this position. The less neuroinvasive strain SA381/00 as well as the laboratory attenuated strains (NY385-99 clone TYP-9376) and (NY385-99 clone 9317B) were glycosylated at this motif. The prototype strain B956 from Uganda and the non-neuroinvasive strains AnMg798 and MRM61C were not glycosalyted. This suggests that virulence is not only determined by the glycosalytion of the envelope protein although it may play a contributing role in virulence determination.

Very few differences existed in the surface proteins between these strains, and none could be identified between the highly neuroinvasive and less neuroinvasive strains. Differences in the strains that caused benign disease, (SPU381/00 and H442), non-fatal meningo encephalitis in patient (SPU93/01) and fatal hepatitis (SPU116/89) were mainly located in the NS proteins. These differences could potentially affect replication efficiency and viral load. No differences were observed in RNA levels in the liver of mice infected subcutaneously with the same amount of these strains (Venter *et al.*, 2005). Since various differences exist between lineage I and II strains it would be difficult to identify specific sites associated with virulence differences between the lineages.

The number of cases of neurological infections seen in the USA relative to South Africa may be attributed to the rapid distribution of a single highly neuroinvasive strain in a highly susceptible population. Nevertheless, the low number of WNV fever or neurological cases reported in South Africa despite the wide distribution of the virus and the presence of highly neuroinvasive strains may also reflect inadequate surveillance and a lack of medical awareness of the disease potential of arboviruses in the country. The importance of WNV in the country may be overshadowed by the presence and impact of other diseases such as HIV-AIDS. The epidemic potential and the potential impact and undescribed outcome that WNV may have on HIV-AIDS patients warrants improved surveillance for arboviruses in humans in South Africa.

To conclude, this study provides the first full genome sequences of highly neuroinvasive lineage II WNV strains. Comparison of the highly neuroinvasive and mild strains suggest that various molecular factors may contribute to the pathogenic

phenotype of these strains, but those mutations in the non structural proteins which encodes the viral replication and protein cleavage mechanisms are the most likely determinants of strain differences.

Chapter Three:

Expression of Recombinant WNV antigens

3.1 INTRODUCTION

West Nile fever (WNV) in humans is usually a febrile, influenza-like disease with an incubation period of 3-6 days associated with myalgia and arthralgia, WNV neuroinvasive disease may also develop in some cases as acute aseptic meningitis, encephalitis or poliomyelitis. Anterior myelitis, hepatosplenomegaly, hepatitis, pancreatitis and myocarditis can also occur (Petersen and Marfin , 2002;Hubálek and Halouzka, 1999). Virus can be recovered from the blood for up to 10 days in the case of immunocompetent febrile patients and as late as 22 to 28 days in immunocompromised patients whereas peak viremia occur 4 to 8 days postinfection (Hubálek and Halouzka, 1999). IgM antibodies can be detected within 8 days after onset of symptoms and IgG antibodies within 12 days. Nucleic acid tests and viral culture can be too insensitive to use for diagnostic and surveillance tools, due to the short viremic phase and low viral load in humans. The alternative is to use an assay that detects WNV-specific antibodies in serum, plasma or cerebrospinal fluid (Prince and Hogrefe, 2005).

Untill 2002 WNV diagnostic assays used tissue culture or mouse-propagated virus as an antigen source (Beasley, 2005; Prince and Hogrefe, 2005). Large amounts of virus have to be produced for production of these assays, leading to a high safety risk. Only specialized BSL-3 laboratories can work with these viruses and even the extracted positive single stranded RNA is still infectious (Center of Disease Control, Beasley, 2005). Therefore, safe methods for producing antigens must be investigated. Alternative methods to produce virus antigen are cloning and expression of recombinant viral antigens in bacterial, baculovirus or mammalian expression systems that allow for the production of large amounts of antigens under safe conditions. IgM antibody capture ELISA (MAC-ELISA) and IgG-ELISA using either purified, inactivated viral or recombinant subviral antigens has already been developed and approved by the FDA for lineage I WNV detection and shown to be successful as discussed in chapter one (Section 1.9) (Beasley, 2005).

All *flaviviruses* are closely related antigenically, which accounts for the serological cross-reactions observed in diagnostic assays. Members of the JE complex are even more closely related, often needing specialized tests to differentiate the infecting flavivirus (Petersen and Roehrig, 2001). Because of the antigenic cross-reactivity between different flaviviruses, WNV ELISA positive samples are retested for

confirmation by plaque reduction neutralization test (PRNT) and nucleic acid testing procedures are used in surveillance studies to unequivocally identify WNV as the causative agent of an outbreak (Briton, 2002; Beasley, 2005). The assumption can be made that for a recombinant antigen to be used successfully in a diagnostic assay, such antigen has to be identical to, or at least a close match to the native antigen

Limited surveillance is being performed at present for WNV in South Africa and the true incidence of WNV infections; geographical distribution and the severity of the disease are unknown factors. A few studies on these factors are e.g. Burt *et al*., 2002, Jupp, 2001 and Jupp *et al*., 1986. Current diagnostic tests make use of inactivated WNV, or recombinant subviral antigens and polyclonal antibodies (Beasley, 2005; Prince and Hogrefe, 2005). The aim of this study was to produce a recombinant antigen that can potentially replace the use of inactivated WNV.

In this study it was attempted to produce a recombinant antigen. This antigen can possibly be applied in a diagnostic enzyme linked immunosorbent assay (ELISA) test for detection of anti-WNV antibodies, which is safer and more economical than available commercial tests and more specific to lineage II strains. This test can potentially be used in epidemiology surveillance of WNV in South Africa. The ELISA assay was chosen because it is the most widely used, safe and cost affective diagnostic method at present for lineage I WNV strain surveillance (Beasley, 2005, Prince and Hogrefe, 2005). The chosen antigen for expression was the envelope protein since it is is the viral hemaglutinin that mediates virus-host cell binding, elicits most of the virus neutralizing antibodies and determines the serological specificity of the virus (Campbell *et al*., 2002). Due to the co-expression nature of the WNV, both the prM and E genes have to be expressed in order to obtain correctly folded immunogenic antigens. It is thought that the prM acts as a shield for the E protein when processed in the ER, thus protecting it from the environment that allows the correct folding of the envelope protein to take place (Section 1.6.1.1) (Petersen and Roehrig, 2001; Konishi and Mason, 1993; Lorenz *et al*., 2002).

3.2 MATERIALS AND METHODS

3.2.1) Virus strain and RNA isolation

South African WNV isolate H442 was used for antigen expression. The isolate was obtained from the Special Pathogens Unit, National Institute for Communicable Diseases, South Afirca, as freeze dried mouse passage 2-4 and replicated by one passage in Vero cells to a titre of $10^{6.25}$ /ml. Viral RNA was then extracted from 140µl Vero cell supernatant (placed in lysis buffer in the BSL-3 laboratory of NICD (SPU)) with the OIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany) according to manufacturer's instructions for microcentrifuge extractions (Section 2.2.4).

3.2.2) Primer design

Primers for the amplification of the prM and E gene as a single unit were designed using DNAMAN (Figure 3.1 and Table 3.1). Strain B956D117B3 (M12294) sequence was used to design primers (this strain is the clone sequence of the Uganda prototype strain B956). Sequence was obtained from GenBank. The complete genome sequence of a West Nile virus lineage II strain (accession number: M12294) was aligned with the E gene sequence of strain H442 (accession number: AF459043). The open reading frame, orientations and translation initiation and stop codons of the genes were taken into consideration when designing the primers. A restriction map of the E gene was also constructed. Primers were designed to bind at the beginning of the prM gene and at the end of the E gene. A sequence primer was also designed to verify the DNA sequence and to confirm that mutations were not generated during the PCR amplification process.

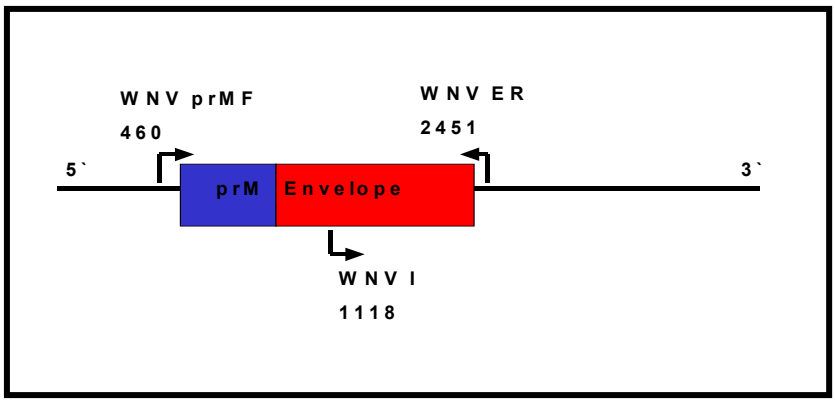


Figure 3.1: Diagram indicating the positions and names of primers used to amplify and sequence the prM and E genes. Positions are based on the WNV strain B956D117B3; GenBank accession number M12294. WNV I was designed and used to sequence the internal part of the envelope gene not sequenced by WNV prMF and WNV ER.

Table 3.1: Nucleotide sequence and annealing temperatures of the PCR and sequence primers used in constructing a recombinant mammalian expression vector.

Primer name	Primer sequence	Position	Length	Tm
WNV prMF	5` GGAGCTGTGAACCTCTCGAAC 3`	460	21	59.6°C
WNV ER	5` GACGTTGACCGAAAGGAAGAG 3`	2451	21	55.6°C
WNV I	5` AACCTCGCAGATGTGCGC 3`	1118	18	58°C
T7 forward	5` TAATACGACTCACTATAGGG 3`	On vector	20	56°C
BHG reverse	5` TAGAAGGCACAGTCGAGG 3`	On vector	18	56°C

3.2.3) First strand cDNA synthesis

Reverse transcription reactions were setup in a final volume of 25 µl. The reaction consisted of 10 mM dNTP's, 4 µg Random Hexanucleotides, 1 X M-MLV Reaction Buffer, 10 µl RNA, 1 µl RNase Inhibitor, 1 µl M-MLV Reverse Transcriptase (20U/µl), and 5 µl DEPC treated H_2O. (All reagents from Promega, Southampton, United Kingdom). The reaction mixture was first incubated for 10 min at 25˚C and then at 42˚C for 60 min. Heating at 95˚C for 5 min stopped the reaction. Random hexanucleotide primers were used because specific primers designed for amplification of the prM/E unit was unsuccessful.

3.2.4) PCR amplification of the E and prM genes as a single unit

The PCR reaction to amplify WNV prM and E genes as a unit was set up in a total volume of 100 µl as follows: 10 µl cDNA template; 1 X Reaction Buffer (10 mM Tris-HCl, 50 mM KCl, 0.1% Triton X-100); 40 pmol of each primer (WNVER; WNVprMF); 25 mM $MgCl_2$; 3.5 U Taq Polymerase; 25 mM dNTPs (All reagents from Promega, Southhampton, United Kingdom); 60.3 µl DEPC H_2O. Thermocycles were performed in a programmable thermocycler (GeneAmp PCR system 3700). The cycles used were; initial denaturation at 94ºC for 5 min; followed by 30 cycles of denaturation at 94ºC for 1 min; annealing at 50ºC for 1 min; extension at 72ºC for 90 seconds; with a final extension step of 72ºC for 10 min. After optimization of the PCR reactions, Taq Polymerase (Promega Corporation) was replaced with 3.75 U of Expand High Fidelity Polymerase (Roche Diagnostics, Mannheim, Germany) due to its proofreading activity and PCR's were performed.

Controls: A negative control was included consisting of all the reagents with no template to verify if any contamination is present. The diagnostic primers CFD2 and FU1 (Kuno *et al*., 1998) were used in the positive control reaction to analyze the reagents performance and to establish if cDNA template were intact. These control reactions were set up in parallel with the PCR to amplify WNV genes.

Analysis: 10 µl of the PCR reaction and 2 µl of gel loading buffer (0.25% Bromophenol Blue in 40% sucrose solution) were mixed and loaded in a 1% agarose gel stained with ethidium bromide (final concentration 0.5 µg/ml). A 1kb DNA ladder was also loaded (Promega, Southhampton, United Kingdom). Electrophoresis was carried out in 1 X TAE buffer (40 mM Tris-HCl; 1 mM EDTA; pH 8.5) at 100 V in a horizontal gel electrophoresis tank.

The Wizard SV gel and PCR clean-up system (Promega, Southampton, United Kingdom) was used to purify the PCR amplicons from an agarose gel after electrophoreses (Section 2.2.6). This was done because multiple bands were obtained.

Analysis: 1 µl of purified PCR product was analyzed by agarose gel electrophoresis as described in Section 3.2.4 in order to visually determine the concentration of the DNA by comparing it to a marker of known concentration.

3.2.5) Molecular cloning of the WNV prM and E genes

3.2.5.1) Cloning of PCR product into the pcDNA3.1/V5-His© TOPO® vector

The pcDNA3.1/V5-His© TOPO® TA Expression Kit ligation protocol (Invitrogen Ltd, Germany) was followed. 40 ng of the purified PCR product and 10 ng of the TOPO vector were ligated in the presence of 1 µl supplied salt solution (1.2 M $NaCl_2$, 0.06 M $MgCl_2$). All the reagents were mixed and incubated for 20 min at room temperature (22-23°C) after which the reaction was incubated on ice. To be noted is that this method doesn't use DNA ligase but topoisomerase that covalently binds to the vector and mediates the binding between the vector's T overhangs at the cloning site and the PCR product A-overhangs (See Figure 3.2).

3.2.5.2) Transformation

One shot TOP 10 chemically competent *Escherichia coli* cells supplied with the TOPO cloning kit was used following the manufacturer's instructions (Invitrogen Life Technologies). Briefly 50 µl Top 10 *E. coli* competent cells and 2 µl of ligation mix were mixed and incubated on ice for 20 min, heat-shocked at 42°C for 30 sec and then cooled on ice. 250 µl SOC medium (2% Tryptone, 0.5% Yeast Extract, 10 mM $NaCl_2$, 2.5 mM KCl 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added to the mixture and incubated with agitation at 37°C for 1 hour. Luria-Bertoni (LB) agar plates (1% Tryptone, 0.5% Yeast extract, 1% NaCl, 1.5% Agar) were prepared as follows; different volumes (25 µl, 100 µl, and 175 µl) were plated out on separate plates. 25 µl of the positive and 25 µl of the negative controls were plated on LB agar plates supplemented with ampicillin (50 µg/ml).

Controls: A negative control, which consisted of only 50 µl Top 10 *E. coli* competent cells, was included to verify if any contamination was present. A positive control was also included from which the transformation efficiency was calculated. The positive control consisted of 50 µl Top 10 *E. coli* competent cells and 10 µl pUC18 (1ng/µl) plasmid.

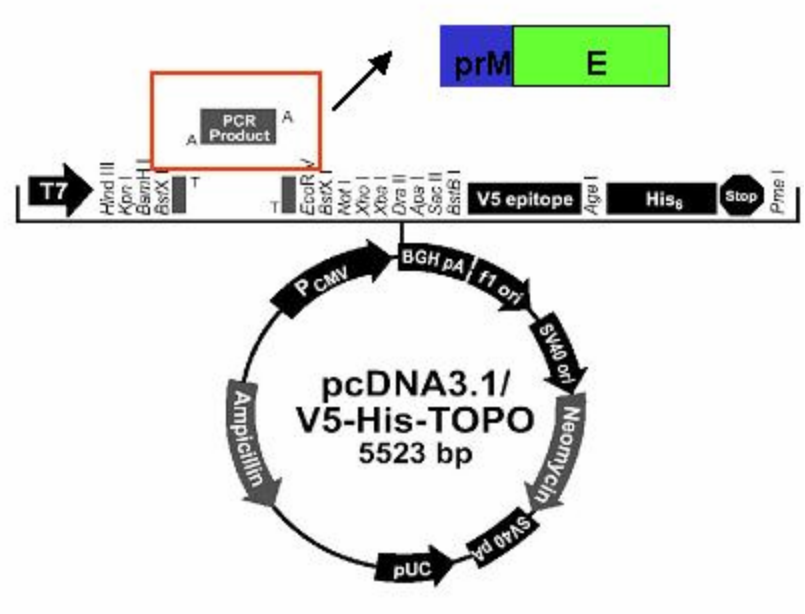


Figure 3.2: pcDNA 3.1/V5-His©TOPO® vector map with the red block indicating where the PCR amplicon prM/E (blue and green blocks) was cloned into. (Vector map modified from Invitrogen TA expression manual version G, 073001, 25-0203)

3.2.5.3) Screening for recombinants

i. Plasmid Isolation

Transformed colonies were picked randomly from LB-agar plates and incubated in LB broth (1% Tryptone, 0.5% Yeast extract, 1% NaCl; pH 7.34) supplemented with ampicillin (50 µg/ml). The inoculated LB-broth was incubated overnight at 37°C with agitation. Plasmid isolation was performed using the alkaline lysis method described in Sambrook *et al*., (1989). Briefly 2.5 ml cell culture was harvested by benchtop centrifuging at 10 000 g for 30 sec. Pellets were resuspended in 100 µl ice-cold solution 1 (50 mM Glucose; 10 mM EDTA; 25 mM Tris; pH8.0) and then 200 µl of solution 2 (0.2 M NaOH, 1% SDS) was added. Tubes were incubated on ice until solutions became clear. Hereafter, 150 µl solution 3 (3 M NaOAc; 5 M Acetic acid) was added and the mixture vortexed immediately and again incubate on ice for 5 min. The solution was then centrifuged at 12 000 g for 15 min in a bench-top microcentrifuge. The clear supernatant containing the plasmids were transferred to a new tube and plasmids precipitated with two volumes of 95% ethanol at room temperature for 30 min. Plasmid DNA was pelleted by centrifugation at 10 000 g for 10 minute. After the supernatant was discarded the pellets were washed with 70% ethanol and air-dried. Finally pellets were resuspended in either nuclease free water or TE buffer (10 mM Tris-HCl; 1 mM EDTA; pH 8.0).

Analysis: 1 µl of the mixture was analyzed by agarose gel electrophoreses as described in Section 3.2.4 to analyze if plasmid DNA was isolated successfully. Plasmids were then further analyzed by restriction enzyme digest and DNA sequencing as described below.

ii. Restriction enzyme analysis

Eco RV (Promega, Southhampton, United Kingdom) was used in restriction enzyme digest of selected plasmids (Figure 3.3). 1 µg DNA and 1U EcoRV were used in the restriction reaction together with 1 X Enzyme buffer D (6 mM Tris-HCL, 6 mM $MgCl_2$, 150 mM NaCl; 1 mM DTT, pH 7.9) and 2.5 µl DEPC water to obtain a final volume of 15 µl. The reaction mixture was incubated at 37ºC for 2 hours. Hereafter, the reaction was terminated at 65ºC for 10 min.

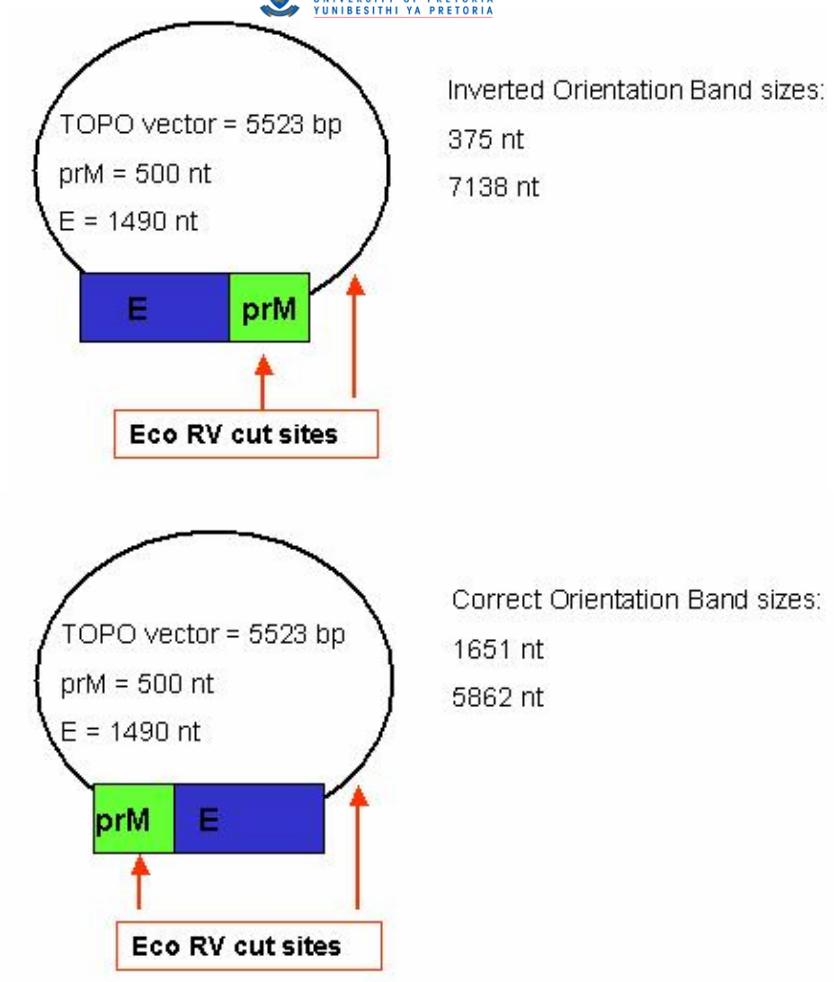


Figure 3.3: Vector map to indicate the orientation of the insert when in the correct/incorrect position as well as to indicate where EcoRV restricts.

Analysis: 1 µl of unrestricted and restricted plasmid were analyzed by agarose gel electrophoresis as described in Section 3.2.4. A 1 kb lambda marker (Promega, Southhampton, United Kingdom) was included on the gel.

iii. DNA sequencing

Selected tramsformed *E.coli* cells containing recombinant plasmid with correctly orientated insert according to the restriction enzyme digest results was cultured overnight in ampicillin supplemented LB broth as described (Section 3.2.5.3 (i)). This was necessary to obtain enough DNA for sequencing analysis and transfection of mammalian cells. The Wizard *Plus* SV miniprep DNA Purification system was used to purify plasmid DNA (Promega, Southhampton, United Kingdom). Procedures according to supplied technical bulletins were followed. Briefly 5 ml bacterial cells were pelleted by centrifugation for 5 min at 10 000 g. Cells were resuspended in 250

µl resuspension solution and 250 µl of cell lysis solution was added. The contents of the tubes were mixed by inverting the tubes 4 times. This was followed by an incubation step of 5 min at room temperature. After incubation 10 µl alkaline protease solution was added and mixed and incubated as before, after which 350 µl of neutralizing solution was added and mixed immediately. Bacterial lysate was centrifuged at maximum speed in a bench-top microcentrifuge for 10 min at room temperature. Cleared lysate was transferred to spin columns and again centrifuged for 1 min at maximum speed. Next columns were washed twice; once with 750 µl washing solution and another with 250 µl and centrifuged as before. Plasmids were eluted with nuclease free water and either stored at –20ºC or –70ºC according to the time period that will elapse before the plasmids will be used for either DNA sequencing analysis or transfection.

Analysis: Gel electrophoresis as described in Section 3.2.4 was used to analyze the purity of the plasmids visually and to determine the concentration of the plasmids after purification.

The T7 sequencing primer, BHG reverse sequencing primer, West Nile Virus prM, E and I primers were used for DNA sequencing (Table 3.1 and Figure 3.1). DNA cycle sequencing using the BigDye Terminator V3.1 kit (Applied Biosystems, Warrington, United Kingdom) was conducted as recommended by the supplier. Briefly 4 µl Ready Reaction Premix, 2 µl BigDye sequencing buffer (5 X), 3.2 pmol primer and 200 ng/µl DNA template was used in a final volume of 20 µl. Sequencing cycles was as follows: 96ºC for 1 min, 25 cycles of 94ºC for 10 sec, 50ºC for 5 sec and 60ºC for 4 min. DNA was purified using ethanol purification. Briefly 100 µl of 60% ethanol was added per reaction tube and incubated at room temperature for 15 min. After incubation tubes were centrifuged in a bench-top microcentrifuge for 20 min at maximum speed to pellet precipitated DNA. The pellet was then washed with 250 µl 70% ethanol, and dried on a heating block at 90°C for 1 min. Sequence reactions were analyzed on an ABI P_{RISM} ® 3100-3130 genetic analyzer at the sequence facility at the Natural and Agricultural Faculty, University of Pretoria, South Africa.

iv. DNA sequencing analysis

Obtained sequence was analyzed using the BLAST function on GenBank to confirm the origin of the sequence was West Nile virus. Hereafter the sequences were

assembled and edited in Vector NTI 9.1.0 (© Invitrogen Corporation, 2004) A ClustalW alignment was performed with obtained sequence and the WNV strain B956D117B3 (GenBank accession number: M12294). Sequence was translated to confirm the open reading frame in order to be sure it will be able to be expressed during consecutive expression studies.

3.2.6) Expression in a mammalian expression system.

3.2.6.1) Maintenance of cells.

General cell culture procedures and aseptic techniques were used in maintaining the cell line according to Freshney, 2000.

BHK 21 cells, passage 125, were received from a private vaccine company, Design Biologix, Pretoria, South Africa. These cells were thawed, sub cultured and cryopreserved in order to maintain a -70°C cell culture line/mother stock. The cells were thawed in a water bath at 37°C. 2 ml of thawed cells was then added to 25 ml complete growth medium consisting of Dulbecco`s modified eagle's medium (DMEM) (GIBCO, Invitrogen, Germany) supplemented with 20% heat inactivated Fetal Bovine serum (FBS) (GIBCO, Invitrogen, Germany) and 1 X antibiotic mixture (PFS) (10 000 µg/ml penicillin, 10 000 µg/ml streptomycin, 25 µg/ml fungizone) (Highveld Biological, South Africa). Cells were centrifuge at 1 200 g for 2 minutes, supernatant discarded and cells gently resuspended in 10 ml complete DMEM growth medium. The cell suspension was transferred to a T25 flask and incubated at 36.5°C supplemented with 4.5% CO_2.

For subculturing, cultures were routinely examined microscopically for morphological deterioration. Cells were sub-cultured (1:4 ratio) when they reached 90% confluency. Briefly medium was removed from the culture flask with a mechanical pipettor, after which monolayer of cells were washed 3 times with 5 ml 1 X trypsin (Sigma-Aldrich, Germany). The flasks were incubated at room temperature for 15 min until the cells dislodged from the floor of the culture vessel. Trypsinized cells were resuspended in pre-warmed complete growth medium using DMEM (GIBCO, Invitrogen, Germany) supplemented with 1 X antibiotic mix PSF and 10% FBS. The cells were subcultured at an appropriate split ratio. BHK-21 cells were maintained in EMEM supplemented with 10% FBS and grown under standard growth conditions.

In order to cryopreserve the cells, cells were grown to a confluency of 90% ($8X10^6$ cells/ml in T75 flask), trypsinized and resuspended in pre-warmed freezing medium (20% FBS, DMEM, 10% DMSO) to a final concentration of $1X10^6$ cells/ml. 1 ml cell suspension was dispersed into cryotubes. Cells were gradually frozen to a temperature of -70°C to avoid damage to cells by crystallization of medium. Cells were either stored at -70°C or in liquid nitrogen.

3.2.6.2) Optimization of transfections

BHK 21 cells were transfected with the recombinant pcDNA3.1/V5-His-TOPO/lacZ vector (Invitrogen, Germany), using SuperFect Transfection reagent (Qiagen, Germany) according to the manufacturer's instructions for use on 6 well plates (Greiner bio one, Germany). Briefly, cells were seeded into a 6 well plate and incubate as described in Section 3.2.6.1 to a confluency of 60% ($5x10^6$ cells). DNA was diluted with DMEM cell culture medium (GIBCO, Invitrogen, Germany) containing no added proteins, serum of antibiotics and to which SuperFect transfection reagent was added. (Table 3.2). The mixture was incubated at room temperature for 10 minutes to allow transfection complexes to form. Cells were washed with 1 X PBS (13.7 mM NaCl, 0.25 mM KCl, 0.43 mM $Na_2HPO_4.2H_2O$, 0.14 mM KH_2PO_4, pH7.3) (Promega, Southampton, United Kingdom) and complete growth medium was added to the mixture containing the transfection-complexes, mixed and transferred to cells in 6 well plates. Plates were then incubated under normal growth conditions for 4 hours after which they were washed again with 1 X PBS. Fresh complete medium containing 10% serum was added to each well and plates incubated under normal growth conditions for 48 hours.

Table 3.2: Ratio of pcDNA3.1/V5-His-TOPO/lacZ control vector DNA to transfection medium is given below as well as the amount of vector DNA used in optimization experiments for transfections.

Ratio of vector DNA to SuperFect transfection reagent	Volume of SuperFect used when 2 µg vector DNA was added	Volume of SuperFect used when 5 µg vector DNA was added
1:1	2 µl SuperFect used	5 µl SuperFect used
1:2	4 µl SuperFect used	10 µl SuperFect used
1:3	6 µl SuperFect used	15 µl SuperFect used
1:5	10 µl SuperFect used	25 µl SuperFect used

3.2.6.3) Analysis of control transfections with β-gal staining

The β-Gal staining kit from Invitrogen was used to monitor transfection by measuring β-galactosidase expression. Manufacturer's instructions were followed. Briefly, cells were fixed by adding 1.5 ml fixation solution (20% Formaldehyde, 2% glutaraldehyde in 10 X PBS) to each well and incubated for 10 min at room temperature. Prepared staining solution, which contains 12.5 µl of staining solution A (400 mM $K_3Fe(CN)_6$), staining solution B (400 mM $K_4Fe(CN)_6.3H_2O$) and staining solution solution C (200 mM $MgCl_2$) was added to each well. Staining solution also contained 62.5 µl 20mg/ml X-Gal in dimethylformamide (DMF) per well and 1.15 ml 1 X PBS per well. Cells were rinsed with 1 X PBS. The 1.25 ml staining solution was added to each well containing the transfected cells and incubated at 37°C for 24 hours. Wells were then analyzed for the production of a blue color that will indicate expression of β-galactosidase.

Controls: A well of cells was left untransfected to monitor conditions of cells through-out the process and a well was transfected with SuperFect transfection reagent but no DNA was added to analyze the effects that the transfection reagent has on the cells.

3.2.6.4) Transfection of BHK 21 cells with recombinant plasmid

Transfection was performed as described in Section 3.2.7.2 with slight modifications; Transfection was performed in a T-75 flasks and not in 6 well plates. 2 µg of DNA with a ratio of 1:3 was best suited and this parameter was used for the transfection. An increase in volume of 10 µg of DNA was used and 30 µl of SuperFect transfection (QIAGEN, Germany) reagent. After a 48-hour incubation cells were harvested and washed with 1 X PBS. Cell pellets were resuspended in 1 X PBS after centrifugation at 1 000 g for 10 min and stored at 4°C until further analysis.

3.2.6.5) Creating a stable cell line

i. Determination of the concentration of geneticin to use

The pcDNA3.1/V5-His© TOPO® vector used to construct the recombinant WNV plasmid for transfection of BHK-21 cells for protein expression contains a neomycin resistance gene to allow selection of stable cell lines using Geneticin®. Geneticin

blocks protein synthesis in mammalian cells by interfering with ribosomal function. Expression in mammalian cells of the bacterial aminoglycoside phospho-transferase gene (APH), derived from Tn5, results in detoxification of Geneticin® selective antibiotic (Southern and Berg, 1982; Invitrogen life technologies instruction manual). Before creating a stable cell line the minimum concentration of Geneticin® necessary to kill untransfected BHK-21 cells need to be determined. The concentration was determined as follows; six-well plates were seeded and grown overnight to a confluency of 25%. Different concentrations of Geneticin® were added to each well (50, 125, 250, 500, 750, 1000 µg/ml Geneticin®). The selective medium was replenished every 3-4 days and percentage of surviving cells was counted regularly to determine the appropriate concentration of Geneticin® that prevents growth of BHK-21 cells within 1-3 weeks.

ii. Creating a stable cell line WNV-prM/E clone.

Mammalian BHK-21 cells were transfected with SuperFect transfection medium and the recombinant pcDNA3.1/V5-His© TOPO® vector. The cell culture medium was supplemented with 375 µg/ml of Geneticin®. This selective medium was replenished every 3-4 days until only transfected cells resistant to Geneticin® survived. BHK-21 cells stably transfected with recombinant vector was cryopreserved and store at -70°C.

3.2.6.6) Analysis of expression by Immunofluorescent antibody testing

Harvested transfected BHK-21 cells were resuspended in 500 µl 1 X PBS. 10 µl cell solution was spotted on a microscope slide and dried at 37°C for 20 min and if not yet dry a hairdryer was used to completely dry slides. Cells were fixed by placing slides in acetone for 15 min. A 1/10 dilution of mouse anti-WNV antibody (SPU, NICD) in 1 X PBS was added to slides and incubated at 37°C for 30 min. Slides were washed three times with 1 X PBS for 3 min and once for 1 min with dH_2O. Slides were again dried with a hairdryer. Anti-mouse FITC antibody (Sigma-Aldrich, Germany) diluted 1/40 in Evans blue (Merck, Germany) was added and incubated for 30 min at 37°C, washed with 1 X PBS and distilled water and dried as before. Mounting fluid and a cover slip were added and results were read under a fluorescent microscope at a 10 X magnification.

3.2.6.7) Analysis of expression by SDS-PAGE

WNV virus proteins expressed in BHK-21 cells were analyzed by protein electrophoresis carried out on SDS-polyacrylamide gels under denaturing conditions (Sambrook *et al.*, 1989). A 12% separating polyacrylamide gel (12% acrylamide, 1% TEMED, 10% ammonium persulphate) and 6% stacking gel (6% acrylamide, 1% TEMED, 10% ammonium persulphate) were prepared from a 30% acrylamide/0.8% bisacrylamide stock solutions. To induce chemical polymerisation, TEMED and ammonium persulfate were added. Transfected BHK-21 cells were first lysed by adding 200 µl lysis buffer (100 mM Tris-HCl, 100 mM NaCl, 0.5% Triton-100) to one T-75 flask of harvested cells. The mixture was vortexed and incubated on ice for 10 min and centrifuged for 10 min at 10 000 g at 4°C. The supernatant was transferred to a new tube. Both the supernatant and pellet were denatured in an equal volume of 2 X Protein Solvent buffer (PSB) (0.125 M Tris, pH 8; 4% SDS; 20% glycerol; 10% 2-mercaptoethanol) and heated at 95°C for 5 min. The denatured samples were separated by SDS-PAGE at 100 V in 1 X TGS buffer (prepared from a 10 X stock 3% Tris; 14.4% Glycine; 1% SDS). Electrophoreses was carried out using a vertical slab gel unit (Hoefer) for 2 hours. A full range rainbow molecular weight marker (Amersham Biosciences) was also loaded. After electrophoresis was completed a gel was stained with Coomassie Brilliant Blue stain solution (0.125% Coomassie Brilliant Blue; 50% methanol; 10% acetic acid) for 20 min. The gel was destained overnight with destaining solution (5% Ethanol, 5% Glacial acetic acid) and agitation. A duplicate gel was used in subsequent Western blot analysis.

3.2.6.8) Analysis of expression by Western Blot analysis

Proteins were transferred from the acrylamide gel to a nitrocellulose membrane electrophoretically by placing the acrylamide gel and the membrane between two layers of filter paper soaked in 1 X transfer buffer (prepared from 10 X stock solution; 250 mM Tris and 192 mM Glycine; 1 X solution contains 20% methanol) in a transfer tank for 2 hours at 100 V. After electrophoresis the membrane was rinsed for 5 min with 1 X PBS. The membrane was then incubated at 4°C overnight in a blocking buffer (1 X PBS and 1% fat free milk powder w/v). After blocking, the membrane was washed in 1 X PBS. Primary mouse anti-WNV antibody (SPU, NICD, Sandringham, South Africa) solution diluted 1:100 in blocking buffer was added. The membrane was incubated with agitation at room temperature for 2 hours and then washed three

times (5 min with agitation) with wash buffer (1 X PBS with 0.05% Tween-20 v/v). Washing steps were repeated with TBST (1 X TBS with 0.1% Tween-20 v/v). The membrane was then incubated for 2 hours with goat-anti mouse-alkaline phosphatase (AP)-conjugated antibody (Sigma-Aldrich, Germany), diluted 1:2 000 in dilution buffer (TBST with 1% fat free milk powder w/v). The membrane was again washed three times with TBST and once with TBS (20 mM Tris, 140 mM NaCl). Fresh substrate solution was prepared and added to the membrane (BCIP/NBT Alkaline Phosphates substrate, Sigma) and incubated with agitation at room temperature for 1 hour in the dark. The color development was stopped by a wash step with distilled water.

3.2.6.9) PCR to confirm the presence of the cloned gene in the stable cell line

Plasmid was extracted from BHK 21 cells stable transfected with recombinant construct with the Wizard Plus SV miniprep DNA purification system as described in Section 3.2.5.3 iii. A PCR was performed on the extracted plasmid to confirm presence of clone gene unit prM/E. The PCR master mix consisted of 3.75U of Expand High Fidelity polymerase (Roche Diagnostics, Mannheim, Germany), 300 µM of each dNTP (Promega, Southampton, United Kingdom), 1 X Buffer (10 mM Tris-HCl, 50 mM KCl, 0.1% Triton X-100), and 30 pmol of each primer (WNV prMF and WNV ER) in a final volume of 50 µl and with 5 µl template. Cycles were as follows: Initial denaturation at 94ºC for 2 min; 10 cycles of denaturation at 94ºC for 10 sec, annealing at 50ºC for 30 sec, extension at 72ºC for 3 min; Followed by 35 cycles of denaturation at 94ºC for 15 sec, annealing at 50ºC for 30 sec, extension at 72ºC for 3 min plus 10 sec per cycle and a final extension step of 72ºC for 7 min. Analysis of the PCR amplicons was perform as described in Section 2.2.6.

3.2.6.10) Recombinant protein purification in BHK 21 cells

For purification of recombinant proteins the ProBond^{TM} Purification System (Invitrogen life technologies) was used following the manufacturer's instructions. Briefly cells were harvested by trypsinization and cell pellets resuspended in 8 ml Guanidinium Lysis Buffer (6 M guanidine HCL, 20 mM NaH_2PO_4 (sodium phosphate), pH 7.8 and 500 mM NaCl). DNA was sheared by passing the preparation through an 18-gauge needle four times. Lysate was centrifuged at 3 000 g for 15 min to pellet the cellular debris and the supernatant transferred to a clean tube. Lysate was stored on ice while the Probond column was prepared. To prepare the column, 2 ml

ProBond resin (50% slurry in 20% ethanol) was added to the column and allowed to settle by gravity for 15 min. After 15 min the supernatant was aspirated and the resin was resuspended with 6 ml sterile, distilled H_2O and allowed to settle again as previously described after which supernatant was removed. Resin was resuspended in 6 ml Denaturing Binding Buffer (8 M urea, 20 mM NaH_2PO_4 pH7.8 and 500 mM NaCl) and allowed to settle and supernatant was again aspirated. This was repeated twice. Eight ml lysate was added to the prepared column and proteins were allowed to bind for 30 min at room temperature with gentle agitation to keep the resin suspended in the lysate solution. Resin was settled by low speed centrifugation at 800 g for 1 min. The column was washed with 4 ml Denaturing Binding Buffer by resuspending the resin, rocking for 2 min, allowing it to settle and removing the supernatant. Washing was repeated twice with 4 ml Denaturing Wash Buffer (8 M urea, 20 mM NaH_2PO_4, pH 6 and 500 mM NaCl). The column was washed four times with 8 ml Native Wash Buffer (50 ml 1 X Purification Buffer (5 X purification buffer; 250 mM NaH_2PO_4 pH8, 2.5 M NaCl) and 335 µl 3 M imidazole (3 M Imidazole, 20 mM sodium phosphate, pH6 and 500 mM NaCl)). The proteins were eluted by adding 8 ml Native Elution Buffer (13.75 ml 1 X Purification Buffer (5 X purification buffer; 250 mM NaH_2PO_4 pH8, 2.5 M NaCl) and 1.25 ml 3 M imidazole (3 M Imidazole, 20 mM sodium phosphate, pH6 and 500 mM NaCl).

3.2.6.11) Concentration of purified proteins

Vivaspin 20 from vivascience (Sartorius, United Kingdom) was used to concentrate purified recombinant proteins following the manufacturer's instructions. In brief, spin column Vivaspin 20 with the membrane 5000 molecular weight cut off (MWCO PES) with a protein molecular weight cut-off 66 000 molecular weight (MW) was used. The column was filled with 20 ml protein containing solution. Vivaspins were placed in a swing bucket centrifuge and centrifuged at 3 000 g for 100 min to obtain a concentrated solution of 1 ml proteins.

3.2.6.12) Recombinant protein precipitation with polyethylene glycol from cell culture medium supernatant

Cell culture medium containing recombinant protein was collected into 50 ml centrifuge tubes and centrifuged at 500 g for 20 min to clarify the medium. 30% polyethylene glycol (PEG) 6000 in 0.4 M NaCl was prepared and added to the medium in a two to one ratio of cell culture medium to PEG. The mixture was

incubated overnight at 4°C with gently agitation. Protein precipitate was collected by 30 min centrifugation at 500 g. The pellet was dried and resuspended in 1 X PBS using 1/100 of the starting medium volume.

3.2.6.13) Indirect ELISA

A WNV indirect ELISA was performed to determine signal strength of recombinant antigens. This ELISA was setup as follows; 50% of the plate was coated with 100 µl of recombinant antigen to be tested (crudely extracted recombinant WNV antigen (transfected BHK 21 cells lysed and lysate used)) and the other 50% with negative antigen (untransfected BHK 21 cells). 2 X dilutions of antigens were used starting with a 1:50 dilution (1:50 to 1:1 600). Antigens were diluted with 1 X PBS. Plates were incubated for 1 hour at 37°C or overnight and then washed three times with wash buffer (1 X PBS and 0.01% Tween 20). Plates were blocked with 100 µl 10% blocking buffer per well (10% fat-free milk powder w/v in 1 X PBS) and incubated for 1 hour at 37°C. The next step was adding the positive and negative human WNV sera to half of the ELISA plate. Mouse sera were added to half the plate as a background control. Sera were diluted 1:400 with 2% dilution buffer (2% fat-free milk powder w/v in 1X PBS). After sera were added the plates were incubated for an hour at 37°C. The plate was washed as before and the conjugated antibody was added. Goat-anti-human horse radish peroxidase (HRP) IgG (Zymed, San Francisco, California, USA) was added in 1:1 000 or 1:5 000 dilutions (antibodies diluted with 2% dilution buffer) to wells containing human sera. Goat anti-mouse-HRP IgG (Zymed, San Francisco, California, USA) was added to wells containing mouse sera in the same dilutions as for the human antibodies. Secondary antibodies were incubated as before and after incubation plates were again washed. Substrate (ABTS) (ABTS Peroxidase substrate (I-Competent) KPL, Goithersburg, USA. Invitrogen) for the horseradish peroxidase enzyme was added. 1 X SDS solution stopped the color reaction.

3.2.6.14) Sandwich ELISA

For the sandwich ELISA the procedures were the same as for the indirect ELISA (Section 3.2.6.13) with the following adjustment. The plate was first coated overnight at 4°C with Mouse anti-WNV SPU 31017031 antibodies (NICD, SPU, South Africa). All the other steps were the same as for the indirect ELISA.

Different combinations of antigens and dilutions of ELISA reagents were used in order to test the antigen in an ELISA system as summarized in Table 3.3.1) Crude cell lysate from transfected BHK 21 cells was used; 2) purified recombinant WNV protein from BHK 21 cells was used and 3) concentrated purified recombinant WNV protein from BHK 21 cells and lastly 4) PEG precipitated recombinant WNV proteins from cell culture medium supernatant.

Table 3.3: Summary of different ELISA conditions performed in this study.

1.	**Indirect ELISA**	**Sandwich ELISA**
Coating Antibody	N/A	Mouse anti-WNV IgG
Dilutions	N/A	1:500, 1:1 000, 1:2 000
Coating Antigen	Crude extracted recombinant WNV protein from BHK 21 Cells. BHK 21 cells as negative antigen	Crude extracted recombinant WNV protein from BHK 21 Cells. BHK 21 cells as negative antigen
Dilutions	1:50 – 1:1600 (2X dilutions)	1:50, 1:100
Test serum	Human positive serum and Human negative serum for anti-WNV antibody Mouse positive and negative serum for anti-WNV antibody as background control	Human positive serum and Human negative serum for anti-WNV antibody
Dilutions	1:400	1:400
Anti species HRPO (Horse radish peroxidase conjugated antibody)	Goat anti-Human IgG-HRPO Goat anti-Mouse IgG-HRPO	Goat anti-Human IgG-HRPO
Dilutions	1:1 000 or 1:5 000	1:1 000 or 1:5 000
2.	**Indirect ELISA**	**Sandwich ELISA**
Coating Antibody	N/A	Mouse anti-WNV IgG
Dilutions	N/A	1:500, 1:1 000, 1:2 000
Coating Antigen	Purified recombinant WNV protein from BHK 21 Cells. BHK 21 cells as negative antigen	Purified recombinant WNV protein from BHK 21 Cells. BHK 21 cells as negative antigen
Dilutions	1:50 – 1:1600 (2X dilutions)	1:50 – 1:1600 (2X dilutions)
Test serum	Human positive serum and Human negative serum for anti-WNV antibody Mouse positive and negative serum for anti-WNV antibody as background control	Human positive serum and Human negative serum for anti-WNV antibody
Dilutions	1:400	1:400

Anti species HRPO	Goat anti-Human IgG-HRPO Goat anti-Mouse IgG-HRPO	Goat anti-Human IgG-HRPO
Dilutions	1:1 000 or 1:5 000	1:1 000 or 1:5 000
3.	**Indirect ELISA**	**Sandwich ELISA**
Coating Antibody	N/A	Mouse anti-WNV IgG
Dilutions	N/A	1:500, 1:1 000
Coating Antigen	Concentrated purified recombinant WNV protein from BHK 21 Cells. BHK 21 cells as negative antigen WNV positive control antigen and WNV negative control antigen	Concentrated purified recombinant WNV protein from BHK 21 Cells. BHK 21 cells as negative antigen WNV positive control antigen and WNV negative control antigen
Dilutions	Undiluted, 1:50, 1:100	Undiluted, 1:50, 1:100
Test serum	Human positive serum and Human negative serum for anti-WNV antibody Mouse positive and negative serum for anti-WNV antibody as background control	Human positive serum and Human negative serum for anti-WNV antibody
Dilutions	1:400	1:400
Anti species HRPO	Goat anti-Human IgG-HRPO Goat anti-Mouse IgG-HRPO	Goat anti-Human IgG-HRPO
Dilutions	1:3 000	1:3 000
4.	**Indirect ELISA**	**Sandwich ELISA**
Coating Antibody	N/A	Mouse anti-WNV IgG
Dilutions	N/A	1:500, 1:1 000
Coating Antigen	PEG purified proteins from cell culture medium supernatant. Complete cell culture medium as negative antigen WNV positive control antigen and WNV negative control antigen	PEG purified proteins from cell culture medium supernatant Complete cell culture medium as negative antigen WNV positive control antigen and WNV negative control antigen
Dilutions	1:50, 1:100, 1:200	1:50, 1:100, 1:200
Test serum	Human positive serum and Human negative serum for anti-WNV antibody Mouse positive and negative serum for anti-WNV antibody WNV as background control	Human positive serum and Human negative serum for anti-WNV antibody
Dilutions	1:400	1:400
Anti species HRPO	Goat anti-Human IgG-HRPO Goat anti-Mouse IgG-HRPO	Goat anti-Human IgG-HRPO
Dilutions	1:3 000	1:3 000

N/A = not applicable, Indirect ELISA's do not have a coating antibody.

3.3 RESULTS

3.3.1) Amplification of the prM and E genes

No primer sets were available for specific amplification of the West Nile virus lineage II prM and E gene section of the virus genome and therefore primers were designed. A full genome sequence of strain B956D117B3 was used (Accession number: M12294) as reference sequence. The open reading frame, orientations of the genes as well as the translation start and stop codons were verified and taken into account before DNAMAN was used to design the primers.

cDNA was produced from isolated viral RNA with random hexanucleotide primers. West Nile virus primers WNVprMF and WNVER were the reaction primers used in a PCR reaction to obtain the prM/E genes for cloning and subsequent expression studies. The PCR reaction was optimized for the West Nile virus prM and E primers. This was performed using different temperature ranges, $MgCl_2$ concentrations, dNTPs concentrations, and template concentrations. Amplification of a specific band sized (1990 nt) was only obtained with temperatures ranging between 45ºC and 55ºC.

After the PCR conditions for the WNV primer set were established, Taq polymerase were replaced with High Fidelity Master polymerase, which is a mixture of Taq and proofreading polymerase (Figure 3.4). This option was chosen in order to minimize mutations that may be incorporated during the PCR process and therefore ensuring that the PCR product is an exact copy of the prM-E genes. This enzyme also generates 3’ An overhangs (Taq polymerase terminal transferase activity that adds non-template single deoxyadenosine to the 3’ ends of PCR products) necessary for cloning in the TOPO vector system.

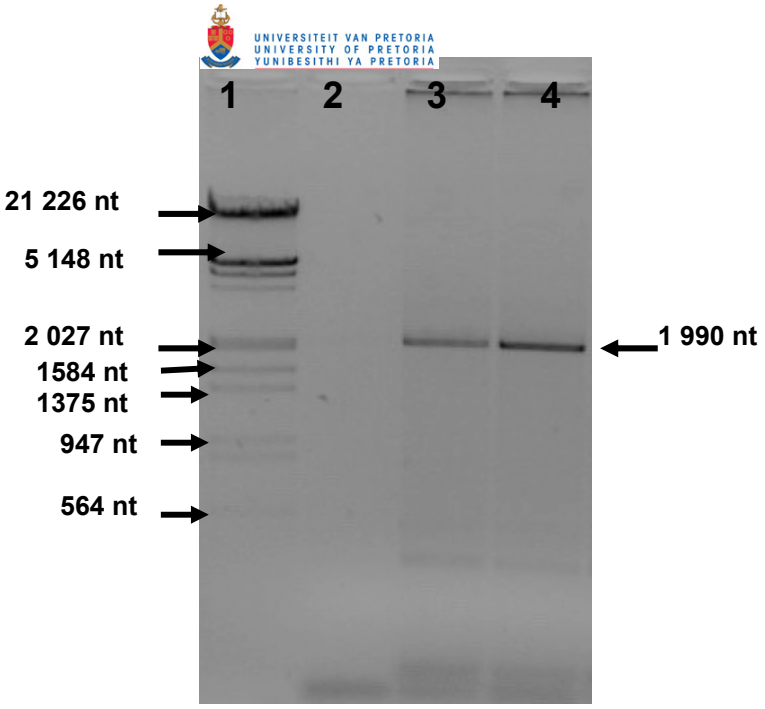


Figure 3.4: PCR results obtained when performing a PCR (primer pair WNV prMF and WNV ER) with Taq Polymerase and Expand High Fidelity PCR master. In lanes 1: DNA molecular marker (lambda EcoRI/Hind III marker from Promega); lane 2: negative control; lane 3: PCR performed with Taq Polymerase; lane 4: PCR performed with Expand High Fidelity PCR master.

The PCR products were purified using the Wizard SV gel and PCR clean-up system and concentration was determined visually by gel electrophoresis. The purified PCR amplicon was used in subsequent cloning steps.

3.3.2) Cloning of the WNV prM and E genes into the mammalian expression vector

Purified PCR product was cloned into the pcDNA3.1/V5-His©TOPO® vector with topoisomerase activity (Section 3.2.6.1). One shot TOP 10 *Escherichia coli* cells were then transformed with these recombinant TOPO vectors and incubated overnight in LB broth (Section 3.2.6.2). A transformation efficiency of 3.6 X 10^5 cells/ml was obtained. This was calculated by using the positive control plate that contained cells transformed with pUC 18 (1ng/ul). No colonies were found on the negative control plate, which only contained One shot TOP 10 competent *Escherichia coli* cells indicating that no contamination was present at this stage of experimentation. Overnight cultured *E. coli* transformants were then used in subsequent screening experiments. The first step was to isolate the plasmids from the bacterial cells. Recombinant plasmids were 7513 nucleotides in size and non-recombinant

plasmids5523 nucleotides in size. Gel electrophoresis analysis shows that all of the plasmid conformations, linear, supercoiled and nicked was present (Figure 3.5).

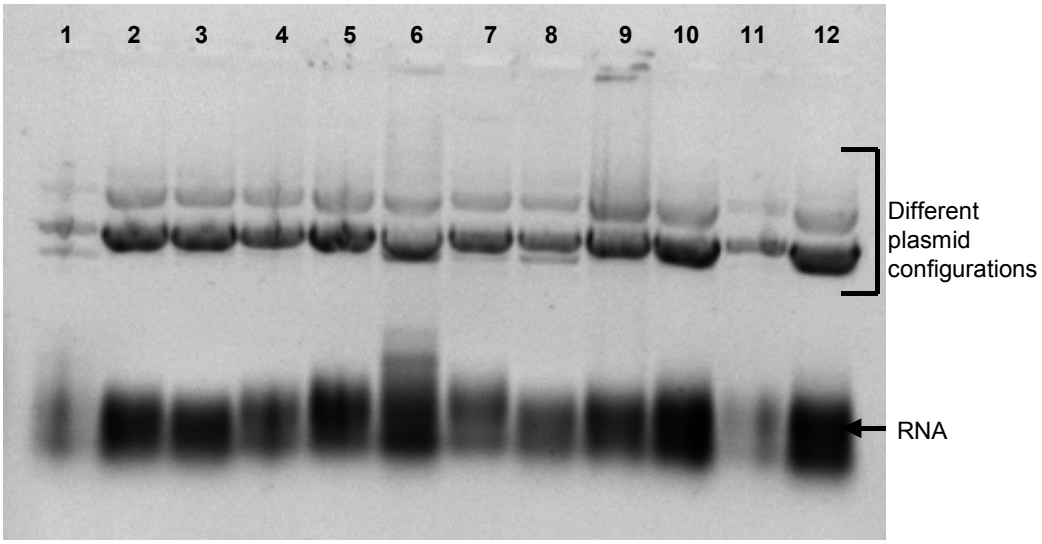


Figure 3.5: Agarose gel electrophoresis analysis of possible recombinant plasmids isolated after transformation. In lane one a control plasmid without insert is shown (non-recombinant plasmid) (5523 bp). In lanes 2-12 the different conformations of possible recombinant plasmids can be seen (7513 bp). RNA is present since Rnase digestion was not done.

Isolated plasmids were restricted with restriction enzyme Eco RV, to determine the insert orientation. Expected results were as follows: correct orientation; two bands of 5 862 nt and 1 651 nt respectively; inverted orientation; two bands of 375 nt and 7 138 nt. 18 plasmids were isolated of which only 10 were subjected to restriction digestion. Of these 10, five had inserts in the correct orientation (Figure 3.6). Plasmid stocks were prepared and purified from three of these plasmids which had the highest concentration.

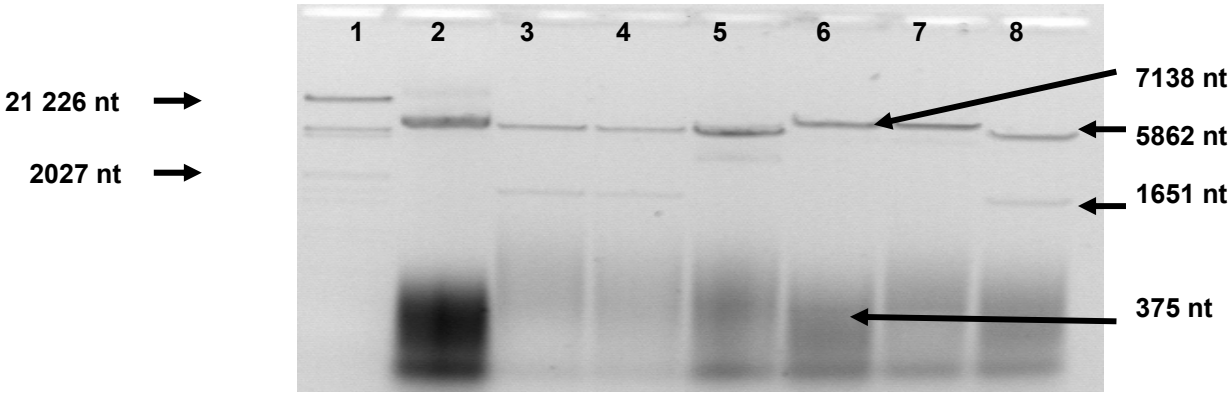


Figure 3.6: Analysis of EcoRV restriction digest of isolated plasmids. In lane 1: DNA molecular marker (lambda EcoRI/Hind III marker from Promega); lane 2: Uncut non-recombinant plasmid. Lanes 3, 4 and 8: Plasmids with correct orientated inserts; lanes 6 and 7: Plasmids with inverted inserts; Lane 5: Plasmid with no insert.

The next screening step for recombinants was DNA sequencing to confirm an open reading frame and codon fidelity of the cloned insert, ensuring correct expression. Primers used to sequence was T7 sequencing primer, BHG reverse sequencing primer, WNVprM, WNVE and WNVI primers (Table 3.1). Sequencing with the above mentioned primers allowed the complete sequencing of the prM and E genes. One of the selected plasmids contained the correct open reading frame and was used in subsequent expression studies (Figure 3.7).

SSDPLVQCGGIALGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVRAMDVGYLCEDTITYE
CPVLAAGNDPEDIDCWCTKSSVYVRYGRCTKTRHSRRSRRSLTVQTHGESTLANKKGAWLDSTK
ATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRVVFAILLLLVAPAYSFNCLGMSNRDFLEGV
SGATWVDLVLEGDSCVTIMSKDKPTIDVKMMNMEAANLADVRSYCYLASVSDLSTRAACPTMGEA
HNEKRADPAFVCKQGVVDRGWGNGCGLFGKGSIDTCAKFACTTKATGWIIQKENIKYEVAIFVHG
PTTVESHGNYSTQIGATQAGRFSITPSAPSYTLKLGEYGEVTVDCEPRSGIDTSAYYVMSVGAKSF
LVHREWFMDLNLPWSSAGSTTWRNRETLMEFEEPHATKQSVVALGSQEGALHQALAGAIPVEFS
SNTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFAGTPADTGHGTVVLELQYTGTDGPCKVPI
SSVASLNDLTPVGRLVTVNPFVSVATANSKVLIELEPPFGDSYIVVGRGEQQINHHWHKSGSSIGK
AFTTTLRGAQRLAALGDTAWDFGSVGGVFTSVGKAIHQVFGGAFRSLFGGMSWITQGLLGALLLW
MGINARDRSIAMTFLAVGGVLLFLSVNVKGNSADIQHSGGRSSLEGPRFEGKPIPNP

Figure 3.7: Amino acid sequence of recombinant plasmid. Blue writing indicates vector sequence and black the insert translated protein sequence.

3.3.3) Expression in a mammalian expression system

Different transfection methods were tested and at different concentration and ratios of transfection medium and DNA, as well as different mammalian cells. Of the methods tested it was found that transfecting BHK cells with SuperFect transfection medium at a ratio 1:3 and 10 µg of DNA for a T75 culture flask gave the best results. Optimization of transfection conditions were performed with the pcDNA3.1/V5-His-TOPO/lacZ vector. Transfection efficiency was calculated by staining transfected cells and counting the amount of blue cells since this vector contains a LacZ gene and blue/white selection can be used for analysis. Transfection efficiency was calculated to be 40% of a complete monolayer of BHK 21 cells (5 different focus points was counted, both blue and unstained cells, Average blue cells were divided by the average total cells and times by 100 to obtain the percentage) (Figure 3.8). This percentage is within the expected ratio as indicated by the manufacturer for transfections efficiency using BHK 21 cells and SuperFect transfection reagent.

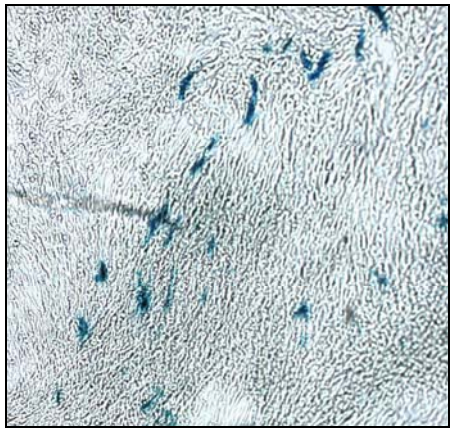
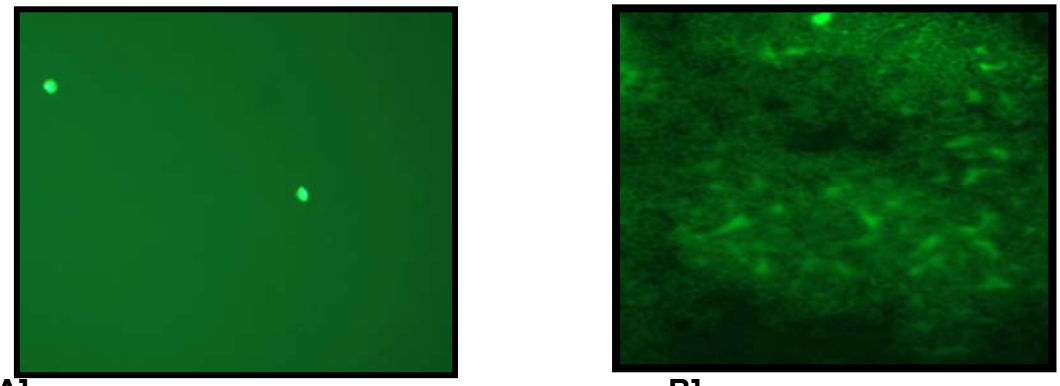


Figure 3.8: β-galactosidase analysis of control transfection of BHK 21 cells with SuperFect transfection medium and pcDNA3.1/V5-His-TOPO/lacZ.

Expression of the prM and E WNV protein was analyzed by either the indirect immunofluorescent antibody assay or a western blot analysis. Positive immunofluorescence was obtained when transfected cells were incubated with WNV polyclonal antibodies and stained with FITC conjugated goat-anti mouse antibodies and evans blue (Figure 3.9).



A] **B]**

Figure 3.9: Indirect immunofluoresence analysis of recombinant WNV proteins. In A) Untransfected BHK 21 cells and B) BHK 21 cells transfected with the recombinant WNV pcDNA3.1/V5-His-TOPO vector.

Western blot analysis was also performed on transiently transfected BHK cells with the TOPO expression vector containing the WNV gene insert. WNV polyclonal antibodies and secondary alkaline phosphates conjugated anti-mouse antibodies were used. The prM/E, prM and E units were visible with SDS-PAGE-gel electrophoresis (Figure 3.10) when these proteins were expressed in mammalian cells; the prM/E was expected to be cleaved because of the presence of host cell proteases which are responsible for cleavages of this junction in natural infections. These proteins will therefore be observed separately on a gel. Expected sizes was the two co-expressed proteins prM/E 83kDa, the envelope protein alone at 58 kDa

and the premembrane protein of 23-25 kDa as well as the mature membrane protein of 13 kDa (Figure 3.11). The SDS-PAGE analyses as well as the Western blot analysis both had a high degree of back ground. Further steps needs to be performed to optimize these systems to be able to indicate with higher degree of confidence that expression of protein did take place. Results obtain from these two gels are not conclusive enough.

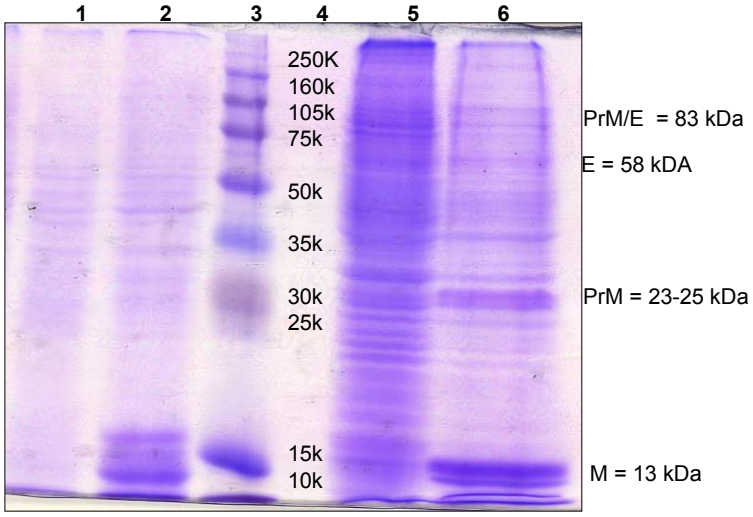


Figure 3.10: SDS-PAGE analysis of BHK 21 cell expression of WNV proteins. Lane one is the supernatant of the uninfected BHK 21 cells. Lane two is the pellet of the lysed uninfected BHK 21 cells. Lanes 5 and 6 is the supernatant and pellet portion of the transfected cells respectively. Lane 3 is a full range rainbow marker from AEC amersham, and lane 4 unloaded.

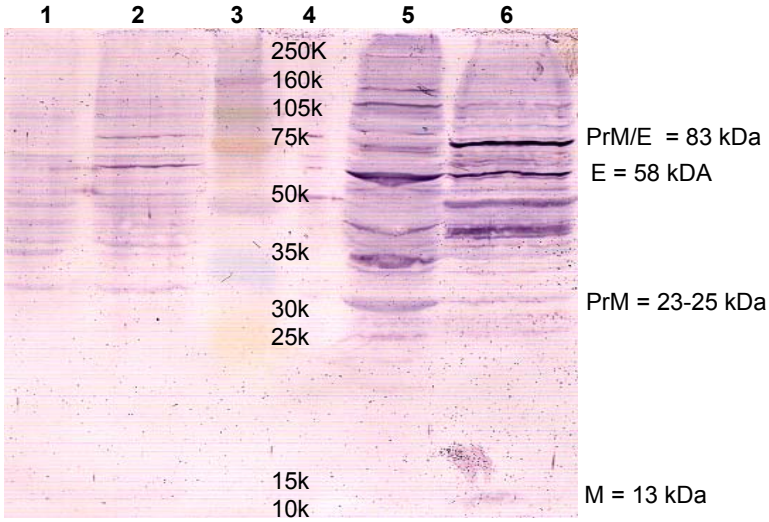


Figure 3.11: Western blot analysis of the BHK 21 cells transfected with a mammalian expression vector containing WNV. Lane one is the supernatant of the uninfected BHK 21 cells. Lane two is the pellet of the lysed uninfected BHK 21 cells. Lanes 5 and 6 is the supernatant and pellet portion of the transfected cells respectively. Lane 3 is a full range rainbow marker from AEC amersham, and lane 4 is a empty lane.

A stable cell line was then constructed using the transient transfected BHK 21 and the antibiotic Geneticin®. A PCR reaction was performed on these cells to confirm the presence of the gene insert in the stable cell line (Figure 3.12).

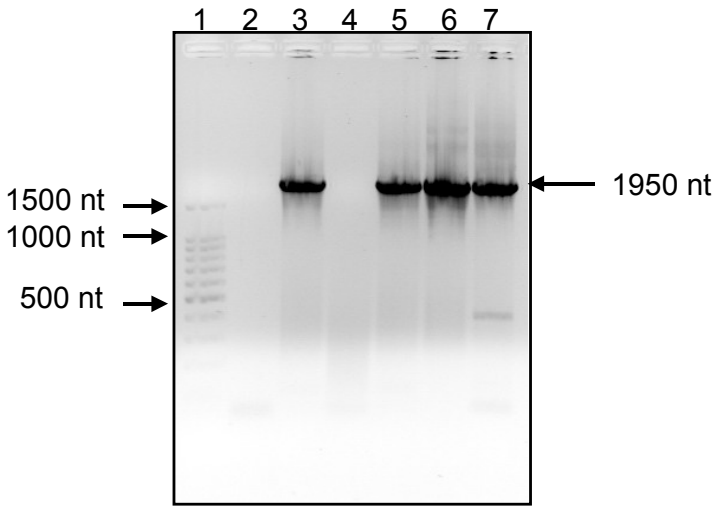


Figure 3.12: PCR results obtained when performing a PCR on isolated plasmids from transient and stable transformed BHK cells with WNV recombinant expression vectors. (primer pair WNV prMF and WNV ER). In Lane 1: DNA molecular marker (100 bp marker from Promega); Lane 2: negative control; Lane 3: PCR performed on plasmids from stable cell line; Lane 4: negative control; Lane 5: PCR performed on plasmids from transient transfected cell line; Lane 6: PCR performed on plasmids used to transfect BHK cells; Lane 7: PCR performed on PCR amplicon used to construct clone with.

Secreted proteins in the cell culture medium supernatant of the stable cell line was precipitated using PEG. These precipitated proteins was analyzed on a SDS-PAGE gel together with a positive (WNV cultured in VERO cells) and negative control (uninfected Vero cells) (Figure 3.13). No Western blot analysis is included in this thesis for the PAGE gel analysis of the stable cell-line due to problems with antibody detection in the detection steps of the Western Blot analysis. Antibodies where suspected to be inactive (Expiring date has expired) because the Western blot membrane showed inconclusive results.

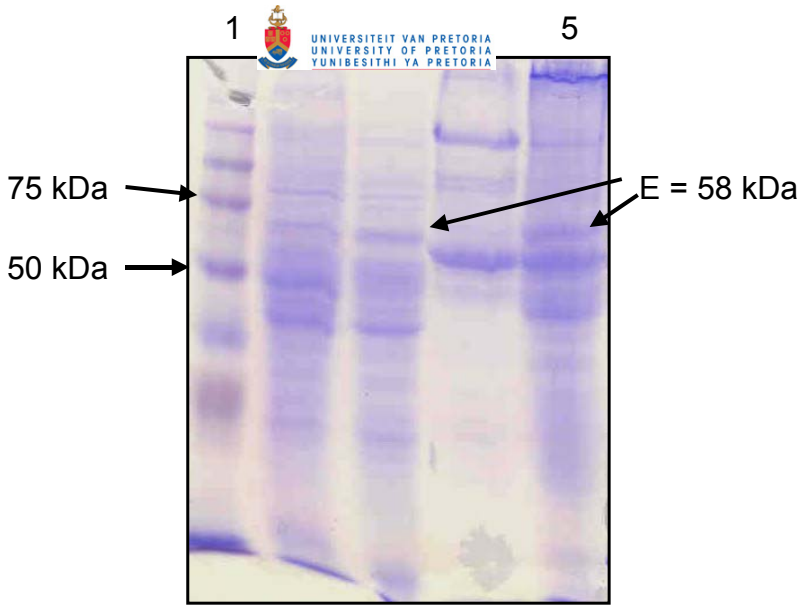


Figure 3.13: SDS-PAGE gel of WNV control antigen and PEG precipitated proteins from stably transformed cell line culture medium. Lane 1: Full range rainbow marker from AEC Amersham; Lane 2: Negative control; Lane 3: WNV antigen positive control; Lane 4 PEG precipitated cell culture medium supernatant; Lane 5: PEG precipitated recombinant protein from cell culture medium supernatant of stable cell line.

3.3.4) Use of recombinant protein antigens in ELISA

Both the indirect and sandwich ELISA format were used. Four repeats with a different antigen were carried out for each of the ELISA's. The first antigen to be used was unpurified lysate from lysed stably transformed BHK 21 cells (no controls included), then purified recombinant WNV proteins from stably transformed BHK 21 cells (no controls included), concentrated purified recombinant WNV proteins from stably transformed BHK 21 cells (included positive and negative controls to confirm to ELISA system works) and finally, PEG precipitated recombinant WNV proteins from cell culture medium supernatant (included positive and negative controls to confirm that the ELISA worked) (Section 3.2.7.14). None of the ELISA's produced high enough OD value differences between the non-recombinant and recombinant antigens to be viewed as clear positives results. To determine the value of the signal obtained in the ELISA, the negative antigen (non-recombinant) with positive serum value needs to be subtracted from the positive antigen (recombinant) with positive serum. This needs to be performed for the negative serum control and the relationship between the value from the positive serum needs to be ten times higher

than the value from the negative serum to be seen as a clear positive result (Crowther, 1995). With the indirect ELISA with crude cell extract the resulting relationship was only 1 to 8 and not the required 1 to 10 (Figure 3.14). In Figure 3.16 and 3.17 the difference between the positive control antigen (inactivated WNV) and negative control antigen (normal mouse ascitic fluid negative for WNV) was as expected for the controls but little to no difference was found between the test recombinant antigen and its test negative antigen. From the results it can be concluded that the ELISA system is successful based on the obtained OD values of the controls (Figure 3.16 and 3.17) but that the recombinant antigen was unable to perform sucessfully as a coating antigen in any of the ELISA tests (Figure 3.14-17).ELISA did not work because of insufficient or no recombinant antigen was present.

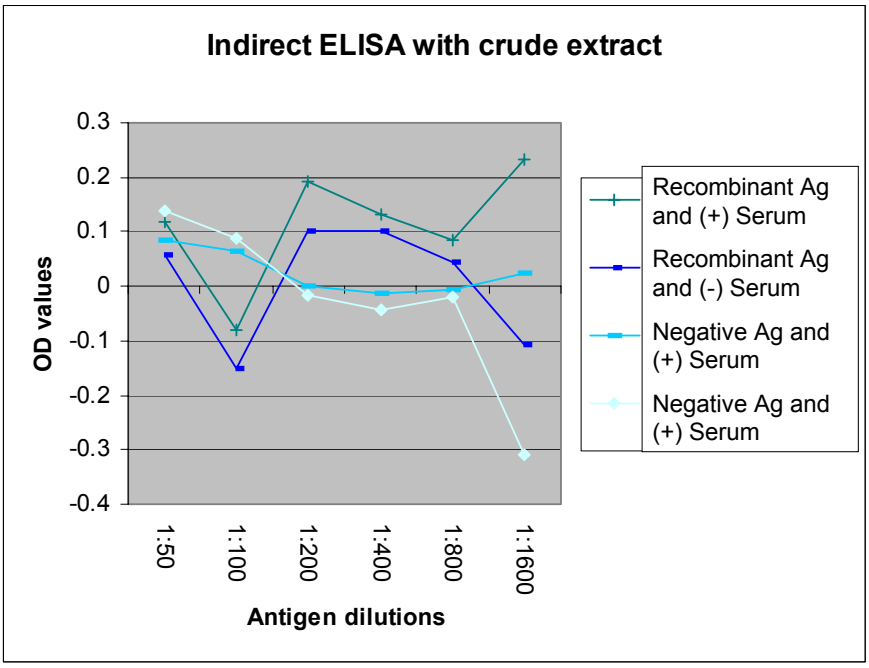


Figure 3.14: Graph constructed with ELISA data obtained when performing an indirect ELISA with crude extract WNV recombinant protein and BHK 21 cells as negative antigen. Human serum positive and negative for WNV was used as test sample and goat anti human-HRPO antibodies as detection system. No controls were included in this ELISA.

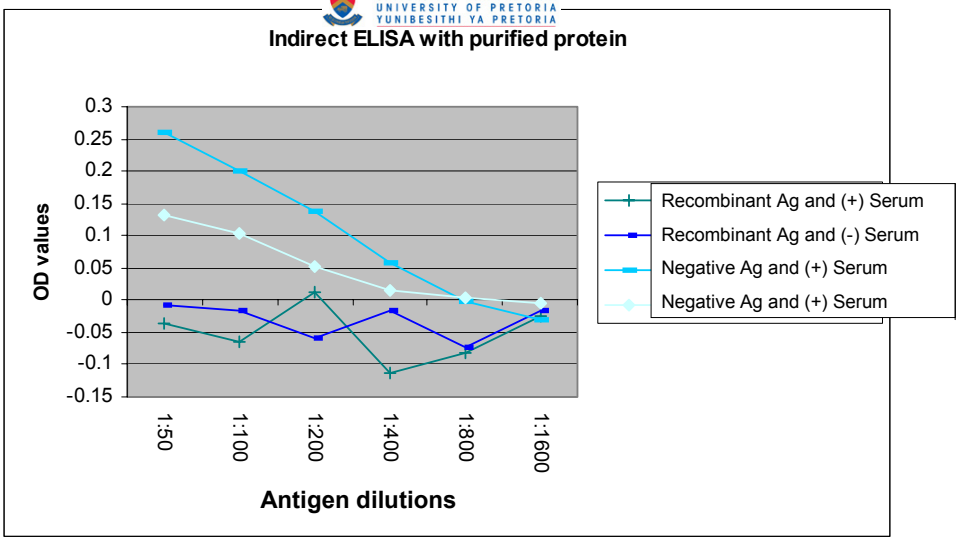
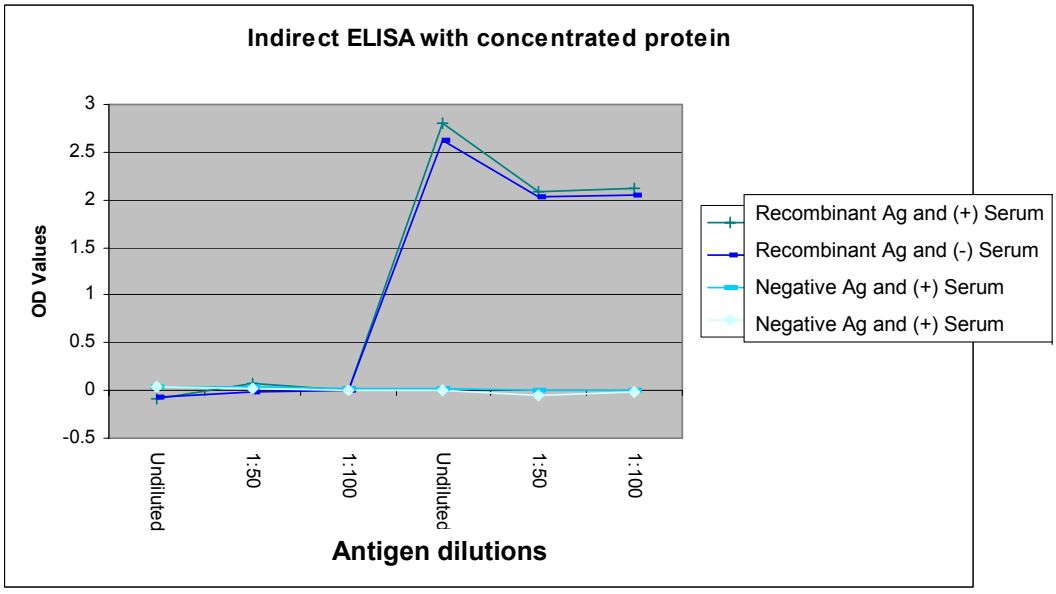
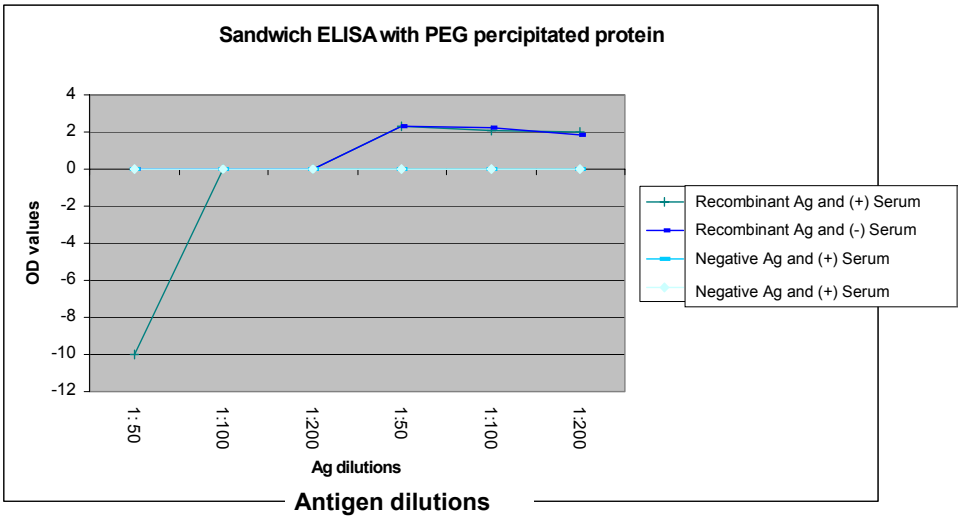


Figure 3.15: Graph constructed with ELISA data obtained when performing an indirect ELISA with purified WNV recombinant protein and BHK 21 cells as negative antigen. Human serum positive and negative for WNV was used as test sample and goat anti human-HRPO antibodies as detection system. No controls were included in this ELISA.



The first three dilution values are for the test recombinant antigen and the last three dilution values for the control antigens incorporated into this ELISA

Figure 3.16: Graph constructed with ELISA data obtained when performing an indirect ELISA with concentrated purified WNV recombinant protein and BHK 21 cells as negative antigen. Human serum positive and negative for WNV was used as test sample and goat anti human-HRPO antibodies as detection system. The first three dilution values on the X-axis are for WNV recombinant antigens which was tested and the last three dilution values for the WNV positive (inactivated WNV) and negative controls (normal mouse ascitic fluid negative for WNV) included in this ELISA



The first three dilution values are for the test recombinant antigen and the last three dilution values for the control antigens incorporated into this ELISA

Figure 3.17: Graph constructed with ELISA data obtained when performing an sandwich ELISA with PEG precipitated WNV recombinant protein and PEG precipitated cell culture medium supernatant of untransfected cells PEG precipitated proteins as negative antigen. Human serum positive and negative for WNV was used as test sample and goat anti human-HRPO antibodies as detection system. The first three dilution values on the X-axis are for WNV recombinant antigens to be tested and the last three dilution values for the WNV positive (inactivated WNV) and negative controls (normal mouse aseptic fluid negative for WNV) included in this ELISA.

3.4 DISCUSSION

This part of the study was dedicated to the production of West Nile virus recombinant antigen for it's possible use as a diagnostic reagent in a diagnostic assay such as an ELISA. A mammalian expression system was chosen as the system to produce the recombinant WNV antigen. The mammalian expression vector contained both the WNV premembrane and envelope genes as a unit (Allison *et al*., 1995; Yoshii *et al*., 2003; Hunt *et al*., 2001, Davis *et al*., 2001). ELISA's using recombinant antigens are not novel and have already been developed by other researchers (Beasley, 2005, Prince and Hogrefe 2005). The difference between previous studies and this study is that an attempt was made to produce recombinant antigen from WNV lineage II strains isolated in South Africa. Expressed lineage II recombinant antigen will be evaluated for its use as a reagent in a diagnostic ELISA. If results should be indicative of these antigens to be successful in an ELISA system, these antigens can then be incorporated into a diagnostic ELISA which can be used for routine diagnostic procedures or for surveillance studies. The main consideration for this part of the study was to provide a safer antigen compared to inactivated virus for use in a diagnostic system (Beasley, 2005). Another consideration was that lineage II recombinant antigen would possibly provide higher specificity for lineage II WNV infections. Commercial ELISA's are not only expensive, but has only been developed for lineage I WNV strains.

To reach our objective of this part of the study the prM and E genes were cloned and their integrity was validated. Recombinant antigens were produced in a mammalian expression system. Both transient and stable cell lines were constructed. The presence of the gene insert was confirmed in the transiently transfected as well as the stable cell line by performing PCR on the transfected cells. SDS-PAGE gel analysis was used to detect expression for both systems. Immunoflourescence analysis and western blot analysis was also performed for the transient cell line. However both the PAGE gel and Western blot membranes had high levels of background. These systems need to be optimized to be able to indicate with a higher degree of confidence the levels of expression. Other method could also have been used to confirm expression of the protein e.g. immunoprecipitation or radiolabelling of the proteins. Shorter storage times for proteins before using them in blot analysis or addition of protease inhibitors to prevent protein degradation could have also assist

in improving the Western Blot results. Another consideration for future experiments is to measure the yield of the protein before proceeding to following experiment.

In testing the ability of the expressed recombinant proteins to be used as diagnostic reagents, four different preparation methods was used to obtain the recombinant WNV antigen from the BHK cells for both indirect and sandwich ELISA's. The four different prepared antigen were as follows: 1) Unpurified lysed transfected BHK cells; 2) purified recombinant WNV antigens from lysed transfected BHK cells 3) then these purified recombinant WNV proteins where concentrated 4) next the cell culture medium was used and secreted recombinant WNV protein was precipitated with PEG and precipitated recombinant WNV proteins used as antigens for ELISA. None of the ELISA systems showed any positive signal. The inclusion of positive controls (whole WNV virus) yielded strong signals in the ELISA (as opposed to the negative controls) and were taken as indicative of the efficiency of the assay when used with WNV antigen. It could thus be concluded that the level or concentration of the recombinant antigen produced in the mammalian expression systems were insufficient for this type of assay.

Optimization needs to be done on the expression of the antigen in order to obtain higher levels of antigen. Other options to be considered when optimizing this expression system can be to use other mammalian cells e.g. COS-1 cells as used in previous studies using mammalian systems (Allison *et al*., 1995; Hunt *et al*., 2001, Davis *et al*., 2001). Other transfection methods can also be considered, possibilities are electroporation as used by Davis *et al*., (2001) in their mammalian expression systems or other transfection medium e.g. transIT-LT1 (Pan Vera, Madison, WI) used by Yoshii *et al*., 2003 in their expression systems. The transfections method that is the least lethal to the cell line being used needs to be found e.g. electroporation would not have any chemical toxic effects on the cells.

CHAPTER FOUR:

CONCLUDING REMARKS

West Nile virus is an avian pathogen with human and equines being incidental hosts that came under the spotlight in 1999 when it first was introduced into Northern America (Petersen and Roehrig, 2001). The 1999 introduction of WNV into America occurred with increased case fatality rates, human and bird deaths. The causative agent of this outbreak was identified as being WNV lineage I strain (Petersen and Roehrig, 2001). Previously it was thought that lineage I was more pathogenic than lineage II WNV strains because of the increased virulence of the lineage I strains in North America and the low numbers of WNV infections reported from South Africa (Beasley *et al*., 2002). However, the National Institute for Communicable Diseases (NICD), Special Pathogens Unit (SPU) isolated pathogenic highly neuroinvasive lineage II strains of WNV in South Africa in recent years, bringing the importance of the disease in South Africa under our attention (Burt *et al*., 2002).

At present we have very little knowledge about the epidemiology of WNV in South Africa. The epizootic potential of lineage II WNV as well as its disease implications in Africa is not fully understood. Currently only a few full genome sequences for lineage II WNV strains are available. The same scenario holds true for the pathogenic determinants of these strains. It thus became important to obtain more information about the lineage II South African strains in order to address pathogenesis and epidemiological questions that the disease may hold for South Africa. To be able to address these problems and obtain more information on WNV in South Africa, sequence information on both pathogenic and non-pathogenic strains were obtained. This sequence information will enable us to do more comprehensive pathogenic studies and assist in combating the disease through development of vaccines and surveillance tools.

In this study the full genome sequences of pathogenic and less pathogenic WNV lineage II strains were determined. Globally, this constitutes the first genome sequences of lineage II strains that exhibited neuroinvasive characteristics. Sequences were compared to known pathogenic and non-pathogenic lineage I and II WNV strains to see if any potential molecular pathogenic markers can be identified. Out of this study it could be shown that phenotype (pathogenicity) of the WNV strain in question is not related to the lineage the WNV strain clusters into phylogenetically but rather by the virus genotype as previously described by Venter *et al*., 2005 and

Burt *et al*., 2002. Genome differences are associated with geographic distribution rather than temporal separation and differences in pathogenic phenotype were related to differences in non-structural proteins e.g. the polymerase gene NS5 than structural proteins exposed to the host immune system.

To address the question of the lack of surveillance in South Africa it was proposed to construct a safer, cheaper and readier available diagnostic ELISA that utilises lineage II strains isolated in South Africa. A diagnostic ELISA produced with recombinant South African WNV strains can potentially be used in epidemiological studies. As a preliminary part of this development of a diagnostic test, the second part of the study attempted to produce a recombinant antigen expressed in a mammalian system. The expressed recombinant WNV strain H442 antigen however could not be successfully used in an indirect or sandwich ELISA test. This is probably due to low levels of expression of the recombinant protein. Levels of expression needs to be determined and Western blot analyses needs to be improved For a new ELISA using a recombinant antigen instead of inactivated lineage II WNV, the expression study needs to be rethought. One of two options needs to be looked at, either one need to test different mammalian cells and transfection methods to optimize expression levels or a new expression system need to be address, which exhibits higher levels of expression.

Two possibilities for other systems are the Baculovirus expression system that has been used with success in the past, or the *Drosophila* system. The Drosophila system was used with great success by Ledizet *et al*., (2005) who expressed a soluble truncated form of the WNV envelope protein in this system. Antibodies from naturally infected horses recognized this recombinant antigen in an ELISA test. This system produced purified recombinant protein yields ranging from 2-20 mg/l of serum-free culture medium. In another study, utilizing the *Drosophila* system to express WNV recombinant proteins, these antigens were used in an ELISA and compared with recombinant antigens expressed in a mammalian expression system using COS-1 cells (Muerhoff *et al*., 2004). A 90 % correlation between the two systems was reported. The advantages that the Baculovirus and *Drosophila* systems have over mammalian expression systems are that the cells grow at 28ºC and do not need CO_2. Mammalian cells on the other hand need incubation at 37ºC and CO_2 supplementation. Furthermore, mammalian cells are adherent and grow in flasks

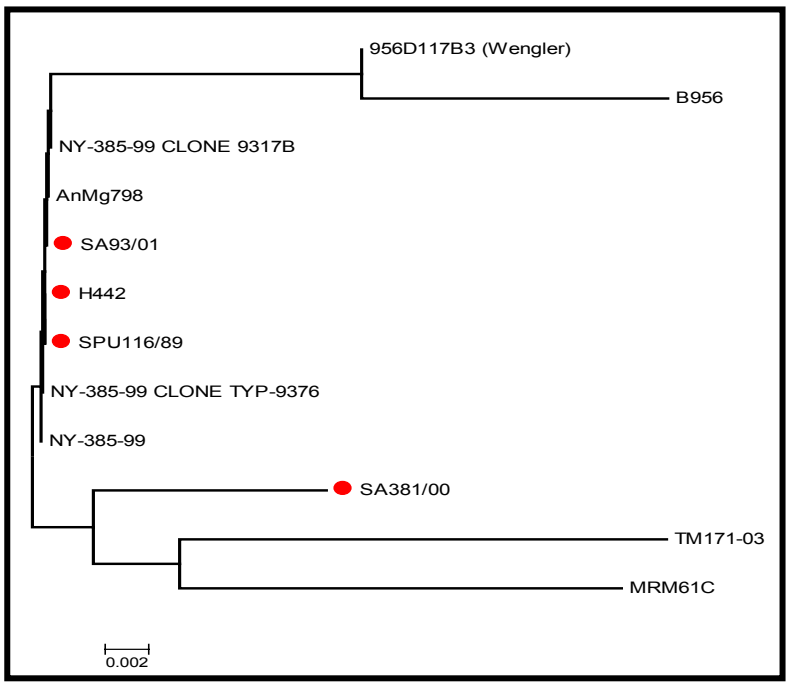
where the insect cells used in the baculovirus and *Drosophila* systems are semi adherent and can grow in flasks or spinner cultures. As a result, these two systems are also known for high expression levels. Even though these two systems have advantages over mammalian systems, the mammalian system was used because of the authenticity of the recombinant proteins expressed with this system (Davis *et al*., 2001; Hunt *et al*., 2001).

APPENDIX

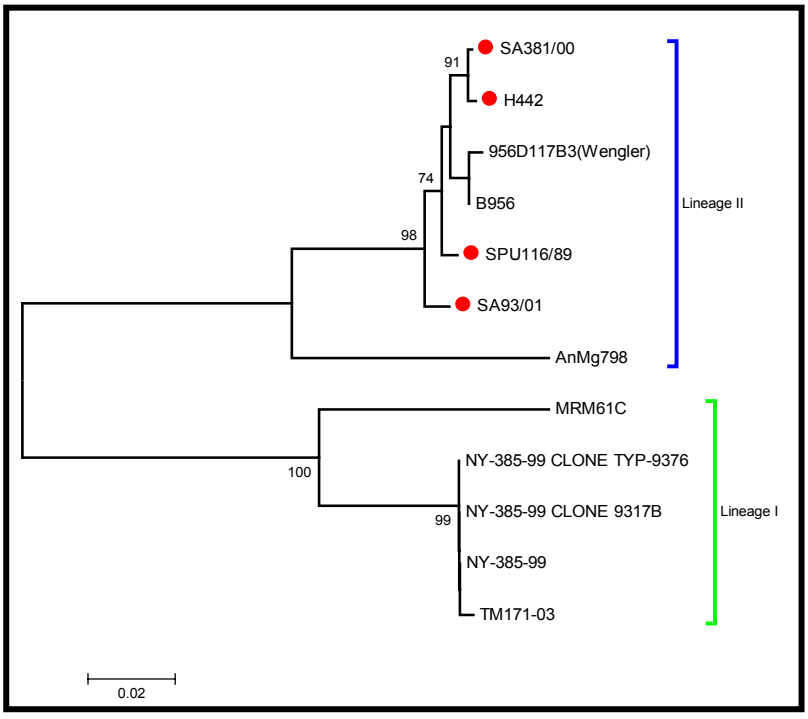
A1) GenBank accession numbers of strains used for analysis.

Strain/Isolate name	Accession number
NY-385-99 clone 9317B	DQO66423
NY-385-99 clone TVP-9376	AY848697
NY-385-99	DQ211652
NY-382-99 FLAM	AF196835
IS-98 STD	AF481864
Mexico-TM171-03	AY660002
TX 2002 2	DQ164205
Goose-Hungary/03	DQ118127
Eg 101	AF260968
Ro97-50	AF260969
Morocco 96-111	AY701412
Italy 1998-Eq	AF404757
KN 3829	AY262283
LEIV-Vlg00-27924	AY278442
PaH001	AY268133
Ast02-696	DQ411035
MRM61C (KUN)	D00246
IND 804994H	DQ256376
AnMg798	DQ176636
SA381/00	EF429199
H442	EF429200
Goshawk-Hungary/04	DQ116961
SPU116/89	EF429197
SA93/01	EF429198
B956D117B3	M12294
ArD76104	DQ318019
Rabensburg isolate 97-103	AY765264
LEIV-krnd 88-190	AY277251
M18370 JEV	M18370
B956	AY532665

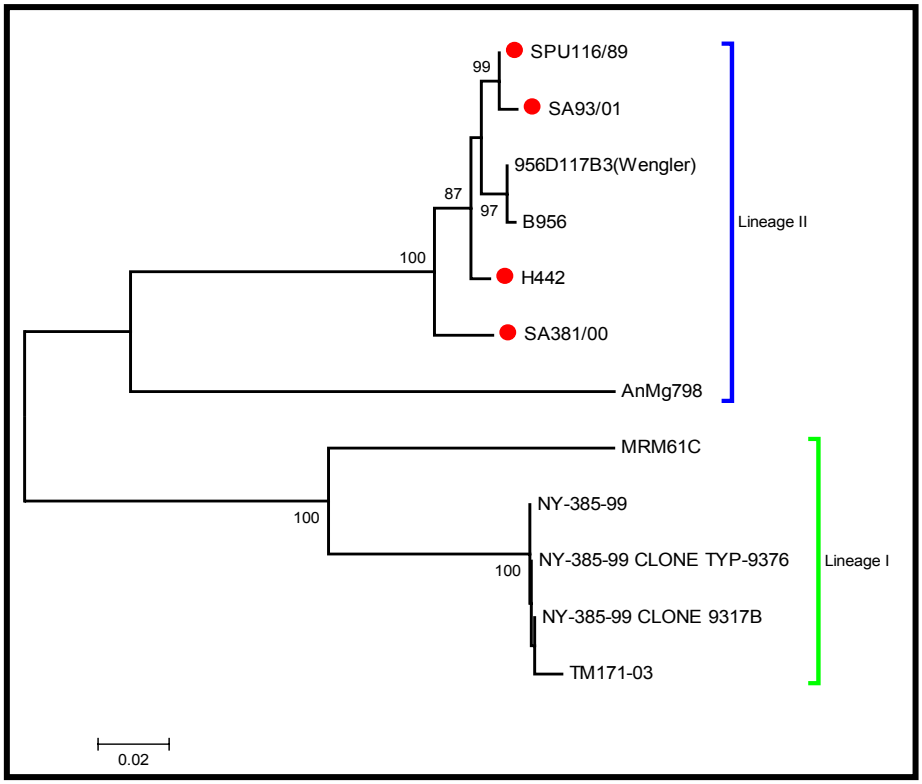
B1) Neighbour joining trees of individual gene nucleotide sequences of selected lineage I and II WNV strains.



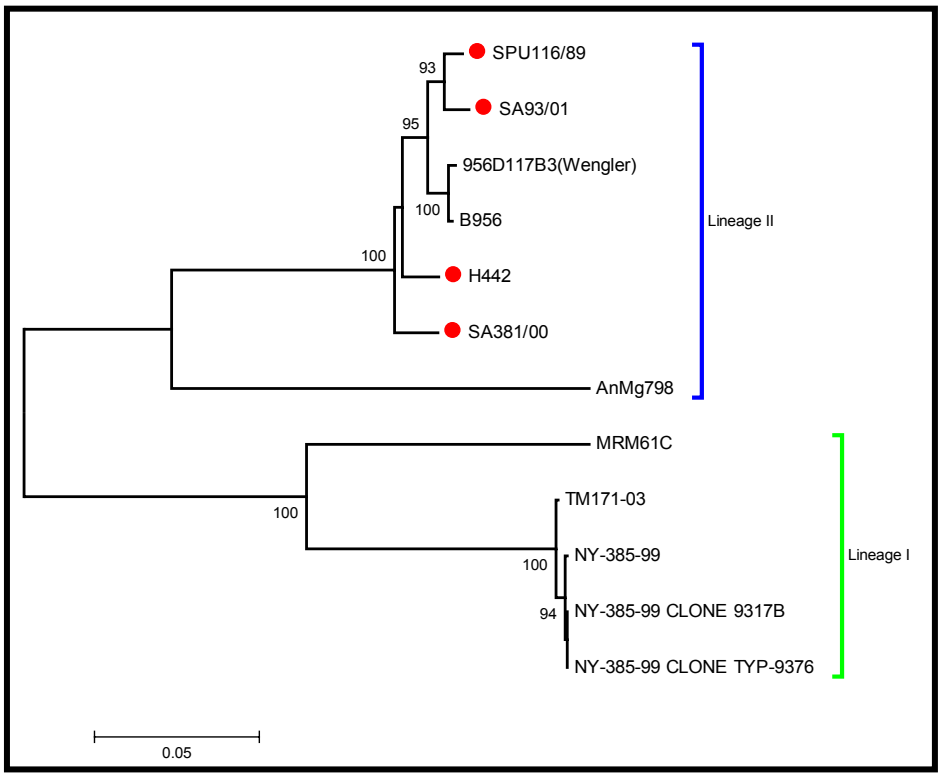
B1.1: Neighbour-joining tree for selected WNV 5'NCR nucleotide sequence**.** The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (Appendix A1 contain GenBank accession numbers)



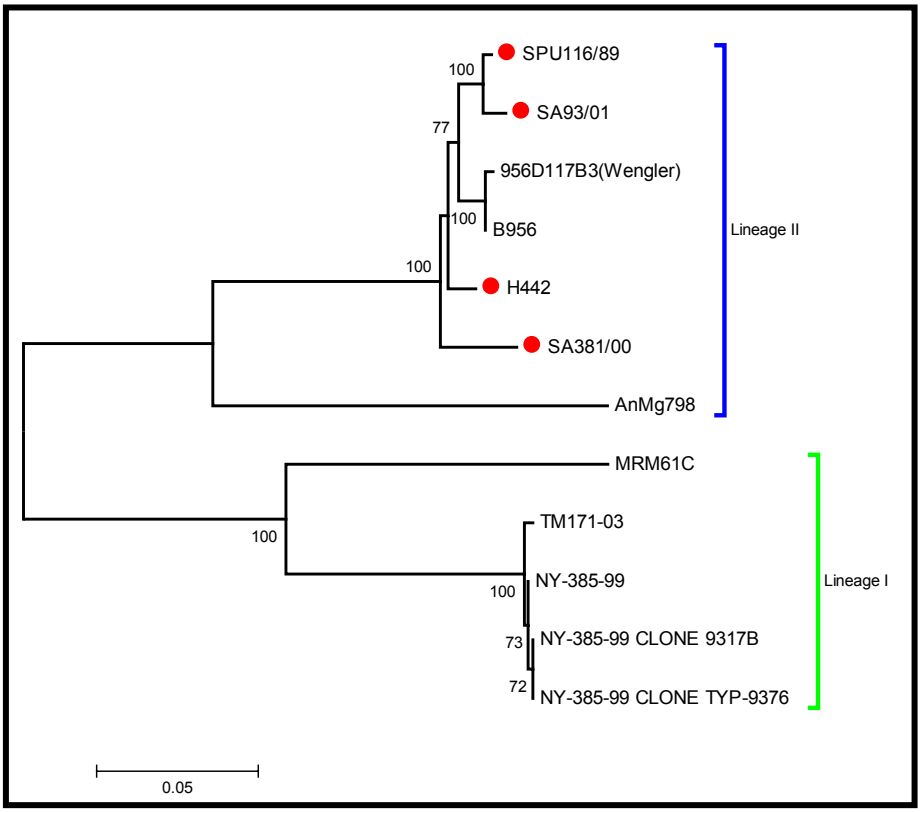
B1.2: Neighbour-joining tree for selected WNV capsid genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



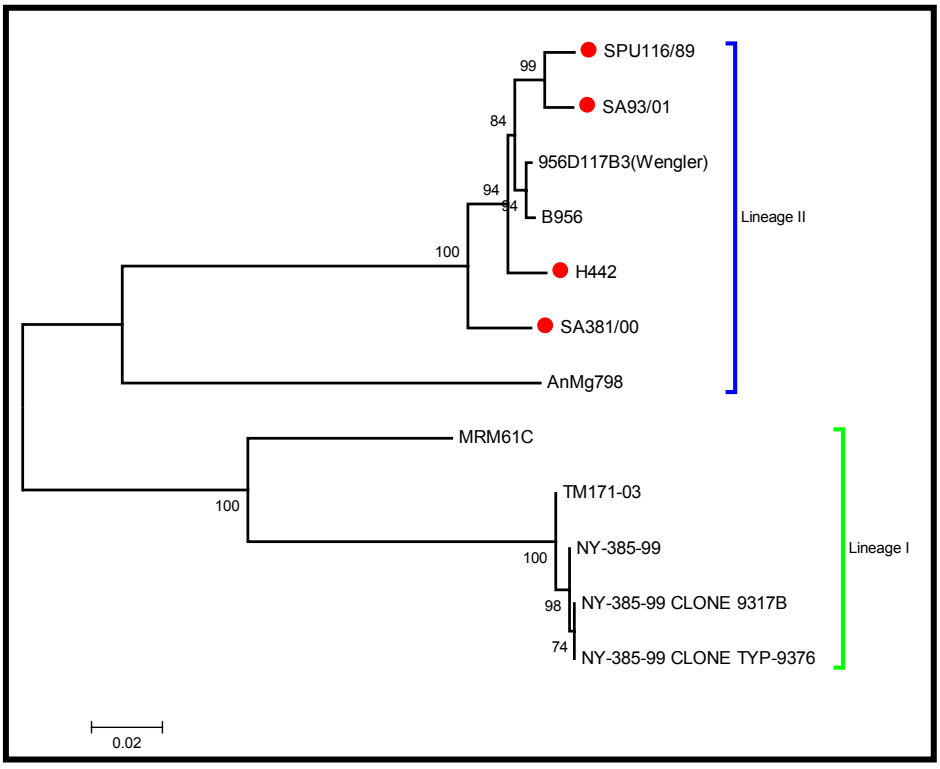
B1.3: Neighbour-joining tree for selected WNV pre-membrane genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. Strains sequences in this study are indicated by red dots. A bootstrap confidence level of 1000 replicates was used. (See appendix A1 for GenBank accession numbers)



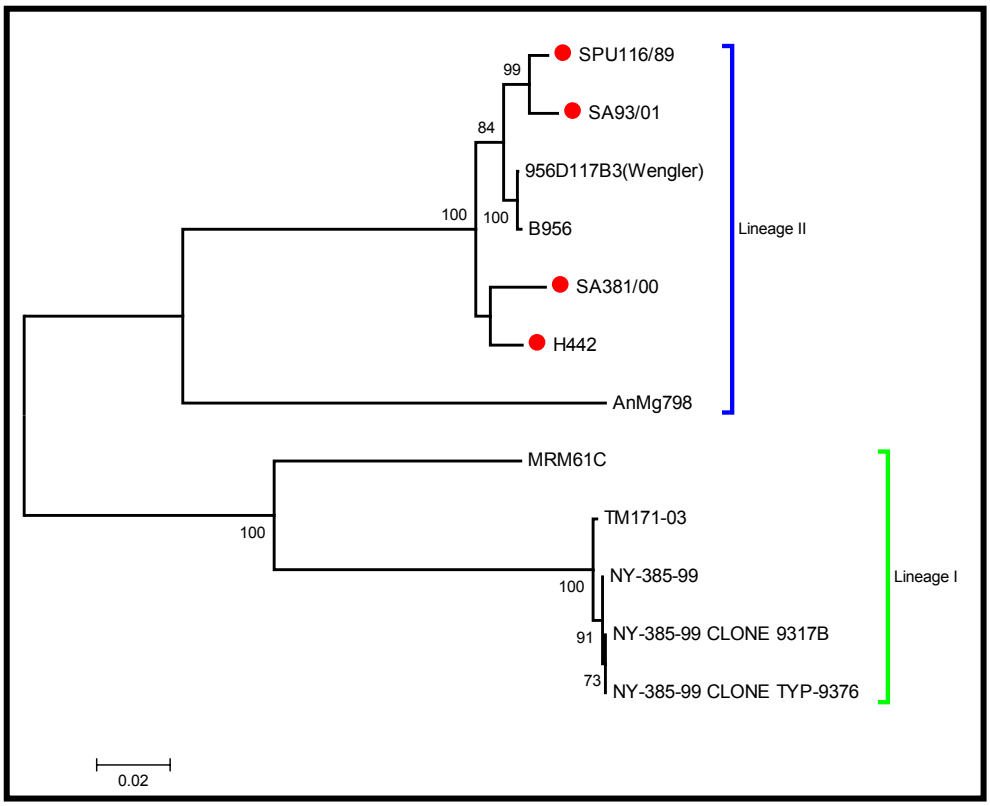
B1.4: Neighbour-joining tree for selected WNV envelope genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



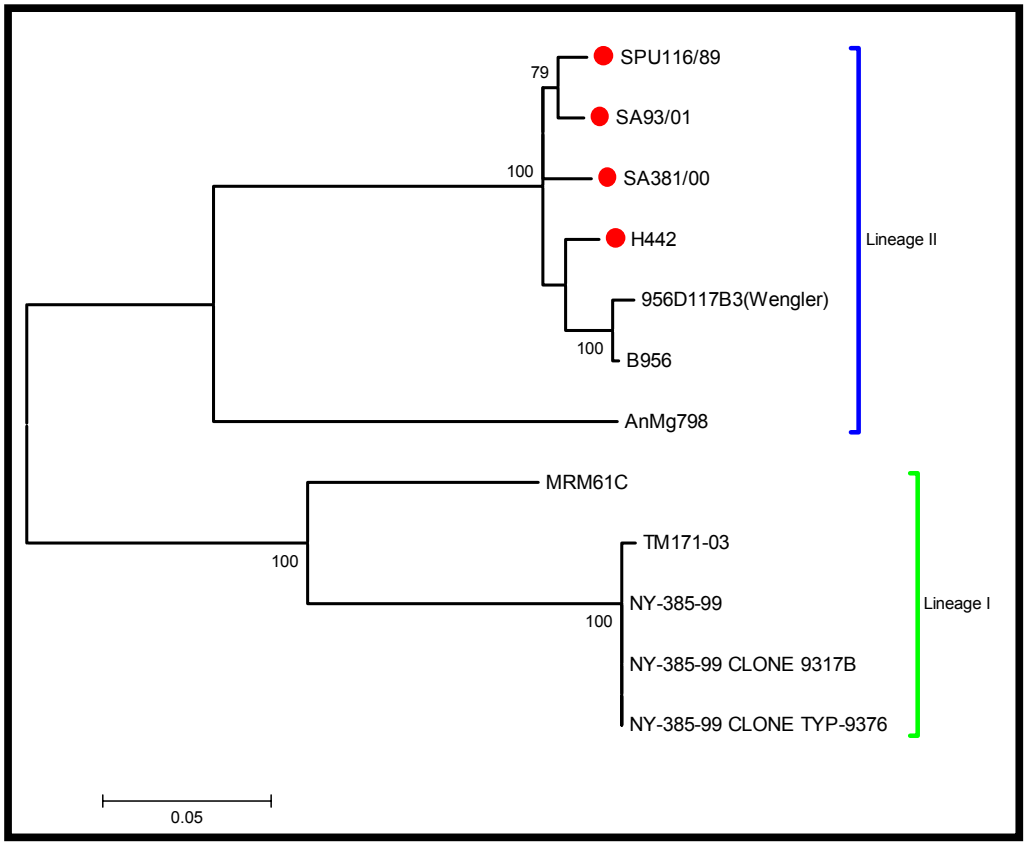
B1.5: Neighbour-joining tree for selected WNV NS1 genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



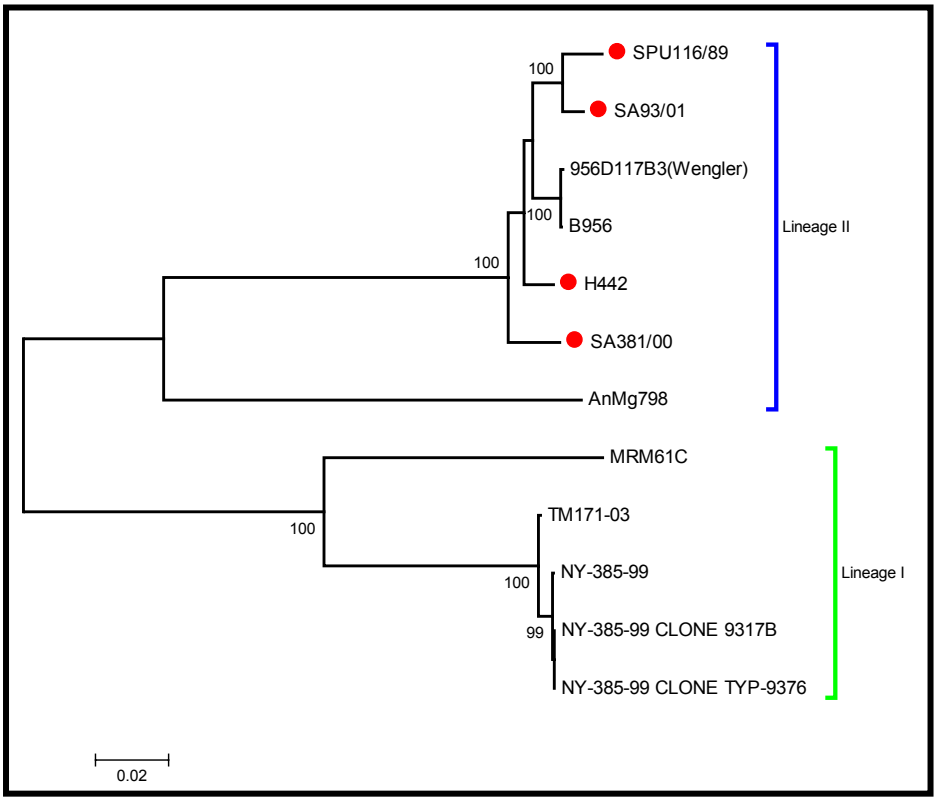
B1.6: Neighbour-joining tree for selected WNV NS2A/B genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



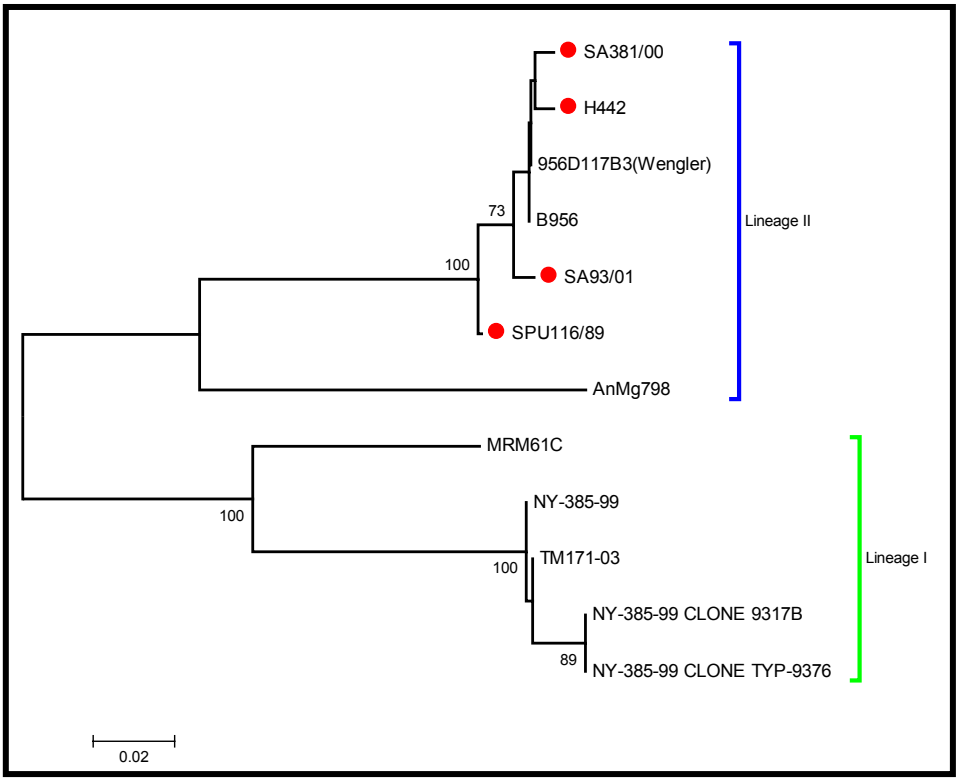
B1.7: Neighbour-joining tree for selected WNV NS3 genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



B1.8: Neighbour-joining tree for selected WNV NS4A/B genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



B1.9: Neighbour-joining tree for selected WNV NS5 genes**.** The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)



B1.10: Neighbour-joining tree for selected WNV 3'NCR genes. The tree was constructed with the program MEGA version 3.1 using the neighbor-joining method with Kimura 2 distance-parameter. A bootstrap confidence level of 1000 replicates was used. Strains sequenced in this study are indicated by a red dot. (See appendix A1 for GenBank accession numbers)

References:

1. Allison Sl, Schalich J, Stiasny K, Mandl CW, Heinz FX. Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. J Virol. 2001; 75(9): 4268-4275.
2. Allison SL, Stadler K, Mandl CW, Kunz C, Heinz FX. Synthesis and sectretion of recombinant Tick-borne encephalitis virus protein E in soluble and particulate form. J Virol. 1995; 69(9): 5816-5820.
3. Allison SL, Stiasny K, Stadler K, Mandl CW, Heinz FX. Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E. J Virol. 1999; 73(7): 5605-5612.
4. Amberg SM, Nestorowicz A, McCourt DW, Rice CM. NS2B-3 proteinase-mediated processing in the yellow fever virus structural region: In vitro and in vivo studies. J Virol. 1994; 68(6): 3794-3802.
5. Bakonyi T, Hubalek Z, Rudolf I, Nowotny N. Novel Flavivirus or new lineage of West Nile virus, central Europe. Emerg Infect Dis. 2005; 11(5): 225-231.
6. Bakonyi T, Ivanics E, Erdelyi K, Ursu K, Ferenczi E, Weissenbock H, Nowotny N. Lineage 1 and 2 strains of encephalitic West Nile virus, central Europe. Emerg Infect Dis. 2006; 12(4): 618-23.
7. Bartholomeusz A, Thompson P. *Flaviviridae* polymerase and RNA replication. J Viral Hepat. 1999; 6(4); 261-270.
8. Beasley DWC, Davis CT, Estrada-Franco J, Navarro-Lopez R, Campomanes-Cortes A, Tesh RB, Weaver SC, and Barrett AD. Genome sequence and attenuating mutations in West Nile virus isolate from Mexico. Emerg Infect Dis. 2004; 10(12): 2221-2224.
9. Beasley DWC, Li L, Suderman MT, Barrett AD. Mouse neuroinvasive phenotype of West Nile virus strains varies depending upon virus genotype. Virology. 2002; 296(1)**:** 17-23.
10. Beasley DWC, Whiteman MC, Zhang S, Huang CY, Schneider BS, Smith DR, Gromowski GD, Higgs S, Kinney RM, Barrett AD. Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. J Virol. 2005; 79(13)**:** 8339-8347.
11. Beasley DWC. Recent advances in the molecular biology of West Nile virus. Curr Mol Med. 2005; 5: 835-850.
12. Bhardwaj S, Holbrook M, Shope RE, Barrett AD, Watowich SJ. Biophysical characterization and vector-specific antagonist activity of domain III of the tick-borne flavivirus envelope protein. J Virol. 2001; 75(8): 4002-4007.
13. Bielefeldt-Ohmann H, Beasley DWC, Fitzpatrick DR, Aaskov JG. Analysis of a recombinant Dengue-2 virus-Dengue-3 virus hybrid envelope protein expressed in a secretory baculovirus system. J Gen Virol. 1997; 78:2723-2733.
14. Bondre VP, Jadi RS, Mishra AC, Yergolkar PN, Arankalle VA. West Nile virus isolates from India: evidence for a distinct lineage. J Gen Virol. 2007; 88: 875-884.
15. Borisevich V, Seregin A. Nistler R, Mutabazi D, Yamshchikov V. Biological properties of chimeric West Nile viruses. Virology. 2006; 349: 371-381.
16. Bork P, Holm L, Sander C. The immunoglobulin fold: Structural classification, sequence patterns and common core. J Mol Biol. 1994; 242(4): 309-320.
17. Briese TJX, Huang C, Grady LJ, Lipkin WI. Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. The Lancet. 1999; 354(9186): 1261-1262.
18. Brinton MA and Dispoto JH. Sequence and secondary structure analysis of the 5`-terminal region of flavivirus genome RNA. Virology. 1988; 162: 290-299.
19. Brinton MA, Fernandez AV, Dispoto JH. The 3`-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology 1986; 153:113-121.
20. Briton MA. The molecular biology of West Nile virus: A new invader in the Western Hemisphere. Annu rev Microbiol. 2002; 56: 371-402.
21. Burt FJ, Grobbelaar AA, Leman PA, Anthony FS, Gibson GV, Swanepoel R. Phylogenetic relationships of southern African West Nile virus isolates. Emerg Infect Dis. 2002; 8(8): 820-826.
22. Calisher CH, Karabatsos N, Dalrymple JM, Shope RE, Porterfield JS, Westaway EG, Brandt WE. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polycolnal antisera. J Gen Virol. 1989; 70: 37-43.
23. Campbell GL, Marfin AA, Lanciotti RS, Gubler DJ . West Nile Virus. Lancet Infect Dis 2002; 2(9)**:** 519-529.
24. Castle E, Nowak T, Leidner U, Wengler G, Wengler G. Sequence analysis of the viral core protein and the membrane-associated proteins V1 and NV2 of the flavivirus West Nile virus and the genome sequence of these proteins. Virology. 1985; 145: 227-236.
25. CDC. 2007. West Nile Virus. 2006 Human cases http://www.cdc.gov/ncidod/dvbid/westnile/surv&controlCaseCount06_detailed.htm
26. Centre of Disease Control (CDC). Laboratory-aquired WNV infections- United states, 2002. MMWR Morb Mortal Wkly Rep. 2002, Dec 20; 51(50): 1133-1135.

27. Chambers TJ, Grakoui A, Rice CM. Processing of the yellow fever virus nonstructural polyprotein: a catalytically active NS3 proteinase domain and NS2B are required for cleavage at dibasic sites. J Virol. 1991; 65: 6042-6050.
28. Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organization, expression, and replication. Annu Rev Micobiol. 1990a; 44: 649-688.
29. Chambers TJ, Halevy M, Nestorowicz A, Rice CM, Lustig S. West Nile virus envelope proteins: Nucleotide sequence analysis of strains differing in mouse neuroinvasiveness. J Gen Virol. 1998; 79: 2375-2380.
30. Chambers TJ, McCourt DW, Rice CM. Production of yellow fever virus proteins in infected cells: Identification of discrete polyprotein species and analysis of cleavage kinetics using region-specific polyclonal antisera. Virology. 1990b; 177: 159-174.
31. Chambers TJ, Nestorowicz A, Amberg SM, Rice CM. Mutagenesis of the yellow fever virus NS2B protein. Effects on proteolytic processing, NS2B-NS3 complex formation, and viral replication. J Virol. 1993; 67: 6797-6807.
32. Chambers Tj, Nickells M. Neuroadapted yellow fever virus 17D: Genetic and biological characterization of a highly mouse-neurovirulent virus and its infectious molecular clone. J Virol. 2001; 75(22): 10912-10922.
33. Chambers TJ, Weir RC, Gralcoui A, McCourt DW, Bazon JF, Fletterick RJ, Rice CM. Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. Proc Nat Acad Sci USA. 1990c; 87(22): 8898-8902.
34. Chang GJ, Hunt Ar, Davis B. A singel intramuscular injection of recombinant plasmid DNA includes protective immunity and prevents Japanese encephalitis in mice. J Virol. 2000; 74: 4244-4252.
35. Chang YS, Liap CI, Tsao CH, Chen MC, Liu CI, Chen LK, Lin YL. Membrane permeabilization by small hydrophobic nonstructural proteins of Japanese encephalitis virus. J Virol. 1999; 73(8): 6257-6264.
36. Clum S, Ebner KE, Padmanabhan R. Cotranslational membrane insertion of the serine proteinase precursor NS2B-NS3 (Pro) of dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. J. Biol. Chem. 1997; 272(49): 30715-30723.
37. Coia G, Parker MD, Speight G, Byrne ME, Westaway EG. Nucleotide and complete amino acid sequence of kunjin virus: definitive gene order and characteristics of the virus-specified protein. J Gen Virol. 1988; 61: 1-21.
38. Cover J, Ortiz A, Allison SL, Schalich J, Heinz FX, Wilschut J. Membrane Fusion activity of tick-borne encephalitis and recombinant subviral particles in a liposomal model system. Virology. 2000; 269(1): 37-46.
39. Crowther JR. ELISA Theory and Practice. Methods in molecular biology. Volume 42. 1995. Humana Press Inc. Totowa, New Jersey
40. Davis BS, Chang GJ, Cropp B, *et al*. West nile Virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays. J Virol. 2001; 75:4040-4047.
41. Davis BS, Gwong-Jen JC, Cropp B, Roehrig JT, Martin DA, Mitchell CJ, Bowen R, Bunning ML. West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses In Vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays. J Virol. 2001; 75(9): 4040-4047.
42. Davis CT, Beasley DWC, Guzman H, Siirin M, Parsons RE, Tesh RB, Barrett ADT. Emergence of attenuated West Nile virus variants in Texas, 2003. Virology. 2004; 330:342-350.
43. Despres P, Dietrich J, Girard M, Bouloy M. Recombinant baculoviruses expressing Yellow fever virus E and NS1 proteins elicits protective immunity in mice. J Gen Virol. 1991; 72: 2811-2816.
44. Ding X, Wu X, Duan T, Siirin M, Guzman H, Yang Z, Tesh RB, Xiao S-Y. Nucleotide and amino acid changes in West Nile virus strains exhibiting renal tropism in hamsters. Am J Trop Med Hyg. 2005; 73(4): 803-807.
45. Falgout B and Markoff L. Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum. J Virol. 1995; 69(11): 7234-7243.
46. Falgout B, Bray M, Schlesinger JJ, Lai CJ. Immunization of mice with recombinant vaccinia virus expressing authentic dengue virus nonstructural protein NS1 protects against lethal dengue virus encephalitis. J Virol. 1990; 64(9): 4356-4363.
47. Falgout B, Miller RH, Lai CJ. Deletion analysis of dengue virus type 4 nonstructural protein NS2B: Identification of a domain required for NS2B-NS3 protease activity. J Virol. 1993; 67: 2034-2042.
48. Feinstein S, Abov Y, Lachmi BE, Lehrer S, Rannon L, Katz D. Determination of human IgG and IgM class atibodies to West Nile virus by enzyme linked immunosorbent assay (ELISA). J med Virol. 1985;63-72.
49. Ferlenghi I, Clarke M, Ruttan T, Allison SL, Schalich J, Heinz FX, Harrison SC, Rey FA, Fuller SD. Molecular organization of a recombinant subviral particle from tick-borne encephalitis virus. Mol Cell. 2001; 7(3): 593-602.
50. Freshney RI. Culture of animal cells: a manual of basic technique. 2000; 4th ed. New York, Wiley.
51. Gaunt MW, Sall AA, De Lamballerie X, Falconar AK, Dzhivanian TI, Gould EA. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. J Gen Virol 2001; 82: 1867-1876.

52. Gollins SW and Porterfield JS. Flavivirus infection enhancement in macrophages: An electron microscopic study of viral cellular entry. J Gen Virol. 1985.; 66:1969-1982

53. Gollins SW and Porterfield JS. The uncoating and infectivity of the flavivirus West Nile on interaction with cells: Effects of pH and ammonium chloride. J Gen Virol. 1986; 67: 1941-1950.

54. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov EM. Two related superfamilies of putative helicase involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acid Res. 1989; 17: 4713-4729.

55. Guirakhoo F, Bolin RA, Roehrig JT. The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles andalters the expression of epitopes within the R2 domain of E glycoprotein. Virology. 1992; 191(2):921-931.

56. Guirakhoo F, Heinz FX, Mandl CW, Holzmann H, Kunz C. Fusion activity of flaviviruses: Comparison of mature and immature (prM-containing) tick-borne encephalitis virions. J Gen Virol. 1991; 72(6): 1323-1329.

57. Guirakhoo F, Hunt AR, Lewis JG, Roehrig JT. Selection and partial characterization of dengue 2 virus mutants that induce fusion at elevated pH. Virology. 1993; 194(1):219-223.

58. Haley M, Retter AS, Fowler D, Gea-Banacloche J, O`Grady NP. The role of intravenous immunoglobulin in the treatment of West Nile virus encephalitis. Clin Infect Dis. 2003; 37: e88-e90.

59. Hall TA. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Wndows 95/98/NT. Nucleic acids Symp. Ser. 1999; 41**:** 95-99.

60. Heinz FX, Allison SL. Structure and mechanisms in Flavivirus fusion. Adv. Virus Res. 2000; 55: 231-269.

61. Heinz FX, Stiasny K, Puscher-Auer G, Holzmann H, Allison SL, Mandl CW, Kunz C. Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. Virology. 1994; 198(1):109-117.

62. Heinz FX. Epitope mapping of flavivirus glycoproteins. Adv Virus Res. 1986. 31(1): 103-168.

63. Henchal EA, Nebchal LS, Schlesinger JJ. Synergistic interactions of anti-NS1 monoclonal antibodies protect passively immunized mice from lethal challenge with dengue 2 virus. J Gen Virol. 1988; 69: 2101-2107.

64. Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighing, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22:4673-4680.

65. Holbrook MR, Li L, Suderman MT, Wang H, Barrett AD, The fernch neurotropic vaccine strain of Yellow fever virus accumulates mutations slowly during passage in cell culture. Virus Res. 2000; 69(1): 31-39.

66. Holbrook MR, Ni H, Shope RE, Barrett AD. Amino acid substitution(s) in the stem-anchor region of Langat virus envelope protein attenuates mouse neurovirulence. Virology. 2001; 286(1): 54-61.

67. Hubálek Z, Haluozka J. West Nile fever-a reemerging mosquito-borne viral disease in Europe. Emerg Infect Dis. 1999; 5(5): 643-650.

68. Hunt AR, Cropp CB, Gwong-Jen J, Chang J. A recombinant particule antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen. J Viro Methods. 2001; 97: 133-149.

69. Hurrelbrink RJ, McMinn PC. Attenuation of Murray Valley encephalitis virus by site-directed mutagenesis of the hinge and putative receptor binding regions of the envelope protein. J Virol. 2001; 75(16): 7692-7202.

70. Hurrelbrink RJ, McMinn PC. Molecular determinants of virulence: The structural and functional basis for flavivirus attenuation. Adv Virus Res. 2003; 60**:** 1-38.

71. Johnson AJ, Martin DA, Karabatsos N, Roehrig JT. Detection of anti-arboviral Immunoglobulin G by using a monoclonal antibody-based capture enzyme-linked immunosorbent assay. J Clin Micobiol. 2000; 38: 1827-1831.

72. Jupp PG, Blackburn NK, Thompson DL, Meenehan GM. Sindbis and West Nile virus infections in the Witwatersrand-Pretoria region. S Afr Med J.1986;70:218-220.

73. Jupp PG. The ecology of West Nile virus in South Africa and the occurrence of outbreaks in humans. Ann NY Acad Sci. 2001; 951:142-152.

74. Keller BC, Federicksen BL, Samuel MA, Mock RE, Mason PW, Diamond MS, *et al*. Resistance to alpha/beta interferon is a determinant of West Nile virus replication fitness and virulence. J Virol. 2006; 80: 9424-9434.

75. Khromykh AA, Sedlak PL, Westaway EG. Trans-complementation analysis of the flavivirus kunjin NS5 gene reveals an essential role for translation of its N-terminal half in RNA replication. J Virol. 1999; 73(11): 9247-9255.

76. Khromykh AA, Varnovski AN, Sedlak PL, Westaway EG. Coupling between and packaging of flavivirus RNA. Evidence derived from the use of DNA based full-length cDNA clones of kunjin virus. J Virol. 2001; 75(10): 4633-4640.

77. Khromyth AA, Kenney MT, Westaway EG. Trans-complementation of flavivirus RNA polymerase gene NS5 by using Kunjin virus replicon-expressing BHK cells. J Virol. 1998; 72(9): 7270-7279.

78. Kinney RM, Huang CY-H, Whiteman MC, Bowen RA, Langevin SA, Miller BR, Brault AC. Avain virulence and thermostable replication of the North American strain of West Nile virus. J Gen Virol. 2006; 87:3611-3622.

79. Konishi E and Mason PW. Proper maturation of the Japanese encephalitis virus envelope glycoprotein requires cosynthesis with the premembrane protein. J Virol. 1993; 67(3): 1672-1675.

80. Koonin EV. Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flaviviruses and lambda 2 protein of reovirus. J Gen. Virol. 1993; 74: 733-740.

81. Koonin EV. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol. 1991; 72:2197-2206.

82. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. 2004; 5**:** 150-163.

83. Kümmerer BM, Rice CM. Mutations in the yellow fever virus nonstructural protein NS2A, selectively block production of infectious particles. J Virol. 2002; 76(10): 4773-4784.

84. Kuno G, Chang G-JJ, Tsuchiya KR, Karabatsos N, Cropp CB. Phylogeny of the genus Flavivirus. J of Virol. 1998; 72(1): 73-83.

85. Kuno G. Serodiagnosis of flaviviral infections and vaccinations in humans. Adv Virus Res. 2003; 61:3-65.

86. Lanciotti RS, Ebel GD, Deubel V, Krest AJ, Murri S, Meyer R, Bowen M, McKinney N, Morrill WE, Crabtree MB, Kramer LD, Roehrig JT. Complete genome sequences and phylogenetic analysis of West Nile virus strains isolated from the United States, Europe, and the Middle East. Virology. 2002; 298:96-105.

87. Lanciotti RS, Roehrig JT, Deubel V, Smith J, Parker M, Steele K, Crise B, Volpe KE, Crabtree MB, Scherret JH, Hall RA, MacKenzie JS, Cropp CB, Panigrahy B, Ostlund E, Schmitt B, Malkinson M, Banet C, Weissman J, Komar N, Savage HM, Stone W, McNamara T, Gubler DJ. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. Science. 1999; 286(5448)**:** 2333-2337.

88. Lanciotti RS. Molecular amplification assays for the detection of flaviviruses. Adv Virus Res. 2003; 61: 67-99.

89. Ledizet M, Kar K, Foellmer HG, Wang T, Bushmich SI, Anderson JF, Fikrig E, Koski RA. A recombinant envelope protein vaccine against West Nile virus. Vaccine. 2005; 23:3915-3924.

90. Lee E, Lobigs M. Mechanism of virulence attenuation of glycosaminoglycan-binding variants of Japanese encephalitis virus and Murray Valley encephalitis virus. J Virol. 2002; 76(10): 4901-4911.

91. Lee E, Lobigs M. Substitions at the putative receptor-binding site of an encephalitic flavivirus alters virulence and host cell tropism and reveal a role for glycosaminoglycans in entry. J Virol. 2000; 74(19): 8867-8875.

92. Lee E, Stocks CE, Amberg SM, Rice CM, Lobigs M. Mutagenesis of the signal sequence of Yellow Fever virus prM protein: Enhancement of signalase cleavage In Vitro is lethal for virus production. J Virol. 2000; 74(1): 24-32.

93. Lee E, Weir RC, Dalgarno L. Changes in the dengue virus major envelope protein on passaging and their localization on the three-dimensional structure of the protein. Virology. 1997; 232(2): 281-290.

94. Levinson SS and Miller JJ. Towards a better understanding of heterophile (and the like) antibody interference with modern immunoassays. Clin Chim Acta. 2002; 325: 1-15.

95. Li W, Brinton MA. The 3'stem loop of West Nile virus genomic RNA can suppress translation of chimeric mRNAs. Virology. 2001; 287: 49-61.

96. Lindenbach BD and Rice CM. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. J Virol. 1999; 73: 4611-4621.

97. Lindenbach BD, Rice CM. Molecular biology of flaviviruses. Adv Virus Res. 2003; 59**:** 23-61.

98. Lobigs M. Flavivirus premembrane protein cleavage and spike heterodimer secretion requires the function of the viral proteinase NS3. Proc Natl Acad Sci USA. 1993; 90: 6218-6222.

99. Lorenz IC, Allison SL, Heinz FX, Helenius A. Folding and dimerization of tick-borne encephalitis virus envelope proteins prM and E in the endoplasmic reticulum. J Virol. 2002; 76(11): 5480-5491.

100. Mackenzie JM, Jones MK, Young PR. Immunolocalization of the dengue virus nonstructural glycoptorein NS1 suggests a role in viral RNA replication. Virology. 1996; 220: 232-240.

101. Mackenzie JM, Khromykh AA, Jones MK, Westaway EG. Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. Virology. 1998; 245(2): 203-215.

102. Mandl CW, Allison SL, Holzmann H, Meixner T, Heinz FX. Attenuation of tick-borne encephalitis virus by structure-based site-specific mutagenesis of a putative flavivirus receptor binding site. J Virol. 2000; 74(20): 9601-9609.

103. Mandl CW, Guirakhoo F, Holzmann H, Heinz FX, Kunz C. Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model. J Virol. 1989A; 63(2): 564-571.

104. Mandl CW, Heinz FX, Stockl E, Kunz C. Genome sequence of tick-borne encephalitis virus (western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. Virology. 1989B; 173(1): 291-301.

105. Martin DA, Biggerstaff BJ, Allen B, Johnson AJ, Lanciotti RS, Roehrig JT. Use of Immunoglobulin M cross-reactions in differential diagnosis of human flaviviral encephalitis infections in the United States. Clin Diagn Lab Immunol. 2002; 9:544-549.

106. Martin DA, Muth DA, Brown T, Johnson AJ, Karabatsos N, Roehrig JT. Standardization of Immunoglobulin M capture enzyme-linked immunosorbent assays for routine diagnosis of arboviral infections. J Clin Microbiol. 2000; 38: 1823-1826.

107. Martins TB, Jaskowski TD, Mouritsen CL, Hill HR. An evaluation of the effectiveness of three Immunoglobulin G (igG) removal procedures for routine IgM serological testing. Clin Diagn Lab Immunol. 1995; 2:98-103.

108. Marx F, Gritsun TS, Grubeck-Loebenstein, Gould EA. Diagnostic immunoassays for Tick-borne encephalitis virus based on recombinant baculovirus protein expression. J Virol Meth. 2001; 91:75-84.

109. Mason PW. Maturation of Japanes encephalitis virus glycoproteins produced by infected mammalian and mosquito cells. Virology. 1989; 169(2): 354-364.

110. Matsuura Y, Miyamoto M, Sato T, Morita C, Yasui K. Characterization of Japanese encephalitis virus envelope protein expressed by recombinant baculoviruses. Virology. 1989; 173: 674-682.

111. Matusan AE, Kelley PG, Pryor MJ, Whisstock JC, Davidson AD, Wright PJ. Mutagenesis of the dengue virus type 2 NS3 proteinase and the production of growth-restricted virus. J Gen Virol. 2001a; 82(7): 1647-1656.

112. Matusan AE, Pryor MJ, Davidson AD, Wright PJ. Mutagenesis of the Dengue virus type 2 NS3 protein within and outside helicase motifs: Effects on enzyme activity and virus replication. J Virol. 2001b; 75(20): 9633-9643.

113. McCown J, Cochran M, Putnak R, Feighny R, Burrous J, Henchal E, Hoke C. Protection of mice against lethal Japanese encephalitis with a recombinant baculovirus vaccine. Am J Trop Med Hyg. 1990; 42(5): 491-499.

114. McIntosh B, Jupp P, Dos Santos I, Meenehan G. Epidemics of West NIle and Sindbis viruses in South Africa and Culex (Culex) univittatus. S. Afr J Sci.1976; 72**:** 295-300.

115. McIntosh BM. The epidemiology of arthropod-borne viruses in southern Africa. [dissertation]. Pretoria, South Africa: University of Pretoria; 1980.

116. McMinn PC, Dalgarno L, Weir RC. A comparison of the spread of Murray Valley encephalitis viruses of high and low neuroinvasiveness in the tissues of Swiss mice after peripheral inoculation. Virology. 1996; 220(2): 414-423.

117. McMinn PC, Lee E, Hartley S, Roehrig JT, Dalgarno L, Weir RC. Murray valley encephalitis virus envelope protein antigenic variants with altered hemagglutination properties and reduced neuroinvasiveness in mice. Virology. 1995; 211(1):10-20.

118. McMinn PC, Sammels L. The molecular pathogenesis of flavivirus encephalitis. Microbiol. Aust. 1997; 18(9): A32.

119. McMinn PC. The molecular basis of virulence of the encephalitogenic flavivirus. J Gen Virol. 1997; 78:2711-2722.

120. Monath T. Flaviviruses (yello fever, dengue, dengue hemorrhagic fever, Japanese encephalitis, St. Louis encephalitis, tick-borne encephalitis). In Mandell GL, Bennett JE, Dolin R, editors. Priciples and Practice of Infectious Diseases. New York, Churchill Livingstone Inc. 1995: 1465-1674.

121. Muerhoff AS. Dawson GJ, Dille B, Gutierrez R, Leary TP, Gupta MC, Kyrk CR, Kapoor H, Clark P, Schochetman G, Desai SM. Enzyme-linked immunosorbent assays using recombinant envelope protein expressed in COS-1 and *Drosophila* S2 cells for detection of West Nile virus immunoglobulin M in serum or cerebrospinal fluid. Clin Diagn Lab Immun 2004; 11(4): 651-657.

122. Murray JM, Aaskov JG, Wright PJ. Processing of the dengue virus type 2 proteins prM and C-prM.J. Gen. Virol. 1993; 74: 175-182.

123. Murthy HM, Clum S, Padmanabhan R. Dengue virus NS3 serine protease: Crystal structure and insughts into interaction of the active site with substrates by molecular modeling and structural analysis of mutational effexts. J Biol Chem. 1999; 274(9): 5573-5580.

124. Muylaert IR, Chambers TJ, Galler RG, Rice CM. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: Effects on virus replication and mouse neurovirulence. Virology. 1996; 222: 159-168.

125. Nicholas KB, Nicholas HB Jr, Deerfield DW II. GeneDoc: Analysis and visualization of genetic variation, EMBNEW.NEWS. 1997; 4:14

126. Nowak T and Wengler G. Analysis of di-sulfates present in the membrane proteins of the West Nile flavivirus. Virology. 1987; 156(1): 127-137.

127. Nowak T, Farber PM, Wengler G, Wengler G. Analyzes of the terminal sequences of West Nile virus structural proteins and of the in vitro translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavage involved in their synthesis.Virology. 1989; 169: 365-76.

128. Palmer DF, Whaley SD. Complement fixation test. In Rose NR, Friedman H, Fahey JL, editors. Manual of Clinical Laboratory Immunology (third edition), Washington, DC: American Society for Microbiology, 1986: 57-66.

129. Parida M, Posadas G, Inoue S, Hasebe F, Morita K. Real-Time reverse transcription loop-mediated isothermal amplification for rapid detection of West Nile virus. J Clin Microbiol. 2004; 42(1):257-263.

130. Peretsen LR, Roehrig JT. West Nile Virus: A reemerging global pathogen. Emerg Infect Dis. 2001; 7(4): 611-614.

131. Petersen LR, Marfin AA. West Nile virus: a primer for the clinician. Ann Int Med. 2002; 137: 173-179.

132. Plentev AG, Bray M, Huggins J, Lai CJ. Construction and Characterization of chimeric tick-borne encephalitis/dengue type 4 viruses. Proc Nat Acad Sci USA. 1992; 89(21): 10532-10536.

133. Pletnev AG, Bray M, Lai CJ. Chimeric tick-borne encephalitis and dengue type 4 viruses: Effects of mutations on neurovirulence in mice. J Virol. 1993; 67(8): 4956-4963.

134. Prince HE, Hogrefe WR. Assays for detection West Nile virus antibodies in human serum, plasma, and cerobrospinal fluid. Clin App Immun Rev. 2005; 5:45-63.

135. Prince HE, Hogrefe WR. Detection of West Nile Virus (WNV)-specific Immunoglobulin M in a reference laboratory setting during the 2002 WNV season in the United States. Clin Diagn Lab Immunol. 2003a; 10:764-768.

136. Prince HE, Hogrefe WR. Performance characteristics of an in-house assay system used to detect West Nile virus (WNV)-specific Immunoglobulin M during the 2001 WNV season in the United States. Clin Diagn Lab Immunol. 2003b; 10:177-179.

137. Puig-Basagoiti F, Tilgner M, Bennett CJ, Yangsheng Z,Munoz-Jordan JL, Garcia-Sastre A, Bernard KA, Shi P-Y. A mouse cell-adapted NS4B mutation attenuates West Nile virus RNA synthesis. Virology. 2007; 361:229-241.

138. Qiao M, Ashok M, Bernard KA, Palacios G, Hong Zhou Z, Lipkin WI, Lianq TJ. Induction of sterilizing immunity against West Nile virus, by immunization with WNV-like particles produced in insect cells. The J Infec Dis. 2004; 190: 2104-2108.

139. Rey FA, Heinz FX, Mandl C, Kunz C, Harrison SC. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. Nature. 1995; 375(6529): 291-298.

140. Rice CM, Lenches EM, Eddy SR, Shin SJ, Sheets RL, Strauss JH. Nucleotide sequence of yellow fever virus: Implications for flavivirus gene expression and evolution. Science. 1985; 229: 726-733.

141. Rice CM, Strauss EG, Strauss JH. Structure of the flavivirus genome. In *The Togaviridae and Flaviviridae*. 1986; 279-326. Edited by S Schlesinger and MJ Schlesinger. New York: Plenum.

142. Rice CM. *Flaviviridae*: the viruses and their replication. In *Fields Virology*, 3rd edn, 1996; 931-959. Edited by BN Fields, DM Knipe, PM Howley. Philadelphia: Lippincott-Raven.

143. Ryan MD, Monaghan S, Flint M. Virus-encoded proteinases of the *flaviviridae*. J Gen Virol. 1998; 79(5): 947-959.

144. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual, second edition. Plainview, New York: Cold Spring Harbor Laboratory Press. 1989.

145. Shi P-Y , Wong SJ. Serologic diagnosis of West Nile virus infection. Expert Rev Mol Diagn. 2003, 2: 733-741.

146. Shurtleff AC, Beasley DX, Chen JJ, Ni H, Suderman MT, Wang H, *et al*. Genetic variation in the 3' non-coding region of dengue viruses. Virology. 2001; 281: 75-87.

147. Smithburn KC, Hughes TP, Burke AW, Paul JH. A neurotropic virus isolated from the blood of a native of Ugandan. Am J Trop Med. 1940; 20: 471-492.

148. Southern PJ, Berg P. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J Molec Appl Gen. 1982; 1:327-339.

149. Stadler K, Allison SL, Schalich J, Heinz FX. Proteolytic activaton of tick-borne encephalitis virus by furin. J. Virol. 1997; 71: 8475-8481.

150. Stocks CE, Lobigs M. Signal peptidase cleavage at the flavivirus C-prM junction: Dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. J Virol. 1998; 72(3): 2141-2149.

151. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuclei Acids Res. 1994; 22:4673-4680.

152. Valle RPC, Falgout P. Mutagenesis of the NS3 Protease of dengue virus type 2. J Virol. 1998; 72(1): 624-632.

153. Venter M, Myers TG, Wilson MA, Kindt TJ, Paweska JT, Burt FJ, Leman PA, Swanepoel R. Gene expression in mice infected with West Nile virus strains of different neurovirulence. Virology. 2005; 342(1): 119-140.

154. Vijay P, Bondre RS, Jadi AC, Mishra PN, Yergolkar, Arankalli VA. West Nile virus isolates from India: evidence for a distinct genetic lineage. J Gen Virol. 2007; 88: 875-884

155. Von Heijne G. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 1986; 14(11), 4683-4690.

156. Wang E, Weaver SC, Shope RE, Tesh RB, Watts DM, Barrett ADT. Genetic variation in yellow fever virus: duplication in the 3` non-coding region of strains from Africa. Virology. 1996; 77: 1035-1042.

157. Wang T, Magnarelli LA, Anderson JF, Hannah Gould L, Bushmich SL, Wong SJ, Fikrig E. A recombinant envelope protein-based enzyme-linked immunosorbent assay for West Nile virus serodiagnosis. Vector Borne Zoonotic Dis. 2002; 2: 105-109.

158. Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA. Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. Nat Med. 2004; 10(12):1366-1373.

159. Wengler G, Wengler G, Gross HJ. Studies on virus-specific nucleic acids synthesized in vertebrate and mosquito cells infected with flaviviruses. Virology. 1978; 89: 423-437.

160. Wengler G, Wengler G. Cell-associated West Nile flavivirus is covered with envelope and premembrane protein heterodimers, which are destroyed and reorganized by proteolytic cleavage during virus release. J Virol. 1989; 63(6): 2521-2526.

161. Wicker JA, Whiteman MC, Beasley DWC, Davis CT, Zhang S, Schneider BS, Higgs S, Kinney RM, Barrett, ADT. A single amino acid substitution in the central portion of the West Nile virus NS4B protein confers a highly attenuated phenotype in mice. Virology. 2006; 349:245-253.

162. Winkler G, Maxwell SE, Ruemmler C, Stollar V. Newly synthesized dengue-2 virus nonstructural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. Virology. 1989; 171(1):302-305.

163. Winkler G, Randolph VB, Cleaves GR, Ryan TE, Stollar V. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. Virology. 1988; 162(1): 187-196.

164. Wu S-C, Lin Y-J, Yu C-H. Baculovirus-insect cell expression, purification, and immunological studies of the full-length Japanese encephalitis virus envelope protein. Enzyme Microb Tech. 2003; 33: 438-444.

165. Wu SJ, Grouard-Vogel G, Sun W, Mascola JR, Biachtel E, Putvatana R, Louder MK, Filgueira L, Marovich MA, Wang HK, Blauvelt A, Murphy GS, Robb ML, Innes BL, Birx DL, Hayes CG, Frankel SS. Human skin Langerhans cells are targets of dengue virus infection. Nature Med. 2000; 6(7): 816-820.

166. Yamshchikov G, Borisevich V, Seregin A, Chaporgina E, Mishina M, Mishin V, Kwok CW, Yamshchikov V. An attenuated West Nile prototype virus is highly immunogenic and protects against the deadly NY99 strain: a candidate for live WNV vaccine development. Virology. 2004; 330(1): 304-312.

167. Yamshchikov VF and Compans RW. Processing of the intracellular form of the West Nile virus capsid protein by the viral NS2B-NS3 ptotease: An in vitro study. J virol. 1994; 68(9): 5765-5771.

168. Yamshchikov VF and Compans RW. Regulation of the late events in flavivirus processing and maturation. Virology. 1993;192:38-54.

169. Yoshii K, Hayasaka D, Goto A, Obara M, Araki K, Yoshimatsu K, Arikawa J, Ivanov L, Mizutani T, Kariwa H, Takashima I. Enzyme-linked immunosorbent assay using recombinant antigens expressed in mammalian cells for serodiagnosis of Tick-borne encephalitis. J Virol Meth. 2003; 108:171-179.

170. Zhou Y, Ray D, Zhao Y, Dong H, Ren S, Li Z, Guo Y, Bernard KA, Shi P-Y, Li H. Structure and function of flavivirus NS5 methyltransferase. J Virol. 2007; 81(8): 3891-3903

Publications

- Article was submitted to Emerging Infectious Diseases on 10 April 2007. Submission number EID-07-0457. Article was accepted on the 9th of October 2007.
- Emerging Infectious Diseases. 2008. February, Volume 14, Issue 2, p222-230. **Genetic Determinants of Virulence in Pathogenic Lineage 2 strains of West Nile Virus.** E.M. Botha[1], W. Markotter[1], M. Wolfaardt[2], J.T. Paweska[3], R. Swanepoel[3], G. Palacios[4], L.H. Nel[1], M. Venter[2]*